Icarus 234 (2014) 109–131

Contents lists available at ScienceDirect

Icarus

journal homepage: www.elsevier.com/locate/icarus

The variability of crater identification among expert and community crater analysts

Stuart J. Robbins [a,*], Irene Antonenko [b,c], Michelle R. Kirchoff [d], Clark R. Chapman [d], Caleb I. Fassett [e], Robert R. Herrick [f], Kelsi Singer [g], Michael Zanetti [g], Cory Lehan [h], Di Huang [h], Pamela L. Gay [h]

[a] Laboratory for Atmospheric and Space Physics, University of Colorado at Boulder, 3665 Discovery Dr., Boulder, CO 80309, United States
[b] Planetary Institute of Toronto, 197 Fairview Ave., Toronto, ON M6P 3A6, Canada
[c] University of Western Ontario, 1151 Richmond Street N., London, ON N6A 5B7, Canada
[d] Southwest Research Institute, 1050 Walnut Street, Suite 300, Boulder, CO 80302, United States
[e] Department of Astronomy, Mount Holyoke College, 50 College Street, South Hadley, MA 01075, United States
[f] Geophysical Institute, University of Alaska Fairbanks, Fairbanks, AK 99775-7320, United States
[g] Department of Earth and Planetary Sciences, McDonnell Center for the Space Sciences, Washington University in St. Louis, 1 Brookings Dr., Saint Louis, MO 63130, United States
[h] The Center for STEM Research, Education, and Outreach at Southern Illinois University, Edwardsville, Edwardsville, IL 62025, United States

ARTICLE INFO

Article history:
Received 15 June 2013
Revised 20 February 2014
Accepted 23 February 2014
Available online 4 March 2014

Keywords:
Cratering
Impact processes
Moon
Moon, surface

ABSTRACT

The identification of impact craters on planetary surfaces provides important information about their geological history. Most studies have relied on individual analysts who map and identify craters and interpret crater statistics. However, little work has been done to determine how the counts vary as a function of technique, terrain, or between researchers. Furthermore, several novel internet-based projects ask volunteers with little to no training to identify craters, and it was unclear how their results compare against the typical professional researcher. To better understand the variation among experts and to compare with volunteers, eight professional researchers have identified impact features in two separate regions of the Moon. Small craters (diameters ranging from 10 m to 500 m) were measured on a lunar mare region and larger craters (100s m to a few km in diameter) were measured on both lunar highlands and maria. Volunteer data were collected for the small craters on the mare. Our comparison shows that the level of agreement among experts depends on crater diameter, number of craters per diameter bin, and terrain type, with differences of up to ∼±45%. We also found artifacts near the minimum crater diameter that was studied. These results indicate that caution must be used in most cases when interpreting small variations in crater size–frequency distributions and for craters $\lesssim$10 pixels across. Because of the natural variability found, projects that emphasize many people identifying craters on the same area and using a consensus result are likely to yield the most consistent and robust information.

1. Introduction

Impact craters are among the most common and numerous features on planetary surfaces in the Solar System. They have been used for decades in various studies, from understanding the dynamics of the Solar System to being a "poor man's drill" by excavating through numerous rock layers. This research relies on a key assumption: Impact craters can be reliably identified. Many applications, especially age estimates (McGill, 1977), also rely on measurements of crater diameter. It is generally assumed that both identification and measurement are trivial, but limited studies have shown this to not always be true; variations in crater identification and diameter measurement on the order of ∼10% between individuals using the same measuring technique have been found (e.g., Gault, 1970; Greeley and Gault, 1970; Kirchoff et al., 2011; Hiesinger et al., 2012).

Gault (1970) had approximately 20 people identify 1.3 million craters using Zeiss particle counters (this device allows the operator to match a pre-set circle size or "size class" of projected light onto a photograph, prick a hole through the photograph at the crater center, and the diameter is automatically registered on the instrument's display). He concluded, "'Calibration' and continuous cross-checks of each individual's work indicate that crater counts by different persons generally agree and/or can be reported within ±20%..." Greeley and Gault (1970) used the same technique and data to further describe the dispersion among researchers.

* Corresponding author.
E-mail address: stuart.robbins@colorado.edu (S.J. Robbins).

http://dx.doi.org/10.1016/j.icarus.2014.02.022
0019-1035/

Measurements were made by five individuals on a single image and showed good agreement for small craters but dispersion in the number of craters among the largest diameters (their Fig. 3). Greeley and Gault (1970) found less than ±20% deviation from the mean for counts of >100 craters in a given size class, "a value that probably represents an irreducible minimum deviation imposed by the subjectivity of the counting." This variation rose to ±100% for counts with <4–5 craters in a given size class. The authors emphasized that a single individual may perform more consistent counts, but individual biases and differences from one day to the next – indeed, one hour to the next – explain why multiple individuals identifying craters on the same terrain are likely to yield the most reliable results.

Kirchoff et al. (2011) provide a more recent comparison with three researchers (two expert, one novice without crater counting experience) from the same lab who used the same technique to identify, measure, and, in this case, classify craters by preservation state. They used *Lunar Reconnaissance Orbiter* Camera Wide-Angle Camera (LROC WAC) images of Mare Orientale. The two experienced analysts had counts that differed by 20–40% in a given diameter range, while the novice counter identified numerous features that are probably not craters, differing from the other two by >100% over some diameter ranges. They also had significant variation among the preservation states attributed to each crater, despite a relatively coarse four-point scale. This work showed that despite common thinking that crater counting is fairly easy and straightforward, there is a learning curve and an individual's crater counts should be discarded during the learning process. It also showed that even well defined crater morphologies may be difficult to classify uniformly.

Hiesinger et al. (2012) also focused on lunar craters, in their case using LROC Narrow-Angle Camera (NAC) images at approximately 0.5 m/px. They were interested in reproducible results for better understanding the lunar cratering flux and performed a single test with two experienced researchers who used the same technique on the same image. The Hiesinger et al. (2012) team found an overall variation of only ±2% between their analysts, a dispersion significantly less than previous work.

What this brief review indicates is that while there has been some discussion in the literature about agreement between different researchers' crater identifications, (a) there has been no thorough discussion on researcher variability, (b) no published study discusses the variability when using different techniques for crater identification and measurement, (c) variation in crater morphology has not been discussed (e.g., sub-km craters appear substantially different at NAC pixel scales when compared with multi-km craters at WAC pixels scales), and (d) expert results have not been extensively compared with how well untrained or minimally trained crater counters do with the identification and measurement process. Given the proliferation of internet crowd-sourcing projects that ask laypeople to help in the data-gathering process, this last point determines if the public can assist in crater counting and produce results that are approximately as reliable as the experts.

For these reasons, the following work was undertaken. Eight researchers, with six to fifty years of experience identifying craters, identified and measured the diameter (D) of craters on a segment of a NAC image. This same region was also analyzed by volunteer "citizen scientists" through CosmoQuest's Moon Mappers ("MM") project which facilitates volunteer identification and measurement of craters and other features that are being studied in a variety of lunar research projects (Robbins et al., 2012; http://cosmoquest.org). In addition, the experts identified and measured craters on a WAC image that covers both lunar mare and highlands. The experts worked independently, with each researcher using their own preferred technique (in total, seven different methods were employed).

These methods and the counting locations are discussed in Section 2 along with terminology, our display techniques, and statistical tests in Section 2.4. Section 3 describes steps taken to ensure that our analysis is based solely on how different people identify craters, including how experts varied (Section 3.1), how well the volunteers compared with experts (Sections 3.2 and 3.3), and how our data reduction process may affect results (Sections 3.4 and 3.5). Section 4 describes the overall crater populations found in each image and the variation among experts and between experts and volunteers. Section 5 moves from the population of craters from Section 4 to how well experts and volunteers agreed on the measurements of individual craters. Section 6 is an analysis of how crater detection depended on preservation state. Section 7 describes artifacts that we found near the minimum crater diameter. Section 8 is a short discussion of likely reasons for differences among experts and between them and volunteers. Section 9 summarizes the work and discusses implications and conclusions. Appendix A provides additional details on each researcher's technique, Appendix B summarizes each researcher's experience in the field, and Appendix C more thoroughly discusses our data reduction methods.

2. Methodology

2.1. Images used

This work was motivated in part by the need to determine how untrained "citizen scientists" compare with experts. The experts in this study were asked to identify and measure craters in the same LROC NAC image that has been most studied by MM volunteers: M146959973L (Fig. 1). A portion of this image has been viewed by every MM volunteer because it is used as a calibration image to assess how well each individual performs. The image has a solar incidence angle of 77°, meaning that useful shadows are present to enhance local topography for better feature identification. It is also of general interest because it contains the *Apollo 15* landing site. MM uses a 4107 × 6651-pixel sub-image of M146959973L that is centered on the *Falcon* lander. The experts in this study were given a sub-image 33% of this size (Fig. 1), 4107 × 2218 pixels, to maximize participation among busy scientists. This sub-image contains on the order of 1000 craters $D \geqslant 18$ pixels (the limit of the MM interface).

The second image in this study was an LROC WAC that encompasses both mare and highland areas to allow comparison of expert crater identification on the two main lunar terrain types. The 1311 × 2802-pixel selected portion of WAC M119455712M covers the southern margin of Mare Crisium and the neighboring rim/highlands to the south (Fig. 2). It has a solar incidence angle of 59°, which is on the boundary of what is considered ideal for crater identification (Wilcox et al., 2005; Ostrach et al., 2011; Robbins et al., 2012).

Each image was downloaded from the Planetary Data Systems (PDS), processed in the USGS's *Integrated Software for Imagers and Spectrometers* (*ISIS*) via standard radiometric and geometric techniques, projected to a local Mercator projection, exported as PNG files, and distributed to each researcher. For several of the crater identification and measurement techniques, GIS-ready files were required. To ensure uniformity, Robbins imported the images into *ArcMap* and exported the GIS-ready files, and he distributed them to the researchers, too.

2.2. Techniques and personnel

Each researcher employed one or more different interfaces and methods, with Robbins using two interfaces and Antonenko using