MOON ZOO CITIZEN SCIENCE PROJECT - ANALYSIS, EVALUATION, AND CONCLUSIONS. R. Bugiolacchi[1,2], S. Bamford [3], P. Tar [4], K. Joy [4,2], I. Crawford [1,2], P. Grindrod [1,2], N. Thacker [4], C.J. Lintott [5,6], and the MoonZoo team. [1]Birkbeck, University of London, roberto.bugiolacchi@ucl.ac.uk; [2]CPS, UCL, London, UK; [3]Un. Of Nottingham, UK; [4]Un. Of Manchester, UK; [5]Oxford, Uk; [6]Adler Planetarium, USA.

Introduction: Moon Zoo is part of the suite of Zooniverse citizen science projects [1], which enlist thousands of science enthusiasts around the world carry out large-scale mapping and cataloguing of astronomical phenomena. The aim of this study is to gain a level of confidence in the Moon Zoo citizen science data to generate reliable crater size-frequency distributions across the lunar surface. We also test the validity of interpreting crater size spread among users as an index of crater erosion, and by implication, age. This work focuses on the statistical analysis of small (<500 m diameter) impact craters surveyed around and beyond the Apollo 17 landing region. In order to assess the reliability of the Moon Zoo output against a 'ground truth', an expert crater survey was carried out. A subset was marked by three other planetary scientists for validation of the larger set. We developed a new method to coalesce crater data annotations (lat, long, radius) from several non-projected, uncalibrated NAC [2] images into single, map-projected entries. Further, we considered the input behavioural pattern of each user in order to allocate individual 'confidence' weighting parameters.

Method and data:

Expert Count Validation. An expert count was carried out by the lead author (RB) on similar left and right NAC image pairs M104311715 and M104318871 for a region of ~400 km^2 in size (Fig. 1). This inventory produced 2,607 craters. In order to assess the accuracy of RB's counting output, we enrolled the help of three volunteer planetary scientists in counting craters in a subset region of 11.62 km^2 (Fig. 1, inset).

Data Filtering. 9,321 Moon Zoo users contributed to generate around 130,000 crater entries. 56% of the volonteers marked fewer than 10 craters, and 91% <30. This shows a low commitment rate by a large proportion of volunteers, including a 17% fraction who only marked one crater, at least based on these four NAC images. It is not unreasonable to question the quality and validity of the entries generated by these citizen scientists. In order to minimise data 'contamination' from unreliable users we applied a simple behavioural threshold: we eliminated all crater data from those users who marked as default sizes more than 50% (P50) or 25% (P25) of their total output. The results are significant: the filtering produces reductions of the number of (pre-clustering) annotations of 81% (P25) and 92% (P50).

Clustering Method. We have also developed a Likelihood-based approach to clustering [3] utilising knowledge of the measurement errors on annotated x, y and diameter parameters. The implementation is very similar to a circular Hough transform, where an x, y and diameter parameter space is populated with Moon Zoo annotations, before being smoothed with a Gaussian with width proportional to the annotation errors (which were measured to be approximately 10% of crater diameter). The smoothing has the effect of coalescing closely adjacent annotations into individual peaks, whilst preserving isolated annotations as single peaks. Each peak in this space is interpreted as an individual candidate crater, with the height of each peak be proportional to the number of annotations around that location (we use the prefix 'M-' in this work).

Crater degradation. This method measures degradation by matching a crater image template to candidate craters using varied levels of image smoothing. An average crater appearance is computed using a selection of verified Moon Zoo craters, with mean illumination subjected to minimise effects of albedo. This template is compared to candidate craters using a normalised dot-product match score. The amount of smoothing required to get the best match between a crater and this template can then be correlated with degradation, as the gradual erosion of craters mimics the appearance of a smoothed image. 16 logarithmic smoothing levels, corresponding to absolute smoothing between 0.1 to 1.9 pixels, were applied.

Results:

Crater Census. Selected resultant Cumulative Crater Frequency curves from (user) filtered data sets are shown in Fig. 2. Disqualifying annotations from default-size centric users improves on the statistical representation distribution of craters, and it tends to rein in the artificial sub-30 m and 120 m diameter peaks.

Crater Degradation. The degradation index broadly correlates with expert degradation classifications (Fig. 3), at least to within a relative ordering. The highest correlations observed ('M-P50') also achieved a relatively linear relationship with the expert. Discrepancies between expert classifications and indices grow at the extremes of the spectrum where the least and most eroded craters are found. Since statistically significant error bars are difficult to derive from subjectively selected classes, we can only speculate that a 10% variation between models to be an acceptable margin of uncertainty and good agreement.

Further Improvements and Work. a] Fixing an alternative minimum crater size approach, to minimise the overuse of the default crater size; b] improve user training; c] only use map-projected base images; d] develop better error analysis in order to prevent over-interpretation.

Conclusions: in this work we have (1) validated a standard expert crater count dataset to use as comparison with Moon Zoo; (2) tested filtering of spoilers/bad data based on users' behavioural pattern in relation to crater default size annotations; (3) compared two different mathematical approaches in clustering multiple crater entries, one developed specifically for the Moon Zoo project; (4) derived and compared crater degradation indexes based on the spread of annotation parameters and smoothing of imaging.

We conclude that the 'in-house' crater clustering and analysis approach ('M-') holds much promise including a derivation of relative erosion classes, especially when combined with prior exclusion by filtering of less reliable users' data.

References:

[1] Lintott, C.J. et al. (2008) *Mon. Not. R. Astron. Soc.* 389, 1179-1189. [2] Robinson, M.S. et al. (2005) *LPS XXXVI*, abstract #1576. [3] Tar, P.D. and Thacker, N.A. and The Moon Zoo *Team. Europ. Planet. Sci. Cong.* Vol. 9, EPSC2014-4.

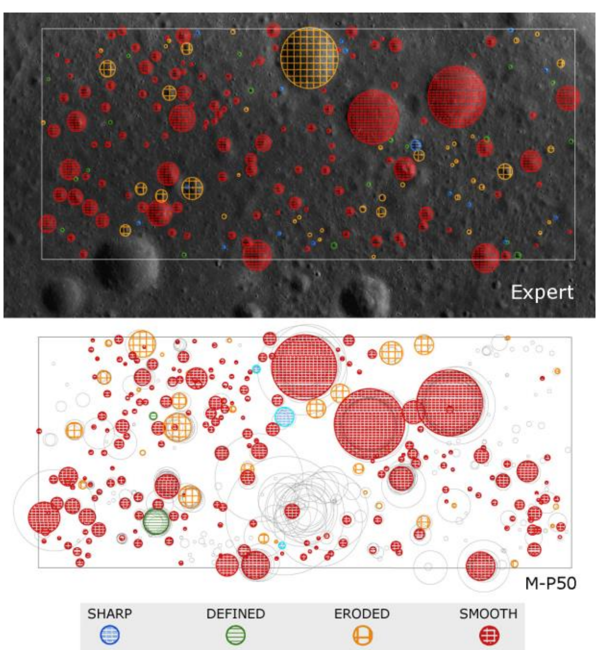

Figure 3. Expert crater survey is mapped including a qualitatitve classification of erosion state.

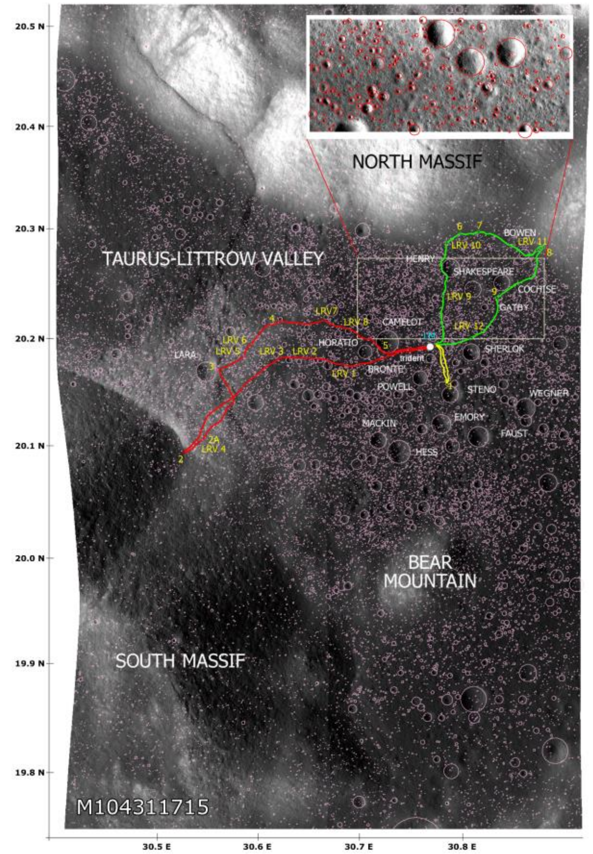

Figure 1. RB crater annotations. Superimposed Apollo 17 landing site and exploration map. Inset shows area used for craer survey comparison.

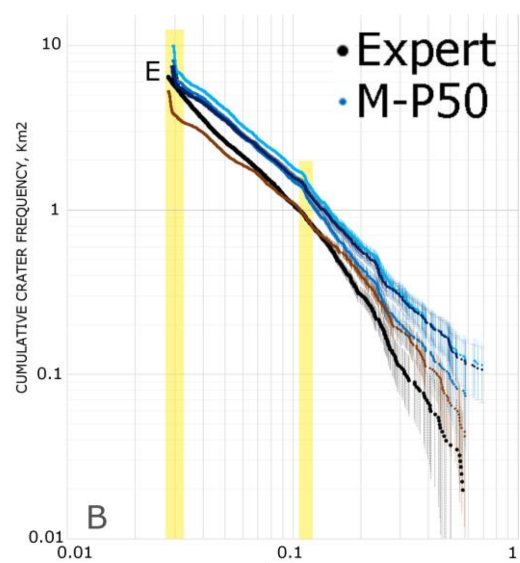

Figure 2. Cumulative Crater Frequency of Moon Zoo annotations, M-P50 (exclusion of users with >50% default size annotations). Black line is Expert (RB) count. Yellow lines highlight default minimum sizes at different zoom levels.