RiverZoo: A Citizen Science and Fractal Analysis Approach to Classifying Interplanetary Drainage Networks

Mariarca D'Aniello[1,2], Maria Raffaella Zampella[3], Andrea Dosi[3], Alvi Rownok[3], Michele Delli Veneri[4], Adriano Ettari[3], Stefano Cavuoti[2,4], Luca Sannino[3], Massimo Brescia[2,3,4], Carlo Donadio[1,5], and Giuseppe Longo[3]

[1] Dipartimento di Scienze della Terra, dell'Ambiente e delle Risorse, Università degli Studi di Napoli Federico II, Naples, Italy

[2] Istituto Nazionale di Astrofisica, Osservatorio Astronomico di Capodimonte, Naples, Italy

[3] Dipartimento di Fisica "Ettore Pancini", Università degli Studi di Napoli Federico II, Naples, Italy

[4] Istituto Nazionale di Fisica Nucleare, Naples, Italy

[5] Stazione Zoologica Anton Dohrn, Naples, Italy

EUROPLANET SCIENCE CONGRESS EPSC2024

INAF ISTITUTO NAZIONALE DI ASTROFISICA

Introduction and objectives

1. RiverZoo survey

- Improve the understanding of terrestrial and extraterrestrial drainage networks through statistical analysis
- Assess the classification accuracy of fluvial networks by human experts
- Broaden the current binary classification (dendritic and non-dendritic) to ten classes, considering all the below (fig.1). The refining of the dataset, including more ground truth, will be beneficial for an **automatic approach to river classification**, using Deep Learning (DL) models

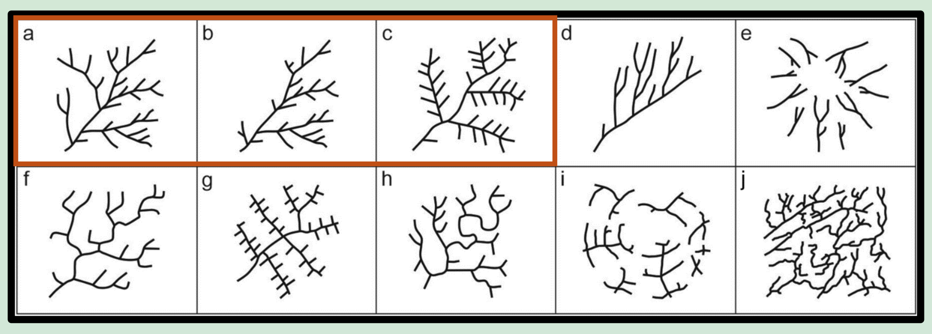


Figure 1 – Different classes of drainage patterns: a) dendritic; b) sub-dendritic; c) pinnate; d) parallel; e) radial; f) rectangular; g) trellis; h) angular; i) annular; j) contorted. (a)–(c) patterns are related to dendritic forms (D), (d)–(j) to non-dendritic ones (ND) [1].

Anyone interested in partecipating at the survey to improve our test campaign, please scan the QR code!

PLANETARY DRAINAGE PATTERNS

- Earth: mainly dendritic
- Mars: dendritic, sub-dendritic, radial, parallel
- Venus: radial, annular, parallel, trellis (volcanic features)
- Titan: rectangular, dendritic and parallel (methane/ethane rivers)

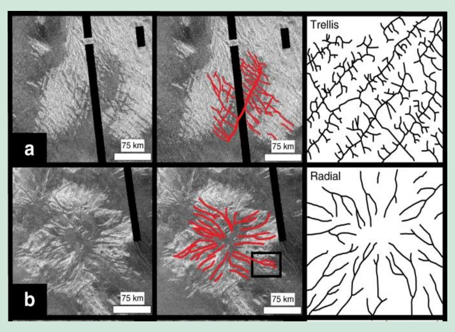


Figure 2 – Example of Venus stream drainage patterns.

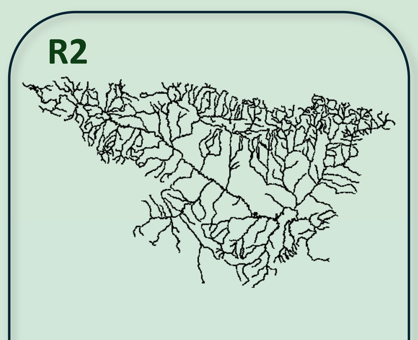
Theoretical background

2. Fractal Dimension (FD) to objectively classify rivers

R1

Indus river - Pakistan
FD ~ 1.38

R2



Ebro river - Spain
FD ~ 1.64

In Maths…

FD quantifies an object's complexity with a decimal value, where higher FD indicates greater complexity. Fractal Dimensions are essential for describing intricate structures beyond the limits of Euclidean geometry, which instead only uses whole numbers[2].

In Planetary Geomorphology…

Fractal Dimension (FD) shows how rivers and drainage networks branch. Comparing FDs over time helps track their history, while the whole and decimal parts of FD can reveal what processes the river has gone through.

In River Analysis…

FD indicates the erosive nature of terrain, with higher values reflecting more complex river branching in softer, easily erodible terrain (dendritic, sub-dendritic), and lower FDs in harder terrains leading to simpler networks (parallel, angular, contorted). Volcanic landscapes can create radial or annular drainage patterns, often seen on planetary surfaces like Venus.
For 2D river images, $1 \leq FD \leq 2$ always holds.

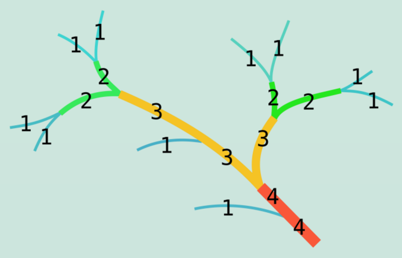
$FD_{R2} > FD_{R1}$ ➡ Ebro is a lot messier

Methodology

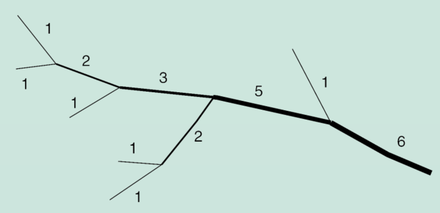
3. Machine Learning (ML) to cluster rivers

A Clustering Approach has been chosen to classify rivers, by considering Fractal dimension (FD) and two morphometric indexes:

- **Strahler Number (St):** represents the hierarchical branching structure of a river, increasing only when branches of equal order meet[3].



- **Shreve Number (Sh):** measures the total connectivity of a river by summing the contributions of all tributaries at each confluence [4].



Using a 3D clustering with the triple $(FD, \frac{St}{Sh}, \frac{Sh}{St})$, rivers are grouped by analyzing their intrinsic properties from a mathematical standpoint.

Two methods to classyfing based on Machine Learning:

1. Deterministic (**K-Means**): *hard clustering* algorithm - each point is assigned to the nearest cluster (class of river), considering Eucledian distance

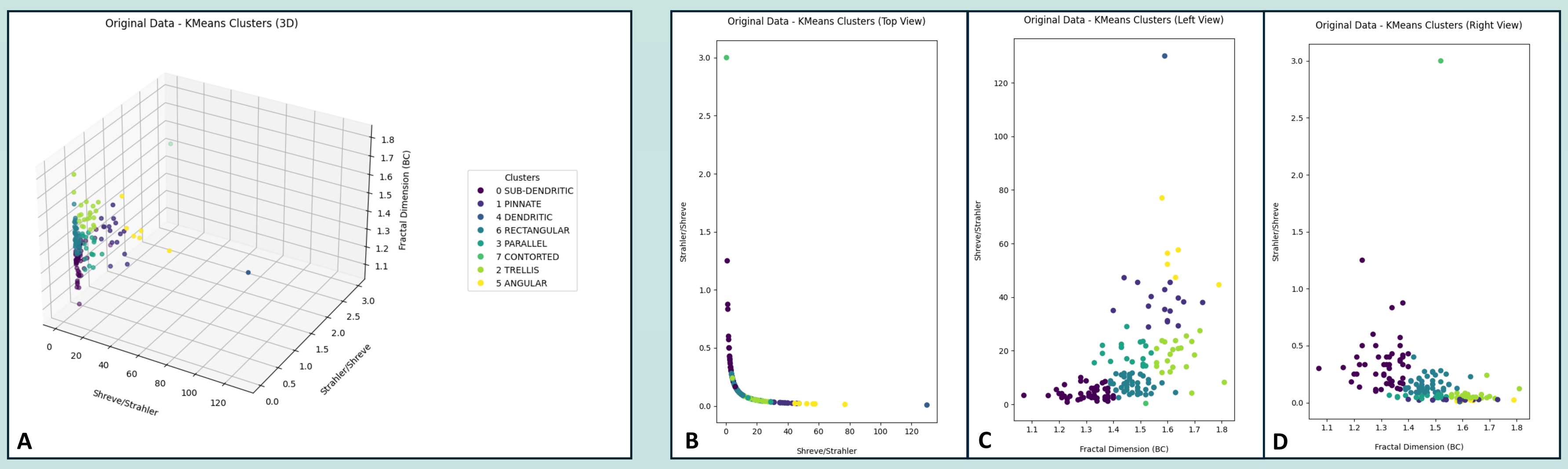


Figure 3.**A** is a 3D representation of the k-means clustering output. **B**, **C** and **D** are the orthogonal 2D projections considering the three axis.

2. Probabilistic (**Fuzzy C-Means**): *soft clustering* algorithm - for each iteration, each point is assigned a probability of belonging to each of the *C* clusters [5]

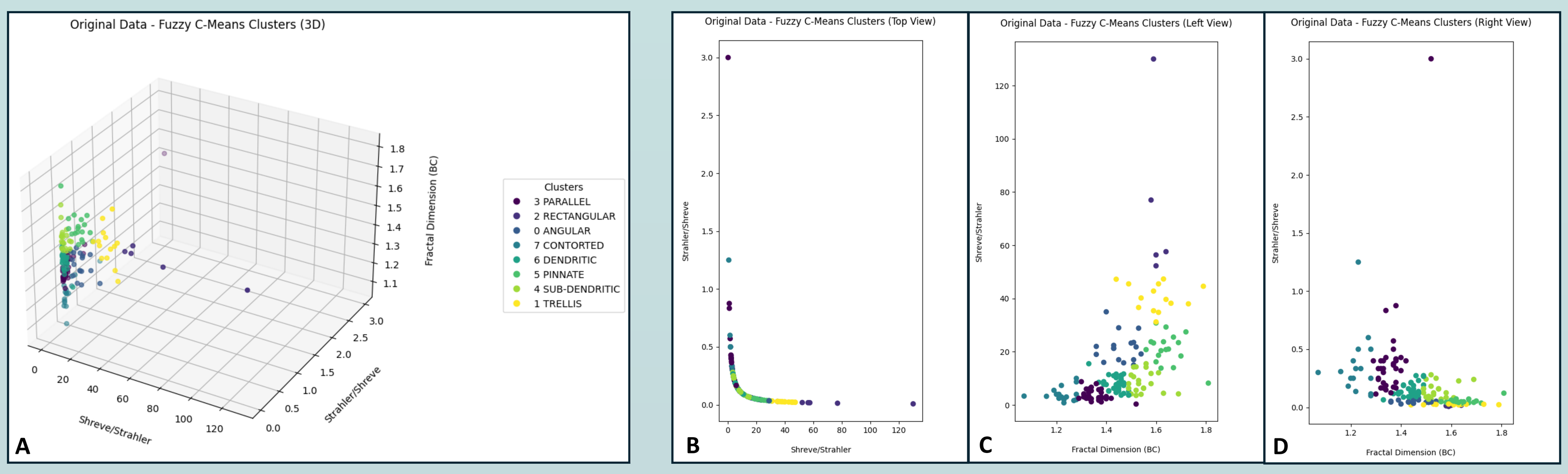


Figure 4. **A** is a 3D representation of the fuzzy clustering output. **B**, **C** and **D** are the orthogonal 2D projections considering the three axis.

Results and discussion

4. Comparison between users' responses and clustering methods

	K-Means and FCM	K-Means and users	FCM and users
Correlation (Pearson's r)	83% ⬆	17% ⬇	14% ⬇
Co-variance p.	3.25	0.78	0.61
Correlation (Cohen's k)	69%	17%	14%

- Reliability of the two methods provide a valid ground truth for classification
- Users and methods disagree, suggesting that users percieve features differently than mathematical approaches do

The clustering methods **k-means** and **fuzzy** have a **strong correlation** between them and low correlation compared to River Zoo users' responses

The strong correlation means that the two methods are reliable and could be taken as ground truth, while the survey answers cannot for two main reasons:

1. There are too few users who responded, compared to the many users of the Galaxy Zoo survey that has been taken as inspiration
2. The answers are very inconsistent with each other

Conclusions

The users were not experts and were not very confident with the morphologies. Additionally, the morphologies themselves are sometimes so similar that they tend to confuse the users, suggesting several improvements for a next project:

- The survey should include questions about the user's expertise, to weigh their responses and refine the confidence metric
- Get more rivers outlines to improve the clustering
- Increase the number of participants to the survey

Finally, an **automatic**, **reliable** and **effective** classification framework has been developed and refined! With more rivers, it will serve as a **ground truth** for the next phases of Earth and space exploration!

TO BE CONTINUED… WITH RIVER ZOO 2!

References

1. Donadio et al. (2021). Scientific Reports 11, 5875.
2. Mandelbrot (1982). W.H. Freeman. ISBN 0-7167-1186-9.
3. Strahler, A. N. (1952). Geological Society of America Bulletin, 63 (11): 1117–1142.
4. Shreve, R., (1966). J. Geol., 74, 17–37.
5. Velmurugan T. (2014). Applied Soft Computing, 19, 134-146.

Info

mariarca.daniello@unina.it
mariarca.daniello@inaf.it