

The theory of cratering statistics: Significant improvements to existing methods for ages, slopes, and plots

Samuel W. Bell ¹, Stuart Robbins ²

¹Planetary Science Institute, Tucson, AZ, USA

²Southwest Research Institute, Boulder, CO, USA

Corresponding author

✉ (sbell@psi.edu)

Received

January 13, 2026

Accepted

April 22, 2026

Published

August 27, 2026

Handling Editor

Mark Wieczorek, Institut de physique du globe de Paris – Université Paris Cité – CNRS, Paris, France

License

This article is distributed under the terms of the Creative Commons Attribution License 4.0 

The peer review report of this article is available on the journal website.

© Bell, S.W., and Robbins, S. 2026

Article summary

We develop improved methods for three cratering statistics problems: (1) Probability Density Functions (PDFs) of model ages of that incorporate multiple sources of error; (2) PDFs of the slopes of log-log plots of crater size-frequency distributions (SFDs); (3) a Slanted Average Shifted Histogram (SASH) to plot crater SFDs.

Abstract

Crater counts provide a rich source of information about the history of our solar system, including age estimates of many geologic terrains. In this work we present substantial improvements to existing statistical methods. Existing methods have only included counting error in age estimates. We have developed mathematical methods to calculate age Probability Density Functions (PDFs) that incorporate the error of identifying craters, measuring diameters, selecting the minimum diameter, and the chronology system. We derive more accurate methods for measuring the slopes of log-log plots of crater size-frequency distributions (SFDs), and we show with synthetic modeling and analytical derivations how existing methods are biased. To measure the full shape of the SFD, we present the Slanted Average Shifted Histogram (SASH), a more accurate method for approximating the SFD from a limited number of observed craters. As examples, we show: (1) how the age PDFs of two counts of Canala Crater become much more consistent when all sources of error are included, (2) a practical guide to the age calculation, (3) a visualization of the math of the age calculation, and (4) how the SASH model reveals that the cratered plains of Mimas, Enceladus, Tethys, Dione, and Rhea display equilibrium saturation. With our methods, future researchers will be able to much more accurately estimate the PDF of age and understand the shapes of crater SFDs.

Cite this article as: Bell, S.W., and Robbins, S. (2026). The theory of cratering statistics: Significant improvements to existing methods for ages, slopes, and plots. *Planetary Research*, **1**(1), 235. <https://doi.org/10.53480/46d8-n314>

1 Introduction: existing methods

The Crater Analysis Techniques Working Group (Arvidson et al., 1979) proposed a set of standardized statistical methods for analyzing and displaying the size-frequency distributions (SFDs) of crater counts. Under these standards, crater SFDs are displayed on four types of log-log plots: cumulative, differential, incremental, and R (relative). For each plot, the Arvidson et al. (1979) standards suggest that analysts assume normally distributed error, with error bars of: $\sigma = \sqrt{N}$, where N is the number of craters.

Using the Arvidson et al. (1979) methods, Neukum (1983) formalized a mathematical framework for using crater counts to date geologic units that has become widely accepted. It centers on two functions: the production function and the chronology function. The production function $p(D)$ shows the expected size-frequency distribution of craters on a cumulative plot as a function of diameter D . The chronology function $\mathcal{C}$ relates T , the age in Ga, to $N(1)$, the cumulative density of $D \geq 1$ km craters per km^2 :

$$\mathcal{C}(T) = N(1) = \frac{\lambda}{A} \left(\frac{p(1\text{km})}{p(D)} \right), \quad (1)$$

where A is area in km^2 , and λ is the model count of craters larger than D (for the relationship between the model count λ and the observed count N , see Section 2.1). By applying the chronology function, the production function could be scaled up or down in crater spatial density space to create isochrons on the cumulative plot for any age. These isochrons could then be fit to the observed data to calculate an age.

Neukum (1983) combined the Neukum et al. (1975a), Neukum and König (1976), and König (1977) production functions with new $D > 5$ km data to create the Neukum Production Function (NPF) for the Moon. After adjustments by Ivanov et al. (2001), the NPF takes the form:

$$\log_{10} \left(p(D) = \frac{\lambda}{A} \right) = \sum_{i=0}^{11} a_i (\log_{10} D)^i, \quad (2)$$

where D is diameter in km, λ is the model count of craters larger than D , A is area in km^2 , and the a_i are empirical constants whose values Ivanov et al. (2001) list. To fit the NPF, Neukum (1983) used least squares fits to Arvidson et al. (1979) plots of his data.

Expanding on work by Neukum et al. (1975b) and Neukum and Wise (1976), Neukum (1983) calibrated the Neukum Chronology Function (NCF) from crater counts on terrains dated by Apollo or Luna samples. After updates by Neukum and Ivanov (1994), it takes the form:

$$\mathcal{C}(T) = 5.44 \times 10^{-14} (e^{6.93T} - 1) + 8.38 \times 10^{-4} T, \quad (3)$$

where T is age in units of Ga. Neither the production nor chronology functions have error models, so the NCF is typically applied as if it were perfectly accurate, even though Neukum (1983) states that there are errors. Other lunar chronology functions (Hartmann et al., 1981; Hartmann et al., 2007; Robbins, 2014; and Werner et al., 2023) disagree with the NCF and each other. The Chang'e 5 and Chang'e 6 missions added two more calibration terrains, but crater counts by different authors differ significantly (e.g. Jia et al., 2020; Qian et al., 2021; Werner et al., 2023; Luo et al., 2024; and Qian et al., 2024).

Computational advancements brought updates to the traditional methods. Michael et al. (2016) replaced the practice of isochron fitting by directly calculating a probability density function (PDF) of age from the observed number of craters and counting statistics. Robbins et al. (2014) measured human error in both crater identification and diameter measurement, but these sources of error have yet to be incorporated into age PDFs. Robbins et al. (2018) replaced least squares fits of single-slope crater SFDs with the Maximum Likelihood Estimator (MLE) of the slope-likelihood distribution of the Pareto distribution, and they also modified Arvidson et al. (1979) plots with a Kernel Density Estimator (KDE) approach. Otherwise, relatively few changes have been proposed to the formalism developed almost half a century ago.

In this work, we provide what we think is a next step towards advancing the proper statistical treatment of impact craters, made possible by faster computers and better mathematical techniques in the literature. We highlight three key applications of cratering statistics. In Section 2, we focus on the Age Problem: calculating ages from crater counts. In Section 3, we focus on the Slope Problem: measuring the slope of crater SFDs in log-log space. In Section 4, we focus on the Plotting Problem: approximating the full form of the SFD from a limited number of craters. In each case, we present significant improvements to existing methods. Section 5 discusses three worked examples, and Section 6 presents our conclusions. Expansive Appendices provide more of the mathematical backing. To make our methods widely available, we have written a Python package, *cratersfd*, which can be downloaded from <https://github.com/samwbell/cratersfd>.

2 The Age Problem

Mathematically, we solve the Age Problem by solving Equation (1):

$$T = \mathcal{C}^{-1}(N(1)) = \mathcal{C}^{-1}\left(\frac{\lambda}{A} \left(\frac{p(1\text{km})}{p(D_{\min}) - p(D_{\max})} \right)\right), \quad (4)$$

where $\mathcal{C}^{-1}$ is the inverse function of the chronology function, $D_{\min}$ is the minimum diameter of craters counted, and $D_{\max}$ is the maximum diameter. (When $D_{\max}$ is infinite, $p(D_{\max}) = 0$.) Critically, however, the values of these variables are not perfectly known. In formal terminology, they are *random variables*. Mathematically, we quantify the probability distribution of different values of a random variable with a probability density function (PDF), a concept that Michael and Neukum (2010) and Michael et al. (2016) introduced to the Age Problem. If the random variable does not have a continuous distribution and instead has discrete possible values (such as a random variable whose possible values are only whole numbers), its probability distribution is a probability mass function (PMF) instead of a PDF.

When we represent random variables as a central value and error bars, this is an approximation. When mathematical operations are applied to random variables, the full form of the PDF matters, and it is not sufficient to apply the mathematical operation to the central value – unlike the cornerstone assumptions of current methods. Instead, there are long-established specific rules for random variable math, which we review in Appendix A. When working with random variables, it is necessary to apply these rules – which existing Age Problem methods have not done. Critically, random variable math also provides a framework to incorporate sources of error beyond counting error. Although existing Age Problem methods acknowledge that sources of error beyond counting statistics apply (e.g. Michael et al., 2025), they did not incorporate other sources of error into age PDFs. Using random variable math, we have further developed the mathematical methods to incorporate multiple sources of error.

We have broken this broader section on the Age Problem into nine smaller sections. In the first seven we discuss different sources of error and how to quantify them: counting error (Section 2.1), human crater identification error (Section 2.2), diameter measurement error (Section 2.3), production and chronology function error (Section 2.4), area error (Section 2.5), $D_{\min}$ selection error (Section 2.6), and additional error factors (Section 2.7). With random variable math, we can incorporate all of these sources of error into the PDF of $N(1)$, and in Section 2.8 we discuss the final step in Equation (4): applying the inverse chronology function $\mathcal{C}^{-1}$ to the $N(1)$ PDF to arrive at the PDF of age. Finally, in Section 2.9 we combine these concepts into our final Age Problem solution.

2.1 Counting error

The Poisson distribution describes the probability P of N craters being observed given a model count λ (Arthur, 1954):

$$P(N) = \frac{e^{-\lambda} \lambda^N}{N!}. \quad (5)$$

In the Age Problem, however, we are looking for the PDF of λ , not the PMF of N . Traditional Age Problem methods (e.g. Michael and Neukum, 2010; Michael et al., 2016; Bell, 2020; Michael et al., 2021; Michael and Liu, 2025) have described the λ PDF as being governed by the Poisson distribution equation. This does produce the correct λ PDF, but it is important to clarify that it is only N that follows a Poisson distribution. The model count λ follows a Gamma distribution, with shape parameter $s = N + 1$ and rate parameter $r = 1$:

$$P(x) = \frac{r^s x^{s-1} e^{-rx}}{\Gamma(s)} \xrightarrow{x=\lambda; s=N+1; r=1} P(\lambda) = \frac{e^{-\lambda} \lambda^N}{\Gamma(N+1)}. \quad (6)$$

(The Greek letters α and β are typically used for the shape and rate parameters. We use s and r to avoid confusion with the α parameter of the Pareto distribution.)

The equation for a PDF will give the relative probability of any of its parameters, but a PDF must be normalized so that its probability integrates to 1, and switching the variable can change the normalization. Here, it replaces the factorial with its continuous extrapolation, the Gamma function Γ defined so that $N! = \Gamma(N+1)$. Recognizing the λ PDF as a Gamma distribution allows us to apply the theorems of the Gamma distribution. For instance, the mode (the maximum likelihood value) is $(s-1)/r = N$, and the mean is $s/r = N+1$. This causes the PDF to be asymmetric at low N and approach symmetry as N rises.

Because the production function varies with diameter, early versions of traditional methods fit production function isochrons to crater SFD plots. In a mathematical proof, Michael et al. (2016) showed that the total N will follow the Poisson distribution (Equation (4)), even though the production function varies with diameter, but this approach was questioned (e.g. Robbins et al., 2018). To test this, we built a synthetic model that breaks the production function into very small, independent diameter increments,

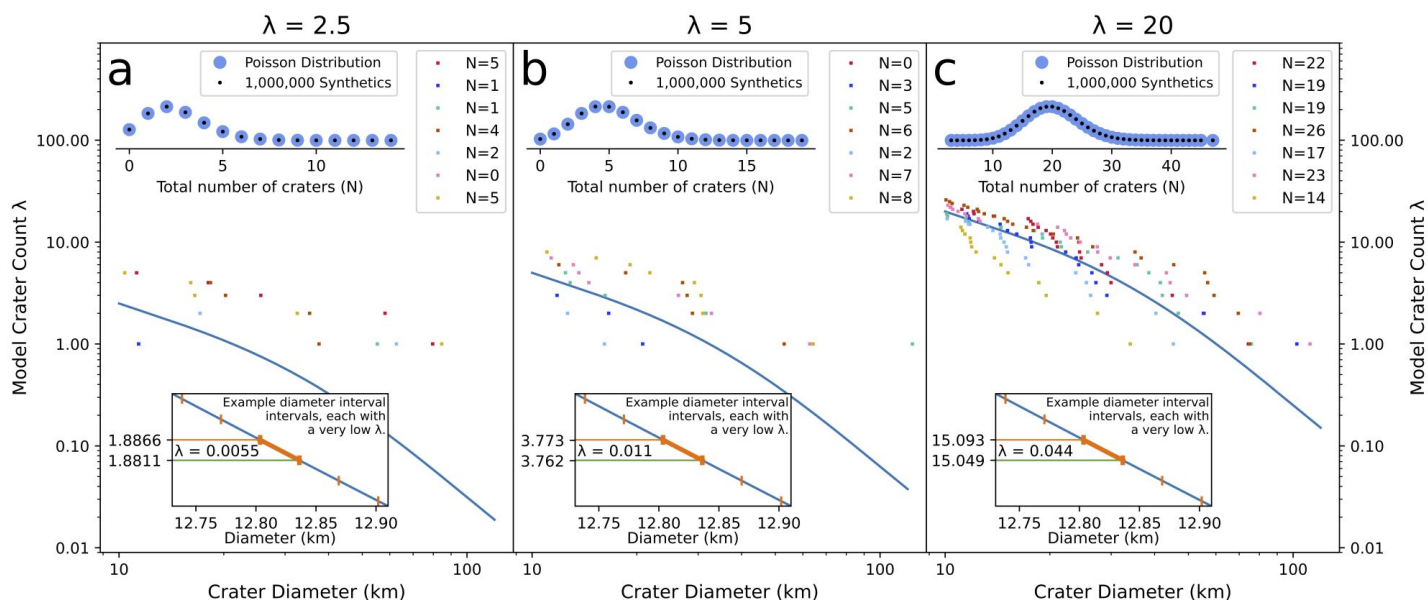


Figure 1 | Synthetic modeling results of the Michael et al. (2016) proof that the total raw N follows the Poisson distribution. We broke the production function (the “new” NPF – Ivanov et al., 2001) into small diameter intervals (example shown in the insets) and randomly generated craters for that interval with the Poisson distribution (Equation (5)). Then, we combined the results for all intervals to produce a synthetic crater dataset. We generated 1 000 000 synthetic datasets and compared the distribution of observed N to the Poisson distribution. At each value of λ , it exactly matches the Poisson distribution, even though we generated craters at different diameters – confirming Michael et al. (2016)’s proof. We show seven randomly selected synthetics with the traditional (Arvidson et al., 1979) form of the cumulative plot without error bars. We use the traditional form because we have yet to introduce our preferred form of the cumulative plot (see Section 4.1 and Figure 6).

then randomly generates craters at those diameters. The total N matched the Poisson distribution, exactly as Michael et al. (2016) predicted (Figure 1). The implications are considerable: in the Age Problem, we can treat each crater in our diameter range equally.

For these reasons, Michael et al. (2016) strongly discourage the practice of isochron fitting because it is unnecessary and introduces additional errors that are completely unneeded. We agree. However, it is still essential to plot the data to ensure that the SFD does, in fact, match the production function (see Section 4). Without plotting the data, this cannot be assumed. If we choose a D_{min} where the SFD is depressed by saturation, erosion, or resolution, the resulting age will be too low. When a terrain has been resurfaced, it can record two different ages at different diameter ranges, which must be measured separately (Michael and Neukum, 2010). Finally, the production function might be meaningfully inaccurate for the data in question.

2.2 Human crater identification error

Traditional methods have not incorporated human error in crater counting, an important shortcoming because it is unrealistic to assume that crater counters always identify every crater and never identify a quasi-circular feature as a crater when it is actually something else. Robbins et al. (2014) quantified these effects, comparing results by different experts on the same two images: a NAC image of the Apollo 15 landing site and a WAC image with both a mare and a highlands terrain. They found that these errors are considerable: at $N = 1$, the error was roughly $\pm 50\%$, but as N increased (and D decreased), the relative error decreased until it plateaued at a minimum level that depended heavily on saturation conditions (see Section 2.7 for a discussion of saturation and the definition of “equilibrium”). For the NAC and highlands terrain, the error plateaued when the SFD reached equilibrium, hitting a minimum level of roughly $\pm 30\%$ in the NAC and $\pm 40\%$ in the highlands. In the mare, human crater identification error plateaued at $\pm 10\%$ before reaching equilibrium, but when the SFD reached equilibrium, the error increased until it reached roughly $\pm 40\%$.

Here, we lay out the math required to incorporate these effects into the Age Problem. We interpret these data as a combination of random and systematic error. Random error will decrease with increasing N as random variations average out, but systematic error will not. This means that if one person systematically counts more vaguely circular depressions as degraded craters than others do, that specific variation will not cancel out with increasing N .

A standard assumption for the PDF shape of a source of error is a normal distribution. Mathematically, however, these errors cannot follow a normal distribution because a negative count is impossible, so we choose a lognormal distribution instead. In a normal distribution, the probability is equal at $\mu + x$ and

$\mu - x$, where μ is the mean. When we apply the natural log function to a normal distribution, we get a lognormal distribution, where the probability is instead equal at μx and μ/x , and the distribution cannot go negative.

We quantify the percentage errors as lognormal random variables with a mean of 1 and a width that reflects the percentage error p . For instance, for 25% error, $p = 0.25$, and the error bars fall at $1/1.25 = 0.8$ and 1.25. The PDF of a lognormal random variable is:

$$P(x) = \frac{e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}}{x\sigma\sqrt{2\pi}}, \text{ where } \sigma = \ln(1 + p), \text{ and } \mu = -\frac{\sigma^2}{2}. \quad (7)$$

This random variable forms an error kernel that is multiplied by the variable in question with random variable math (see Appendix A). In the case of human crater identification error, the error kernel is applied to the model count λ .

In Appendix A.1 we derive how to combine random and statistical error into a single lognormal human crater identification error kernel:

$$\mathcal{E}_h = \mathcal{E}_s \frac{\sum_{i=1}^N \mathcal{E}_{r_i}}{N}, \quad (8)$$

where $\mathcal{E}_s$ is the systematic error kernel, the $\mathcal{E}_{r_i}$ are the random error kernels for each crater, and $\mathcal{E}_h$ is the combined human crater identification error kernel. If we assume the $\mathcal{E}_{r_i}$ are identical, the percentage error of $\mathcal{E}_h$ is p_h , which can be found by the very close approximation:

$$p_h = e^{\sqrt{\ln\left(\frac{e^{(\ln(1+p_r))^2} - 1}{N} + 1\right) + (\ln(1+p_s))^2}} - 1, \quad (9)$$

where p_r is the random error percentage of a single crater, and p_s is the systematic error percentage. The Robbins et al. (2014) data match a model where $p_r \approx 50\%$, and $p_s \approx 10\%$ in the unsaturated WAC mare, $p_s \approx 30\%$ in the saturated NAC mare, and $p_s \approx 40\%$ in the saturated WAC highlands. Because the Age Problem must be applied to unsaturated data, the Robbins et al. (2014) results would imply values of $p_r \approx 50\%$ and $p_s \approx 10\%$, but they miss an additional source of human crater identification error: the image used for counting.

Robbins et al. (2014) used the same image on a single planetary body, and image characteristics can introduce additional error factors due to variations in lighting conditions and resolution. Because the sharp rims of fresh craters are easier to see without the shadows of low sun angles than the subtler rims of degraded craters, the effect of incidence angle varies significantly with saturation and crater degradation conditions. Examining a distribution almost entirely in equilibrium, Robbins et al. (2025) found significant effects that faded as diameters increased towards leaving equilibrium, and Ostrach (2013) found that these effects continue to fade as D increases past the rollover into diameters that have not reached equilibrium. Examining the same region of Mare Imbrium at 87° , 82° , 71° , and 50° , Ostrach (2013) found larger rollover diameters at lower incidence angles but equivalent ages for all four incidence angles with a conservative selection of D_{min} safely above the rollover. Although resolution effects have been assumed to become negligible above a completeness diameter calculated from the resolution-induced SFD rollover, Robbins et al. (2025) found effects continuing above the apparent completeness diameter.

Considering these data cumulatively, we interpret image to image variation as only slightly increasing human error for the Age Problem, so we recommend assuming $p_r \approx 60\%$ and $p_s \approx 12\%$ (increased by 20% from $p_r \approx 50\%$ and $p_s \approx 10\%$ implied by the Robbins et al. (2014) results). However, these are only the default values. If resolution is low, the incidence angle is below $\sim 50^\circ$, or there are other factors complicating crater identification, we recommend estimating higher values. Ideally, a researcher would perform their own experiment to quantify their own human variation, but that is not always feasible. Finally, if multiple counts by the same person are being compared, we recommend assuming no systematic error.

2.3 Diameter measurement error

Humans do not measure crater diameters with perfect accuracy, and craters are almost never perfect circles. There will always be some error, but these errors have never before been incorporated into the Age Problem. Comparing expert counts of the same terrains, Robbins et al. (2014) found $\sim 10\%$ diameter measurement error on lunar maria close to where the SFD rolls over to equilibrium, decreasing to $\sim 5\%$ as diameters increase further away from the rollover. When he compared his global $D > 1$ km lunar crater database with the Losiak et al. (2009), Head et al. (2010), and Povilaitis et al. (2018) databases, Robbins (2019) found an average random error of $\sim 4.9\%$ and an average systematic error of 0.83% .

Based on these results, we recommend default assumptions for diameter measurement error of $p_{D_s} \approx 1\%$ systematic error and $p_{D_r} \approx 5\%$ random error when D_{min} is chosen far from the equilibrium saturation rollover, rising to $p_{D_s} \approx 2\%$ and $p_{D_r} \approx 10\%$ when it is close to the rollover. We again recommend assuming a lognormal distribution. These values will change with various conditions, so we recommend estimating them for your own data. (Once again, if multiple counts by the same person are being compared, we recommend assuming no systematic error.)

With diameter measurement error estimates, we urge accuracy because diameter measurement error does not simply increase the uncertainty: it also shifts the result. When humans make diameter measurement errors, craters that are slightly smaller than D_{min} can get shifted above D_{min} , increasing the observed N , and craters that are slightly larger than D_{min} can get shifted below D_{min} , decreasing the observed N . Because of the negative slope of the production function, there are typically more slightly smaller craters available to be shifted above D_{min} (where they count towards N) than slightly larger ones available to be shifted below D_{min} (where they would no longer count towards N). This means that random diameter measurement error typically shifts more $D < D_{min}$ craters above D_{min} than it shifts $D > D_{min}$ craters below D_{min} . As a result, random diameter error typically increases the observed N and apparent age, and adjusting for it typically yields an age PDF with a smaller mean age. However, actual crater diameters differ with random chance, and these effects depend on the specific data in question.

To demonstrate the effect of diameter measurement error, we built a synthetic model (Figure 2) focusing on a steep portion of the production function, where diameter measurement error has a greater effect. For each run, we start with a synthetic crater count with 10 000 $D > 0.0077$ km craters generated from the NPF, then we count the number of craters larger than $D_{min} = 30$ m to measure the true N . To simulate a human crater counter not quite perfectly measuring crater diameters, we randomly add in measurement error for each crater (see Appendix A.2). Then, we count the number of craters larger than $D_{min} = 30$ m to measure the observed N . Repeating this process 10 000 times, we calculate the distribution of observed N to true N ratios that diameter measurement error would produce.

The results of the synthetic model are clear (Figure 2). On average, diameter measurement error artificially skews observed N and apparent ages higher, but the effect depends on the data in question, and sometimes diameter measurement error can lower N values and apparent ages. In the synthetic model, our default values of 5% random error and 1% systematic error produce a ratio of observed N to true N of 1.009 ± 0.044 (Figure 2). In less ideal conditions, the effect can be considerable, with 20% random error and 3% systematic error producing a ratio of observed N to true N of 1.121 ± 0.121 in our synthetic model (Figure 2).

In Appendix A.2, we lay out the math for correcting for the bias and incorporating diameter measurement error, which we have validated with a synthetic model (Figure A1). Although the math is relatively complex, especially for random error, we present a code that can efficiently calculate it as part of the *cratersfd* package.

2.4 Production and chronology function error

Traditional Age Problem methods do not incorporate production or chronology function errors for two reasons: both production and chronology functions lack robust error models, and the math for incorporating chronology system error has not been developed. Because these errors are considerable, by ignoring them, crater analysts have been producing ages with dramatically understated error bars. Calling attention to this problem, Michael et al. (2016) recommended applying a μ operator to ages as a “perpetual reminder” of chronology system uncertainty.

Instead of simply saying this error is unknown, we recommend making an estimate: applying an error factor to the $N(1)$ PDF. Lagain et al. (2021) recommended a factor of ~ 2 on Mars, but we believe the sources of error are, cumulatively, sufficient to justify somewhat wider error bars: calibration terrain crater counts vary by roughly a factor of $\sim 50\%$ between different researchers (e.g. Neukum, 1983; Robbins, 2014; Jia et al., 2020; Qian et al., 2021; Werner et al., 2023; Luo et al., 2024; Qian et al., 2024). Different researchers have also disagreed about which calibration terrains can be robustly correlated with returned samples (e.g. Neukum, 1983; Robbins, 2014; Werner et al., 2023). Additionally, potential spatial variations in the cratering rate may further increase these uncertainties (e.g. Morota and Furumoto, 2003; Le Feuvre and Wiczorek, 2011; Ghent et al., 2024).

Moreover, all the additional sources of error beyond counting error we describe here apply to the calibration counts (e.g., human crater identification error also applies to whoever constructed the calibration curves). Although the primary effect of our error model is to increase uncertainty, it also corrects bias, and there is likely a small average bias correction. (We are unsure of the direction of the average bias correction because of the uncertainty of the average effect of the D_{min} PDF.) Because this average bias is not removed from the chronology function calibration data, it adds to the effective chronology function uncertainty.

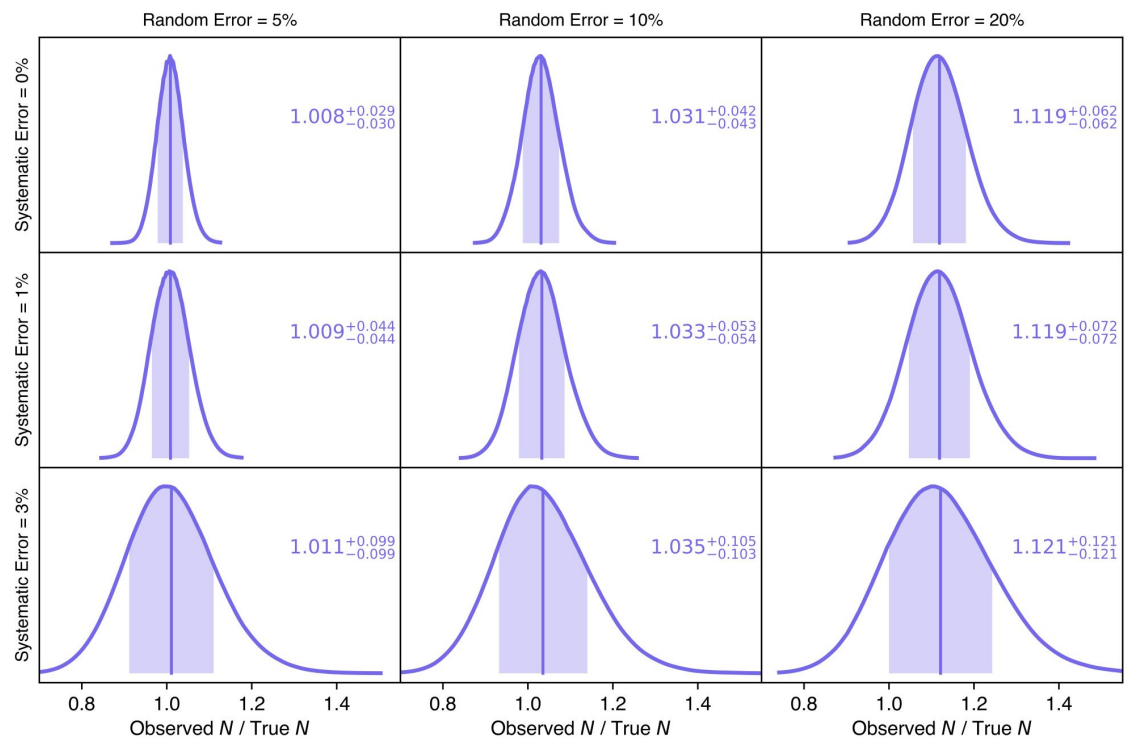


Figure 2 | A synthetic model of diameter error, simulating measuring N at $D_{\min} = 30$ m from 10 000 $D \geq 10$ m NPF synthetics with diameter error added in (see Appendix A.2 and Section 2.3 for details). These results confirm that, on average, random diameter error biases the observed N higher because of the steep exponential shape of the SFD. It also increases the uncertainty, and even a small amount of systematic error significantly increases the uncertainty. Note: We show values to 3 decimals to illustrate offsets and biases; we do not recommend a researcher actually show this many significant figures.

Perhaps most importantly, none of these models incorporate production function error. As formalized by Neukum (1983) and Ivanov et al. (2001), the NPF has no error model, and it is typically used without one. However, during the development of the NPF, Neukum et al. (1975a) and König (1977) *did* include error estimates, and from those estimates, we have extrapolated an NPF error model (Appendix B). From this NPF error model, there is ~90% uncertainty in $N(1)$ between ~10 m, where the recent calibration terrains are best measured, and ~500 m, where the older calibration terrains are measured.

Crucially, this NPF error model does not include target property effects. Although the NPF (Neukum, 1983) preceded the development of crater scaling theory, Neukum and Ivanov (1994) accepted the scaling law of Holsapple (1993), clarifying that crater size does depend on the material properties of the target, especially in the strength regime. Because younger calibration terrains like North Ray, South Ray, and Cone craters are unconsolidated fresh crater ejecta deposits, and older terrains are mostly basaltic maria, target property effects are especially important for chronology function calibration. Production function error matters for more than just the chronology function calibration. It also affects the Age Problem calculation for a crater count at a diameter different from those used to calibrate the chronology function.

Finally, there is the possibility that the cratering rate and production function may vary on timescales of several hundred Myr or less. Although the NPF assumed a constant cratering rate since ~3 Ga, Neukum et al. (2001) estimated that the cratering rate could vary within a factor of ~2 on shorter timescales over the past three billion years. Major asteroid collisions and dynamical spreading through the Yarkovsky effect are predicted to cause impact showers (e.g. Bottke et al., 2001; Bottke et al., 2007). Because the Yarkovsky effect migrates larger asteroids more slowly (e.g. Bottke et al., 2007), these showers are predicted to cause fluctuations in the production function as well as the cratering rate. Crucially, these fluctuations affect both the date of an individual terrain and the chronology function calibration.

Taken together, these sources of error are considerable. In general, very rough estimates for the magnitudes of these effects range from ~50% to a factor of ~2 (100%), so when we combine them all, they should exceed a factor of ~2. At the same time, there are reasons to think a factor of ~3 might be too high on the Moon: not every source of error applies in every case. Moreover, errors do not add linearly. For instance, under the lognormal quadrature rule (Equation (A7)), two 100% (factor of 2) errors combine to make a ~170% (factor of ~2.7) error, not a 200% (factor of ~3) error. We therefore feel justified in recommending a lognormal factor of ~2.5 for the Moon. For Mars, Venus, and Mercury, we recommend a factor of ~3 to incorporate the error of scaling the chronology system from the Moon.

Because chronology function error affects all terrains on a body equally, it should not be included when determining whether one terrain is older than another. As a result, we suggest it is reasonable to show an age without chronology system error *in addition to* an age that does include it. However, we caution that ages measured at different diameters will still be subject to production function error. To calculate this error, in the absence of a better model, we recommend using our derivation of the NPF error model in Appendix B, but we note that it may dramatically underestimate the error in the strength regime due to target property effects. For many geologic questions, ages without chronology system error still have deep explanatory power, but the true ages are the true ages. For questions where the true age matters, it is important to acknowledge the reality of these wide error distributions.

2.5 Area error

The area of the count region can often be measured very precisely, and its error can frequently be ignored, but this is not always the case – especially when dating a diffuse unit. For instance, a continuous ejecta deposit might have seven visible craters larger than D_{min} , but it might not be clear exactly what the count area around those craters is. Here, we recommend estimating the error as a lognormal factor. We also note that many area calculation algorithms, such as the ArcMap-based method used by CraterTools (Kneissl et al., 2011), can be imprecise for large regions because the lines between vertices change when individual vertices are reprojected into an equal area projection. Geodesic methods that avoid re-projection, such as those available in ArcGIS Pro, only partially mitigate this effect. Geodesic methods still present problems for measuring large count areas drawn in a projected coordinate system because the lines between vertices still change when they go from straight lines in the projected coordinate system to geodesics that trace the shortest distance between two vertices on an oblate spheroid. To avoid these problems, we recommend densifying the lines by inserting more points between vertices (or using a gridded approach of very small rectangles whose area is known and are counted if they are within the polygon). As an example, Yue et al. (2020) calculated the area of the Orientale ejecta as 1.68×10^6 km with CraterTools, and we find an area of 1.85×10^6 km with both line densification and an integration of very small rectangles. In *cratersfd*, we present code to incorporate an area PDF into ages and safely calculate the area of a large, irregular region.

2.6 D_{min} selection error

Finally, we must confront the error in the *selection* of D_{min} . To address equilibrium saturation, resolution limits (Robbins et al., 2014, 2025; Williams et al., 2018; Wang et al., 2020), and erosion, a human counter typically selects the smallest value of D_{min} not affected by the rollover at smaller diameters. As Robbins et al. (2014, 2018) and Michael and Liu (2025) warned, N must be measured at a D_{min} where the SFD represents the production function and is not deflected by saturation, resolution, or erosion. This critical part of the Age Problem is not standardized – Robbins et al. (2014) could find no commonality amongst the eight researchers for defining D_{min} .

Typically, D_{min} is chosen by visually estimating the precise diameter of an often subtle rollover of the SFD from one of the Arvidson et al. (1979) plots. However, the binned Arvidson et al. (1979) plots provide limited resolution for how the differential slope changes. With our improved SASH method for making differential plots (Section 4.2), the SFD can be resolved much more accurately, enabling a more accurate estimate of D_{min} . (However, we discuss D_{min} here because it is relevant to the Age Problem.)

This will still be an estimate. As with any estimate, there is a range of plausible values. We recommend estimating this range and quantifying it as a PDF of D_{min} (Figures 10d and 12u). For the math to incorporate the PDF of D_{min} , see Appendix A.3. When a researcher picks a single value of D_{min} , that choice can have a considerable effect on the age, so choosing a single value of D_{min} effectively has the researcher choosing the age. This makes human biases difficult to exclude – even for the most conscientious analyst. Choosing a PDF of D_{min} that quantifies the range of values they might pick significantly mitigates this bias. (For examples, see Sections 5.1–5.3 and Figures 10–12.) Researchers should take care to only include plausible values in the D_{min} PDF, excluding values they consider likely to be saturated. Including saturated values could introduce a greater source of error than the selection bias the D_{min} PDF exists to mitigate.

2.7 Additional error factors

In Sections 2.1–2.8, we have only included well-established sources of error that can be reasonably constrained and mathematically quantified. Many other potential error drivers exist, but the nature of these effects is poorly constrained: There is potential bias in count area selection and geologic unit identification. Secondary crater exclusion methods are not standardized, and different choices frequently lead to different results. For instance, differences in secondary exclusion methods were one of the factors that led Jia et al. (2020) and Qian et al. (2021) to reach different crater counts for the Chang'e 5

landing site. If diffuse “field” secondaries that might be hard to identify as clear clusters or rays (e.g. Shoemaker, 1965; Wilhelms et al., 1978; McEwen et al., 2005; Xiao and Strom, 2012; Minton et al., 2019) contribute substantially to observed crater counts, they will not be independent events, which would increase error.

Saturation, the effect of new craters erasing previous craters, presents significant complications. Saturation falls into two regimes: (1) the equilibrium regime (Gault, 1970) when the production function’s slope on a cumulative plot is steeper than ~ -2 , and (2) the quasi-equilibrium regime (Chapman and McKinnon, 1986) when the production function’s slope on a cumulative plot is shallower than ~ -2 .

In the equilibrium regime, the observed SFD follows the production function until the crater density reaches a rollover density, where the SFD enters equilibrium and rolls over to a cumulative plot slope that is shallower than ~ -2 (Xiao and Werner, 2015). As more craters are added, the density of the equilibrium SFD will not increase (Gault, 1970), and in fact, it will often decrease (Xiao and Werner, 2015). The rollover density has a cumulative plot slope of ~ -2 , and it typically has a value of roughly 0.1–0.3 on an R plot (Richardson, 2009).

In the quasi-equilibrium regime, on the other hand, the SFD follows the production function at low crater densities, but when the density increases to the point where large craters begin overprinting small craters, the SFD becomes shallower (Chapman and McKinnon, 1985). Because each new overprinting crater becomes a blank slate for future cratering, the SFD does not reach a stable equilibrium; instead, it reaches a quasi-equilibrium that jumps around when overprinting events occur (Chapman and McKinnon, 1985).

For the Age Problem, we need a count where saturation has not decreased the crater density. In the quasi-equilibrium regime, saturation effects begin at significantly lower crater densities than the equilibrium regime’s rollover density (Richardson, 2009), creating significant complications for the Age Problem. Even when a crater falls in the equilibrium regime, but the production function has a quasi-equilibrium branch at higher diameters, the Richardson (2009) model predicts that quasi-equilibrium saturation can lower N for diameter ranges that have yet to reach equilibrium. This is important in the inner solar system. The “new” NPF (Ivanov et al., 2001) has a $6 \text{ km} \leq D \leq 52 \text{ km}$ shallow branch in the quasi-equilibrium domain. A lunar crater count that contains some craters larger than $\sim 6 \text{ km}$ may experience quasi-equilibrium effects even if the minimum diameter is smaller than $\sim 6 \text{ km}$ and falls in the equilibrium domain.

In the equilibrium regime, the Age Problem is traditionally done by measuring the crater density at a diameter above the rollover. However, models (Richardson, 2009; Hirabayashi et al., 2017; Minton et al., 2019) of equilibrium saturation suggest that N may be depressed out to higher diameters than the visible rollover diameter. Conceptually, these effects must take place to some degree. When a large crater forms, it does erase craters, either directly overprinting them or burying them beneath the continuous ejecta. Rays of discontinuous ejecta drive erosion of other craters beyond the extent of the continuous ejecta (Minton et al., 2019). When the surface has yet to reach equilibrium, these two effects may not dominate, but they do clearly *sometimes* happen. However, because these effects have yet to be quantified empirically, it remains unclear precisely how significant they are.

Then, there is target property variation. Significant differences between solid rock and ejecta targets have been observed empirically, (e.g. van der Bogert et al., 2017), and the precise material properties of ejecta and solid rock targets are poorly constrained and may vary to a meaningful degree. Moreover, the subsurface lithology may differ from the surface. For instance, the Em4 terrain where Chang’e 5 landed has at least two layers of buried regolith beneath its thin layer of surface basalt (Du et al., 2022).

Finally, there is the possibility of additional effects, unknown unknowns. Empirically speaking, sub-areas within broader count areas have been observed with more variation than can be explained by human crater identification and diameter measurement error at the Orientale ejecta (Yue et al., 2020), the Chang’e 5 landing site (Qian et al., 2021), and the Chang’e 6 landing site (Luo et al., 2024). Kirchoff (2018) observed that the lunar maria are more clustered than they would be were cratering fully random, implying that additional error factors are at work.

Taken collectively, we believe these additional error factors cannot be ignored. To address these cumulative potential errors, we recommend including an additional lognormal uncertainty kernel on the model count λ , $\mathcal{E}_a$. Even in favorable conditions, we estimate that count area selection bias, geologic unit identification error, secondary crater exclusion error, saturation effects outside of equilibrium, and target property variability each probably have average effects of at least a few percent. Considering all these factors as well as unknown unknowns, we recommend a minimum $\mathcal{E}_a$ value of 10%, but this value depends on the data and should be raised when there is special reason for concern.

With a method to incorporate an estimate of the additional sources of error that have yet to be precisely quantified, we have now addressed each source of error. We can use random variable math to calculate a PDF of $N(1)$ that incorporates all of these sources of error, but there is one step left in Equation (4): applying the inverse chronology function to the $N(1)$ PDF to calculate the PDF of age.

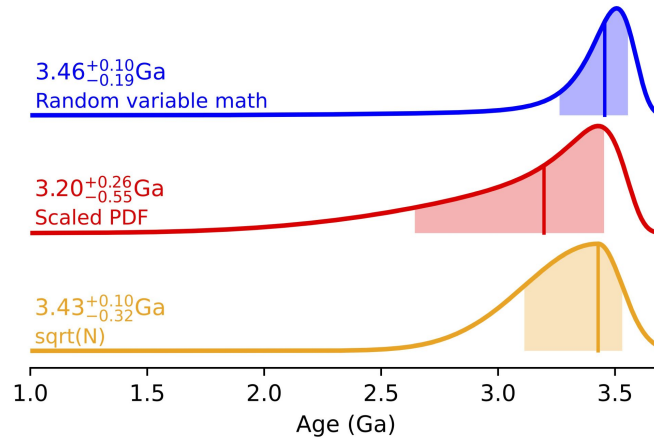


Figure 3 | An example of the effect of properly applying the inverse chronology function $\mathcal{C}^{-1}$ with random variable math. Here, we show the age calculation for a lunar terrain with an N of 10, an area of 2500 km², a $D_{\min}$ of 1.0 km, the new NPF (Ivanov et al., 2001), the NCF (Neukum and Ivanov, 1994), and no sources of error beyond counting error under three methods: applying $\mathcal{C}^{-1}$ with random variable math, scaling the PDF by $\mathcal{C}^{-1}$ (Michael et al., 2016), and applying $\mathcal{C}^{-1}$ to the central value and bounds of $\sqrt{N}$ error bars.

2.8 Applying the inverse chronology function to a random variable

From isochron fits (e.g. Neukum, 1983) to the Michael et al. (2016) λ PDF method, traditional Age Problem methods have not applied the inverse chronology function $\mathcal{C}^{-1}$ to the random variable $N(1)$. In many cases, $\mathcal{C}^{-1}$ was applied to a single $N(1)$ value and the upper and lower bounds of its error bars, not the full PDF. Michael et al. (2016) did apply their calculation to the full PDF, but instead of applying $\mathcal{C}^{-1}$ to $N(1)$, they scaled the $N(1)$ PDF by $\mathcal{C}^{-1}$, such that $P_T(t) = P_{N(1)}(\mathcal{C}(t))$. (See Appendix A for details on applying functions to random variables and scaling a random variable by a function.) To apply $\mathcal{C}^{-1}$ to $N(1)$, the equation is:

$$P_T(t) = P_{N(1)}(\mathcal{C}(t)) \frac{d(\mathcal{C}(t))}{dt}. \quad (10)$$

If $\mathcal{C}$ is nonlinear (e.g., >3 Ga for the NCF, or any portion of the Robbins, 2014 or Hartmann et al., 2007 chronology), scaling $N(1)$ by $\mathcal{C}^{-1}$ will not be equivalent to applying $\mathcal{C}^{-1}$ to $N(1)$. As an example, let us apply the inverse NCF to a lunar crater count of $N = 10$, $D_{\min} = 1.0$ km, $A = 2500$ km², and no other sources of error beyond counting statistics. Here, the age PDF from properly applying $\mathcal{C}^{-1}$ with Equation (10) has a median of $3.46^{+0.10}_{-0.19}$ Ga, but the scaled age PDF would have a median of $3.20^{+0.26}_{-0.55}$ Ga, and the age from applying $\mathcal{C}^{-1}$ to the central value and bounds of $\sqrt{N}$ error bars would be $3.43^{+0.10}_{-0.32}$ (Figure 3).

Our interpretation of the reason why Michael et al. (2016) did not include the derivative term (in technical mathematical terminology: the Jacobean) is that Michael et al. (2016) imposed an initial prior assumption that all ages have equal probability. Effectively, Michael et al. (2016) assumed that $N(1)$ values associated with higher ages were far less likely. With this assumption, for instance, we would conclude from the observation of 10 craters in our example that the model count was $\lambda = 6.25^{+1.35}_{-1.22}$, not $\lambda = 10.0^{+1.3}_{-1.4}$. Effectively, we would be assuming that our observation was an outlier. Although this assumption may seem neutral, in the vast majority of cases it is not realistic to how crater counting works. Ages of planetary terrains are not equally scattered throughout time. Instead, ages typically concentrate towards the beginning of the solar system for a simple reason: cratering is the dominant geologic process. On the Moon, for instance, most terrains being dated are craters or volcanism induced by cratering, such as the formation of maria in large impact basins. Even when formed by other types of geologic activity, such as fluvial activity on Mars (Fassett and Head, 2008), many terrains are likelier to date from earlier times when the cratering rate was higher.

As a result, we prefer to report what the data directly indicate, rather than imposing a structure on them with an assumption of equal probability of all ages. However, as we discuss in Appendix A, if a specific prior age assumption is being used, then the age PDF should be scaled by $\mathcal{C}^{-1}$ instead of having $\mathcal{C}^{-1}$ applied to it with random variable math.

As a result of this issue, previous methods produce skewed ages whenever the cratering rate is not constant. In the *cratersfd* package, we include the code to apply $\mathcal{C}^{-1}$ to an $N(1)$ PDF (or scale it by $\mathcal{C}^{-1}$). With this final step, we are ready to bring together all we have discussed to produce a solution for the Age Problem that incorporates all the sources of error with random variable math.

2.9 Final Age Problem solution

Combining all the error factors (Table 1), the solution for the Age Problem is:

$$T = \mathcal{C}^{-1} \left(\frac{\left(\mathcal{E}_\lambda = \mathcal{E}_a \mathcal{E}_s \left(\frac{\sum_{i=1}^{N_{\text{observed}}} \mathcal{E}_{r_i}}{N_{\text{observed}}} \right) \right) \lambda}{A} \left(\frac{p(1\text{km})}{p(\mathcal{E}_{D_s} D_{\min})} \right) \mathcal{E}_c \right), \quad (11)$$

where T is age, $\mathcal{C}^{-1}$ is the inverse chronology function, p is the production function, λ is the model count, N_{observed} is the observed count, A is the area, $\mathcal{E}_{D_s}$ is the systematic diameter measurement error kernel, $\mathcal{E}_c$ is the chronology function error kernel, $\mathcal{E}_a$ is the additional error factors kernel, $\mathcal{E}_s$ is the systematic human crater identification error kernel, the $\mathcal{E}_{r_i}$ are the individual random human crater identification errors for each crater, and $\mathcal{E}_\lambda$ is the combined error kernel for λ . (For our recommended values of the error kernels, see Table 1.) This equation is evaluated using random variable math (Appendix A) because λ , the error kernels, and sometimes A are random variables.

Assuming that the $\mathcal{E}_{r_i}$ are identical, p_λ , the percentage error of $\mathcal{E}_\lambda$, can be found by:

$$p_\lambda = e^{\sqrt{\ln \left(\frac{e^{(\ln(1+p_r))^2} - 1}{N_{\text{observed}}} + 1 \right) + (\ln(1+p_s))^2 + (\ln(1+p_a))^2}} - 1, \quad (12)$$

where p_r , p_s , and p_a are the percentage errors for random human crater identification error, systematic human crater identification error, and additional error respectively. The PDF of λ , which can be found with Equation (6), depends on the PMF of model N , which in turn depends on N_{observed} and the random diameter measurement kernel $\mathcal{E}_{D_r}$ (see Appendix A.2).

To help the reader understand how to implement our methods, we show worked examples in Sections 5.1–5.3. To enable researchers to apply our methods to their data, we once again refer the reader to our *cratersfd* code, available at <https://github.com/samwbell/cratersfd>. In *cratersfd*, we have incorporated the math in this section into an `age_pdf(ds, area, dmin)` function, which takes an array of diameters, the area A , and $D_{\min}$ as required inputs and calculates the PDF of the random variable age. The *cratersfd* code stores random variables in `RandomVariable` objects whose PDF can be plotted with a `.plot()` method, and `age_pdf()` can take `RandomVariable` objects for both A and $D_{\min}$. With keyword arguments of `age_pdf()`, users can specify the lognormal error kernels, the production function, and the chronology function.

With one line of code, researchers can calculate the full PDF of age from crater counts, incorporating the sources of error we describe above: Poisson counting statistics, human crater identification error, diameter measurement error, chronology system error, area error, $D_{\min}$ selection error, and additional sources of error. This represents a significant improvement over traditional methods (e.g. Neukum, 1983; Michael et al., 2016; Robbins et al., 2018), which treated counting statistics as the only source of error and made the math error we describe in Section 2.8.

Our Age Problem methods still rely on plotting the data to select $D_{\min}$ and ensure that N is not decreased by saturation, erosion, or resolution effects. More fundamentally, it is essential to first plot the data to understand them, and plots of crater SFDs contain critical information about sources of impactors and the bombardment histories of planetary bodies. To understand our solution to the Plotting Problem, however, we must first address the Slope Problem. Fundamentally, the Plotting Problem is the problem of approximating an SFD from limited samples. Before we tackle the more complex case of an SFD whose slope on a log-log plot changes, we must first tackle the simpler case where the slope on a log-log plot is constant. We therefore turn first to the Slope Problem.

Source of error	Variable	Distribution	Recommendation	Section
Counting statistics	λ	Gamma	Equation (6)	2.1
Random human crater identification error	$\mathcal{E}_{r_i}$	Lognormal	60%	2.2
Systematic human crater identification error	$\mathcal{E}_s$	Lognormal	12%	2.2
Random diameter measurement error	$\mathcal{E}_{D_r}$	Lognormal	5% (10% near equilibrium)	2.3
Systematic diameter measurement error	$\mathcal{E}_{D_s}$	Lognormal	1% (2% near equilibrium)	2.3
Chronology system error	$\mathcal{E}_c$	Lognormal	Moon: factor of 2.5; Mars: factor of 3	2.5
Area error	A	Lognormal	If present, estimated by user	2.6
$D_{\min}$ selection error	$D_{\min}$	Defined by user	Estimated by user	2.7
Additional sources of error	$\mathcal{E}_a$	Lognormal	10%	2.8

Table 1 | A summary of sources of error in the Age Problem. For each source of error, we show: the random variable that quantifies it, the distribution it takes, our recommendation, and the section where we discuss it. For our recommendation, we show the default values, but we encourage researchers to estimate these values for their data, guided by the discussion in the appropriate section of this work.

3 The Slope Problem

In the Slope Problem, we ask the question: What is the slope of a crater SFD on a log-log plot? In Section 3.1, we begin with a simple case: an open interval stretching from D_{min} to infinity with one constant slope (a straight line in log-log space). In Section 3.2, we address the trickier case of a closed interval from D_{min} to D_{max} . Finally, in Section 3.3, we address sources of error.

3.1 An open interval

Crater SFDs are themselves distributions. They give the PDF of the random variable diameter. With a constant slope in log-log space over an open interval, a crater SFD follows a well-known distribution in mathematics called the Pareto distribution.

The PDF of the Pareto distribution of diameter D takes the form:

$$P(D) = \frac{\alpha D_{min}^\alpha}{D^{\alpha+1}}. \quad (13)$$

Starting from the Pareto distribution, in Appendix D we derive an explicit expression for the PDF of α that depends only on the observed crater diameters D_i :

$$P(\alpha) = \frac{\alpha^N \left(\sum_{i=1}^N \ln \left(\frac{D_i}{D_{min}} \right) \right)^{N+1} \prod_{i=1}^N \left(\frac{D_{min}}{D_i} \right)^\alpha}{\Gamma(N+1)}. \quad (14)$$

This expression follows a Gamma distribution (Equation (6)), with:

$$x = \alpha, \quad s = N + 1, \quad \text{and} \quad r = \sum_{i=1}^N \ln \left(\frac{D_i}{D_{min}} \right). \quad (15)$$

From the theorems of the Gamma distribution, we can calculate the mean:

$$\mu(x) = \frac{s}{r} \rightarrow \mu(\alpha) = \frac{N+1}{\sum_{i=1}^N \ln \left(\frac{D_i}{D_{min}} \right)} \quad (16)$$

and the mode (maximum likelihood value):

$$\hat{x} = \frac{s-1}{r} \rightarrow \hat{\alpha} = \frac{N}{\sum_{i=1}^N \ln \left(\frac{D_i}{D_{min}} \right)}. \quad (17)$$

To quantify the improved accuracy of our PDF approach, we performed synthetic modeling on least squares fits to Arvidson et al. (1979) cumulative plots and the Robbins et al. (2018) method (Figures 4 and 5). With least squares fits to cumulative plots, we found a significant bias towards shallow slopes, especially in the “binned” presentation. The shallow slope bias was especially strong when all points were treated equally. Incorporating error bars reduced it, but it remained significant at low N . (For details on the synthetic modeling, see Appendix E.)

Instead of the full PDF, Robbins et al. (2018) approximated it with a normal distribution around the full PDF’s mode (maximum likelihood), which they referred to as the Maximum Likelihood Estimator (MLE), $\hat{\alpha}$. Like Arvidson et al. (1979), Robbins et al. (2018) used the Central Limit Theorem to calculate symmetric error bars of $\sigma = \hat{\alpha}/\sqrt{N}$. Their MLE method produced the lowest random error (Figure 5a). However, because the PDF is not symmetric at low N , and the MLE does not match the mean at low N , the MLE is biased toward steeper slopes at low N (Figures 4 and 5b). In Appendix F, we derive this bias analytically and show that it exactly predicts the synthetic results (Figure 5b).

The apparent mismatch between Figures 4 and 5 and Robbins et al. (2018)’s Figure 11 merits a brief comment. In their figure, they showed that MLE was significantly less biased than least-squares. However, that work was done with a differential SFD rather than cumulative SFD; additionally, they included a finite D_{max} , while here we do not. Those differences can account for this apparent contradiction between these two works.

To avoid the biases inherent in approximating the PDF with a central value and error bars, where possible, we recommend plotting the full Equation (14) PDF. If a single value must be chosen, we recommend indicating the asymmetry with asymmetric error bars (see Appendix C.1).

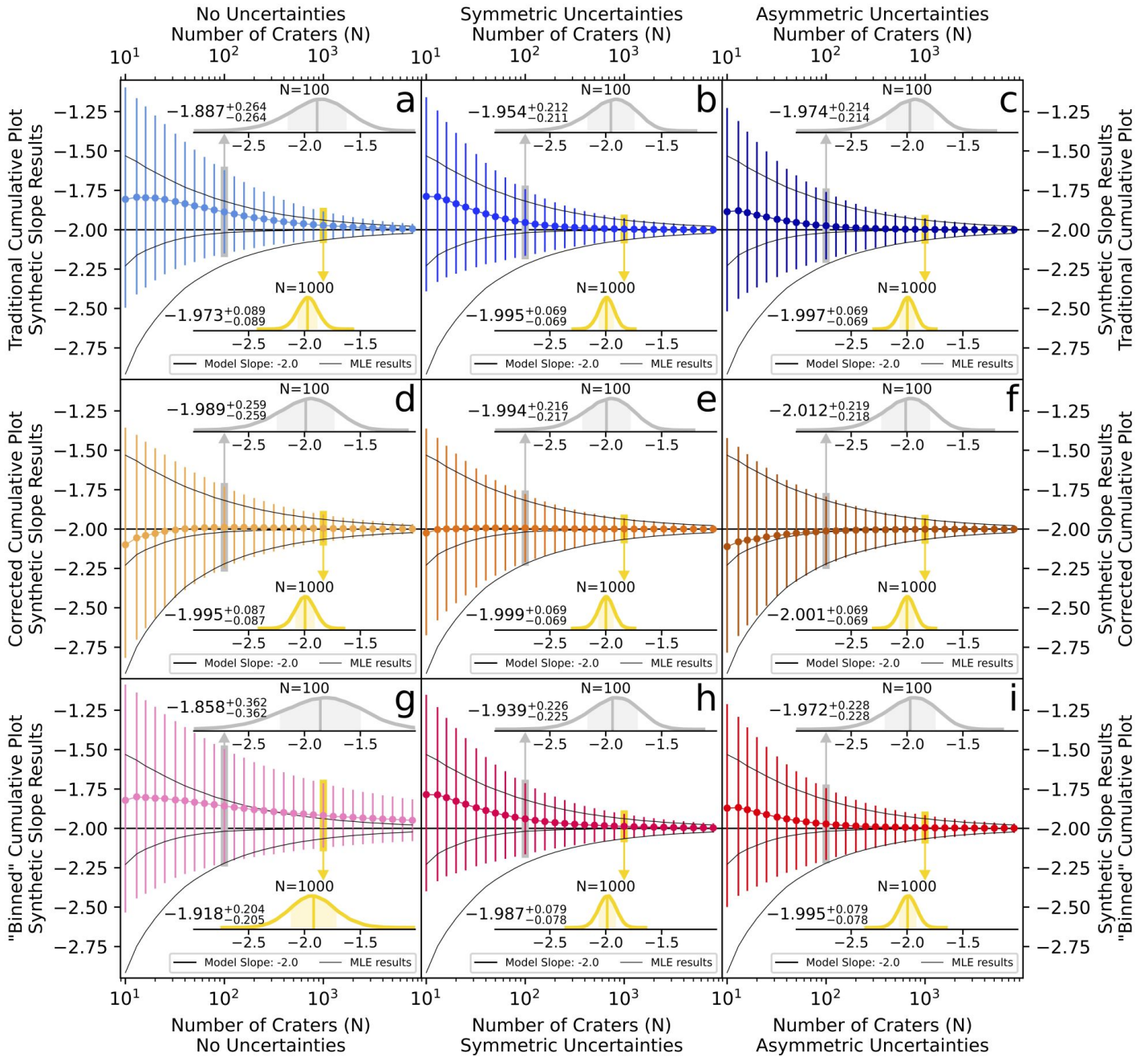


Figure 4 | A grid of synthetic results, showing the biases and uncertainties in measuring slopes from least squares fits of nine versions of the cumulative plot, compared with the results from the MLE method of Robbins et al. (2018). Each point shows the mean slope measured from 100 000 synthetic datasets of N craters, with 1σ -equivalent error bars (see Appendix C). At $N = 100$ and $N = 1000$, we show the full form of the empirical PDF of the 100 000 slope results. In each case, the model cumulative slope is -2.0 ($\alpha = 2$), and there is no maximum diameter. Note how the MLE produces a bias towards steeper slopes, due to the difference between the maximum likelihood slope and the mean slope (see Appendix F). Because of its biased sampling (Section 4.1 and Figure 6), the traditional Arvidson et al. (1979) unbinned cumulative plot produces artificially shallow slopes (a). Symmetric error bars reduce this bias (b), and asymmetric error bars reduce it further (c), but they do not eliminate it. When we correct the biased sampling (Figure 6f), we see steeper slopes (d). With symmetric error bars on the corrected plot (e), the biases nearly cancel out, and with asymmetric error bars (f), the corrected plot gets very close to the MLE results. Binned cumulative plots produce the most biased results with the highest uncertainty (g), although error bars (h and i) reduce the bias and uncertainty.

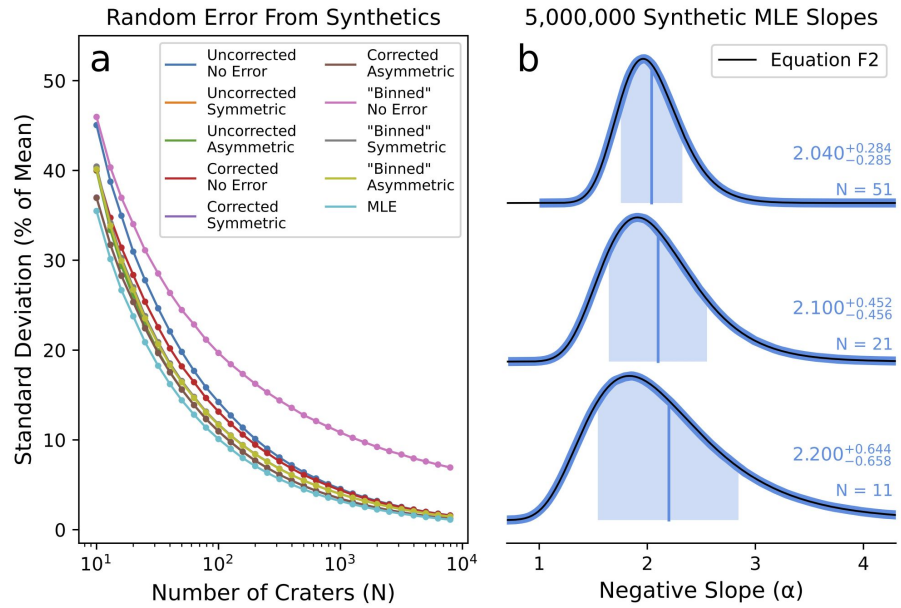


Figure 5 | (a) The random error of the results in Figure 4. For each N , we show the standard deviation of the results of each method as a percentage of the mean. The MLE method of Robbins et al. (2018) has the lowest random error – consistent with their figure 11 – while the method almost all papers published in the last 50+ years have used is shown in the pink “Binned” No Error” data. (b) The empirical PDFs of MLE results at three different N values, showing how Equation (F2) exactly matches the results. Note how the mean of the distribution of MLE results is $\alpha N / (N - 1)$ (Equation (F4)).

3.2 A closed interval

For a closed interval from D_{min} to D_{max} , the problem becomes slightly more complex. When α is constant over an interval, the crater SFD on a differential plot follows the PDF of the Truncated Pareto distribution:

$$P(D) = \frac{\left(\frac{\alpha}{D}\right) \left(\frac{D_{min}}{D}\right)^\alpha}{1 - \left(\frac{D_{min}}{D_{max}}\right)^\alpha} \text{ for } \alpha \neq 0, \text{ with } P(D) = \frac{1}{D \ln\left(\frac{D_{min}}{D_{max}}\right)} \text{ at } \alpha = 0. \quad (18)$$

Here, the expression for the relative probability distribution of α is (Appendix D):

$$P(\alpha) \propto \frac{\alpha^N \prod_{i=1}^N \left(\frac{D_{min}}{D_i}\right)^\alpha}{\left(1 - \left(\frac{D_{min}}{D_{max}}\right)^\alpha\right)^N}. \quad (19)$$

Formally, PDFs are normalized so that the integrated probability sums to 1. Because Equation (19) does not match a well-known distribution like the Gamma distribution, and its integral does not have a simple closed-form solution, we normalize it with numerical integration. We refer to this slope estimation method as the Truncated Pareto method.

As with the (untruncated) Pareto method’s α PDF, Robbins et al. (2018) approximated the Truncated Pareto method’s α PDF with symmetric error bars around the mode (or MLE). When the number of craters is small or the interval is narrow, the form of the Truncated Pareto method’s α PDF over a closed interval can be much more asymmetric than the Pareto method’s α PDF, so the truncated MLE can have a much stronger bias and much greater uncertainty. Although they did not report these effects, Robbins et al. (2018) ran a synthetic model that showed much higher MLE bias and uncertainty over a truncated interval. As a result, with the Truncated Pareto method, it is especially important to plot the full PDF instead of approximating it. If a single value and error bars must be chosen, we recommend using the mean method (see Appendix C).

3.3 Sources of error

Currently, Slope Problem solutions only address the error of counting statistics. However, each of the additional sources of error that apply to the Age Problem affects the Slope Problem, too. Incorporating

them into the Slope Problem matters. While it lies beyond our scope here, we strongly encourage future work to develop the necessary math to address this critical issue.

In addition to the sources of error that affect the Age Problem, the Slope Problem relies on a critical assumption: the SFD follows a Pareto (or Truncated Pareto) distribution across the diameter interval. To assess the validity of this assumption, we recommend plotting the SFD. As Clauset et al. (2009) noted, statistical tests like goodness-of-fit metrics require high N and can miss meaningful deviations from a Pareto (or Truncated Pareto) distribution. Testing this assumption by plotting the SFD, however, requires a plotting method that reliably can image deviations from a straight line on a log-log plot without requiring very high N .

We therefore turn to the Plotting Problem, where we use our Slope Problem methods to develop a much improved plotting method, the SASH model (Section 4.3).

4 The Plotting Problem

We begin our treatment of the Plotting Problem by examining two solutions from traditional methods: the cumulative plot (Section 4.1) and the binned differential plot and its close cousin, the binned R plot (Section 4.2). In each case, we document significant drawbacks to these traditional methods. In Section 4.3, we present our solution: the SASH model. In Section 4.4, we discuss the error of the SASH model. Finally, in Section 4.5, we discuss how to understand the SASH model as it extends beyond the largest crater.

4.1 The cumulative plot

Although frequently used, Arvidson et al. (1979) cumulative plots have considerable problems. Crater SFDs are distributions of the random variable diameter. When we want to understand the distribution of a random variable, typically what we want is the PDF, but cumulative plots do not show the PDF. As Chapman and Haefner (1967) observed, because each part of the cumulative plot depends on the portion to its right, slope variations distort the plot so that it does not show the value of the PDF at each diameter. As a simple thought experiment to illustrate one way the cumulative SFD distorts shape, one could take their craters, create a cumulative SFD, and then remove the largest 10 craters and overplot the cumulative SFD for the revised craters. The slope at largest few bins will be different in the new, $N-10$ sample than it is at those diameters for the full sample. For this reason, we discourage the use of cumulative plots.

Our synthetic results (Figure 4), which show a consistent bias in the cumulative plot towards shallow slopes, raise deeper concerns. This shallow slope bias is especially noteworthy given that least squares, as an MLE method, should show the steep slope bias of the Robbins et al. (2018) method. To explain these results, we have identified a significant bias in the Arvidson et al. (1979) version of the cumulative plot. A cumulative count exists at all diameters, not just the diameters where there are craters. The true cumulative function (Figure 6c) forms a step function that covers all diameters, not just the crater diameters. However, Arvidson et al. (1979) visualized the cumulative plot as a series of points, with each point shown at the right-hand side of each step, equivalent to the crater diameter (Figure 6d). The right-hand side of each step is biased to the right. For larger craters with wider steps between the craters, the bias towards larger diameters will be larger than for the smaller craters. As a result, a slope fit to these points will be artificially shallow (Figure 6e). In our Slope Problem synthetic model, when we correct for this effect by sampling the cumulative function at the middle of each step (which is the geometric mean because the plot is in log-log space) (Figure 6f), the shallow slope bias disappears (Figure 4f).

Frequently, the Arvidson et al. (1979) cumulative plot is used in what Arvidson et al. (1979) term a “binned” form. Because the cumulative plot is cumulative, this does not involve grouping the data into bins. Instead, the cumulative function is sampled at regular intervals in log space, so it would be more accurate to refer to it as “sampled” rather than “binned.” This distinction matters. Calling the plot “binned” implies that it is showing only the data within the bins, when it is actually just giving the cumulative count at a particular point. While this form avoids the biased sampling of the “unbinned” form, it performs even worse in our synthetic tests, with the bias towards shallow slopes continuing to very high N (Figure 4). Adding in error bars (Figures 4h and 3i) reduces the bias, especially at higher N , but it is still more biased than the traditional “unbinned” plot with symmetric error (Figure 4b). The bias in the “binned” plot is heavily due to the standard practice of not including “bins” with no additional craters in them, which effectively removes points that would make the slope steeper (Figure 6j), producing shallower results (Figure 6k). When the zero-crater points are added back in (Figure 6l), this bias is reduced (Figure 6m). The “binned” presentation also loses the information within the “bins”, leading to a wider error distribution (Figures 4 and 6).

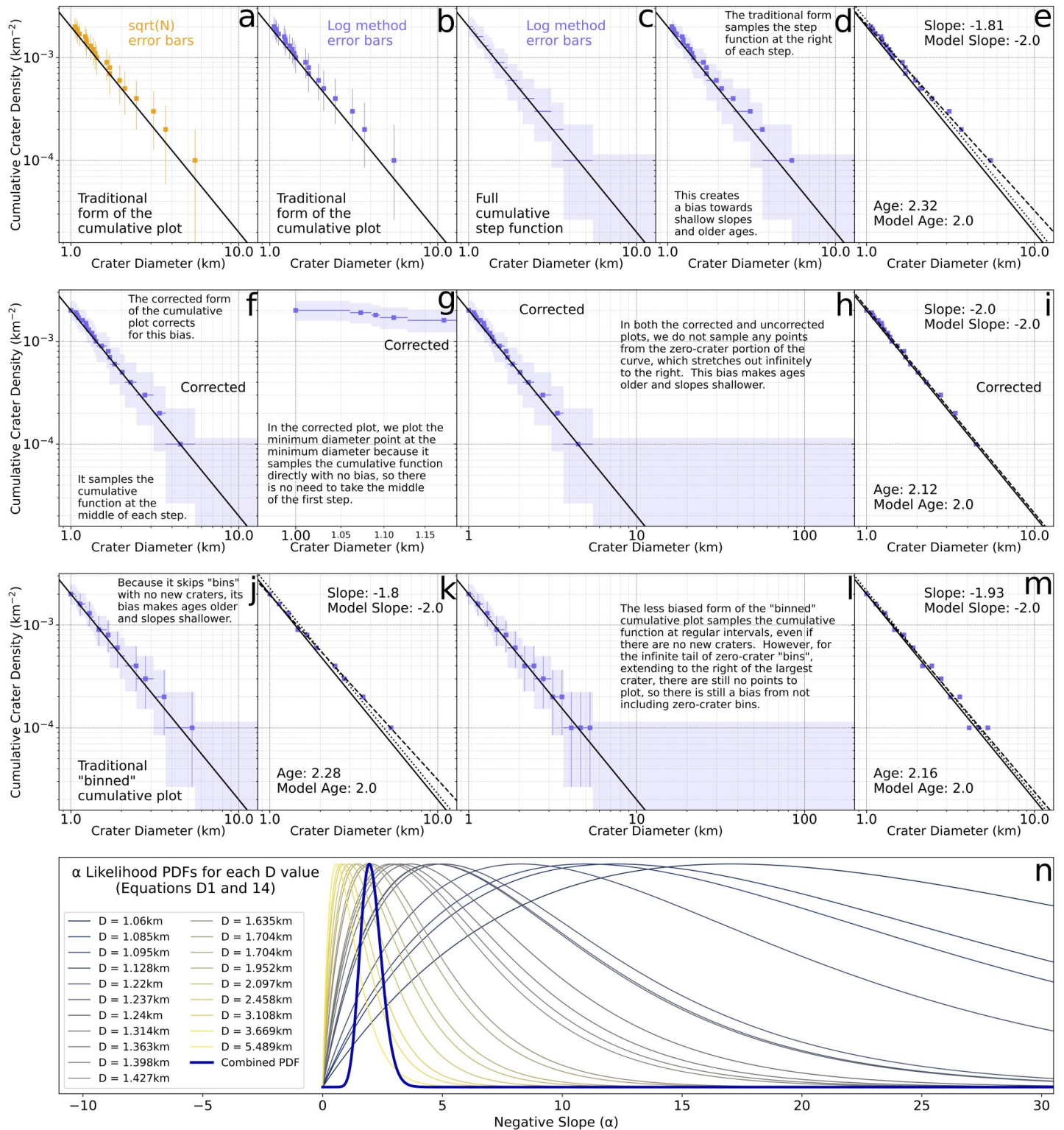


Figure 6 | Different methods for plotting the data from an example synthetic. (a) The traditional form of the cumulative plot with $\sqrt{N}$ approximation error bars. (b) The traditional form with log method error bars. (c) The full cumulative step function (our recommended form of the cumulative plot). (d) The biased sampling of the Arvidson et al. (1979) version of the cumulative plot, where the point is taken at the right of each step. (e) An example slope and isochron fit, showing a shallow slope bias and older age bias. (f) Sampling in the middle of each step to correct the bias. (g) Showing how there is a point at the minimum diameter, not the middle of the first step. (h) Showing how the cumulative plot extends indefinitely to infinity. (i) A fit to the corrected version of the cumulative plot, reducing the slope bias. (j) How the "binned" plot is constructed, skipping zero-crater samples. (k) The biased slope fit to the "binned" plot. (l) The "binned" plot with zero-crater bins included. (m) The fit to the "binned" plot often improves when the zero-crater samples are included. (n) Finding the α (negative slope) likelihood PDF by combining the α likelihood PDFs of the individual observations (Equation (D1) gives the unnormalized form, and Equation (D7) gives the normalized form) to produce a combined PDF (Equation (14)).

If cumulative plots are used, we strongly recommend plotting the full continuous step function, with error bars shown through shading (Figure 6c). For the error bars, we discourage the $\sqrt{N}$ approximation of Arvidson et al. (1979) and instead recommend the far more accurate log method we have developed (see Appendix C; included in our code distribution). However, we do not encourage the cumulative presentation. Instead, we recommend the differential plot, which shows the shape of the PDF.

4.2 Binned differential and R plots

The differential plot avoids many of the challenges of a cumulative plot. Mathematically, the differential plot is the PDF. However, instead of being normalized so that it integrates to 1, it is normalized to integrate to N/A , the total number of craters divided by the count area. Unlike the cumulative plot, the differential plot does show the probability of observing a crater of a particular diameter. The steeper slopes of the differential plot can make it harder to make out slope changes visually, but Arvidson et al. (1979) successfully addressed this issue with the R (relative) plot. In the R plot, the differential plot is multiplied by the diameter cubed, so a straight line represents a differential slope of -3 ($\alpha = 2$). In an R plot, shallow ($\alpha < 2$) SFDs slope upwards from small to large craters, and steep ($\alpha > 2$) SFDs slope downwards from small to large craters, making it easier to see slope changes with geological implications.

To plot the PDF, we have to estimate it. In the Arvidson et al. (1979) approach, the differential plot is a simple histogram. It breaks the SFD into bins and plots the number of craters in each bin, divided by the bin width, divided by area. Although it is shown with individual points at each log space bin center (the geometric mean of the bin edges), these points are calculated with the assumption that the PDF is constant across the bin – and it is not. In fact, because of the exponential shape of crater SFDs, the PDF often varies dramatically across the bin, and this affects the value at the bin center (Michael, 2013). If we know α – and assume that α is constant across the bin – then the PDF follows the Truncated Pareto distribution (Equation (18)), and we can calculate the value at the bin center. However, α is rarely constant across the bin, and we cannot know α without already knowing the SFD shape.

Moreover, binned differential and R plots are biased at the large crater end. In the thirteen NPF synthetics in Figures 7 and 8, the bins on the right consistently plot above the production function. Zero-crater bins create this bias. When there is no crater in the bin, it cannot be plotted on a log-log plot, so that bin must be omitted. When that bin is omitted, however, the surrounding points plot higher than they would have if the information from the zero-crater bin had been averaged in. Because there are always zero-crater bins to the right of the largest-diameter bin, the final bin in every binned differential and R plot is biased upwards. In Figure 13c, we show how much lower the final bin would plot if we extended it to include the zero-crater bins.

4.3 The Slanted Average Shifted Histogram (SASH) model

The differential plot shows an unnormalized PDF. Constructing a PDF from a limited number of samples is a well-known problem in math, so we turn to modern statistical methods for approaching this problem. Robbins et al. (2018) proposed using the Kernel Density Estimator (KDE), but they cautioned that this method broke down both for low- N bins and at the minimum diameter. To address these problems, we present an improved statistical method, the Slanted Averaged Shifted Histogram (SASH).

The standard Averaged Shifted Histogram (ASH) algorithm (Scott, 1985) is straightforward. First, a histogram is constructed. Then, the bins are shifted multiple times, and the resulting histograms are averaged together. Each histogram starts out as a bar graph, with a constant value across each bin, but when they are averaged together, they approximate the true functional form of the PDF. Although a powerful method, ASH still loses out on the information gained from the distribution of the points within the bin. In the SASH algorithm, we capture this information by using the Truncated Pareto method to estimate α for each bin, averaging slanted histograms instead of bar graphs. Then, we iterate four more times, recalculating the SASH model using the results of its previous iteration to provide more accurate estimates of the bins' α values (see Appendix G and Figure 7 for details of the SASH algorithm).

In synthetic tests, the SASH model significantly outperforms both binned differential plots (Arvidson et al., 1979) and the KDE (Robbins et al., 2018), plotting the full PDF curve and reliably imaging features in low- N bins that previous methods miss (Figures 7 and 8). We strongly recommend using the SASH model to plot the full continuous form of differential and R plots.

To make the SASH model easy to use, we have wrapped the algorithm into two functions in *cratersfd*, `plot_sash()` and `plot_sash_R()`. We intend *cratersfd* to provide a comprehensive system for analyzing crater counts, so it also allows the user to plot crater counts with Arvidson et al. (1979) legacy plots. As a Python package, *cratersfd* can be easily integrated into users' codes, and its plots can be customized with Matplotlib and exported to common graphics formats.

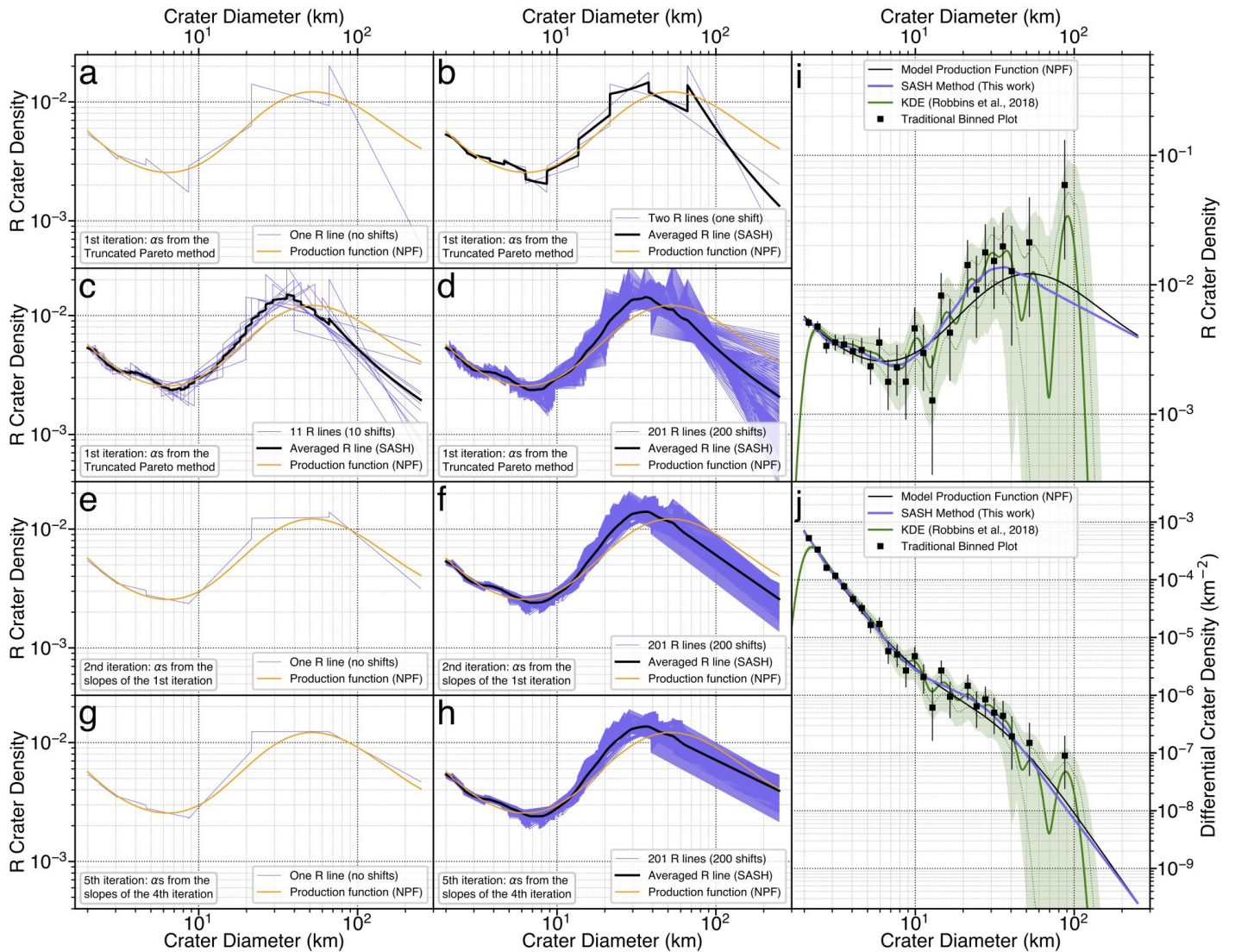


Figure 7 | The SASH model applied to synthetic crater counts generated from the “new” NPF (Neukum, 1983; Ivanov et al., 2001) with parameters: $N = 500$, $D_{\min} = 2$ km, and $D_{\max} = 250$ km. Here, the SASH model uses an initial bin width of 18 per decade and a bin growth rate of $2^{1.3}$. See Appendix G for the details of this process. (a) The line of R plot values for one set of bins in the SASH model. Because this is the first iteration, we determine the a value from the mean of the a likelihood distribution from the Truncated Pareto method (Equation (19)). (b) The R lines from two sets of bins and their average. (c) With 11 sets of bins (ten shifts), the average matches the production function much better than a single set of bins. (d) The average of 201 sets of bins after 200 shifts. (e) The R line from the same bins as in (a) but in the second iteration, where we determine the a values by measuring the slope of the results from the first iteration. These slopes provide a more accurate match to the production function. (f) The results of the second iteration after all 200 shifts. (g) The R line from the same bins as in (a) and (e) in the fifth and final iteration, where we determine the a values by measuring the slope of the results from the fourth iteration. (h) The final SASH model after all 200 shifts in the fifth iteration. (i) A comparison of the SASH model, a traditional Arvidson et al. (1979) binned R plot, and the Robbins et al. (2018) Kernel Density Estimator (KDE) method. The SASH model resolves the SFD much more clearly, producing an SFD that much more closely matches the production function. (j) The same as (i) but in differential plot form.

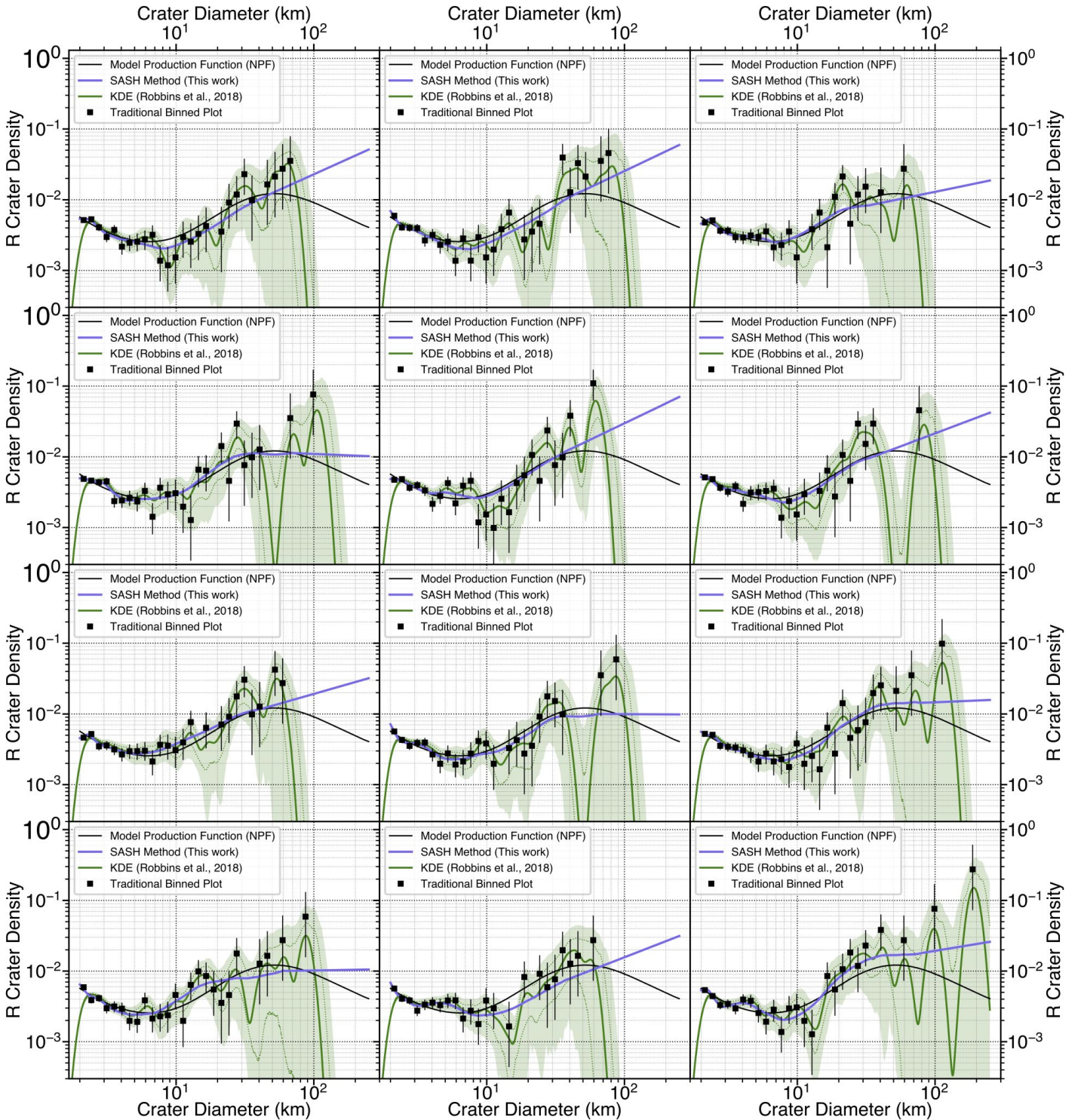


Figure 8 | A comparison of the SASH model, a traditional Arvidson et al. (1979) binned R plot, and the Robbins et al. (2018) Kernel Density Estimator (KDE) method for twelve randomly selected synthetic crater counts generated from the “new” NPF (Neukum, 1983; Ivanov et al., 2001) with parameters: $N = 500$, $D_{\min} = 2$ km, and $D_{\max} = 250$ km. As in Figure 7, the SASH model uses an initial bin width of 18 per decade and a bin growth rate of $2^{1.3}$. In each case, the SASH model consistently recovers the model production function much more accurately than previous methods. However, the results are still limited by the data. For instance, without enough $D \geq 52$ km craters, the SASH model cannot detect the NPF’s $D_{\max} = 250$ km steep branch, so it assumes that the SFD continues past the largest crater at a shallower slope projected from the NPF’s $6 \text{ km} \leq D \leq 52 \text{ km}$ shallow branch. The information from the absence of craters between the largest crater and $D_{\max}$ can only provide an upper constraint, not a lower one.

4.4 Error of the SASH model

Because it shows the full form of the curve, not discrete points at individual bins, the SASH model by definition cannot have a definitive error model. The possibility of a deviation from the SFD shape the SASH model observes depends heavily on the shape of the potential deviation. We cannot give the error of the SFD at any given diameter because it depends on the universe of possible SFD curve shapes.

As formulated by Arvidson et al. (1979), error bars on binned differential plots rely on the assumption that the SFD is constant across the bin. With Michael (2013)'s correction, they rely on the assumption that α is constant within the bin. Most importantly, they assume that the error confines itself neatly to one bin and does not vary within that bin. This is virtually never the case. Fundamentally, error analysis is about testing the chance of a different result, and error analysis of an SFD shape is about testing the chance of a different SFD shape. For instance, let's say we observe a steep branch at the low- N large-crater end of a distribution in equilibrium saturation. Instead of assuming that it has left equilibrium, now reflects the production function, and can be used for dating, it is important to test whether it could be a spurious feature, and the SFD might actually continue at an equilibrium slope of $\alpha = 2$. Such a steep branch typically stretches across multiple bins. We also often want to test whether a portion of the SFD matches a production function – and here the perturbations from the SFD would stretch across all the bins. Rarely if ever do we want to test a model where the SFD shape is perturbed from the observed SFD equally across a specific diameter range that happens to exactly match the bin in question.

Because binned differential and R plots assume a bin width, they assume a known width of possible perturbations. If narrower bins were used, the error bars would be wider. If wider bins were used, the error bars would be narrower. With the SASH model, error is highly width-dependent. A narrow perturbation is much likelier than a wider one. If we want to know if a bump on the SFD from 10 km to 15 km was robustly observed, the error bars will be much larger than if we want to know whether a shallow branch between 1 km and 50 km was robustly observed.

For this reason, any method that draws an error envelope will not accurately capture the error, leading to misleading interpretations. This is why we discourage plotting the range of SASH model results, which we show in Figure 7h for the sole purpose of explaining how the SASH model works. It could be confused for an error envelope, which would be misleading for many reasons. It would show spurious spikes that the SASH model averages out, and it would not capture the error of counting statistics or slope (α) estimation. Most fundamentally, it would not address the shape of a perturbation to the SFD, so it would erroneously treat narrow perturbations the same as wider ones of the same magnitude.

Because the error depends on the specific shape of the perturbed SFD, we have to measure the error for the specific shape of the perturbed SFD. To understand the error of the SASH algorithm, we therefore recommend testing possible other SFD shapes with a synthetic model. To do this, first we define a possible alternative for what the true SFD could be. From this alternative, we randomly generate a large number of synthetic datasets of possible observed craters. Then, we apply the SASH model to each synthetic dataset, producing a synthetic dataset of possible SFDs that the SASH model would have observed. From the values of these synthetics at each diameter, we then generate an error envelope of confidence intervals ("error bars") at 1σ -equivalent percentiles. We are effectively creating a Monte Carlo envelope of what the SFD would be if a proposed model were true, given a finite number of craters. (See Appendix G.1 for details on the synthetic modeling, Appendix E for details on how our code generates synthetics, and Section 5 for examples of how to apply this synthetic modeling approach.)

Because it allows us to test specific hypotheses for what the SFD might be, our synthetic modeling approach represents a powerful improvement. With previous methods (Arvidson et al., 1979), error bars would be made for individual bins, and analysts would have to make a rough visual assessment of whether a hypothesis was consistent with the error bars. This rough visual assessment can often be highly inaccurate. For instance, because narrower bins have wider error bars, a statistically robust feature that stretches across many bins could erroneously appear to fall within the margin of error. An SFD without the feature might plot within the wider error bars of individual bins – but these error bars would only be wider because they cover a much narrower diameter range than the feature being tested. By testing hypotheses directly, instead of comparing them to error bars, we can much more accurately assess whether they are consistent with the data.

When interpreting SASH model SFDs, we strongly recommend running and plotting a synthetic model. Before concluding that a feature is robust, test if it could be spurious. Before concluding that a feature is absent, test if it could have failed to show up even if it were actually there. Most importantly, to ensure that others can run synthetic models to test their own hypotheses, we strongly recommend making the raw data available online free of charge in an independent, easy-to-access repository (as journals and funders are increasingly requiring).

Currently, our synthetic model only includes Poisson counting error, and it is critical to remember this limitation when interpreting the results. With future work, other error factors should be included. As we discuss in Appendix G.1, the math we develop in Section 2 for the Age Problem can be applied to the synthetic model, but this creates computational challenges that require future work to address.

4.5 The SASH model beyond the largest crater

For diameters larger than the largest crater, we must take special care. Because the information from the absence of craters affects the SFD at the large crater end, we calculate the SASH model beyond the largest crater. As we discuss in Section 4.3, excluding this information biases binned plots towards artificially high values for the bin with the largest crater. However, we must be very careful about interpreting parts of the SFD where we do not have any craters.

By default, the `calc_sash()` function in *cratersfd* assumes that the largest bin extends out to a D_{max} of ten times the diameter of the largest crater – effectively assuming an open-ended interval. In most practical applications of crater counting, however, we can only observe an absence of craters out to a maximum diameter. At the extreme end, we are unlikely to observe a crater whose diameter exceeds the dimensions of the count area, or for global counts, a diameter that significantly exceeds the diameter of the body. More subtly, we are unlikely to choose a count area that is dominated by one large crater – and if we did, it would probably skew the SFD through overprinting.

Although D_{max} is typically negligible in the Age Problem, it can affect the Plotting Problem in a non-negligible way. To address this reality, we strongly encourage estimating D_{max} when measuring the shape of the SFD at the large crater end and specifying it with the `d_max` keyword argument if it is smaller than `calc_sash()`'s default. To estimate D_{max} , it is important to assess the crater counter's tolerance for including a larger crater in their count areas. There are no fixed rules for how to make this judgment call, and in practice this tolerance level varies wildly.

Between the largest crater and D_{max} , the PDF that the SASH model estimates largely assumes that the same slope observed at smaller diameters continues out to larger diameters, but the error is asymmetric. The information from the absence of craters constrains an upper limit but not a lower limit. The fact that we do not observe craters can be strong evidence that the SFD does not turn upwards beyond the largest crater, but it cannot provide a lower bound. As Figure 8 shows, the SFD can turn downwards into a steep branch beyond the largest crater. To properly model these errors, we strongly recommend the synthetic modeling approach we describe in Section 4.4.

A downside of the SASH model is that it does not show the specific crater diameters in the plot. This information can be especially meaningful at the large crater end. To address this, we have created an optional form of the SASH model plot where the craters are shown with tick marks (Figures 10a–10b). Because these tick marks can crowd the plot at lower diameters, sometimes it is clearer to only show the largest ~10 craters (Figures 13d–13i).

With this final option, the SASH model represents a clear improvement over existing solutions to the Plotting Problem, allowing a much more accurate assessment of a crater SFD from a limited sample of craters. Theoretical solutions, however, are not enough. To fully demonstrate our methods, we turn now to practical examples.

5 Examples

In Section 5.1, we begin with an Age Problem example, showing how the age PDFs from two crater counts of Canala Crater become much more consistent when sources of error beyond counting statistics are incorporated. Because the Age Problem begins with plotting the data and analyzing them, this also serves as a SASH problem example. In Section 5.2, we provide a practical guide to the calculation. In Section 5.3, we visualize the random variable math of the Age Problem. In Section 5.4, we show how to use the SASH model to analyze crater SFDs with an example that demonstrates its explanatory power: the cratered plains of the moons of Saturn.

5.1 Two ages for Canala Crater

Our goal with Age Problem examples is two-fold: (1) to demonstrate how to perform the counts and (2) to show how why it is necessary to include sources of error beyond counting error. For our first Age Problem example, we turn to Canala Crater, a young, fresh $D = 11.9$ km crater near Labeatis Fossae on Mars (Figure 9a). S. J. Robbins collected two counts of craters superimposed on the ejecta deposits (Figure 10a): one from a CTX image of the whole ejecta and one from a higher-resolution HiRISE image that only overlaps part of the ejecta (Figure 9b).

Canala presents a number of challenges that increase the error of crater identification. To exclude buried craters that predate the impact (and therefore cannot be used for dating), S. J. Robbins visually

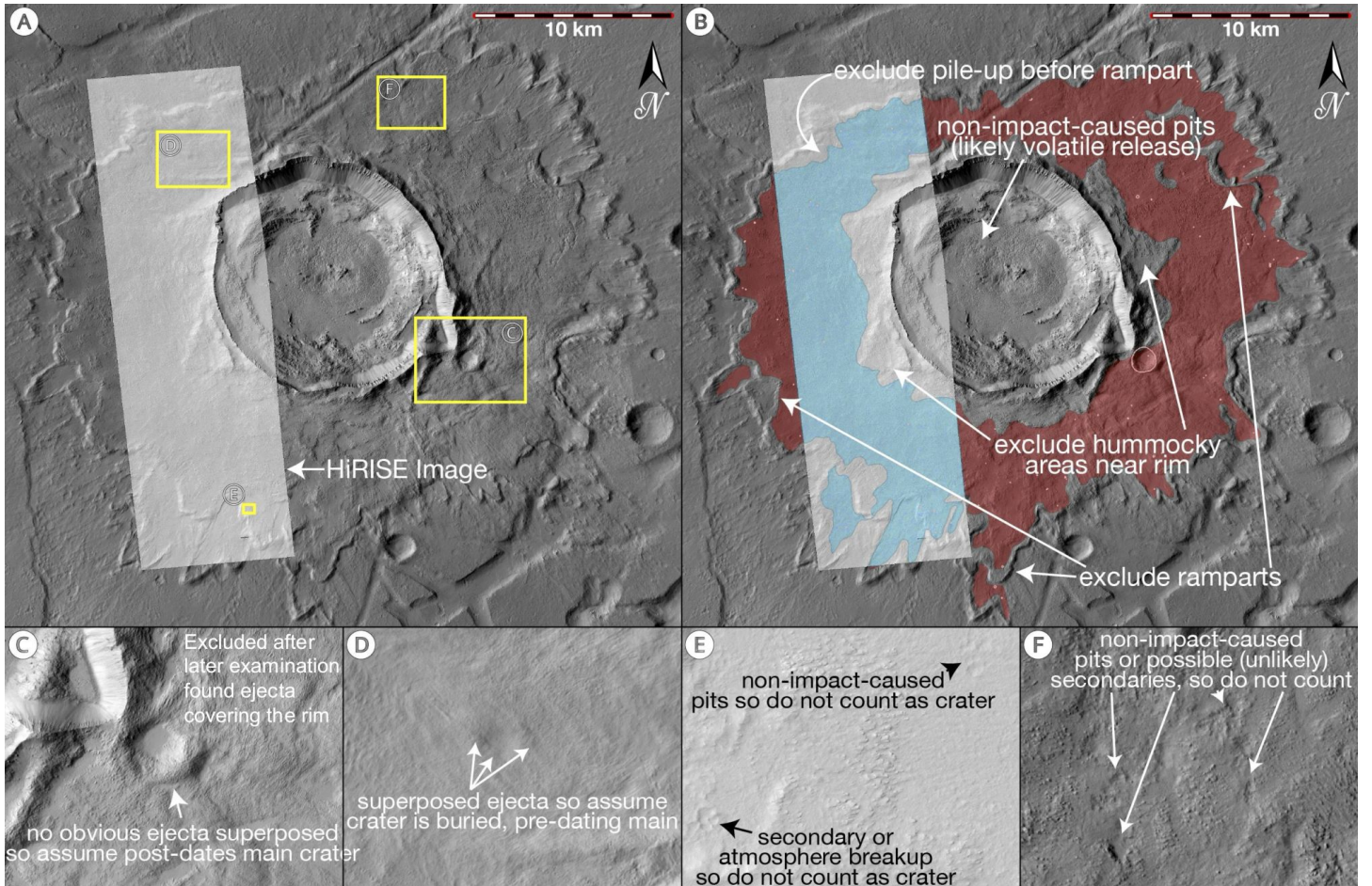


Figure 9 | Canala Crater (24.35°N 279.93°E, $D=11.9$ km). (a) CTX basemap with the boundaries of the HiRISE image shown. (b) The area we used for counting craters. (c) A large outlier crater initially included but later excluded due to ejecta covering the rim. (d) A clear example of an excluded buried crater. (e) Examples of excluded features: non-impact pits and a secondary or atmospheric breakup crater. (f) More examples of excluded features: depressions without the clear form of a primary crater, which we assess as probably non-impact pits and possibly self-secondaries.

examined each crater to assess whether Canala ejecta overlaid it (Figures 9c and 9d). This determination can be difficult and subjective. For instance, the initial counts included a $D = 1.2$ km crater on the CTX image (Figure 9c) – much larger than the $D = 0.13$ km next largest crater. However, after S. Bell examined it because it formed such an outlier and concluded that the ejecta covered the rim, we chose to exclude it. Additionally, the Canala ejecta contain volatile release pits that must be distinguished from impacts (Figures 9e and 9f). To incorporate these increased errors, we raised random human crater identification error to $p_r = 80\%$ and systematic human crater identification error to $p_s = 15\%$. Otherwise, we used the default values in Table 1.

For our production function, we selected the Mars version of the “new” NPF (Ivanov, 2001). The observed SFDs show considerable deviations from this production function. In light of the challenges Canala poses, these deviations raise potential concerns that the data might not reflect the production function. To test whether these SFD features could have been observed due to random chance, we constructed synthetic models, modeling a SFDs that transitioned to the production function at D_{min} . The results are clear: these deviations are not statistically significant. The CTX SFD plots fully within the 1σ synthetic error envelope (Figure 10b), and the HiRISE SFD mostly falls within the error envelope (Figure 10c). Although the HiRISE SFD does plot slightly below the error envelope for the largest two craters, this represents a very plausible outlier, especially given that the synthetic model only includes counting error.

We get median ages of $2.57^{+5.44}_{-1.75}$ Ma from the CTX data and $2.00^{+4.20}_{-1.36}$ Ma from the HiRISE data (Figure 10f). Without chronology system error, the median CTX age would be $4.72^{+1.54}_{-1.19}$ Ma, and the median HiRISE age would be $3.68^{+1.10}_{-0.87}$ Ma (Figure 10f). As this example shows, chronology system error can substantially shift the median age, even though the mean ages remain the same (Figure 10f).

These two ages are consistent, but if we had used only counting error, the ages would not have matched nearly as well, with a median CTX age of $5.42^{+1.02}_{-0.91}$ Ma and a median HiRISE age of $3.83^{+0.65}_{-0.58}$ Ma (Figure 10d). It is certainly not impossible for these ages to represent the same age. Outliers outside of 1σ happen 31.7% of the time, and the 4.48 Ma upper error bar of the HiRISE age almost overlaps with the

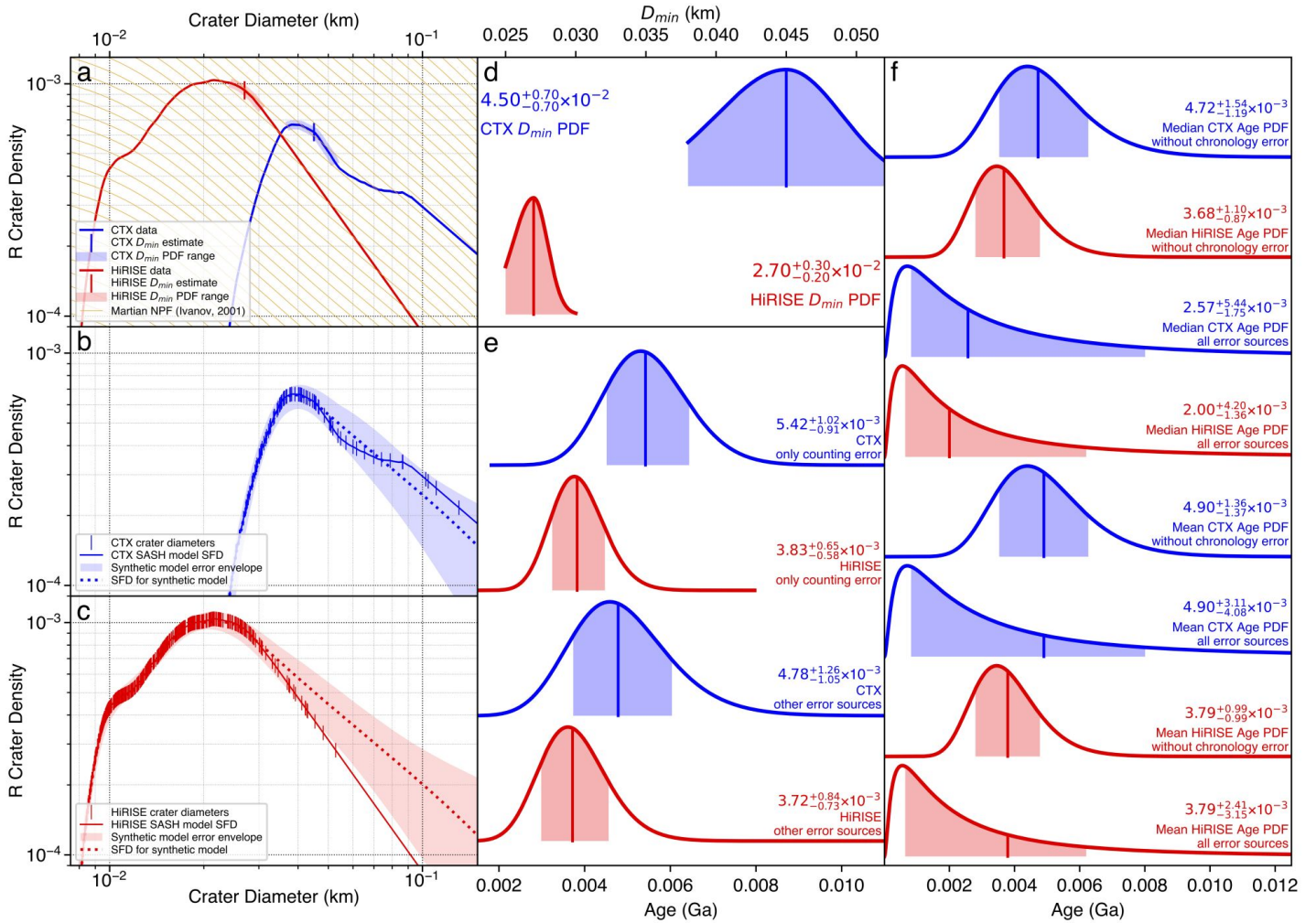


Figure 10 | Two ages for Canala Crater, one from HiRISE data (red) and one from CTX data (blue). (a) The SFDs plotted with the SASH model, showing the D_{min} estimate and the range of the D_{min} PDF. (We used the default bin growth rate of $2^{1.2}$, and we did not set a D_{max} because we estimated it to be 3 km – larger than the default of 10 times the diameter of the largest crater.) To help estimate D_{min} , we plot isochrons of the Martian NPF (Ivanov, 2001). (b) A synthetic model testing whether the CTX data are consistent with the production function. The SFD for the synthetic model follows the observed SFD for $D \leq D_{min}$ and follows the production function for $D > D_{min}$. (c) The same as (b) but for the HiRISE data. (d) The D_{min} PDFs. (e) The age PDFs with only counting error and other error sources: $p_{D_r} = 5\%$, $p_r = 80\%$, $p_a = 5\%$, and the D_{min} PDFs in (d) (systematic errors are excluded because the same counter was used). With other sources of error, the age PDFs are much more consistent with each other. (f) The full age PDFs with $p_{D_r} = 5\%$, $p_{D_r} = 1\%$, $p_r = 80\%$, $p_s = 15\%$, $p_a = 5\%$, and the D_{min} PDFs in (d), both with and without a factor of 3.0 chronology system error. To show how the mean ages remain the same when chronology system error is included, we show the mean ages as well, although in general we follow the convention set by Michael et al. (2016) of giving the median error.

4.51 Ma lower error bar of the CTX age. However, they do match much better when other error sources are taken into account. For this calculation, because one counter performed both counts, we assumed no systematic human crater identification or diameter measurement error; because we were comparing two craters of similar ages measured at similar diameters, we assumed no chronology system error; and because the images overlapped, we assumed additional error sources of $p_a = 5\%$ instead of the default value, $p_a = 10\%$. With these assumptions, the age PDFs much more strongly overlapped, with median values of $4.78^{+1.26}_{-1.05}$ Ma (CTX) and $3.72^{+0.84}_{-0.73}$ Ma (HiRISE).

Critically, we note that this stronger match is only partially due to broader PDFs; primarily it is caused by the CTX age shifting younger when other sources of error are factored in. This emphasizes an important point: failure to incorporate other sources of error does not just lead to excessively narrow error bars. It can also lead to an incorrect central value, too. In this case, the primary reasons for this shift are random diameter measurement error and D_{min} selection error.

At Canala, we must choose D_{min} to exclude the rollover caused by resolution limits in the CTX image and atmospheric shielding in the HiRISE image. For our single likeliest D_{min} value (the peak of the D_{min} PDF), we chose the points where the slope of the SFD begins to roll over away from the production function slope (CTX: 45 m, HiRISE: 27 m). While this is a standard way to choose D_{min} , it does raise the risk of selection bias, especially at low N . Because a bump on the SFD can cause this slope change, selecting D_{min} this way would run the risk of picking D_{min} at an outlier. Alternatively, one could pick D_{min} at a lower

diameter where the resolution rollover is more clearly observed or choose a larger diameter more safely distant from resolution effects. Choosing between these options effectively places the researcher in the position of choosing the age, so a D_{min} PDF is necessary to mitigate this bias, and incorporating in the PDF of D_{min} choices shifts the median age lower. A closer analysis of the specific crater diameters (Figure 10b) shows an increased density of craters very slightly larger than the 45 m we chose for our likeliest D_{min} , which supports the hypothesis that it may have been an outlier. When random diameter measurement error is factored in, the chance that some of these craters may have actually been smaller also shifts the median age younger.

As Canala Crater demonstrates, errors beyond counting error cannot be ignored. These other sources of error matter. With our methods, they can finally be incorporated.

5.2 Practical guide to doing the calculation

We want our methods to be used, so here we provide a guide showing how to actually do the calculation, using the Canala CTX data as an example. Although the math may be complex, our *cratersfd* code makes the process simple and straightforward. In Figure 11, we show screenshots of the code to make clear what the user should be actually doing. There are four steps:

1. **Plot the SFD with the SASH model.** For an R plot, use the `plot_sash_R()` function; for a differential plot, use the `plot_sash()` function. To help with Step 2, it is often helpful to overlay production function isochrons with the `plot_pf_isochrons()` function.
2. **Analyze the plot to estimate the D_{min} PDF.** To help the user quantify the D_{min} PDF as a RandomVariable object, *cratersfd* has a `get_dmin_pdf()` function. This function treats D_{min} as an asymmetric two-sided normal distribution (see Appendix C) with different σ values on the left and right sides and optional upper and lower bounds. The *cratersfd* code can also handle any arbitrary shape for the D_{min} PDF.
3. **Estimate the error kernels and calculate the age PDF.** To do this, use the `age_pdf()` function. By default, `age_pdf()` assumes the values in Table 1, except that it assumes no chronology system error. Users can specify different values using keyword arguments. (These keyword arguments are for error factors, so in Figure 11 we specify `random=1.8` for $p_r = 80\%$ random human crater identification error.) By default, `age_pdf()` assumes the lunar versions of the “new” NPF (Ivanov et al., 2001) and the inverse NCF (Neukum and Ivanov, 1994). Users can specify other production functions and inverse chronology functions with the `pf` and `cf_inv` keyword arguments. In Figure 11, we use the built-in `npf_mars` and `ncf_mars_inv` functions, and users can also define their own production and inverse chronology functions. The `age_pdf()` function returns a RandomVariable object (see Appendix A). To plot its PDF, use its `.plot()` method.
4. **Calculate the age without chronology system error.** This is the same as Step 3, except the `cf_error` keyword argument is left at its default value of a factor of 1.0, indicating no error.

5.3 Visualizing the math of the Age Problem

In this example, we use the Robbins et al. (2014) clustered (aggregate) counts of NAC image M146959973L to visualize the math behind the Age Problem, showing the PDFs of the random variables within Equation (11) and how the age PDF evolves as different sources of error are added (Figure 12). By showing plots of random variables’ PDFs within Equation (11), we visually emphasize what is happening with random variable math: These are the same mathematical operations as we know from the math of perfectly known variables, but here the math is applied to variables whose values are defined by PDFs.

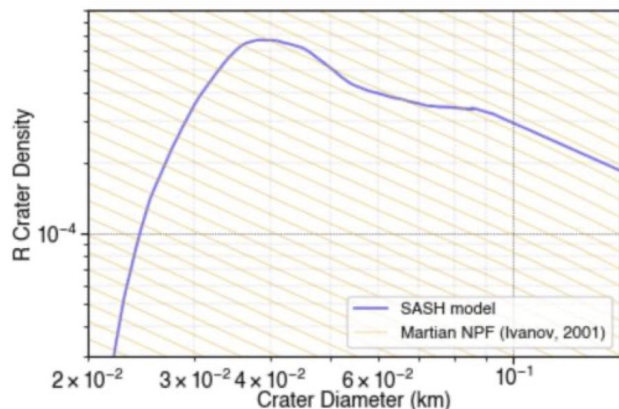
Figure 12 also shows the shape of a lognormal error kernel. For the 2% systematic diameter measurement error kernel (Figure 12e), the PDF closely resembles a normal distribution. For the 10% random diameter measurement error kernel (Figure 12d), it becomes slightly asymmetric. For the 60% random human crater identification kernel (Figure 12i), it becomes more asymmetric, and for the 150% chronology system error identification kernel (Figure 12m), it becomes highly asymmetric. For each kernel, the mean remains 1.0, but the mode shifts lower and lower as the kernel widens. The shape of the Gamma distribution of λ from Equation (6) (Figure 12f) is very similar to a Lognormal distribution, including (Figure 12h) when it incorporates the N PMF (Figure 12g) created by diameter measurement error.

When we only include counting error, the median age is $1.70^{+0.90}_{-0.67}$ Ga (Figure 12n). When we add diameter measurement error, it becomes $2.04^{+0.99}_{-0.81}$ Ga (Figure 12o). When we add human crater identification error and additional error factors, it becomes $1.96^{+1.15}_{-0.85}$ Ga (Figure 12p). When we add chronology system error, it becomes $1.96^{+1.15}_{-0.85}$ Ga (Figure 12q). When we add D_{min} selection error, it becomes $2.01^{+1.29}_{-1.06}$ Ga without chronology system error (Figure 12r) and $1.27^{+2.06}_{-0.89}$ Ga with it (Figure 12s).

Step 1: Plot the data with the SASH model.

```
plot_sash_R(diameter_array, area)
plt.xlim([0.02, 0.15])
plt.ylim([3E-5, 9E-4])
plot_pf_isochrons(npf_mars_R)
labels = ['SASH model', 'Martian NPF (Ivanov, 2001)']
plt.legend(labels, fontsize=8)
```

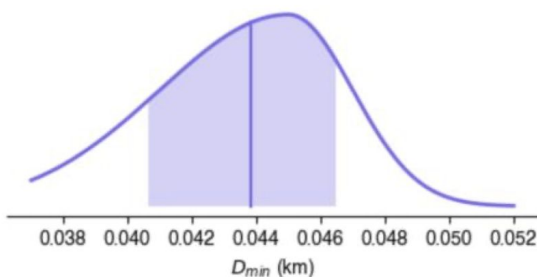
Calculation time: 0.38216 seconds
<matplotlib.legend.Legend at 0x130ebe420>



Step 2: Analyze the plot to estimate the D_{min} PDF.

```
ctx_dmin = 0.045
ctx_dmin_min = 0.037
ctx_dmin_max = 0.049
ctx_dmin_rv = get_dmin_pdf(
    ctx_dmin, left_std=0.005, right_std=0.0015,
    left_edge=ctx_dmin_min, right_edge=ctx_dmin_max
)
dmin_rv.plot()
plt.xlabel(rf'$D_{\{min\}}$ (km)')
```

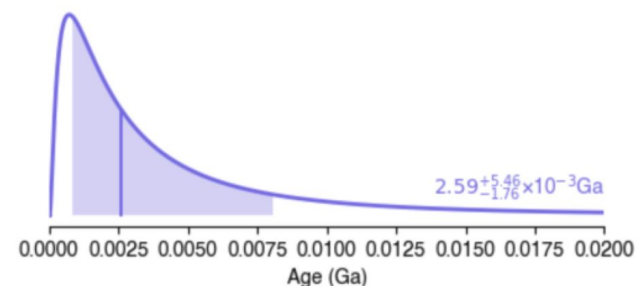
Text(0.5, 0, '\$D_{\{min\}}\$ (km)')
<Figure size 460x300 with 0 Axes>



Step 3: Estimate the error kernels and calculate the age PDF.

```
age_pdf(
    diameter_array, area, dmin_rv,
    random=1.8, systematic=1.15, additional=1.1,
    d_random=1.05, d_systematic=1.01, cf_error=3.0,
    pf=npf_mars, cf_inv=ncf_mars_inv
).plot(label=True, xlim=[0, 0.02], unit='Ga')
plt.xlabel('Age (Ga)')
```

Text(0.5, -7.277777777777782, 'Age (Ga)')
<Figure size 460x300 with 0 Axes>



Step 4: Calculate the age without chronology system error.

```
age_pdf(
    diameter_array, area, dmin_rv,
    random=1.8, systematic=1.15, additional=1.1,
    d_random=1.05, d_systematic=1.01, cf_error=1.0,
    pf=npf_mars, cf_inv=ncf_mars_inv
).plot(label=True, xlim=[0, 0.02], unit='Ga')
plt.xlabel('Age (Ga)')
```

Text(0.5, -7.277777777777782, 'Age (Ga)')
<Figure size 460x300 with 0 Axes>

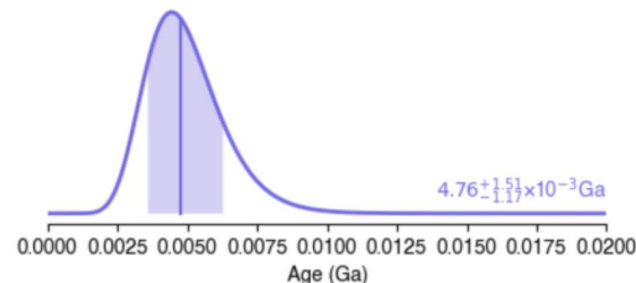


Figure 11 | A four-step guide to how to do the Age Problem calculation with our *cratersfd* code (see Section 5.2 for details and Section 5.1 for a discussion of the choice of error kernels and the D_{min} PDF). Example data are the CTX data for Canala Crater. The *cratersfd* code is a Python package, and these examples are created in a Jupyter notebook.

These data, the Robbins et al. (2014) clustered (aggregate) counts of NAC image M146959973L, also allow us to demonstrate how to apply our methods to a terrain in equilibrium saturation where only a few craters are not in equilibrium, emphasizing the power of the SASH model and its synthetic error model. When we apply the SASH model (Figures 12a and 12b), the SFD rolls over from a shallow slope of $\alpha \approx 1.4$ for $20 \text{ m} < D < 100 \text{ m}$ to a steep slope for $D > 130 \text{ m}$ – the classic pattern of equilibrium saturation (Gault, 1970). This pattern is especially clear on an R plot (Figure 12a), and it can also be seen on a differential plot (Figure 12b).

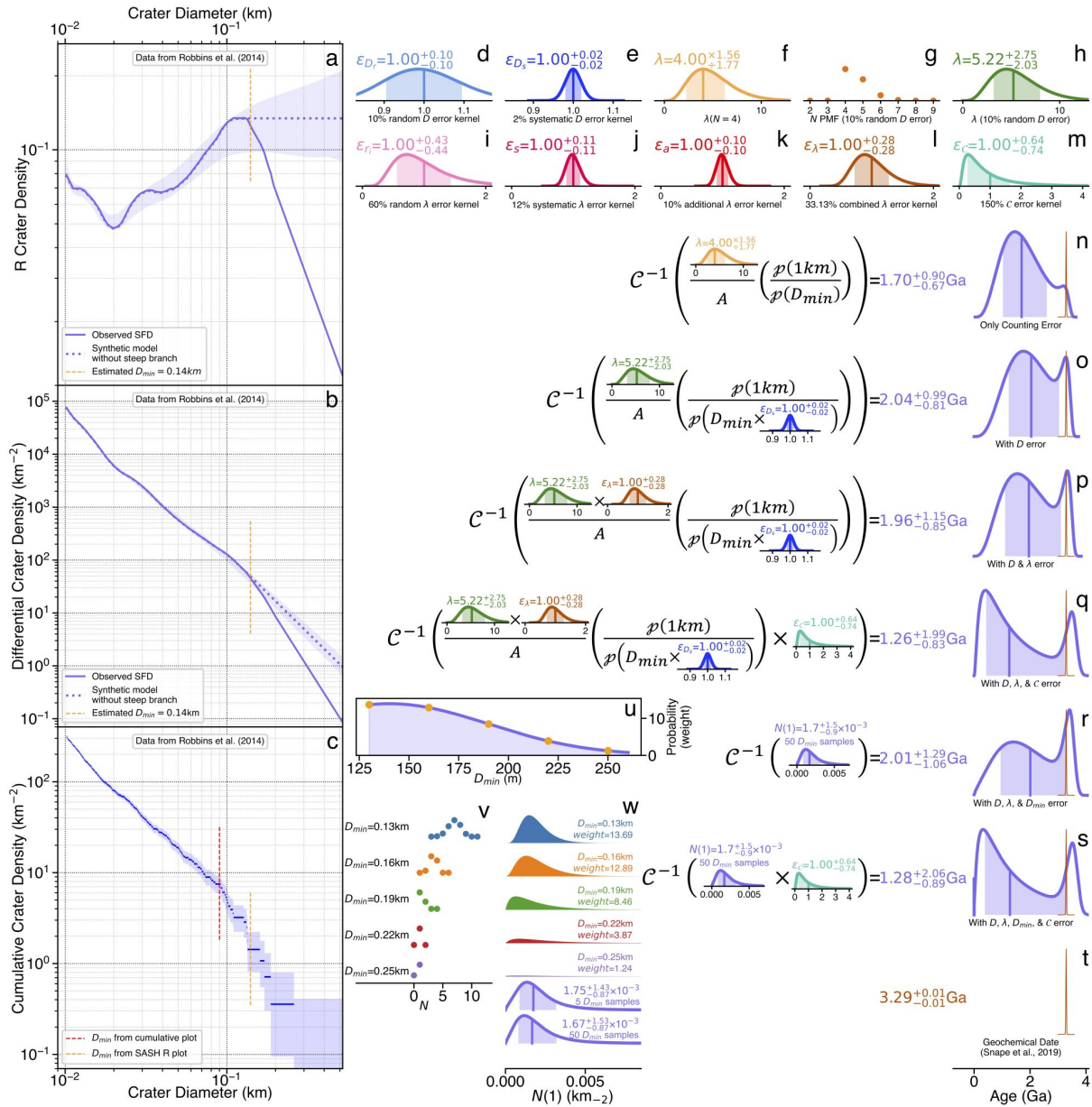


Figure 12 | An example Age Problem calculation for a low- N count, using the Robbins et al. (2014) data from the Apollo landing site (see Section 5.3 for details). The goal of this example is to demonstrate that we cannot assume that counting error is the only source of error even when N is very low. (a) A SASH R plot of the data (initial bin width: 18 per decade. Bin growth rate: $2^{1.2}$. Number of synthetics: 1000. $D_{max} = 0.93$ km.). In lighter blue, we show the error envelope of a synthetic model of how the SFD might appear if no steep branch were present and the SFD continued at a saturated slope of $\alpha=2$ over the whole diameter range (see Section 4.4). (b) Same as (a) but in differential plot form. (c) A cumulative plot of the same data, showing that the D_{min} we estimated from the SASH R plot is conservative compared to a D_{min} that could be estimated from the cumulative plot. (d) The 5% random diameter measurement kernel E_{D_r} . (e) The 1% systematic diameter measurement error kernel E_{D_s} . (f) The PDF of the model count λ given an observed N of 4 at $D_{min}=0.14$ km. (g) The PMF of N incorporating diameter error with the E_{D_r} in (d). (h) The PDF of λ from the N PMF in (g). (i) The 60% individual random human crater identification error kernel for each crater E_{r_i} . (j) The 12% systematic human crater identification error kernel E_s . (k) The 10% additional error factors kernel E_a . (l) The combined λ error kernel E_λ from Equation (10). (m) The 150% (factor of 2.5) E_c chronology system error kernel. (n-s) Age PDFs incorporating various sources of error, compared to the geochemical date (Snape et al., 2019). For each age PDF, we show how it is calculated from Equation (11), plotting the PDFs of the random variables. (t) The known age from returned Apollo 15 samples (Snape et al., 2019). (u) The user-estimated PDF of D_{min} . Because D_{min} changes N , the D_{min} PDF cannot be simply substituted into Equation (11). Instead, we apply Equation (11) to weighted samples of the PDF. For explanatory purposes, here we show an example with five samples. For each sample, the weight is the value of the PDF. (v) The N PMFs for the D_{min} samples in (u), given the 5% E_{D_r} in (d). (w) The $N(1)$ PDFs for each D_{min} sample, calculated from Equation (11) with the N PMFs in (v). The probability values of each $N(1)$ PDF are multiplied by the weight in (u), and then they are added together to get the final $N(1)$ PDF. We emphasize that five samples are not enough to robustly sample the PDF. We recommend 50 samples, which we used in the calculations of (r) and (s).

However, we must ensure this is a statistically robust feature, not an artifact of random chance. Here, to test the hypothesis that the terrain is saturated over the whole diameter range, and the steep branch is a statistical artifact, we construct a hypothetical SFD that continues at a slope of $\alpha = 2$ instead of rolling over for the steep branch. Because the observed slope of the equilibrium branch is much shallower than $\alpha = 2$, this is a conservative assumption, but we want to be sure we have resolved the steep branch. From this hypothetical SFD, we randomly generate 1000 synthetic crater datasets with the same N as the Robbins et al. (2014) data. The results are clear (Figures 12a,b): it is a robustly resolved feature and unlikely to be due to random chance because the steep branch plots well below the envelope of synthetic results.

Now that we can be confident that the steep branch is real, we must estimate D_{min} . To address D_{min} selection error, we recommend treating D_{min} as a random variable and estimating its PDF. The correct method for selecting D_{min} is an area where more research is needed. In this case, we choose a D_{min} of 140 m, slightly above the 130 m beginning of the steep branch (Figure 12a). To broaden our single-diameter estimate to a PDF, we take a normal distribution around 140 m with a standard deviation of 50 m and truncate this PDF at the 130 m beginning of the steep branch and the 260 m diameter of the largest crater.

Compared to a potential traditional reading of the cumulative plot, this ± 50 m is a conservative approach (“conservative” as in large uncertainty). The cumulative plot (Figure 12c) transitions from the saturated branch to the steep branch at a fairly sharp inflection point at ~ 90 m. There is no consensus about how to estimate D_{min} from a cumulative plot, and some analysts might prefer a more conservative larger value, but one potential approach is to use the inflection point, especially when it is sharp. With so few craters, however, such a sharp inflection point could not be robustly resolved if it were present. Although higher N might well allow a lower D_{min} to be resolved, we must confine ourselves to what we can resolve from the data at hand.

The effect of the D_{min} PDF is especially significant at low N because the choice of D_{min} has a significant effect on the results, as this example shows. The cumulative plot (Figure 12c), which shows the specific crater diameters, highlights the subjective dangers in a single D_{min} estimate. At $D_{min} = 140$ m, $N = 4$. At $D_{min} = 160$ m, N is still 4, so the user’s choice of D_{min} directly affects the age PDF. At $D_{min} = 130$ m, N jumps to 8. The age depends heavily on the subjective choice of a single number in these cases of low N , and that single number is far from perfectly known. Broadening this estimate to a random variable defined by a PDF does not eliminate this bias, but it does significantly mitigate it.

Next, we estimate values for the several lognormal error kernels (Sections 2.1–2.9) and choose a chronology system. Here, we largely use the default values in Table 1 (because we are close to the equilibrium saturation rollover, we assume 10% for p_{D_s} and 2% for p_{D_c}). For explanatory purposes, we will be using the Neukum (1983) system, the NPF and NCF. We do not choose them because we endorse them as the most accurate, but we choose these because they remain the standard in the field, we expect readers to be most familiar with them, and we want to focus on the math of incorporating multiple error factors to build an age PDF.

Finally, these data demonstrate why it is necessary to incorporate sources of error beyond counting error. Because these counts were taken from the Apollo 15 landing site, we can check our results against a known, radiometrically derived age. For the geochemical age from returned Apollo samples, we use the Snape et al. (2019) estimate of 3.29 ± 0.01 Ga. This known age falls at the 1σ -equivalent upper bound of the $2.01^{+1.29}_{-1.06}$ Ga age PDF that incorporates all sources of error except for chronology system error. Because the Apollo 15 landing site was a chronology system calibration point, we should expect the age PDF without chronology system error to be consistent with the known age, and it is, with the known age again falling at the upper bound. With counting error alone, however, the $1.70^{+0.90}_{-0.67}$ Ga age PDF does not match the known age.

With an N of 4, the Robbins et al. (2014) counts of the Apollo 15 landing site are something of an extreme case. If there were any scenario where the traditional assumption of counting error as the sole source of error might appear reasonable, it might be such an extremely low- N case. When the only counting error model fails here, too, it makes the case especially strong. Crater counting ages must incorporate other sources of error. For decades, the lack of the mathematical methods necessary has prevented the field from incorporating other sources of error, but with our methods, it is not only possible but straightforward.

5.4 Cratered plains SFDs: the Saturn system and the farside of the Moon

As an example of the explanatory power of the SASH model, we applied it to the cratered plains distributions of the moons of Saturn. In cratered plains, impact craters cover the whole area, and no underlying geologic unit boundaries can be detected. Cratered plains SFDs matter because they reveal the history of the oldest terrains on a body, but heavy saturation can mean that only the largest craters are unsaturated. As a result, analyzing these terrains requires carefully resolving the shape of the SFD at the low- N , high-diameter end.

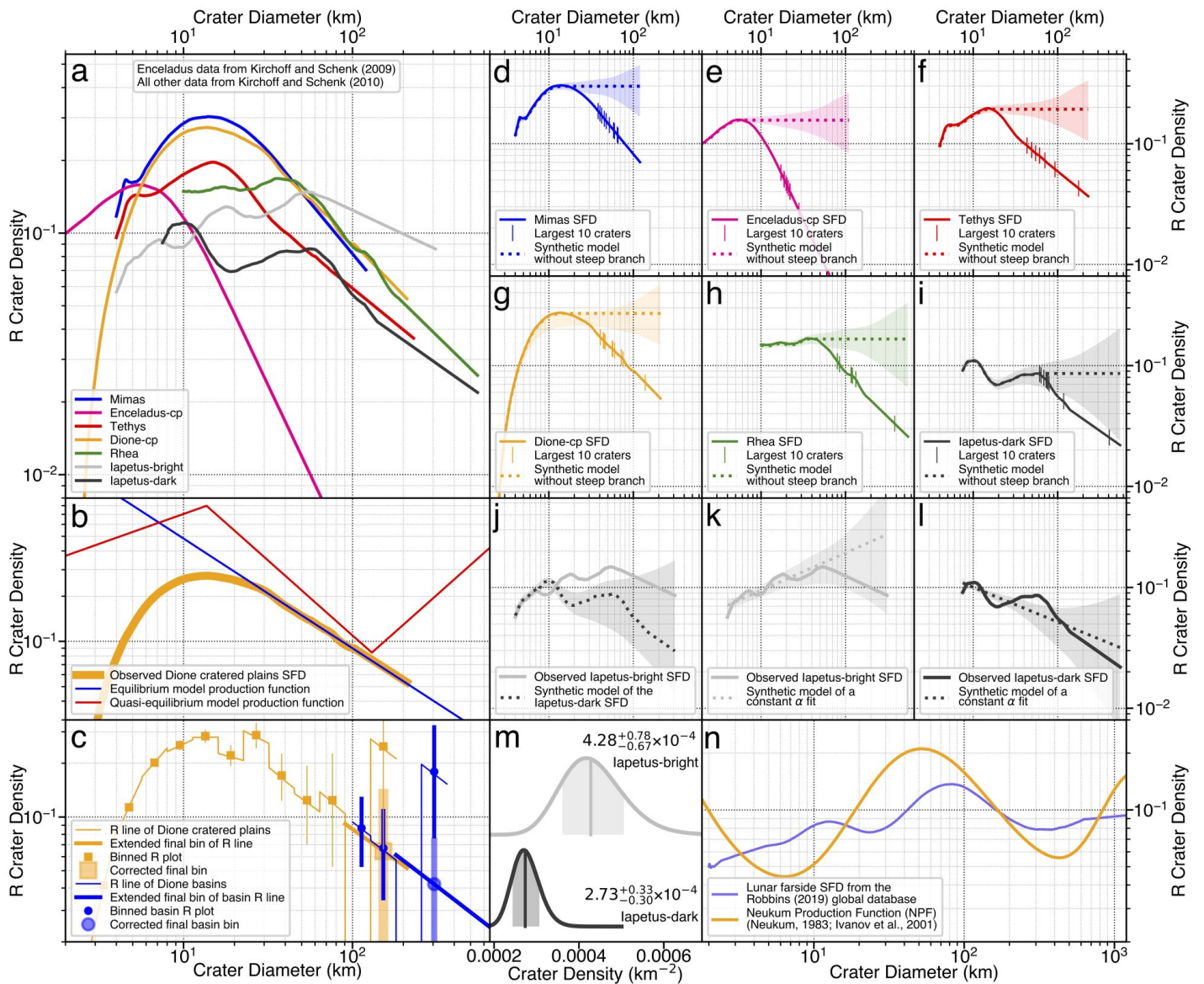


Figure 13 | (a) The SASH model applied to the Kirchoff et al. (2009) and Kirchoff et al. (2010) crater counts of the cratered plains of the moons of Saturn, revealing steep branches at the large crater end. (Initial bin width: 18 per decade. Bin growth rate: $2^{1.25}$. D_{max} : Mimas, 120 km; Enceladus, 110 km; Tethys, 230 km; Dione, 210 km; Rhea, 550 km; Iapetus-bright, 310 km; Iapetus-dark, 550 km.) (b) A schematic of the equilibrium saturation model (Lissauer et al., 1988) and the quasi-equilibrium saturation model (Kirchoff and Schenk, 2010), showing potential production functions each model might imply for the Dione cratered plains SFD. (c) A reproduction of Figure 4 from Kirchoff and Schenk (2010), showing the Dione data as an example of how the largest bin in their cratered plains (orange squares) and global basins (blue circles) distributions consistently plotted above the steep branch. This led Kirchoff and Schenk (2010) to conclude that the SFD had a shallow branch for very large craters, but we attribute this effect to the bias caused by missing zero-crater bins (see Section 4.3). To correct for this effect, we show the line of R values implied from the binned plot, with the slopes across each bin calculated from the SASH model results (thin lines). For both the cratered plains and the global basins data, the final bin (the bin with the largest crater) has a zero-crater bin to its left, as well as zero-crater bins to its right extending out to D_{max} . We extend the final bin over these zero-crater bins, showing its line of R values in a thick line. At the diameter where the points for the final bins were plotted, we show what the values would have been if the zero-crater bins had been incorporated. (d)-(i) Synthetic models testing whether the steep branches could have been observed if they were not present (1000 synthetics each). Except for the dark terrain of Iapetus, each steep branch plots well outside the error envelope, showing that the features are robustly resolved and unlikely to be due to random chance. (j) A synthetic model testing whether the observed SFD for the bright terrain of Iapetus could have been observed if the true SFD matched the dark terrain. (k)-(l) Synthetic models comparing the bright and dark terrains to straight line fits to test if their features are statistically robust. (m) PDFs of $D \geq 12$ km crater density incorporating 60% random human crater counting error and 5% random diameter measurement error, as well as a 10% additional error kernel, showing that the difference between the dark and light terrains remains robust. (Because the same counter counted both terrains, we assumed no systematic error.) (n) The SASH model applied to the whole lunar farside from the Robbins (2019) global database, compared to the “new” NPF (Neukum, 1983; Ivanov et al., 2001). It shows that quasi-equilibrium saturation does not cause the SFD to still reflect the production function, contrary to the predictions of the Richardson (2009) model.

Results from previous methods have led to differing interpretations. From Voyager data, Lissauer et al. (1988) interpreted the Rhea and Iapetus SFDs under an equilibrium saturation model, with a steep $\alpha \approx 2.7$ large crater slope (reflecting the production function) and shallow slopes at smaller craters (reflecting equilibrium) (Figure 13b). At Mimas, Lissauer et al. (1988) again found a steep large crater slope, but they noticed that the largest crater on Mimas, Herschel, plotted far above this line on an R plot. With higher-quality Cassini data, Kirchoff and Schenk (2010) noticed that this pattern was consistent: on each moon, the largest bin in the cratered plains distribution plotted well above the steep branch. Moreover, when Kirchoff and Schenk (2010) added in the data from the global distributions of large basins, the pattern reappeared, with the largest one or two bins plotting well above the steep branch. As a result, Kirchoff and Schenk (2010) concluded that, as diameter increases, the steep branch rolls over again to a shallow branch for very large craters.

Having found one inflection point in the production function, Kirchoff and Schenk (2010) proposed another. They proposed – and Kirchoff et al. (2018) concluded – that the rollover from shallow slopes to the steep branch was a real feature of the production function, not the result of equilibrium saturation. In the Richardson (2009) saturation model, when a production function rolls over from a branch in the $\alpha \leq 2.0$ quasi-equilibrium domain to a branch in the $\alpha \geq 2.0$ equilibrium domain, slopes on both sides of the inflection point become shallower, but the diameter of the inflection point is preserved (Figure 13b). Kirchoff and Schenk (2010) proposed that this effect explains the shallow branches observed in the Saturn system. As a result of this quasi-equilibrium model (Figure 13b), much work has been devoted to interpreting the SFDs of the cratered plains of the moons of Saturn at diameters smaller than the inflection points, treating them as reflecting the production function (e.g. Ferguson et al., 2020; Bottke et al., 2024; Robbins et al., 2024).

Other studies have found ambiguous results. With binned cumulative plots, Bell (2020) reinterpreted the Kirchoff and Schenk (2010) data as potentially reflecting a production function with single slope of $\alpha \approx 2.2$, just barely in the $\alpha \geq 2.0$ equilibrium domain, with inflection points around ~12 km caused by equilibrium saturation. However, the slope was just barely in the equilibrium domain, so Bell (2020) could not rule out the quasi-equilibrium model. Robbins et al. (2024) could not resolve the steep branch with the Robbins et al. (2018) KDE method, which breaks down at low N , but they did find evidence for it in MLE slope calculations.

With the SASH model, we reexamined the Kirchoff and Schenk (2009, 2010) data, revealing that the cratered plains of Mimas, Enceladus, Dione, Tethys, and Rhea each have a steep ($\alpha \geq 2.5$) branch on the large crater end. To rule out the hypothesis that these steep branches are spurious artifacts of random chance, we used the synthetic model we describe in Section 4.4. For each cratered plains terrain, we created a model SFD that matched the observed SFD up to the inflection point but continued with a slope of $\alpha = 2$ instead of having a steep branch. From each of these SFDs, we randomly generated 1000 synthetic crater datasets and reran the SASH model, plotting the results in 1σ -equivalent error envelopes. The results of the synthetic models (Figures 13d–13h) are clear: each steep branch plots well below the observed error envelope. These features are statistically robust. They are highly unlikely to be artifacts of random chance.

The steep branches would have been steeper if we had not estimated D_{max} values (Figure 13), demonstrating the importance of estimating D_{max} . Kirchoff and Schenk (2010) were unusually willing to include large craters, including two craters whose diameters were more than half the width of the count area – Penelope (177 km) on Tethys and Tirawa (383 km) on Rhea. As a result, in those cases, we chose D_{max} values unusually close to the largest crater. In each case except for Enceladus and the bright terrain of Iapetus, we chose a D_{max} that was less than twice the diameter of the largest crater. On Mimas, in the context of Kirchoff and Schenk (2010)'s willingness to include large craters, we estimated that the substantial size of the count area would imply a D_{max} of 200 km. However, because the count area did not include the 140 km Herschel crater, we chose a D_{max} of only 120 km to account for the possibility that Kirchoff and Schenk (2010) deliberately excluded it. This is likely a conservative assumption – in other terrains, Kirchoff and Schenk (2010) showed loyalty to avoiding the selection bias created by avoiding large craters when drawing count terrains – but we wanted to err on the side of choosing smaller D_{max} values out of an abundance of caution. We can be sure that the steep branches are not artifacts of choosing D_{max} values that are not small enough because they would have still been observed even if we had chosen D_{max} to be only slightly (5%) larger than the largest crater, as Supplemental Figure S1 shows.

The resolution of steep branches that clearly fall in the $\alpha \geq 2.0$ equilibrium domain supports the equilibrium model. In fact, much of the evidence for the quasi-equilibrium model was due to statistical artifacts of previous methods. For instance, in our SASH model results (Figure 13a), there is no evidence of a return to a shallow branch for very large craters, and we believe it was an artifact of binned plots. Kirchoff and Schenk (2010) proposed the shallow branch because they consistently found that the largest bin plotted well above the steep branch (Figure 13c). For instance, Figure 13c reproduces Figure 5 from Kirchoff and Schenk (2010), showing how the largest bins in both the Dione cratered plains distribution and the global basins distribution appear well above the steep branch on a binned R plot. However, as

we discussed in Section 4.2, the right-hand bin is biased upwards because binned plots omit zero-crater bins. When we correct for this bias by extending the right-hand bins over the zero-crater bins to their sides, the evidence for the shallow branch for very large craters disappears, and the right-hand bins fall right in line with the steep branch. This simple explanation confirms the results of Robbins et al. (2021), who examined this question for Mimas specifically with a more complex statistical analysis, concluding that the largest crater, Herschel, is not an outlier.

Bell (2020) could not rule out the quasi-equilibrium model because he found shallow ($\alpha \approx 2.2$) slopes that were only barely in the equilibrium domain. This was due to the shallow slope bias of binned cumulative plots (Figure 4), which Bell (2020) exacerbated by making plots with median method error bars. Moreover, because of the limitations of binned cumulative plots in resolving the true shape of the SFD, which prevented him from resolving Rhea's inflection point at all, Bell (2020) incorrectly concluded that these terrains had consistent inflection point diameters around ~ 12 km. However, the SASH model shows that the inflection points clearly fall at *different* diameters. We estimate them at: ~ 22 km for Mimas, ~ 5 km for Enceladus, ~ 13 km for Dione, ~ 15 km for Tethys, and ~ 45 km for Rhea. Each of these inflection points falls at an absolute density consistent with equilibrium saturation. In equilibrium saturation, the rollover density generally falls in the $0.1 \leq R \leq 0.3$ range, but it can be highly variable (Richardson, 2009), especially when lighting conditions vary (Ostrach, 2013), as they do for Cassini data. These findings strongly support the equilibrium model. If the quasi-equilibrium model were correct instead, if the rollovers were a real feature of the production function, the moons would have to be cratered by very different populations whose SFDs each just happened to roll over at a diameter consistent with equilibrium saturation.

This has major implications. Under the equilibrium model, the portions of the SFD below the inflection point are saturated and do not reflect the production function. This matters because these shallow small crater portions of the SFD have been heavily interpreted as reflecting the production function. For instance, they have been cited both as evidence of planetocentric cratering by bodies orbiting Saturn (e.g. Ferguson et al., 2020; Ferguson et al., 2022; Ferguson et al., 2024) and heliocentric cratering by bodies orbiting the sun (Bottke et al., 2024).

Empirical saturation observations further strengthen the case that these portions of the curve do not reflect the production function. For instance, Ferguson et al. (2020) cited the small diameter ends of these SFDs, where the slope rolls over to become much shallower than $\alpha \approx 2.0$, as evidence of the quasi-equilibrium saturation model. However, SFDs in equilibrium often roll over to very shallow slopes in this exact way (Xiao and Werner, 2015). For instance, Xiao and Werner (2015)'s SFD for the Cayley Plain on the Moon reached equilibrium at $D \approx 850$ m and rolled over to a very shallow slope of $\alpha \approx 1$ for $D \leq 500$ m. Spatial statistics also support equilibrium saturation. Equilibrium saturation causes craters to become more uniformly spaced, and quasi-equilibrium saturation causes craters to become less evenly spaced (Kirchoff, 2018). The cratered plains of Mimas, Tethys, and Dione have very uniformly spaced $D \geq 6$ km craters, indicating equilibrium saturation (Kirchoff, 2018).

To explore how the SFD would behave if the quasi-equilibrium theory were correct, we applied the SASH algorithm to the Robbins (2019) global database of $D > 2$ km lunar craters to resolve an R plot of the whole farside of our Moon in unprecedented accuracy. (Because the boundaries of the highlands on the nearside are heavily affected by the largest craters, a human-defined unit tracing the boundaries of the highlands might suffer from selection bias, so the whole farside SFD provides a more neutral view.) According to the “new” NPF (Ivanov et al., 2001), the lunar production function has a steep branch with $\alpha > 2$ for $D \leq 6$ km, enters a shallow branch in the $\alpha < 2$ quasi-equilibrium domain for $6 \text{ km} \leq D \leq 52$ km, and then returns to the $\alpha > 2$ equilibrium domain with a steep branch above an inflection point at $D \approx 52$ km (Figure 13n). Under the Richardson (2009) model, although the slopes should become shallower, the portions of the SFD below the inflection point should still reflect the production function.

In the observed farside distribution, however, this does not happen. Even though the $D \leq 6$ km steep branch is a robust feature of the NPF, it does not survive saturation in the farside distribution. Instead, there is a shallow branch over $2 \text{ km} < D \leq 13$ km. For $13 \text{ km} \leq D \leq 24$ km, the whole farside distribution shows a steep branch – a feature entirely absent from the production function. With an N of 276 692, these features are very robustly resolved. The farside SFD makes clear that the SFD below an inflection point from a shallow $\alpha < 2$ branch to a steep $\alpha > 2$ branch can deviate dramatically from the production function. Even if the Kirchoff and Schenk (2010) quasi-equilibrium theory were correct, the portions of the SFD more than a factor of ~ 2 below the inflection points could not be interpreted as reflecting the SFD.

The improved SFD resolution from the SASH method also reveals a different picture of the Kirchoff and Schenk (2010) Iapetus cratered plains data. Multiple studies have interpreted these data as reflecting the production function (e.g. Bell, 2020; Bottke et al., 2024), but the SASH model results cast doubt on this interpretation. With the coarser resolution of binned R plots, Kirchoff and Schenk (2010) concluded that the bright terrain and dark terrain have similar SFDs. The SASH model reveals that the SFDs do not, in fact, match (Figure 13a). For $D \geq 12$ km, the dark terrain is less heavily cratered than the bright terrain

(Figure 13a), a result that holds when human error and additional error factors are included (Figure 13m). In a synthetic model, we tested whether the bright terrain SFD could be observed if the true SFD matched the dark terrain SFD. The results (Figure 13j) show that the mismatch is probably not due to random chance.

The dark terrain does show a steep branch above ~60 km, but the SASH model does not resolve it definitively. As with other steep branches, we ran a synthetic model of the SFDs that could have been observed if the steep branch were not present (Figure 13i). The steep branch only plotted slightly outside the 1σ -equivalent error envelope. Considering that the synthetic error model only includes counting error, and 1σ is a rather low bar for error, the steep branch on the dark terrain could be a real feature, but we cannot rule out the possibility that it may be a statistical anomaly. In the light terrain, no steep branch is evident, but the synthetic test of the dark terrain SFD at the N value for the bright terrain shows that a steep branch could easily be present (mechanically: the SASH model of the dark terrain was used as the PDF from which random draws were made to test between it and the light terrain).

For both terrains, we ran synthetic models of a straight line fit (in log-log space), calculating the mean α from the Truncated Pareto method. This synthetic model shows that the bright terrain SFD's bumps and wiggles are not statistically robust (Figure 13k). In the dark terrain, with its higher N , the synthetic model suggests that features like a steep branch from ~11 km to ~19 km (Figure 13l) are probably robust.

The mismatch between the bright and dark terrain SFDs is likely due to the fact that the dark terrain count area contains the 409 km Falsaron Crater, parts of the even larger Turgis Crater, and their ejecta blankets, as well as some areas farther away. As a result, it combines younger resurfaced areas with densely cratered plains. One potential interpretation of the dark terrain data could be that the steep branch above ~60 km reflects the production function, the ~60 km rollover reflects equilibrium saturation in the cratered plains, and the steep branch from ~11 km to ~19 km reflects where the production function can be observed on the younger resurfaced areas.

However, we urge caution. As our synthetic tests show, the data are not sufficient to conclude this definitively. Because the steep branch cannot be definitively resolved, we caution against interpreting any part of the Iapetus SFDs we have generated from the Kirchoff and Schenk (2010) data as reflective of the production function.

With the clearer resolution of the SFDs provided by the SASH model, a very different picture of the moons of Saturn emerges. Mimas, Enceladus, Tethys, Dione, and Rhea have statistically robust steep branches at the large crater end. These results strongly support the equilibrium saturation model, where the shallow branch is in equilibrium saturation, and the steep branch follows the production function (Figure 13b). Because of the limitations of previous statistical methods, many studies (e.g. Bell, 2020; Ferguson et al., 2020; Ferguson et al., 2022; Wong et al., 2023; Bottke et al., 2024) misinterpreted saturated data as reflecting the production function. However, our results also provide robust avenues for future work. With the steep branches resolved by the SASH model, there is now a portion of the SFDs that *can* be interpreted as reflecting the production function.

6 Conclusions

In our investigation into the theory of cratering statistics, we have laid out the fundamental math behind the Age Problem, the Slope Problem, and the Plotting Problem – much of which has never been used in planetary science. We confirm the insights of previous work: both the Slope Problem (Robbins et al., 2018) and the Age Problem (Michael et al., 2016) can be directly solved from the underlying math instead of plot fitting, and plot-based methods will introduce problematic errors and biases. However, we have identified and addressed several errors and biases with previous methods.

Previous Age Problem methods did not apply the rules of random variable math when applying the inverse chronology function to the crater count PDF, or in the case of Michael et al. (2016), they applied an unrealistic prior assumption. As a result, they produced skewed ages whenever the cratering rate is not constant. More fundamentally, they excluded all sources of error except for Poisson counting error. In our Age Problem method, we properly apply random variable math to calculate the correct age PDF.

With counting error, we have identified the PDF of model count λ as a Gamma distribution. We have developed the math to incorporate random and systematic human crater identification error, random and systematic diameter measurement error, area error, estimated chronology system error, D_{min} selection error, and additional error factors. Without our corrections, diameter measurement error typically produces artificially high ages because of the steep exponential slope of most production functions. Reviewing the literature, we have developed practical recommendations for estimates of these error factors. To demonstrate our recommended method, we provide an example Age Problem calculation.

For the Slope Problem over an open interval, we derive a fully normalized slope PDF from the Pareto distribution and prove that it follows the Gamma distribution. With synthetic modeling, we show that least squares fits to cumulative plots and the Robbins et al. (2018) MLE approximation are biased at low N . For the Robbins et al. (2018) MLE approximation, we analytically derive this bias, exactly reproducing

the synthetic results. For the Slope Problem over a closed interval, we derive a numerically normalized slope PDF from the Truncated Pareto distribution.

For the Plotting Problem, we have developed a much more accurate method for making differential plots. Our Slanted Average Shifted Histogram (SASH) method plots the full form of the differential plot function (which is an unnormalized PDF). In synthetic tests, it significantly outperforms both Arvidson et al. (1979) binned plots and the Robbins et al. (2018) Kernel Density Estimator, providing a much closer match to the production function used to generate synthetic data. Because many other fields measure the shape of roughly Pareto PDFs, the SASH method also has significant applications in other fields, including the mathematical problem of tail density estimation.

We also identify and correct biased sampling in the Arvidson et al. (1979) version of the “unbinned” cumulative plot, and we recommend avoiding cumulative plots because they do not show the full form of the PDF. Instead, we recommend making differential or R (relative) plots with the SASH model, which can much more clearly resolve the form of the SFD, especially at the low- N , high-diameter end.

To demonstrate the explanatory power of the SASH model, we apply it to crater counts of cratered plains distributions of the moons of Saturn (Kirchoff and Schenk, 2009; Kirchoff and Schenk, 2010). With clearer resolution of the SFD, the SASH method reveals that the cratered plains of Mimas, Enceladus, Tethys, Dione, and Rhea show steep $\alpha \geq 2.5$ branches at the large crater end that roll over to shallower slopes for craters smaller than an inflection point. The diameters of these inflection points vary significantly, ranging from ~5km at Enceladus to ~45 km at Rhea, and synthetic modeling shows that these features are statistically robust. Because of methodological limitations with previous statistical methods, previous work (Kirchoff and Schenk, 2010; Bell, 2020; Robbins et al., 2024) was unable to clearly and unambiguously resolve these critical features. Because this observation supports equilibrium saturation, it implies that the heavily interpreted shallow slopes of these SFDs below the inflection points are saturated and do not reflect the production function.

We have written a Python package to allow others to implement our methods. It is available online at: <https://github.com/samwbell/cratersfd>, complete with examples, including the code used to produce the results and figures we present here. Although there are further improvements to be made, our methods free crater counts from many of the methodological problems that have long obscured their full explanatory power.

Acknowledgements. We thank Caleb Fassett for analysis of the error bar method of Kreslavsky (2007) and Kreslavsky et al. (2015) and encouraging us to undertake this work. We thank Michelle Kirchoff for providing the Kirchoff and Schenk (2009, 2010) data and for a conversation about the results of our analysis of their data. We thank Sheng Guo for providing the shapefile for the Yue et al. (2020) counts of the Orientale ejecta and for conversations about area calculation methods. Artificial intelligence was used for code debugging and mathematical research.

Open science statements

Author contributions The primary author of the manuscript, S. Bell also wrote the *cratersfd* code and developed the mathematical methods. The second author of the manuscript, S. J. Robbins provided essential big picture thinking and many helpful specific edits and suggestions for the text, figures, and methods. S. J. Robbins also performed the Canala crater counts.

Code availability We have produced a Python package, *cratersfd*, to allow others to implement our methods. It is available for download free of charge at github.com/samwbell/cratersfd. In this GitHub repo, we also include Jupyter notebooks with the code we used to generate the figures in this manuscript. The code version used in this work has been frozen and preserved as a GitHub release tied to the Bell and Robbins (2026a) Zenodo archive, which is available at: <https://doi.org/10.5281/zenodo.18917772>

Data availability We have produced a Python package, *cratersfd*, to allow others to implement our methods. It is available for download free of charge at github.com/samwbell/cratersfd. In this GitHub repo, we also include Jupyter notebooks with the code we used to generate the figures in this manuscript. To ensure that the code can be run without repeating time-intensive calculations, we have uploaded datasets of saved calculation results that can be placed in the *calc/saved* directory of the repo, allowing the figures to be generated from these saved results. We have also uploaded the full figure files used in this work, including original PDF versions prior to JPG conversion. The full figure files and datasets of saved calculation results can be found at the Bell and Robbins (2026b) Zenodo archive: <https://doi.org/10.5281/zenodo.13917307>

Funding S. J. Robbins acknowledges partial support from NASA SSERVI award 80NSSC23M0176.

Competing interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Arthur, D. W. G. (1954). The distribution of lunar craters. *Journal of the British Astronomical Association*, 64, 127-132.
- Arvidson, R. E., Boyce, J., Chapman, C., et al. (1979). Standard techniques for presentation and analysis of crater size-frequency data. *Icarus*, 37, 467-474. [https://doi.org/10.1016/0019-1035\(79\)90009-5](https://doi.org/10.1016/0019-1035(79)90009-5)
- Bell, S. W. (2020). Relative crater scaling between the major moons of saturn: Implications for planetocentric cratering and the surface age of Titan. *Journal of Geophysical Research: Planets*, 125, e2020JE006392. <https://doi.org/10.1029/2020JE006392>
- Bell, S. W., & Robbins, S. J. (2026a). samwbell/cratersfd: v1.0.2 (v1.0.2). Zenodo. <https://doi.org/10.5281/zenodo.18917772>
- Bell, S. W., & Robbins, S. J. (2026b). Saved calculations for the Python package cratersfd. Zenodo dataset. <https://doi.org/10.5281/zenodo.13917307>
- Bottke Jr, W. F., Vokrouhlický, D., Rubincam, D. P., Broz, M., & Smith, D. E. (2001). Dynamical evolution of asteroids and meteoroids using the Yarkovsky effect. *Science*, 294(5547), 1693-1696. <https://doi.org/10.1126/science.1066760>
- Bottke, W. F., Vokrouhlický, D., & Nesvorný, D. (2007). An asteroid breakup 160 Myr ago as the probable source of the K/T impactor. *Nature*, 449(7158), 48-53. <https://doi.org/10.1038/nature06070>
- Bottke, W. F., Vokrouhlický, D., Nesvorný, D., Marschall, R., Morbidelli, A., Deienno, R., et al. (2024). The bombardment history of the giant planet satellites. *The Planetary Science Journal*, 5(4), 88. <https://doi.org/10.3847/PSJ/ad29f4>
- Chapman, C. R., & Haefner, R. R. (1967). A critique of methods for analysis of the diameter-frequency relation for craters with special application to the Moon. *Journal of Geophysical Research*, 72(2), 549-557. <https://doi.org/10.1029/JZ072i002p00549>
- Chapman, C. R., & McKinnon, W. (1986). Cratering of planetary satellites. In Burns, J.A., Matthews, M. (eds.), *Satellites*, The University of Arizona Press, pp. 492-580. <https://doi.org/10.2307/j.ctv1v3gr3r.15>
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data, *SIAM Review*, 51(4), 661-703. <https://doi.org/10.1137/070710111>
- Du, J., Fa, W., Gong, S., Liu, Y., Qiao, L., et al. (2022). Thicknesses of mare basalts in the Chang'e-5 landing region: Implications for the late-stage volcanism on the Moon. *Journal of Geophysical Research: Planets*, 127(8), e2022JE007314. <https://doi.org/10.1029/2022JE007314>
- Fassett, C. I., & Head III, J. W. (2008). The timing of Martian valley network activity: Constraints from buffered crater counting. *Icarus*, 195(1), 61-89. <https://doi.org/10.1016/j.icarus.2007.12.009>
- Ferguson, S. N., Rhoden, A. R., & Kirchoff, M. R. (2020). Small impact crater populations on Saturn's moon Tethys and implications for source impactors in the system. *Journal of Geophysical Research: Planets*, 125(9), e2020JE006400. <https://doi.org/10.1029/2020JE006400>
- Ferguson, S. N., Rhoden, A. R., & Kirchoff, M. R. (2022). Regional impact crater mapping and analysis on Saturn's moon Dione and the relation to source impactors. *Journal of Geophysical Research: Planets*, 127(6), e2022JE007204. <https://doi.org/10.1029/2022JE007204>
- Ferguson, S. N., Rhoden, A. R., & Kirchoff, M. R. (2024). Mimas: A middle-aged moon of Saturn?. *Earth and Planetary Science Letters*, 642, 118859. <https://doi.org/10.1016/j.epsl.2024.118859>
- Gault, D. E. (1970). Saturation and equilibrium conditions for impact cratering on the lunar surface: Criteria and implications. *Radio Science*, 5, 273-291. <https://doi.org/10.1029/RS005i002p00273>
- Ghent, R. R., Costello, E. S., & Parker, A. H. (2024). The population of young craters on the moon: New catalog and spatial and temporal analysis. *The Planetary Science Journal*, 5(4), 89.
- Hartmann, W. K., Quantin, C., & Mangold, N. (2007). Possible long-term decline in impact rates 2. Lunar impact-melt data regarding impact history. *Icarus*, 186, 11-23. <https://doi.org/10.1016/j.icarus.2006.09.009>
- Hartmann, W. K., Strom, R. G., Grieve, R. A. F., Weidenschilling, S. J., Diaz, J., et al. (1981). *Chronology of Planetary Volcanism by Comparative Studies of Planetary Craters, in Basaltic Volcanism on the Terrestrial Planets*, Pergamon Press, 1050-1127.

- Head III, J. W., Fassett, C. I., Kadish, S. J., Smith, D. E., Zuber, M. T., Neumann, G. A., & Mazarico, E. (2010). Global distribution of large lunar craters: Implications for resurfacing and impactor populations. *Science*, 329(5998), 1504-1507. <https://doi.org/10.1126/science.1195050>
- Hirabayashi, M., Minton, D. A., & Fassett, C. I. (2017). An analytical model of crater count equilibrium. *Icarus*, 289, 134-143. <https://doi.org/10.1016/j.icarus.2016.12.032>
- Holsapple, K. A. (1993). The scaling of impact processes in planetary sciences. *Annual review of earth and planetary sciences*, 21(1), 333-373. <https://doi.org/10.1146/annurev.earth.21.050193.002001>
- Ivanov, B. A. (2001). Mars/Moon cratering rate ratio estimates. *Space Science Reviews*, 96(1), 87-104. <https://doi.org/10.1023/A:1011941121102>
- Ivanov, B. A., Neukum, G., & Wagner, R. (2001). Size-frequency distributions of planetary impact craters and asteroids. In *Collisional processes in the solar system*, Springer, 1-34. https://doi.org/10.1007/978-94-010-0712-2_1
- Jia, M., Yue, Z., Di, K., Liu, B., Liu, J., & Michael, G. (2020). A catalogue of impact craters larger than 200 m and surface age analysis in the Chang'e-5 landing area. *Earth and Planetary Science Letters*, 541, 116272. <https://doi.org/10.1016/j.epsl.2020.116272>
- Kirchoff, M. R. (2018). Can spatial statistics help decipher impact crater saturation? *Meteoritics & Planetary Science*, 53(4), 874-890. <https://doi.org/10.1111/maps.13014>
- Kirchoff, M. R., & Schenk, P. (2009). Crater modification and geologic activity in Enceladus' heavily cratered plains: Evidence from the impact crater distribution. *Icarus*, 202(2), 656-668. <https://doi.org/10.1016/j.icarus.2009.03.034>
- Kirchoff, M. R., & Schenk, P. (2010). Impact cratering records of the mid-sized, icy saturnian satellites. *Icarus*, 206(2), 485-497. <https://doi.org/10.1016/j.icarus.2009.12.007>
- Kirchoff, M. R., Bierhaus, E. B., Dones, L., Robbins, S. J., Singer, K. N., Wagner, R. J., & Zahnle, K. J. (2018). Cratering histories in the Saturnian system. In *Enceladus and the icy moons of Saturn*, The University of Arizona Press, pp. 267-284. https://doi.org/10.2458/azu_uapress_9780816537075-ch013
- Kneissl, T., van Gasselt, S., & Neukum, G. (2011). Map-projection-independent crater size-frequency determination in GIS environments-New software tool for ArcGIS. *Planetary and Space Science*, 59(11-12), 1243-1254. <https://doi.org/10.1016/j.pss.2010.03.015>
- König, B. (1977). Study of primary and secondary impact structures of the Moon and laboratory experiments to study the ejection of secondary particles. Dissertation, University of Heidelberg.
- Kreslavsky, M. A. (2007). Statistical Characterization of Spatial Distribution of Impact Craters: Implications to Present-Day Cratering Rate on Mars. Seventh International Conference on Mars, 1353, 3325.
- Kreslavsky, M. A., Ivanov, M., & Head, J. (2015). The resurfacing history of Venus: Constraints from buffered crater densities. *Icarus*, 250, 438-450. <https://doi.org/10.1016/j.icarus.2014.12.024>
- Lagain, A., Benedix, G. K., Servis, K., Baratoux, D., Doucet, L. S., et al. (2021). The Tharsis mantle source of depleted shergottites revealed by 90 million impact craters. *Nature Communications*, 12(1), 6352. <https://doi.org/10.1038/s41467-021-26648-3>
- Le Feuvre, M., & Wieczorek, M. A. (2011). Nonuniform cratering of the Moon and a revised crater chronology of the inner Solar System. *Icarus*, 214(1), 1-20. <https://doi.org/10.1016/j.icarus.2011.03.010>
- Lissauer, J. J., Squyres, S. W., & Hartmann, W. K. (1988). Bombardment history of the Saturn system. *Journal of Geophysical Research: Solid Earth*, 93(B11), 13776-13804. <https://doi.org/10.1029/JB093iB11p13776>
- Liu, J., Yue, Z., Di, K., Gou, S., & Lin, Y. (2023). New lunar crater production function based on high-resolution images. *Remote Sensing*, 15(9), 2421. <https://doi.org/10.3390/rs15092421>
- Losiak, A., Wilhelms, D. E., Byrne, C. J., Thaisen, K. G., Weider, S. Z., et al. (2009). A new lunar impact crater database. Lunar and Planetary Science Conference Abstract.
- Luo, F., Xiao, Z., Wang, Y., Ma, Y., Xu, R., Wang, S., et al. (2024). The production population of impact craters in the Chang'e-6 landing mare. *The Astrophysical Journal Letters*, 974(2), L37. <https://doi.org/10.3847/2041-8213/ad821a>

- McEwen, A. S., Preblich, B. S., Turtle, E. P., Artemieva, N. A., Golombek, M. P., Hurst, M., et al. (2005). The rayed crater Zunil and interpretations of small impact craters on Mars. *Icarus*, 176, 351-381. <https://doi.org/10.1016/j.icarus.2005.02.009>
- Michael, G. G. (2013). Planetary surface dating from crater size–frequency distribution measurements: Multiple resurfacing episodes and differential isochron fitting. *Icarus*, 226(1), 885-890. <https://doi.org/10.1016/j.icarus.2013.07.004>
- Michael, G., & Liu, J. (2025). Planetary surface dating from crater size–frequency distribution measurements: Interpretation of small-area and low number counts. *Icarus*, 431, 116489. <https://doi.org/10.1016/j.icarus.2025.116489>
- Michael, G.G., & Neukum, G. (2010). Planetary surface dating from crater size–frequency distribution measurements: Partial resurfacing events and statistical age uncertainty. *Earth and Planetary Science Letters*, 294, 223-229. <https://doi.org/10.1016/j.epsl.2009.12.041>
- Michael, G.G., Kneissl, T., & Neesemann, A. (2016). Planetary surface dating from crater size-frequency distribution measurements: Poisson timing analysis. *Icarus*, 277, 279-285. <https://doi.org/10.1016/j.icarus.2016.05.019>
- Michael, G. G., Yue, Z., Gou, S., & Di, K. (2021). Dating individual several-km lunar impact craters from the rim annulus in region of planned Chang'e-5 landing: Poisson age-likelihood calculation for a buffered crater counting area. *Earth and Planetary Science Letters*, 568, 117031. <https://doi.org/10.1016/j.epsl.2021.117031>
- Michael, G., Zhang, L., Wu, C., & Liu, J. (2025). Planetary surface dating from crater size–frequency distribution measurements: Sequence probability and simultaneous formation. Did the close Chang'E-6 mare units form simultaneously? *Icarus*, 438, 116644. <https://doi.org/10.1016/j.icarus.2025.116644>
- Minton, D., Fassett, C., Hirabayashi, M., Howl, B., & Richardson, J. (2019). The equilibrium size-frequency distribution of small craters reveals the effects of distal ejecta on lunar landscape morphology. *Icarus*, 326, 63-87. <https://doi.org/10.1016/j.icarus.2019.02.021>
- Morota, T., & Furumoto, M. (2003). Asymmetrical distribution of rayed craters on the Moon. *Earth and Planetary Science Letters*, 206(3-4), 315-323. [https://doi.org/10.1016/S0012-821X\(02\)01111-1](https://doi.org/10.1016/S0012-821X(02)01111-1)
- Neukum, G. (1983). Meteoritenbombardement und Datierung Planetarer Oberflächen, Habilitation Dissertation for Faculty Membership, University of Munich, 186 pp.
- Neukum, G., & Ivanov, B.A. (1994). Crater Size distribution and Impact Probabilities on Earth from Lunar, Terrestrial-planet, and Asteroid Cratering Data, in T. Gehrels (ed.), *Hazards due to Comets and Asteroids*, The University of Arizona Press, pp. 359-416. <https://doi.org/10.2307/j.ctv23khmpv.18>
- Neukum, G. & Wise, D.U. (1976). Mars: A standard crater curve and possible new time scale. *Science*, 194, 1381-1387. <https://doi.org/10.1126/science.194.4272.1381>
- Neukum, G., König, B., & Arkani-Hamed, J. (1975a). A Study of Lunar Impact Crater Size distributions. *The Moon* 12, 201-229. <https://doi.org/10.1007/BF00577878>
- Neukum, G., König, B., Fechtig, H., Storz, D. (1975b). *Cratering in the earth-moon system: Consequences for age determination by crater counting*. Proceedings of the 6th Lunar Science Conference, 2597-2620.
- Neukum, G., & König, B. (1976). Dating of individual lunar craters. *Proceedings of the 7th Lunar Science Conference*, 2867-2881.
- Neukum, G., Ivanov, B.A., & Hartmann, W.K. (2001). Cratering records in the inner solar system in relation to the lunar reference system. *Space Science Reviews*, 96, 55-86. <https://doi.org/10.1023/A:1011989004263>
- Ostrach, L. R. (2013). Impact-related processes on Mercury and the Moon. Doctoral dissertation, Arizona State University.
- Povilaitis, R. Z., Robinson, M. S., Van der Bogert, C. H., Hiesinger, H., Meyer, H. M., & Ostrach, L. R. (2018). Crater density differences: Exploring regional resurfacing, secondary crater populations, and crater saturation equilibrium on the moon. *Planetary and Space Science*, 162, 41-51. <https://doi.org/10.1016/j.pss.2017.05.006>

- Qian, Y., Xiao, L., Head, J. W., van der Bogert, C. H., Hiesinger, H., & Wilson, L. (2021). Young lunar mare basalts in the Chang'e-5 sample return region, northern Oceanus Procellarum. *Earth and Planetary Science Letters*, 555, 116702. <https://doi.org/10.1016/j.epsl.2020.116702>
- Qian, Y., Head, J., Michalski, J., Wang, X., Van der Bogert, C. H., Hiesinger, H., et al. (2024). Long-lasting farside volcanism in the Apollo basin: Chang'e-6 landing site. *Earth and Planetary Science Letters*, 637, 118737. <https://doi.org/10.1016/j.epsl.2024.118737>
- Richardson, J. E. (2009). Cratering saturation and equilibrium: A new model looks at an old problem. *Icarus*, 204(2), 697-715. <https://doi.org/10.1016/j.icarus.2009.07.029>
- Robbins, S.J. (2014). New crater calibrations for the lunar crater-age chronology. *Earth and Planetary Science Letters*, 403, 188-198. <https://doi.org/10.1016/j.epsl.2014.06.038>
- Robbins, S. J. (2019). A new global database of lunar impact craters > 1-2 km: 1. Crater locations and sizes, comparisons with published databases, and global analysis. *Journal of Geophysical Research: Planets*, 124(4), 871-892. <https://doi.org/10.1029/2018JE005592>
- Robbins, S.J., Antonenko, I., Kirchoff, M.R., Chapman, C.R., Fassett, C.I., et al. (2014). The variability of crater identification among expert and community crater analysts. *Icarus*, 234, 109-131. <https://doi.org/10.1016/j.icarus.2014.02.022>
- Robbins, S.J., Riggs, J.D., Weaver, B.P., Bierhaus, E.B., & Chapman, C.R. (2018). Revised recommended methods for analyzing crater size-frequency distributions. *Meteoritics and Planetary Science*, 53, 891-931. <https://doi.org/10.1111/maps.12990>
- Robbins, S. J., Riggs, J. D., & Parker, A. H. (2021). Is the diameter of Herschel crater, Mimas, an outlier? A mathematical framework for analyzing planetary feature size-frequency distribution anomalies. *Geophysical Research Letters*, 48(14), e2021GL093247. <https://doi.org/10.1029/2021GL093247>
- Robbins, S. J., Bierhaus, E. B., & Dones, L. (2024). Crater populations of the Saturnian satellites Mimas, Rhea, and Iapetus. *Journal of Geophysical Research: Planets*, 129(1), e2023JE007941. <https://doi.org/10.1029/2023JE007941>
- Robbins, S. J., Kirchoff, M. R., & Ostrach, L. R. (2025). Crater detection dependence on resolution, incidence angle, emission angle, and phase angle. *Geophysical Research Letters*, 52(4), e2024GL110570. <https://doi.org/10.1029/2024GL110570>
- Rytgaard, M. (1990). Estimation in the Pareto distribution. *ASTIN Bulletin: The Journal of the IAA*, 20(2), 201-216. <https://doi.org/10.2143/AST.20.2.2005443>
- Schenk, P. M., & Zahnle, K. (2007). On the negligible surface age of Triton. *Icarus*, 192(1), 135-149. <https://doi.org/10.1016/j.icarus.2007.07.004>
- Scott, D. W. (1985). Averaged shifted histograms: effective nonparametric density estimators in several dimensions. *The Annals of Statistics*, 1024-1040.
- Shoemaker, E. M. (1965). Preliminary analysis of the fine structure of the lunar surface in Mare Cognitum [Jet Propulsion Laboratory Technical Report No. 32-700]. In W. N. Hess, D. H. Menzel, & J. A. O'Keefe (Eds.), *The nature of the lunar surface*, Johns Hopkins Press, 23-77.
- Snape, J. F., Nemchin, A. A., Whitehouse, M. J., Merle, R. E., Hopkinson, T., & Anand, M. (2019). The timing of basaltic volcanism at the Apollo landing sites. *Geochimica et Cosmochimica Acta*, 266, 29-53.
- Van der Bogert, C. H., Hiesinger, H., Dundas, C. M., Krüger, T., McEwen, A. S., Zanetti, M., & Robinson, M. S. (2017). Origin of discrepancies between crater size-frequency distributions of coeval lunar geologic units via target property contrasts. *Icarus*, 298, 49-63. <https://doi.org/10.1016/j.icarus.2016.11.040>
- Wang, Y., Xie, M., Xiao, Z., & Cui, J. (2020). The minimum confidence limit for diameters in crater counts. *Icarus*, 341, 113645. <https://doi.org/10.1016/j.icarus.2020.113645>
- Werner, S. C., Bultel, B., & Rolf, T. (2023). Review and revision of the lunar cratering chronology-lunar timescale Part 2. *The Planetary Science Journal*, 4(8), 147. <https://doi.org/10.3847/PSJ/acdc16>
- Wilhelms D. E., Oberbeck V. R., Aggarwal H.R. (1978) Size-frequency distributions of primary and secondary lunar impact craters. *Proceedings of the Lunar and Planetary Science Conference*, 9, 3735-3762.

- Williams, J. P., van der Bogert, C. H., Pathare, A. V., Michael, G. G., Kirchoff, M. R., & Hiesinger, H. (2018). Dating very young planetary surfaces from crater statistics: A review of issues and challenges. *Meteoritics & Planetary Science*, 53(4), 554-582. <https://doi.org/10.1111/maps.12924>
- Wong, E. W., Brasser, R., Werner, S. C., & Kirchoff, M. R. (2023). Saturn's ancient regular satellites. *Icarus*, 406, 115763. <https://doi.org/10.1016/j.icarus.2023.115763>
- Xiao, Z., & Strom, R.S (2012). Problems determining relative and absolute ages using the small crater population. *Icarus*, 220, 254-267. <https://doi.org/10.1016/j.icarus.2012.05.012>
- Xiao, Z., & Werner, S. C. (2015). Size-frequency distribution of crater populations in equilibrium on the Moon. *Journal of Geophysical Research: Planets*, 120(12), 2277-2292. <https://doi.org/10.1002/2015JE004860>
- Yue, Z., Yang, M., Jia, M., Michael, G., Di, K., Gou, S., & Liu, J. (2020). Refined model age for Orientale Basin derived from zonal crater dating of its ejecta. *Icarus*, 346, 113804. <https://doi.org/10.1016/j.icarus.2020.113804>

Appendix

A Rules of random variable math

Here, we lay out the rules of random variable math we use in this work. Note: we did not invent these rules; they are well-established theorems we are restating for the reader's convenience.

The equation for applying a function f to a random variable x is:

$$P_y(y = f(x)) = P_x(x = f^{-1}(y)) \left| \frac{df^{-1}(y)}{dy} \right|, \quad (\text{A1})$$

where $y = f(x)$, P_x is the PDF of x , P_y is the PDF of y , and f^{-1} is the inverse function of f . For instance, when applying the inverse chronology function $\mathcal{C}^{-1}$ to an $N(1)$ random variable, the inverse of the inverse chronology function is the chronology function $\mathcal{C}$, so the equation for the PDF of age t when the inverse chronology function is applied to a PDF of $N(1)$ value $n(1)$ is:

$$P_T(t = \mathcal{C}^{-1}(n(1))) = P_{N(1)}(n(1) = \mathcal{C}(t)) \frac{d(\mathcal{C}(t))}{dt}. \quad (\text{A2})$$

Equation (A1) can also be applied to any given mathematical expression. For instance:

$$P_Y(y = x^a) = \frac{y^{\frac{1}{a}-1}}{|a|} P_X\left(y^{\frac{1}{a}}\right) \text{ and } P_Y(y = a^x) = \frac{1}{y \ln a} P_X(\log_a y). \quad (\text{A3})$$

If the derivative or the inverse function cannot be computed analytically, we have implemented methods in *cratersfd* to calculate them numerically for any arbitrary function. (For functions whose first derivative changes sign, *cratersfd* will automatically find the inflection points, calculate multiple inverse functions, and combine the probability values.)

To add or subtract two random variables together, the formulas are:

$$P_{C=A+B}(c) = \int_{-\infty}^{\infty} P_A(x) P_B(c-x) dx \quad \text{and} \quad P_{C=A-B}(c) = \int_{-\infty}^{\infty} P_A(x) P_B(c+x) dx \quad (\text{A4})$$

where x is an arbitrary integration variable, and A , B , and C are random variables. (In mathematical terms, Equation (A4) is the same as the convolution operation used in signal processing, and it can be calculated efficiently with Fast Fourier Transforms.) The formulas for multiplying and dividing two random variables are:

$$P_{C=AB}(c) = \int_{-\infty}^{\infty} \frac{P_A(x) P_B\left(\frac{c}{x}\right)}{|x|} dx \quad \text{and} \quad P_{C=A/B}(c) = \int_{-\infty}^{\infty} P_A(x) P_B(cx) |x| dx. \quad (\text{A5})$$

In practice, as long as both random variables are always positive, it is typically much more computationally efficient to apply the log function to the random variables, add or subtract them, then apply the function $f(x) = 10^x$ to the result. This is mathematically equivalent to multiplication or division, and it allows the use of Fast Fourier Transforms.

Because we were unable to identify a Python package for performing random variable math, we built one. With *cratersfd*, users can create RandomVariable objects. Functions generate RandomVariable objects for use cases like λ , a , crater SFDs, and ages. Users can create their own custom RandomVariable objects with X and P arrays that define the PDF. To plot PDFs, there is a `.plot()` method. To apply a function, there is an `.apply()` method. Simple arithmetic happens symbolically. For instance, if $rv1$, $rv2$, and $rv3$ are all RandomVariable objects, users could type: $(rv1 + rv2) / rv3$. Error bars can be calculated for any of the error methods in Appendix C, and RandomVariable objects have a long list of methods that provide many other functionalities.

If both random variables follow a normal distribution, when they are multiplied together, the mean will be the product of the two means, and the σ parameter of the normal distribution (which equals the standard deviation in the case of the normal distribution) will follow what is known as the quadrature rule:

$$\sigma_{AB} = \sqrt{\sigma_A^2 + \sigma_B^2}. \quad (\text{A6})$$

Under the lognormal quadrature rule, when two lognormal random variables are multiplied together, Equation (A6) applies to the $\sigma = \ln(1 + p)$ parameter of the lognormal distribution (which is not the same as the standard deviation of the lognormal distribution). For two lognormal variables with a percentage error, where p is percentage (e.g. $p = 0.3$ for 30%), the percentage error of the multiplied lognormal random variables will be:

$$p_{AB} = e^{\sqrt{(\ln(1+p_A))^2 + (\ln(1+p_B))^2}} - 1. \quad (\text{A7})$$

To combine the information from two PDFs of the same variable, one representing our prior knowledge and one representing the likelihood information we gain from additional data, we use Bayes's Rule:

$$P_{\text{posterior}}(x) = \frac{P_{\text{likelihood}}(x)P_{\text{prior}}(x)}{\int_{-\infty}^{\infty} P_{\text{likelihood}}(t)P_{\text{prior}}(t)dt}. \quad (\text{A8})$$

We multiply together the probability values of the two PDFs at every value of the variable and then normalize the resulting PDF so that it integrates to 1.

For the Age Problem, if the prior is quantified in terms of age, then $\mathcal{C}^{-1}$ should not be applied to the $N(1)$ PDF with Equation (A1). Instead, the $N(1)$ PDF should be scaled by $\mathcal{C}^{-1}$, as Michael et al. (2016) did. However, if an age prior is used, we strongly discourage a uniform age prior, which biases the model in a way that rarely reflects prior information. Instead, if age analysis is being done with Bayesian priors, we recommend constructing specific priors based on prior information specific to the problem at hand. When this is done, we strongly recommend incorporating the prior information that the cratering rate itself provides: craters and geologic features linked to cratering are likelier to have formed when the cratering rate is higher.

A full discussion of priors and Bayesian analysis lies outside this scope of this work. What we propose here instead is a practical consistent standard. Instead of imposing a specific prior structure on the data, we suggest simply applying random variable math to the data we observe. (Implicitly, this is equivalent to a uniform prior in crater count space.) If priors are used, we strongly recommend that priors be carefully constructed and clearly disclosed.

A.1 Incorporating human crater identification error

In Section 2.2, we model the random human crater identification error kernel with a lognormal error kernel with percentage p_r around each individual crater. Here, we derive an equation for the percentage error p_h of human crater identification error assuming that p_r is identical for each crater. To calculate $\mathcal{E}_R$, the error kernel around the model count λ , we need the sum of each of the individual random error kernels $\mathcal{E}_{r_i}$ divided by the total count N (Equation (8)):

$$\mathcal{E}_R = \frac{\sum_{i=1}^N \mathcal{E}_{r_i}}{N}. \quad (\text{A9})$$

To calculate the sum, we apply the theorem for the variance (Equation (C2)) of a lognormal random variable X :

$$\text{Var}(\mathcal{E}_R) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}. \quad (\text{A10})$$

Because $\mu = -\frac{\sigma^2}{2}$, and $\sigma = \ln(1 + p)$ (Equation (7)), this reduces to:

$$\text{Var}(\mathcal{E}_{r_i}) = e^{(\ln(1+p_{r_i}))^2} - 1, \quad (\text{A11})$$

where p_{r_i} is the percentage error for $\mathcal{E}_{r_i}$. The Bienaymé Identity gives the variance of the sum of independent variables X_i with weights a_i :

$$\text{Var}\left(\sum_{i=1}^N a_i X_i\right) = \sum_{i=1}^N a_i^2 \text{Var}(X_i). \quad (\text{A12})$$

Thus:

$$\text{Var}(\mathcal{E}_R) = \text{Var}\left(\frac{\sum_{i=1}^N \mathcal{E}_{r_i}}{N}\right) = \sum_{i=1}^N \frac{\text{Var}(\mathcal{E}_{r_i})}{N^2}. \quad (\text{A13})$$

Assuming the $\mathcal{E}_{r_i}$ are identical:

$$\text{Var}(\mathcal{E}_R) = \sum_{i=1}^N \frac{\text{Var}(\mathcal{E}_{r_i})}{N^2} = \frac{N \text{Var}(\mathcal{E}_{r_i})}{N^2} = \frac{\text{Var}(\mathcal{E}_{r_i})}{N}. \quad (\text{A14})$$

Applying Equation (A11):

$$\text{Var}(\mathcal{E}_R) = \frac{e^{(\ln(1+p_r))^2} - 1}{N}. \quad (\text{A15})$$

Next, we apply the Fenton-Wilkinson Approximation, approximating $\mathcal{E}_R$ as a lognormal random variable. Although the sum of lognormal random variables is not *quite* lognormal, this common approximation is very good – almost certainly more accurate than the approximation that human crater identification error is perfectly lognormal. Because $\mu = -\frac{\sigma^2}{2}$, Equation (A10) becomes:

$$\text{Var}(\mathcal{E}_R) = e^{\sigma^2} - 1. \quad (\text{A16})$$

Then, we solve for the $\sigma = \ln(1 + p)$ parameter of the lognormal random variable $\mathcal{E}_R$:

$$e^{\sigma^2} - 1 = \frac{e^{(\ln(1+p_r))^2} - 1}{N} \rightarrow \sigma = \sqrt{\ln\left(\frac{e^{(\ln(1+p_r))^2} - 1}{N} + 1\right)}. \quad (\text{A17})$$

Applying the lognormal quadrature rule (Equations (A6) and (A7)), when we combine $\mathcal{E}_R$ with the kernel for systematic counting error, we produce a final equation for the percentage error p_h of the combined human crater identification error kernel:

$$p_h = e^{\sqrt{\ln\left(\frac{e^{(\ln(1+p_r))^2} - 1}{N} + 1\right) * (\ln(1+p_s))^2}} - 1. \quad (\text{A18})$$

A.2 Incorporating diameter measurement error

Correcting for diameter measurement error presents several complications, each of which we have solved. To begin with, there is a difference between the PDF that describes the error in the forward problem, measuring a crater, and the PDF of possible values of a crater's diameter given a known measurement error, the reverse problem. This distinction is similar to the difference between the observed count N and the model count λ . To find the reverse problem PDF, we simply express the lognormal PDF equation in terms of the true diameter D instead of the observed diameter D_{observed} :

$$P(D) \propto \frac{e^{-\left(\frac{\ln D_{\text{observed}} - \ln D + \frac{(\ln(1+p))^2}{2}}{2(\ln(1+p))^2}\right)^2}}{D_{\text{observed}} \sqrt{2\pi \ln(1+p)}}, \quad (\text{A19})$$

where p is the percentage error.

As we describe in Section 2.3, the shape of the production function means that diameter measurement error causes an artificially high N at any given diameter. To correct for this effect, we remember that diameter measurement is only one source of information about crater diameter. The SFD itself is another. Before the measurement is made, the SFD provides us with prior expectations about crater diameter. Both the measurement uncertainty and the SFD are PDFs of diameter. In Bayesian terminology, the SFD is the prior, and the measurement uncertainty is the likelihood, so we incorporate them together into one PDF with Bayes's Rule (Equation (A8)). To find the SFD, we use the SASH algorithm (Section 4.2). Once we have combined the SFD and the diameter measurement kernel, we have a diameter error kernel we can apply to any diameter.

The mathematical treatment of diameter measurement error depends on whether it is random or systematic. Systematic error is simpler to address because it shifts the whole curve the same way, so it is equivalent to the curve being measured at a different D_{min} . To incorporate systematic error, therefore, we apply the systematic diameter error kernel $\mathcal{E}_{D_s}$ to D_{min} , creating a D_{min} PDF. Then, we simply apply random variable math. (If there is a D_{max} , we apply the function $f(\mathcal{E}_{D_s}) = p(D_{\text{min}}\mathcal{E}_{D_s}) - p(D_{\text{max}}\mathcal{E}_{D_s})$ to $\mathcal{E}_{D_s}$.)

With random error, however, we must address each crater individually. When we apply the combined diameter error PDF to each individual diameter, it is no longer certain whether it falls above or below D_{min} . If it is close to D_{min} , the chance that it lies between D_{min} and D_{max} (and should be included in the count of N for the Age Problem) equals the integrated probability of the PDF from D_{min} to infinity (or D_{max}). Each crater, then, has a probability of counting towards N . Because N is a discrete variable which must have whole number values, the distribution of N values will be a Probability Mass Function (PMF) instead of a Probability Density Function (PDF). To calculate the PMF of N , we turn to the Poisson-Binomial distribution, which describes the PMF of the sum of individual trials, each with a probability of counting towards the sum:

$$P(N) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-iNt} \prod_{k=1}^n ((1 - p_k) + p_k e^{it}) dt = \mathcal{F}^{-1} \left(\prod_{k=1}^n ((1 - p_k) + p_k e^{it}) \right), \quad (\text{A20})$$

where n is the total number of individual probabilities p_k . By expressing Equation (A20) in terms of the inverse Fourier operator $\mathcal{F}^{-1}$, we can evaluate it efficiently with Fast Fourier Transforms.

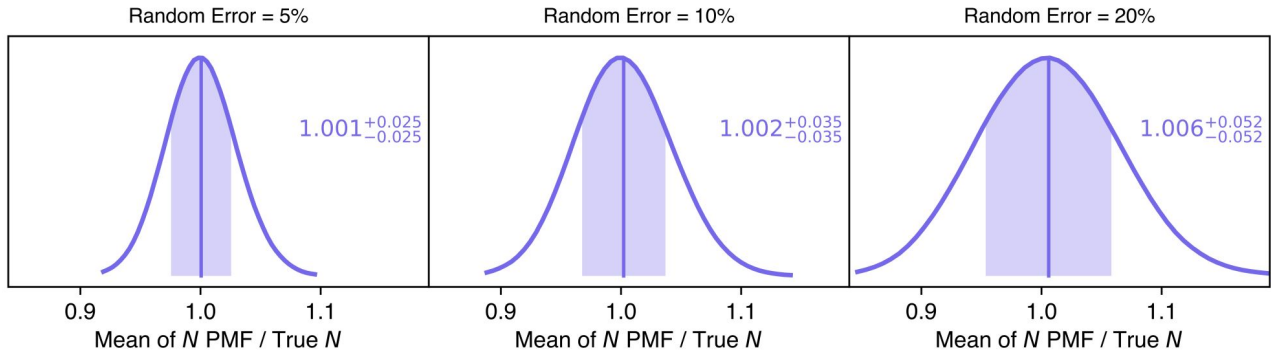


Figure A1 | A synthetic model testing our method for correcting the bias from random diameter measurement error (see Section 2.3 and Appendix A.2). Here, we simulated measuring N at $D_{\min} = 30$ m from 2000 $D \geq 10$ m NPF synthetics with diameter error added in. For each synthetic, we apply our corrections and calculate a PMF of predicted N values, then we calculate the ratio of our mean estimate of N to the true N . These results show that we can successfully remove nearly all the bias. For instance, our method reduces the 11.9% bias at 20% in Figure 2 to just 0.5%. The primary reason for the very small remaining bias is that diameter measurement error slightly biases the shape of the SFD.

Once we have the PMF of N , we use it to calculate the PDF of λ . When random variable math is applied to a discrete random variable, the resulting PDF (or PMF) becomes a weighted sum of the PDFs (or PMFs) calculated from each possible value of the discrete random variable. To calculate the PDF of λ from the N PMF, we calculate a PDF of λ for each possible value of N with Equation (6) and combine these PDFs in a weighted sum:

$$P(\lambda; N) = \sum_{i=1}^k w_i P(\lambda; N_i), \quad (\text{A21})$$

where the weight w_i is the probability of N_i under the N PMF. This is not a sum of random variables; the individual probability values are summed, and then the resulting PDF is normalized. Figures 12u–12w show an example of this calculation when the $D_{\min}$ PDF is approximated as a PMF (see Appendix A.3). Finally, we take the λ PDF that incorporates the N PMF from random diameter error and apply Equation (11) to calculate the age.

To test our method, once again, we generate a random set of craters and apply random diameter error to each crater to produce a synthetic dataset with random error. For each synthetic, we apply our corrections and calculate a PMF of predicted N values, then we calculate the ratio of our mean estimate of N to the true N . As Figure A1 shows, our method accurately estimates the true N , correcting the bias of random error in diameter measurement.

While our method is highly accurate, making it perfectly accurate will require future work. For instance, in our synthetic trial of 20% random error, our method retains a small 0.6% bias after correcting for the 11.9% bias. The primary reason for this discrepancy is that diameter measurement error slightly biases the shape of the SFD. With future work to correct the SASH model for random diameter error, this small discrepancy should shrink still further. Future work is also needed to incorporate the uncertainty of the SFD shape measured by the SASH model into our random diameter measurement error problem. This is a nontrivial problem, but the SASH model is accurate enough that these effects are marginal, except at low N .

The math for incorporating our method is fairly complex. To make our diameter measurement error methods easy to use, we have written Python functions to make these calculations and included them in *cratersfd*. When doing an age calculation, *cratersfd* automatically assumes our default recommendations of 5% random and 1% systematic diameter measurement error, and it allows users to specify their own values for their data.

A.3 Incorporating $D_{\min}$ selection error

Because the choice of $D_{\min}$ affects N , incorporating the $D_{\min}$ PDF into the Age Problem calculation is more complex than a simple random variable multiplication. First, we sample the $D_{\min}$ PDF at regular intervals to approximate the PDF as a PMF of discrete values (Figure 12u). (By default, we use 50 samples, but more can be used for higher precision.) The PDF of $N(1)$ then becomes:

$$P(N(1); D_{\min}) = \sum_{i=1}^k w_i P(N(1); D_{\min_i}), \quad (\text{A22})$$

where the weight w_i is the probability of the $D_{\min}$ PDF sample $D_{\min_i}$.

Next, we incorporate random diameter measurement error to calculate the PMF of N for each sampled D_{min} value, as described in Appendix A.2 (Figure 12v). Then, we calculate an $N(1)$ PDF from each D_{min} value and its corresponding N PMF:

$$N(1) = \frac{\left(\mathcal{E}_\lambda = \mathcal{E}_a \mathcal{E}_s \left(\frac{\sum_{i=1}^{N_{observed}} \mathcal{E}_{r_i}}{N_{observed}} \right) \right) \lambda}{A} \left(\frac{p(1km)}{p(\mathcal{E}_{D_s} D_{min})} \right), \quad (A23)$$

calculating λ from the N PMF as described in Appendix A.2 (Figure 12w). These $N(1)$ PDFs are then weighted by the probability of that D_{min} in the D_{min} PDF (Figure 12w). Once the $N(1)$ PDFs are weighted, they are combined with an element-wise addition (Equation (A22)) to produce a final $N(1)$ PDF (Figure 12w). Finally, we apply the inverse chronology function $\mathcal{C}^{-1}$ to the final $N(1)$ PDF to calculate the age PDF (Figure 12r), first multiplying by $\mathcal{E}_c$ if we are including chronology system error (Figure 12s).

B NPF Error Model

To construct an error model for the NPF, we used a linear interpolation in log space between the error estimates of Neukum et al. (1975a) and König (1977) at individual diameters: Neukum et al. (1975a) estimated the error as $\pm 50\%$ at $D = 300$ m $\pm 10\%$ at $D = 800$ m, no error at $D = 1$ km, $\pm 10\%$ at $D = 3$ km and $\pm 25\%$ at $D = 10$ km, and $\pm 50\%$ at $D = 20$ km. With updated data on the lower D end, König (1977) shifted the diameter for $\pm 50\%$ error to $D = 100$ m. Because Neukum (1983) and Ivanov et al. (2001) did not update the large D error estimates, we must estimate the error model improvements from these works. Acknowledging that any choice is arbitrary, we choose to replace the $\pm 25\%$ at $D = 10$ km and $\pm 50\%$ at $D = 20$ km points with a $\pm 50\%$ point at $D = 75$ km.

Neukum et al. (1975a) and König (1977) phrased their error model in terms of error in $N(1)$ at a particular diameter, but for the Age Problem, what matters is *the error relative to the diameters where the chronology function was calibrated*. To calculate the relative error between two diameters, we turn to the lognormal quadrature rule. If the two diameters are on opposite sides of $D = 1$ km, then they are added in lognormal quadrature with Equation (A7). If they are on the same side, then they are subtracted in lognormal quadrature:

$$p_{AB} = e^{\sqrt{(\ln(1+p_{furtherest}))^2 - (\ln(1+p_{closest}))^2}} - 1, \quad (B1)$$

where $p_{closest}$ is the percentage error for the diameter closest to 1 km, and $p_{furtherest}$ is the percentage error for the diameter furthest from 1 km.

C Error bar methods

In this appendix, we have three goals. First, we review the theory of error bars to derive the $\sqrt{N}$ approximation, which Arvidson et al. (1979) simply asserted. Second, we address the specific question of the best error bars to use on Arvidson et al. (1979) plots of crater SFDs for the purpose of slope measurement in the synthetic modeling of cumulative plots (see Section 3.1, Section 4.1, and Appendix E). We review existing methods, show that they are inadequate, and propose a more accurate method. Third, we define the methods we use in the rest of this work and discuss the general problem of representing a PDF as a single value with error bars.

C.1 Derivation of the $\sqrt{N}$ approximation

In an era when less computational power was readily available, Arvidson et al. (1979) standardized error bar methods with an approximation: assume the error is normally distributed with a mean of N and a standard deviation of $\sqrt{N}$. Although this $\sqrt{N}$ approximation has been widely adopted, Arvidson et al. (1979) never justified why it should be used. Here, we derive this approximation and prove that it breaks down when N is not high.

From the mathematics of random variables, any random variable can be described with a series of moments. The first moment is the *mean*, μ :

$$\mu = \int_{-\infty}^{\infty} x f(x) dx, \quad (C1)$$

where x is the random variable, and $f(x)$ is its PDF. The second moment is the *variance* σ^2 :

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx. \quad (C2)$$

By convention, standard error bars in crater counting and many other fields are set equal to the square root of the variance, the *standard deviation*, σ . Even if a distribution is asymmetric, the standard deviation is symmetric. The third moment, the *skewness*, γ , quantifies asymmetry:

$$\gamma = \frac{\int_{-\infty}^{\infty} (x - \mu)^3 f(x) dx}{\sigma^3}. \quad (C3)$$

From the theorems of the Gamma distribution, the mean (μ), standard deviation (σ), and skewness (γ) of the λ likelihood PDF are:

$$\mu = \frac{s}{r}, \quad \sigma = \frac{\sqrt{s}}{r}, \quad \text{and} \quad \gamma = \frac{2}{\sqrt{s}} \xrightarrow{s=N+1; r=1} \mu = N + 1, \quad \sigma = \sqrt{N + 1}, \quad \text{and} \quad \gamma = \frac{2}{\sqrt{N + 1}}, \quad (C4)$$

where $s = N + 1$ and $r = 1$ are parameters of the generic form of the Gamma distribution (Equation (5)). As N becomes large, $N + 1 \rightarrow N$, $\mu \rightarrow N$, and $\sigma \rightarrow \sqrt{N}$. If we take the limit as N approaches infinity, the skewness approaches zero, and the PDF approaches symmetry:

$$\lim_{N \rightarrow \infty} \gamma = \frac{2}{\sqrt{N + 1}} = 0. \quad (C5)$$

For large N , the distribution will approach a symmetric distribution with a mean of N and a standard deviation of $\sqrt{N}$. For lower values of N , however, the approximation does not hold. At lower N , the skewness is not negligible, and the PDF is definitively asymmetric. Moreover, at low N , the mean, $N + 1$, cannot be treated as N , and the standard deviation, $\sqrt{N + 1}$, cannot be treated as $\sqrt{N}$. Although valid at high N , the Arvidson et al. (1979) $\sqrt{N}$ approximation is not valid at low N , where counting error is the most significant.

C.2 Existing methods

Despite its inaccuracy at low N , no consensus has emerged on the best alternative to the $\sqrt{N}$ approximation, and it continues to be used for plots. Proposed alternatives (Kreslavsky, 2007; Fassett and Head, 2008; Michael and Neukum, 2010; Kreslavsky et al., 2015; Michael et al., 2016; Robbins et al., 2018; Bell, 2020) use percentiles for the upper and lower bounds. The 15.87th and 84.13th percentiles fall at the 1σ bounds in a normal distribution, and we can use these percentiles to approximate 1σ error bars for a PDF of any shape. For instance, for the Gamma distribution λ PDF from an observation of $N = 4$, these values are 7.16 and 2.84. To determine error bars, we must also pick a central value to represent the PDF, and we will refer to each method by its central value: if it is the mean, we call it the *mean method*; if it is the median, we call it the *median method*; and if it is the mode (maximum likelihood value), we call it the *mode method*.

Kreslavsky (2007) and Kreslavsky et al. (2015) proposed a version of the mode method that used 90% confidence intervals and calculated the lower bound using $N - 1$ instead of N . Neither Kreslavsky (2007) nor Kreslavsky et al. (2015) explain this anomalous choice of lower bound, and we believe it to be a typo. Robbins et al. (2018) rejected the Poisson assumption and generated an error distribution from the bootstrap, and they used the mode method to reduce it to a single value and error bars.

To calculate error bars on ages, Michael and Neukum (2010) and Michael et al. (2016) proposed using the median method, but they continued to use the $\sqrt{N}$ approximation for plots. Bell (2020) applied it to plots, but we disagree with this decision. When applied to plots, the median, mean, and mode methods can be highly misleading – especially when used to measure SFD slopes (as Bell, 2020 did).

The median value on an observation of 1 is 1.68, and the median value on an observation of 20 is 20.67. On a log-log plot, using the median value would artificially deflect low- N bins upwards by much more than higher- N bins, creating artificially shallow slopes. This would only increase the shallow slope bias of fits to cumulative plots (Figure 4).

C.3 Our proposed method: fitting an asymmetric normal distribution

To these existing methods, we suggest an alternative approach. Here, we make the mode (maximum likelihood value) the central value. To determine the error bars, we split the PDF at the mode. For each side, we fit half of a normal distribution. The error bars are the standard deviations of the half normal distribution fits on each side. We call this the *linear method*. To represent the error bars for log-log plots, we scale the PDF into log space before fitting it to two half normal distributions. We call this the *log method*.

Although we do not recommend using Arvidson et al. (1979) plots, if one were to fit a slope to Arvidson et al. (1979) plots, we would recommend log method error bars. In our synthetic tests, log method error bars improve the results significantly (Figure 4). Least squares fits assume a normal distribution, and our asymmetric least squares model assumes the asymmetric normal distribution we fit in the log method.

D Derivation of the open interval slope PDF

As we described in Section 2.1, when we have the equation for a PDF of one variable, the PDF equation will give the relative likelihood of other variables. For the absolute PDF of another variable, we must normalize it so that the PDF integrates to 1 for that variable. This can be done numerically, but if we can show that the new PDF follows a known distribution, we can use the theorems of that distribution to derive the normalized PDF. When we derived the PDF of λ from the Poisson distribution (Equation (6)), we normalized it by proving that it follows a Gamma distribution. When we derive the likelihood PDF of negative slope a from the Pareto distribution, we will follow a similar approach.

First, we write the Pareto distribution (Equation (13)) in terms of a to give the relative likelihood PDF of a :

$$P(a) \propto \frac{a D_{\min}^a}{D^{a+1}}. \quad (D1)$$

Because Equation (D1) gives the relative probability as a function of a , we can remove the D in the denominator because D does not depend on a :

$$P(a) \propto \frac{a D_{\min}^a}{D^a}. \quad (D2)$$

As Equation (6) gave the likelihood PDF of λ from an observation of N , Equation (D2) gives the likelihood PDF of a given an observation of D . This might seem odd: How can a single diameter provide evidence of a slope? The answer is that for every observation of D , there is also a second piece of information: $D_{\min}$. If an individual crater is observed much larger than the minimum diameter, it is evidence of a shallower slope. If it is very close to the minimum diameter, it is evidence of a steeper slope.

Note that Equations (D1) and (D2) are no longer normalized. The normalization of the Pareto PDF assumes D as the variable, not a . To derive the normalization, we reformat Equation (D2):

$$P(a) \propto \frac{a D_{\min}^a}{D^a} \rightarrow P(a) \propto a \left(e^{\ln\left(\frac{D_{\min}}{D}\right)} \right)^a \rightarrow P(a) \propto a e^{-a \ln\left(\frac{D}{D_{\min}}\right)}. \quad (D3)$$

Equation (6) shows the generic form of the Gamma distribution of variable x with shape parameter s and rate parameter r . As is standard for PDFs, it is normalized so that it sums to 1 when integrated to infinity. If we reformat Equation (6) as the relative probability of variable x , then we can remove the normalization because it does not depend on x :

$$P(x) = P(x) = \frac{r^s x^{s-1} e^{-rx}}{\Gamma(s)} \rightarrow P(x) \propto x^{s-1} e^{-rx}, \quad (D4)$$

revealing the unnormalized Gamma distribution. If we compare Equations (D3) and (D4), we can see that, as reformatted, Equation (D3) now takes the form of a Gamma distribution, where:

$$x = a, s = 2, \text{ and } r = \ln\left(\frac{D}{D_{\min}}\right). \quad (D5)$$

Equation (6) gives the normalized form of the Gamma distribution for given values of s and r , so a Gamma distribution formatted as Equation (D4) can be normalized by applying a factor of:

$$\frac{r^s}{\Gamma(s)}. \quad (D6)$$

Applying the s and r values (Equation (D5)), the normalized form of Equation (D3) becomes:

$$P(\alpha) = \alpha \left(\ln \left(\frac{D}{D_{\min}} \right) \right)^2 \left(\frac{D}{D_{\min}} \right)^{-\alpha}. \quad (D7)$$

With Equation (D7) we can calculate the likelihood PDF of negative slope α from a single observation of a crater with diameter D . One crater, of course, is not enough to robustly determine the slope, so we must combine the information from multiple observations. Applying Bayes's Rule, we multiply together the values of the α likelihood PDF at every value of α for each observed diameter D_i . Figure 6n shows the α likelihood PDF for each crater, as well as the combined PDF from all the craters.

While we can calculate it numerically, it is more efficient to calculate the combined PDF analytically. Beginning in relative probability, we multiply together the PDF values from Equation (D2) for every D_i , producing a combined likelihood PDF from every observation:

$$P(\alpha) \propto \alpha^N \prod_{i=1}^N \left(\frac{D_{\min}}{D_i} \right)^\alpha. \quad (D8)$$

Robbins et al. (2018) derived Equation (D8), but their version contains a typo where $(1 + D_i)$ is used instead of D_i .

Next, we apply an algebraic transformation:

$$\prod_{i=1}^N \left(\frac{D_{\min}}{D_i} \right)^\alpha \rightarrow \prod_{i=1}^N \left(\frac{D_i}{D_{\min}} \right)^{-\alpha} \rightarrow e^{\ln \left(\prod_{i=1}^N \left(\frac{D_i}{D_{\min}} \right)^{-\alpha} \right)} \rightarrow e^{\sum_{i=1}^N \ln \left(\frac{D_i}{D_{\min}} \right)^{-\alpha}} \quad (D9)$$

and reformat Equation (D8) into an unnormalized Gamma distribution (Equation (D4)):

$$P(\alpha) \propto \alpha^N e^{\sum_{i=1}^N \ln \left(\frac{D_i}{D_{\min}} \right)^{-\alpha}} \rightarrow P(\alpha) \propto \alpha^N e^{-\alpha \sum_{i=1}^N \ln \left(\frac{D_i}{D_{\min}} \right)}, \quad (D10)$$

where:

$$x = \alpha, \quad s = N + 1, \quad \text{and} \quad r = \sum_{i=1}^N \ln \left(\frac{D_i}{D_{\min}} \right). \quad (D11)$$

Applying the normalization (Equation (D6)), the combined likelihood PDF of α for all observations of D_i becomes:

$$P(\alpha) = \frac{\alpha^N \left(\sum_{i=1}^N \ln \left(\frac{D_i}{D_{\min}} \right) \right)^{N+1} \prod_{i=1}^N \left(\frac{D_{\min}}{D_i} \right)^\alpha}{\Gamma(N + 1)}. \quad (D12)$$

With Equation (D12), we can calculate the explicit PDF of negative cumulative slope α directly from the observed diameters. Like the PDF of model count λ , the PDF of α follows a Gamma distribution.

E Details of synthetic modeling

The synthetic model in Section 2.1 generated crater counts with varying N from a known model count λ . In other applications – such as when we measure the bias of least squares fits to cumulative plots (Sections 3.1 and 4.1) or study the error of the SASH model with synthetic modeling – we have a fixed N . With a fixed N , the computation is much more straightforward. Because crater SFDs are random variables, and the differential plot is an unnormalized PDF, generating a random dataset of N craters from a production function is equivalent to drawing N samples from a PDF.

This is a well-studied problem with many available methods. Like Robbins et al. (2018), we use the Inverse Transform Sampling method for computational efficiency. For this, we leverage the fact that the cumulative plot form of the PDF gives the integrated probability. We generate a D array of 10 000 diameter points from $D_{\min}$ to $D_{\max}$ (for open intervals, we choose a large $D_{\max}$). Then, we apply the cumulative production function to D to create a C array. To normalize the C array, we divide every point by the C array's maximum value, the value at $D_{\min}$. When normalized, the C array ranges from 0 to 1. Then, we generate N random numbers from 0 to 1. To convert each random number to a random crater diameter, we interpolate it from the normalized C array to the D array. For a graphical representation of this method, see Figure 13 of Robbins et al. (2018).

To measure the accuracy of least squares fits to cumulative plots and the Robbins et al. (2018) MLE method, we use a production function with a cumulative slope of -2 ($\alpha = 2$) and a range of N values. Because the Pareto distribution gives the PDF of crater diameter for a distribution with a constant α , when α is constant, we use an even more computationally efficient process to generate synthetics: we

draw N samples from an $\alpha = 2$ Pareto distribution with the `random.pareto()` function in the *NumPy* Python package. For each N , we generate 100 000 craters, measuring an empirical slope PDF from the histogram of each method's results. For these PDFs, we use the mean error bar method.

E.1 Error bars in synthetic modeling of Slope Problem methods

In our least squares synthetic models, we apply three forms of error bar: no error bars, symmetric, and asymmetric. For the asymmetric error bars, we choose the log method, since it most closely approximates the asymmetric normal distribution that our asymmetric least squares method uses. For the symmetric error bars, we averaged the asymmetric error bars (in log space).

In most cases, in the traditional Slope Problem method of fitting the data on a cumulative or differential plot with least squares, the error bars receive very little, if any, discussion. From Neukum (1983) to Schenk and Zahnle (2007) to Liu et al. (2023), plots are shown with Arvidson et al. (1979) $\sqrt{N}$ approximation error bars, but the authors do not indicate what, if any, method they used to incorporate the error bars into the least squares fit, even though different formulations of least squares have different assumptions and will produce different results (e.g., orthogonal versus vertical residual calculations, weighted or unweighted). Without guidance in the literature as to what was actually used, we assume it was no error bars or some symmetric version of $\sqrt{N}$ approximation error bars. Complicating the matter further, if symmetric error bars were used, there is no clear way to convert $\sqrt{N}$ approximation error bars into symmetric error bars, since an average at $N = 1$, where the lower error bar is $-\infty$ in log space, would be infinite. If symmetric error bars were used, perhaps averages were used, with a specialized treatment at $N = 1$, or perhaps only the upper error bars were used. Most symmetric versions of $\sqrt{N}$ error bars should produce broadly similar results to the symmetric versions of log method error bars we use in our symmetric fits, but there will be slight differences.

The least squares method dominates curve fitting methods. Because it assumes normally distributed error, its error bars are assumed to be symmetric. To create an asymmetric version of least squares, we use an iterative approach. First, we average the error bars and apply symmetric least squares to calculate an initial estimate for the best fit line. Then, we pick which error bar to use based on whether the point is above or below the line and apply least squares with the chosen error bars.

For instance, for a point with an x value of $\log D_i = -1$ and a y value of $\log p_i = -3_{-0.3}^{+0.3}$, we would initially guess an error bar of $\sigma = 0.4$. Imagine that we get values of $m = -3.1$, and $b = -6.0$ from symmetric least squares, predicting that the value at our point will be $mx + b = -2.9$. Because -2.9 falls above -3 , we would then select the upper error bar, setting $\sigma = 0.3$. We iterate this process until it no longer changes an error bar. The model converges extremely quickly. In many cases, switching to the initial asymmetric error bars does not cause any points to cross the line. We call this method *asymmetric least squares*.

F Derivation of the MLE bias found in synthetic results

In our synthetic model of fitting methods, the Robbins et al. (2018) MLE method (Figure 4) and the corrected asymmetric least squares fits to the cumulative plot (Figure 4f) show a bias towards steeper slopes. This result requires explanation. It is due to the asymmetry of the PDF. Each method, least squares and Robbins et al. (2018) MLE, both center around finding the best fit model. For any PDF, the MLE (or mode) gives the best fit, but the Gamma distribution is asymmetric, so the MLE is offset from the mean. To derive this bias analytically, we rearrange Equation (17):

$$\sum_{i=1}^N \ln \left(\frac{D_i}{D_{\min}} \right) = \frac{N}{\hat{\alpha}}, \quad (\text{F1})$$

substitute it into Equation (14), and rearrange it into a normalized Inverse Gamma distribution, with $x = \hat{\alpha}$, shape parameter $s = N$, and rate parameter $r = \alpha N$:

$$P(\hat{\alpha}) \propto \frac{\alpha^N \left(\frac{N}{\hat{\alpha}} \right)^{N+1} e^{-\frac{\alpha N}{\hat{\alpha}}}}{\Gamma(N+1)} \rightarrow P(\hat{\alpha}) \propto \frac{\alpha^N N^N e^{-\frac{\alpha N}{\hat{\alpha}}}}{N \Gamma(N) \hat{\alpha}^{N+1}} \rightarrow P(\hat{\alpha}) = \frac{(\alpha N)^N e^{-\frac{\alpha N}{\hat{\alpha}}}}{\Gamma(N) \hat{\alpha}^{N+1}}. \quad (\text{F2})$$

The generic form of the Inverse Gamma distribution is:

$$P(x) = \frac{(r)^s e^{-\frac{r}{x}}}{\Gamma(s)x^{s+1}}. \quad (F3)$$

We now have an expression for the PDF of the observed $\hat{a}$ that will be produced by a given model α . As Figure 5b shows, it *exactly matches the observed results of our synthetic model*. From the theorem for the mean of an Inverse Gamma distribution, the mean of $\hat{a}$, $\mu(\hat{a})$, is:

$$\mu(x) = \frac{\beta}{\alpha - 1} \xrightarrow{x=\hat{a}; \alpha=N; \beta=\alpha N} \mu(\hat{a}) = \frac{N}{N-1} \alpha, \quad (F4)$$

which we see in Figure 8b, where the mean of $\hat{a}$ for $\alpha = 2$ is 2.2 at $N = 11$, 2.1 at $N = 21$, and 2.04 at $N = 51$. (Note: Rytgaard, 1990 derived Equation (F4) – but not Equation (F2) – in a more complex proof with a summation of exponential distributions.)

As this derivation shows, the Robbins et al. (2018) MLE method is biased for the same reason that the Arvidson et al. (1979) $\sqrt{N}$ approximation is biased: the Gamma distribution is asymmetric, and its mode does not equal its mean. At high N , it will approach a symmetric distribution around its mode, but the approximation will break down at low N . Because least squares methods also find the mode, they show the same bias once we correct for asymmetric error bars and the biased sampling of the cumulative plot (Figure 4f).

G Details of the SASH model

Here, we describe the SASH model in more detail. First, we split the data into an initial set of bins. In order for the bins to properly capture the low- N , high-diameter part of the curve, we recommend that the bins increase in size with diameter. Our default assumption is to start with an 18 per decade “Neukum” bin at the lowest diameter end. With each subsequent bin, the bin width then grows by our default growth rate factor of $2^{1.2}$. Although the model performs well on most data with these default values, for optimal performance, both the growth rate and the initial bin width should be altered depending on the data.

Next, we shift the bins to create $n_{shifts} + 1$ different sets of bins, where n_{shifts} is the number of shifts we perform. We shift each internal bin edge (not the minimum and maximum diameters) to the left so that the shifted bin edges stretch evenly in log space across the original bin to the left of the bin edge. To prevent the smallest bin from getting too small, instead of shifting its right-hand edge all the way towards the minimum diameter, we leave a buffer of 30% in log space. For n_{shifts} , we suggest a default value of 200. This could also be varied to optimize the model, but we did not vary it in the SASH models we show here (except in Figure 7). Whenever any bin has no craters, we merge it into the bin to its left. For the right-hand edge of the plot, this creates a large bin containing the largest craters that stretches out to D_{max} . By default, *cratersfd* assumes that D_{max} is ten times the diameter of the largest crater, effectively creating an open-ended interval for the highest-diameter bin, but we urge users to accurately estimate D_{max} (Section 4.5).

Then, we use Equation (19) (the α likelihood distribution of the Truncated Pareto distribution) to calculate the α PDF for each bin, and then we numerically calculate the mean of the α PDF for each bin. Next, we plot the differential plot line over each bin’s interval (Figure 7a): with the mean α for each bin, we plot the Truncated Pareto distribution PDF (Equation (18)), multiplied by the number of craters in the bin, N , divided by the terrain area, A . (This works because the differential plot is, by definition, the PDF multiplied by N/A .) Together, the lines from each individual bin make up a differential plot line for each set of bins. This gives us a set of $n_{shifts} + 1$ differential plot lines, and we average them together to produce the first run of the Slanted Average Shifted Histogram (Figure 7d).

Because the SASH model produces a more accurate slope estimate than the mean of the α PDF from the Truncated Pareto method, we then repeat the process, but this time we measure the slope directly from the SASH plot, taking the rise over run of the endpoints in log space (Figure 7f). Then, we repeat the process three more times for a total of five iterations (Figure 7h). The iterations produce a smoother plot that more faithfully represents the SFD.

Before the bins are shifted, we adjust the final bin’s left-hand edge to prevent two types of anomalous results: (1) When the final bin begins right before the largest crater, it can create an anomalous bump in the SASH model’s SFD results. To avoid this problem, we place a maximum diameter on the left-hand edge of the final bin. The maximum diameter is the geometric mean of the diameters of the two largest craters or 75% of the largest crater’s diameter, whichever is smaller. To prevent the maximum diameter from creating a very small penultimate bin, we cap the maximum diameter at the original penultimate bin’s center in log space (the geometric mean of its edges). (2) The final bin is not shifted, so the SASH model produces a straight line in log space above the beginning of the final bin, missing the benefits of

bin shifting. When the final bin begins too far below the largest craters, the final SFD misses meaningful information about SFD variation across the final bin. To address this problem, when the final bin edge falls below half the diameter of the largest crater, we impose a minimum diameter: the geometric mean of the diameters of the third-largest and fourth-largest craters.

To prevent anomalous results, when we calculate the mean a from Equation (19), we truncate the a PDF at a minimum of 10^{-5} and a maximum of 10. Critically, this truncation affects the calculation of the mean because the probability of a falling outside this range is no longer included in the mean. In the context of crater counting, where a rarely exceeds 4, the upper limit of 10 is unlikely to create any issues. (When the SASH model is applied to problems where a could exceed 10, this upper limit should be raised.)

We assume a lower limit right above 0 because a fundamental assumption of the SASH model is that the PDF is roughly Pareto with a negative slope (positive a) on a differential plot. With crater counting data, this assumption is nearly always valid, but there is one prominent exception: the rollover at low diameters from erosion or resolution limits. In these cases, it becomes less likely to have a smaller crater than a larger one, and the SFD rolls over to a positive slope (negative a) on a differential plot at the small crater end. At these rollovers, the PDF has a curved shape that is not well approximated by a Truncated Pareto distribution, but once it has rolled over to an upward-sloping curve, it often roughly follows a Truncated Pareto distribution where a is negative (officially known as a Truncated Power Law distribution).

Even though the SASH model assumes that a is positive in the initial run, the iterations do allow negative slopes. Once it has reached five iterations, the SASH model actually provides a quite accurate estimate of the portions of a PDF with a negative a . However, the default binning scheme is optimized to assume the data are roughly Pareto, so we recommend running the SASH model on the portions of the curve with positive a (and negative slope).

With future work, the SASH model can be optimized to further improve its performance. A weakness of the current model is that it requires user input on the binning scheme to optimize the results because the optimal binning scheme depends on the SFD. This problem could be addressed by using an initial run of the SASH model to help determine the optimal binning scheme. An initial run could also be used to automatically identify portions of the curve that are not roughly Pareto. For these portions – and for applications in other fields – the SASH model could be extended to infer another distribution, such as a Linear distribution, from the samples within the bins. The current SASH model is optimized to the specific problem of crater SFDs, but with a binning scheme approximated from an initial run, it could be applied to any generic distribution without imposing assumptions about the underlying shape of the PDF.

G.1 Details of synthetic modeling of SASH model error

Computational efficiency provides a limiting factor on synthetic modeling of the error of the SASH model. Because we want our recommended methods to be easy to use, we have avoided recommending methods that take an hour or more to run.

Computationally, we calculate these synthetic SFDs as arrays that give the differential plot value for each value in an array of diameters. (By default, the diameter array has 10 000 points evenly spaced in log space from D_{min} to D_{max} .) For each value in the diameter array, we obtain a list of the differential plot values for each synthetic.

The absence of error factors beyond counting error allows a simplification that greatly improves the accuracy of the synthetic model at a lower number of runs. Because counting error was the only factor, Equation (6) governs, and the distribution of synthetic SFD values at any given diameter very closely resembles a Gamma distribution. For each diameter in the diameter array, we fit the list of synthetic differential plot values to a Gamma distribution with SciPy's `gamma.fit()` function, which fits a Gamma distribution from discrete samples. From this Gamma distribution, we then calculate 1 σ -equivalent percentiles for each diameter point. Because we do this for every point in the 10 000-point diameter array, we calculate the error envelope at every point with no need for interpolation.

If we had simply calculated the 1 σ -equivalent percentiles from the raw histogram of the synthetic SFD values, the results would be much, much noisier. By fitting a Gamma distribution, we can reach stable results for the error envelope with only 1000 synthetics. By comparison, for the synthetic modeling of least squares fits to cumulative plots, we calculated the percentiles from the raw histogram, and we needed 100 000 synthetics to produce stable results.

For $N \leq 5000$, the SASH model runs in ~0.5 seconds or less on a MacBook Pro laptop. For published plots, we recommend running 1000 synthetics, which would take ~8.3 minutes at ~0.5 second per synthetic. This is too slow to be practical for quick interpretation of the data. Fortunately, the Gamma distribution approximation allows a fairly accurate assessment of the error envelope at just 100 synthetics, so we recommend running the synthetic model at 100 synthetics for casual interpretation of the data or

a quick check of a potential model SFD.

Conceptually, it is fairly straightforward to apply the math we develop in Section 2 for several of the sources of error beyond counting statistics to synthetic modeling of SASH model results, but computational efficiency presents a challenge that will require future work to solve. For instance, in our synthetic model of diameter measurement error (see Figure 2, Section 2.3, and Appendix A.2), we already laid out how to add diameter measurement error to a synthetic model. However, applying these effects to each synthetic would increase the computation time per synthetic. Moreover, incorporating additional sources of error would produce a distribution of synthetic SFD values that no longer matches the Gamma Distribution, so many more synthetics would be needed. With future work to solve the computational problem, our synthetic modeling approach can easily be adapted to incorporate additional sources of error, leveraging the math we develop in Section 2.3.

Supplemental figure

From: Bell, S.W., and Robbins, S. (2026). The theory of cratering statistics: Significant improvements to existing methods for ages, slopes, and plots. *Planetary Research*, **1**(1), 235. <https://doi.org/10.53480/46d8-n314>

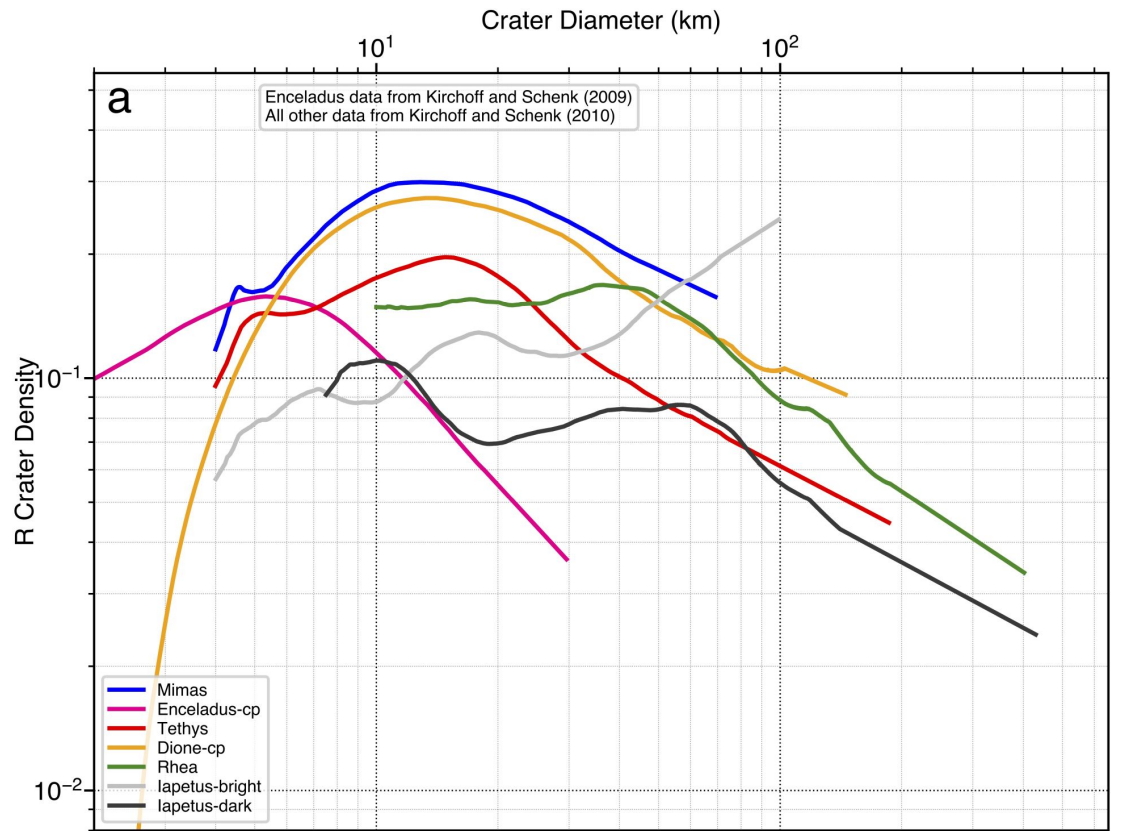


Figure S1 | A reproduction of Figure 13a where D_{max} is chosen to be only slightly larger (5%) than the largest crater. Although they are shallower, especially in the case of Mimas, the steep branches remain for all terrains except Iapetus-bright, showing that they do not depend on information from the absence of craters larger than the largest crater.