A CRITIQUE OF METHODS FOR ANALYSIS OF CRATER SIZE-FREQUENCY DISTRIBUTIONS.

Clark R. Chapman[1], [1]Dept. of Space Studies, Southwest Research Institute, Suite 300, 1050 Walnut St., Boulder CO 80302 USA, cchapman@boulder.swri.edu.

Introduction: Since my publication 48 years ago of a *JGR* paper with a similar title [1], I have studied crater size-frequency distributions (SFDs) on many planets, satellites, and asteroids, and I have watched the development of this major specialty in planetary science. Such research can yield fundamental information on many issues, including relative stratigraphic ages of geological units, absolute ages and chronology, processes of landform degradation and resurfacing, collisional break-up of asteroids and evolution of impactor SFDs, surface attributes that affect crater scaling, and straightforward cataloging of what are often the dominant topographical features on these bodies.

Unfortunately some valuable methodologies developed decades ago remain underutilized and some methodological errors discussed decades ago persist today and unnecessarily hold back the full potential of these studies. I also discuss the power and limitations of some newer approaches to studying crater statistics (e.g. automatic crater recognition and measurement). Unfortunately, there remain a few issues that will prove difficult or impossible to fully resolve. But we can do better, for example, in addressing issues of contamination of primary crater populations by far-field secondary craters. Many such methodological issues play into unrecognized uncertainties in published crater SFDs (via both biases and statistical uncertainties) that can lead to overconfidence or false interpretations.

Recognition and Measurement of Craters: Recognition of impact craters in images is especially difficult on a planet with complex geology where impact craters must be distinguished from numerous endogenic circular depressions (drainage pits, volcanic craters, and landforms of many types). Even on surfaces where there is a near certainty that circular depressions must be impact craters, experienced crater counters can vary by more than a factor of two in the numbers of craters they recognize [2]. Since a factor of >2 can be the basis for strong scientific disagreements (e.g. about the age of the Rheasilvia basin on Vesta [3]) it is important to understand what is real and what is not…and what reasonable uncertainties are. Part of these differences can be ascribed to different philosophies, which may relate to the purpose of the research: is the goal to be virtually *certain* that a feature is an impact crater, to record any *possible* impact crater, or to catalog all craters having >50% likelihood of being real and having an impact origin? Even then, the visual scale, contrast, solar illumination, pixel resolution, and other factors can dramatically affect recognition of craters. I recommend that crater counters experiment more than most do with varying contrast, comparing their results on the same terrain as seen at different resolutions, and so on in order to better calibrate the uncertainties of their results.

Special issues involve recognition of secondary craters. There are important scientific issues involving both primary and secondary (and occasionally sesquinary) craters, but they require distinguishing the populations. Since secondary craters tend to be small, any real morpohological differences that distinguish them from primaries (which may not even exist for far-field secondaries) often cannot be resolved. Thus one must rely on valid statistical tests of spatial randomness [4]. The common practice of omitting just “obvious” secondaries, while practically convenient, addresses only a minor fraction of the real contamination. Results from common age-dating methodologies such as the cumulative density of craters > 1 km ($N(1)$) may be spurious for the Moon and Mars due to unrecognized secondaries [5, 6]. Even $N(10)$ may suffer from secondary contamination for older terrains on Mercury [7].

Cumulative Plots, CraterStats, and Error Bars: Cumulative SFDs are one of the approved ways to plot crater frequencies as a function of size [8]. But, as pointed out in my 1967 paper [1], they are highly misleading and error bars are treated improperly in the widely used black-box crater tool “Craterstats2” (http://hrscview.fu-berlin.de/software.html). A cumulative *count* of, say all craters > 1 km diameter ($N(1)$) is a valid way to obtain a crater density (although it is dominated by the numbers of craters near the lower diameter cut-off, which may preferentially suffer from biases due to issues previously discussed), and the use of $\sqrt{N}$ as an estimate of statistical uncertainty is valid. But if the goal is to study the *shape* of the SFD, then a cumulative plot can be highly misleading and use of $\sqrt{N}$ as error bars for each diameter bin is just plain wrong. Each value of N – by definition – incorporates all values of N in bins for greater diameters, so the appropriate error bars for identifying the significance of slopes, kinks, and other attributes of curve shape are $\sqrt{N}$ for the incremental values in each bin. But Craterstats2 uses the $\sqrt{N}$ for each cumulative number! Thus there are countless examples in the literature where kinks are discussed as being due to “resurfacings” while the kinks have no statistical significance. (Truly significant kinks, of course, can result from many other causes than the resurfacings usually simplistically invoked, such as horizontal layers of different strengths in the targets, kinks in the SFD of the impactor population, etc.)

Cumulative SFDs look less "noisy" than the inherent incremental data from which they are computed, but they thus hide vital features. In particular, users often fail to detect incompleteness in counts near the smallest resolvable diameter, which would be obvious in a differential or incremental plot. Also, cumulative plots can introduce spurious slopes in certain diameter ranges; for instance, a stochastic excess by a few large craters will introduce a spurious shallower slope over a range of diameters as its importance diminishes towards cumulative numbers for smaller diameter craters. A differential plot might well show a realistic slope over that range, with just the single anomaly at the size of the excess of large craters.

Craterstats2 has other problems. Its black-box output generates model ages with one-to-two more significant figures than are statistically justified. Use of the word "model" is an insufficient caveat. Other researchers interested in results for chronology, and even many crater specialists, get misled by such overly precise ages, which ignore errors due to biases and other Bayesian uncertainties in addition to misrepresenting the simple statistical uncertainties in the data.

I encourage young people entering this field to avoid Craterstats2 and cumulative plots. The most useful and intuitively straightforward way to plot crater data is clearly the R-plot (also approved [8]) which is essentially differential but is divided by D^{-3} so that slopes are moderate and deviations are easily seen. It has the very useful properties that the vertical axis represents spatial density of craters and the density for crater saturation is represented by a horizontal line. Error bars can be calculated correctly by $\sqrt{N}$ and linear or other curve fits to the log-log data can be done in a straightforward manner.

Other Considerations. In 1967, Newell Trask introduced a method of quickly assigning craters a morphological class from 1 = fresh to 4 = highly degraded and plotting SFDs of the four classes. This approach [9] can provide great insight into the nature of processes that degrade and erase populations of craters of different sizes, which often reflects how other landform types are being degraded. The classifications are very quickly assigned (faster than measuring depth/diameter ratios) and a theory of how to interpret the SFDs has been developed (it was applied, for example, to demonstrate that Mars underwent an apparent episode of landform obliteration early in its history [10]). The methodology has been confused by people who have inverted the scale so that class 1 = degraded craters, and it has been underutilized. Of course, it is no substitute for detailed measurements of crater morphologies (peak rings, ejecta blankets, etc,).

Advances have been made in the last two decades in developing algorithms that recognize craters and measure diameters and other morphological parameters. They have not yet achieved the quality of an expert human crater counter in completeness and avoidance of false positives. But they have the virtue of being extremely fast and objectively making the same detections and measurements every time (unlike the jaded human analyst). Such algorithms should be employed, very carefully, as the first-step in analyzing an image: let the human analyst spend efforts *correcting* the mistakes of the automatic routine and *adding* missed craters.

In conclusion, consider the widely adopted philosophy of my good friend and scientific opponent, the late Gerhard Neukum. Neukum and many of his followers adopted some *assumptions* that underlie many of their results…assumptions that are sometimes wrong or are at least subject to scientific dispute. In many cases, things are assumed that, in fact, are what we are trying to learn about. For instance, with few exceptions, Neukum believed that the same population of bodies – asteroids – dominated cratering of all solar system bodies from Mercury to the satellites of the outer planets (the latter is wrong). He further believed that this population's SFD was invariant. Thus, if an observed crater SFD departed from that expected Neukum Production Function, it was never ascribed to a possible change in the shape of the asteroid SFD or other possible causes, but always to "resurfacing episodes" on the body's surface. Neukum also believed that a smoothly declining curve for the bombardment rate he derived for the Moon applied to all other bodies. In reality, the existence and functional form of a Late Heavy Bombardment, and how it might have varied across the Solar System, is unresolved, so a declining curve cannot simply be assumed. We must never let our beliefs and assumptions about cratering prejudge our research results.

References: [1] Chapman, C. R. (1967) *JGR, 72,* 549-557. [2] Robbins, S. J. et al. (2014) *Icarus, 234,* 109-131. [3] Schenk, P. et al. (2015) *LPSC 46th,* Abs. 2309. [4] Bierhaus, E. B. et al. (2005) *Nature, 437,* 1125-1127. [5] Robbins, S. J. & Hynek, B. M. (2014) *E&PSL 400,* 66-76. [6] Ostrach, L. R. et al. (2015) *LPSC 46th*, Abs. 1082. [7] Strom, R. G. et al. (2008) *Science, 321,* 79-81. [8] Crater Analysis Techniques Working G. (1979) *Icarus, 37,* 467-474. [9] Trask, N. J. (1967) *Icarus, 6,* 270-276. [10] Chapman, C. R. & Jones, K. L. (1977) *Ann. Rev. EPS, 5,* 515-538.