On the statistical distribution of dust devil diameters

Ralph Lorenz
Johns Hopkins University, Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723, USA

Abstract
Dust devil diameters, like many features in nature, have skewed distributions. I summarize, and present in a unified manner, literature values of dust devil optical diameters from seven terrestrial surveys, three rover-based surveys on Mars and two orbital surveys on Mars. The problems of appropriately treating these data are analogous to those for impact craters, and similar display and binning approaches are suggested. Remarkably, the Mars dust devil population remains better-known than Earth’s. The theoretical justifications for possible log-normal, and (truncated) exponential and power-law descriptions of dust devil properties are discussed, and the challenges of discriminating between these candidate distributions with finite (and often, coarsely-binned) observation sets are noted: the best-sampled datasets so far appear well-described by power laws. Data required for advances in model discrimination are discussed: data binned in four or fewer ranges are useless for this purpose. Caution must be exercised in applying the notion of an ‘average’ dust devil and in calculating population-integral properties such as dust flux.

1. Introduction
Dust devils (e.g. Balme and Greeley, 2006) are a familiar feature of terrestrial arid region meteorology, and play an important role in lofting dust into the atmosphere, with implications for air quality and sediment balance. Dust devils can be a nuisance for outdoor activities and can occasionally cause structural damage or fire. A review of air accident statistics (Lorenz and Myers, 2005) indicates some 100 damaging – and in some cases fatal – incidents to aircraft that are attributable to dust devils. Dust devils are also the most prominent dynamic phenomena observed on the surface of Mars. They are major agents of surface change, forming tracks which not only are individually visible, but may lead to accumulated albedo changes over large regions. Additionally, in the thin martian atmosphere, they are a principal mechanism of dust-raising, with a resultant influence on the entire martian climate. The operations of solar-powered vehicles on the martian surface can be significantly affected both by the dust lofted planetwide, and by local dust removal events. Thus it is important that the frequency and behavior of dust devils be understood.

Dust devils, like many other natural phenomena, have a highly skewed size distribution, with small devils being much more abundant than large ones. Use of appropriate models of these skewed distributions is required for accurate assessment of encounter frequency with devils of a given size (e.g. for hazard assessment) and for the calculation of process rates such as dust-lifting or triboelectric chemistry. The highly skewed distributions indicated by the data call into question the notion of an ‘average’ dust devil and the accuracy of scaling such an average by some number of dust devils to deduce integral dust-raising rates.

In this paper we first summarize prior work on defining populations of dust devil diameters, and review the statistical properties of candidate functions (notably power-law, exponential and log-normal) and point out some aspects of the binning and display of skewed data. We then present several dust devil surveys in cumulative and differential plots, with statistical errors, and discuss the extent to which the data permit (or not) discrimination between candidate functions. We then discuss some theoretical aspects of the different population models, and close with some conclusions and requirements on number and resolution or binning for field measurements in order to make progress on these questions.

2. Prior descriptions of dust devil optical diameters
It has been traditional to record dust devil diameters in optical surveys in rather coarse size categories, largely because of the challenges of real-time measurement of diameters in field conditions. This binning results in loss of information, but the earliest reports – and indeed many since – presented size statistics without any attempts at analytic description, and thus the inadequacies of data for discriminating models may not have been recognized.

While many workers have noted the minimum, maximum and mean in their respective surveys, these numbers have only modest utility for describing skewed and incompletely-sampled populations. The first useful analytic description of a dust devil size distribution appears to be that of Kurgansky (2006) who suggested that
dust devils might form a truncated exponential distribution. He argued for an exponential population (i.e. the probability $p$ of the diameter appearing in the range $x$ to $x + dx$ is given by $p(x) = \exp(-x/kx) dx$) on information-theoretic grounds, namely that (reasonably) following Jaynes’ statistical concepts, the distribution with the highest entropy ($S = -\sum p(x) \log p(x))$ subject to any provided constraints is the most probable. For a distribution of values in the range zero to infinity, the exponential distribution has the highest entropy. Kurgansky then goes onto note that the population must be truncated – there is a minimum size of dust devil defined by a fluid-dynamical limit. Kurgansky suggested that the mean dust devil diameter should be related to the Obukhov length: this hypothesis was advanced by Hess and Spillane (1990). Kurgansky introduces the minimum limit into the algebraic description by a multiplier, leading to a Weibull function of the form $P(x) = \exp(-x/kx)$, where $k$ is a shape parameter and $x_o$ is a scaling diameter (close to the median). The function (upper case) $P(x)$ is the cumulative probability that a dust devil of size $x$ or larger, the Weibull function, used widely in terrestrial wind energy studies, was advanced by Lorenz (1996) as a convenient description of Viking measurements of Mars surface windspeeds – the shape parameter which leads to a Rayleigh distribution for $k = 2$. Kurgansky (2006) then goes onto compare the shape of this distribution, via linear histograms without discussion of statistical errors, with a few surveys and finds the function to be a satisfactory fit.

Lorenz (2009), noting the coarse binning of those surveys, examined the more finely-histogrammed data of Greeley et al. (2006) of dust devil diameters on Mars (Gusev) recorded in MER images, plotting the data with counting errors on a double-logarithmic plot (apparently the first time dust devil sizes were presented this way). The data formed a straight line, suggesting a power law: the data appeared inconsistent with an exponential which would have too convex a shape (declining too steeply at large sizes) to fit. This was an appealing result, because power laws are used to describe many self-organized phenomena in nature, as discussed later in this paper. Lorenz (2009) also noted that the size distribution, coupled with a fixed angular size detection threshold, might explain the apparently-discrepant dust devil densities noted in studies of different survey areas.

However, Pathare et al. (2010), found that two sets of (coarsely-binned) new terrestrial dust devil observations did not appear to fit a $-2$ power law, but were superficially consistent with an exponential distribution. That work, prompting the present paper, raises the questions of whether the observed Gusev size distribution is indeed fundamentally different from the two observed on Earth (and if so, why, whether a real population difference or an observational bias, or if statistical inadequacies in the datasets in fact do not permit after all a formal statement on whether the populations are different and/or are described by given analytic functions. First, let us review some possible functions.

3. Power laws, log-normal and other distributions

Statistics as a field is dominated historically by Gaussian (‘normal’) distributions, which can be generated by the addition of random processes and are characterized by a symmetric distribution about a mean value. The bulk of the data lie in a range, described by the variance of the population, that is a modest multiple of the mean value. However, as was noted for sand sizes by Bagnold (1941), inspection of data on logarithmic axes can represent more instructively the distribution of many quantities, and reveals the presence of significant numbers of extreme values (which tend to be suppressed on linear axes). In modern times, it has become fashionable to refer to these values as forming a ‘long tail’ (usually the extremes are in one direction, so there is only one tail, although sometimes it is referred to as ‘fat’ rather than ‘long’).

Such data often lie on a straight line in a log–log plot, and thus define a function of the form $p(x) = x^{-\alpha}$, or a power law with exponent $-\alpha$. As discussed in a lucid and illuminating review by Newman (2005), early workers in a variety of fields made discoveries of power-law behavior in incomes (Pareto), coastline geometry and the size of wars (Richardson), word frequency and city size (Zipf) and earthquake magnitudes (Gutenberg-Richter). More recently, the association of these distributions with ‘fractals’ (Mandelbrot), and the discovery of self-organized criticality (SOC, Bak, 1996) has prompted searches for power-law distributions in ever-wider fields. However, stimulated in part by practical application to information systems engineering (e.g. filesizes and web traffic where the accuracy and predictive utility of the relevant functions is of economic, and not merely academic, interest) closer scrutiny is now being applied to the detailed shape of distributions in data, e.g. Clauset et al. (2009).

Specifically, Clauset et al. (2009) argue that while plotting a histogram of data on doubly-logarithmic axes and finding a straight line is a necessary condition for finding a power law, it is not sufficient. As also discussed by Mitzenmacher (2004), log-normal and power law distributions can be equivalent over significant ranges. Thus, especially if the population is itself truncated by size limits, and more so if the population is incompletely measured by finite observations, it may be impossible in many cases to formally discriminate a power law from a log-normal distribution, or a Pareto, or exponential, or other function. Some care should be exercised in the choice of bin sizes – too large bins lose information, yet too small bins can lead to counting (Poisson) errors. As discussed in Newman (2005), logarithmically-scaled bin sizes (the boundary value between successive bins having a constant ratio) are often convenient in providing adequately large bins at large sizes to mitigate counting errors, and have the aesthetically-pleasing property of a uniform width when plotted on logarithmic axes. Further, if simple counts are presented, rather than counts normalized to the width of the bin, arbitrary bin sizes may distort the appearance of the data.

Fig. 1 shows these candidate distributions in a number of forms. In fact, these challenges in examining populations of dust devils are not new to the planetary sciences (nor, indeed to aeolian studies, since Bagnold devotes an entire chapter of his 1941 book to plotting grain sizes and advocating the use of logarithmic axes). Like dust devils, impact craters on planetary surfaces have similarly skewed distributions, and a variety of methodologies have been evolved to appropriately treat these population data (e.g. Crater Analysis Techniques Working Group, 1979; Melosh, 1989). The former paper in particular lists recommendations that are echoed in the present work. Data should be presented in tabular form, and on doubly-logarithmic axes, displaying both cumulative and relative (difference) size–frequency plots. Data, if binned at all (note that it is not necessary to bin data when presenting it in a cumulative form), should be binned in $\sqrt{2}$-factor or ranges with absolute numbers of craters (devils) indicated in each bin, together with $1\sigma$ confidence intervals.

4. Dust devil datasets and model fits

In this section, we present data from the literature in a consistent set of formats, following the guidelines above. Balme and Greeley (2006) present a convenient summary of the dust devil size data from optical surveys on Earth by Snow and McClelland (1990), Sinclair (1969) and Ryan and Carroll (1970) (see also Carroll and Ryan, 1970). Additionally, size data was reported for two surveys by Pathare et al. (2010). A recent survey...
in the Atacama by Kurgansky et al. (2010) presents good time-of-day statistics, but unfortunately not size data.

On Mars, Fisher et al. (2005) examined Mars Orbital Camera (MOC) data from the Mars Global Surveyor (MGS), listing diameters and heights of 28 and 20 dust devils respectively. Note that the observation geometry for orbital observations is rather different from that for landers, influencing the resolution and thus size threshold. (A very good survey of seasonal activity observed by MOC is reported by Cantor et al. (2006), however this work does not present size data.)

Stanzel et al. (2008) present a survey of dust devils observed by the High Resolution Stereo Camera (HRSC) on the ESA Mars Express (MEx) spacecraft. This work identifies the diameter and heights, and in many cases speeds of motion, of some 205 dust devils.

From the surface, the first comprehensive optical survey is that from Mars Pathfinder, by Ferri et al. (2003) – it may be recalled that while the Viking landers recorded the signatures of dust devils in in situ meteorological data (Ryan and Lucich, 1983; Ringrose et al., 2003), the Viking imagers were single-pixel raster-scan devices that were not well-suited to detecting moving objects, as famously demonstrated with tortoises by Carl Sagan. However, the Ferri et al. (2003) survey documents only about 20 devils, too small a sample for useful statistical analysis.

Greeley et al. (2006) present a histogram (with usefully populated bin sizes) of dust devil diameters observed in imaging data from the Spirit rover at Gusev on Mars. Greeley et al. (2010) present similar data for two following years.

We plot selected datasets as binned in the literature, or summarized with modest binning where raw data are supplied, in Fig. 2. Here raw counts, with 1-σ counting error bars, are shown: the square root relationship between count and error is evident. For the convenience of other workers, the data plotted are also listed in Table 1. These plots allow us to inspect the utility of the data. In almost all the terrestrial cases, the counting statistics, especially for the smallest bin, are exquisite, yet the bin can be an order of magnitude wide. Thus the data have been binned in a very nonoptimal manner, losing information (whether in the field it is practical to bin the data better is another question). In contrast, the Gusev data are finely binned, although counting statistics at the larger size range become poor (which can be addressed simply by adopting wider bins, since information has not been lost). Note that in making the plots, a somewhat arbitrary bin width is assumed for the minimum size when only a maximum is given. This should not affect the perceived goodness of fit of any function since little weight should be given to the smallest size bin as this bin is most likely affected by poor detection efficiency anyway.

An interesting effect may be noted in the White Sands and Gusev data, which cover their respective sites for multiple years. In each case there is a substantial fluctuation in the relative size of the smallest bin. Unfortunately, it is difficult to say whether this represents a change in the underlying population due to different meteorological circumstances, or perhaps more likely a change in detection efficiency at the smallest sizes (see Section 7 and Fig. 5.)

Fig. 3 re-plots the same data, but expressed in a way that permits intercomparison between datasets, namely in a normalized...
bin width (factor of $\sqrt{2}$). Thus the counts are rescaled by a factor ($\sqrt{2}$) (min/max). Candidate power law and exponential curves are shown (these are not formal fits, merely shape indicators, since proper fits require introduction of other free parameters).

Visually, it appears that the Mojave diameter data by Carroll and Ryan (1970) and the Gusev season 1 (Greeley et al., 2006, 2010) data are very linear on a log–log plot and may be best described by power laws. The case is less clear-cut for the other Gusev data. The Eloy data of Pathare et al. (2010) is clearly convex and so appears better described by an exponential. However, given that real distributions must be modified by a minimum size (power law and exponential) and a maximum size (power law), these additional degrees of freedom could be chosen in order to secure an adequate power law fit for all datasets (given that the Pathare et al. (2010) datasets have only four points, the four variables in a truncated power law only provide as many degrees of freedom as there are datapoints). At this point, then, it seems difficult to conclude that either a power law or exponential are universal descriptions of the observed or real populations (and analogy with the literature on rain cell sizes suggests there is little hope for convergence on the matter). As a single simple function (in that it effectively embodies a maximum value and does not climb to infinity for small sizes) an exponential may be convenient: a power law on the other hand is a more useful relationship for discussing detection efficiency and for integrating population parameters such as dust-raising. We discuss some of the theoretical justifications for functions in the next section.

Fig. 4 plots the datasets in cumulative form. Although such plots are more aesthetically-pleasing, especially for sparse data (since the integration of the data gives smooth curves), these plots are actually less useful at visually discriminating populations or functions. However, by plotting the data on a common x-axis, it is easy to see the basis for the commonly-offered assertion (not usually backed by data) that Mars dust devils are bigger than Earth’s. It certainly appears that the largest dust devils (>100 m) represent a greater fraction of the observed population on Mars than on Earth, although the question of whether this is due to higher detection efficiency of small dust devils on Earth remains to be studied.

Fig. 2. Binned data (thin solid lines) plotted as differential size distribution (i.e. number per bin). Thin vertical bars denote 1-σ counting errors (in some cases invisibly small). Note the large width of the bins used in most terrestrial surveys. Data are also presented in Table 1.
Fig. 5 illustrates these effects – an observation sample may not capture the maximum of the underlying population, and detection inefficiency at small sizes may introduce a bias.

As discussed by Clauset et al. (2009), the best approach to formally discriminate populations from each other, or to identify the viability of an analytic description (e.g. power law) of the population, is the Kolmogorov–Smirnov (K–S) test. This test, which makes no assumptions about the distribution, uses the maximum distance in the cumulative histograms to estimate the probability that differences between the two have arisen at random, with a 5% or 10% level usually being adopted as the threshold for asserting that the two curves represent the same population. The distance that corresponds to this probability threshold depends on the number of points on the curve: at least four points are required for the statistic to be meaningful at all – the Pathare et al. (2010) data, and indeed all the terrestrial data except perhaps those of Ryan and Carroll (1970) are therefore of borderline significance.

Table 1

Table 1 presents measured populations of dust devil diameters from the literature, as plotted in Figs. 2–4. Complete original data are presented in Fisher et al. (2005) and Stanzel et al. (2008) but have been binned here. In most cases a population maximum and minimum has been assumed.

5. Theoretical considerations

We now set aside the somewhat inconclusive data, and consider the theoretical arguments in support for candidate functions as described in Section 3. First, we note that the real dust devil population is likely to be bounded by two finite values – a maximum and a minimum.

This is significant, because the principal theoretical justification offered for an exponential distribution by Kurgansky (2006) is that this function is the maximum-entropy (most probable) distribution. This maximum-entropy argument is only true if the distribution is allowed to span the range of zero to infinity. In fact, if there are finite limits to a distribution, the maximum entropy distribution subject to those constraints is a uniform one (i.e. dust devils of all sizes are equally abundant). The same logic can be applied successfully to the scores on a six-sided die – these are bounded in the range 1–6, and the distribution with maximum entropy is that with each side having a 1/6 probability of being face-up.

Clearly, dust-devil populations do not follow a uniform distribution. This reductio ad absurdum argument is not to dismiss a truncated exponential distribution as a candidate description of dust devil statistics – we have shown that it adequately describes several empirical observations (sometimes better, sometimes less well than a power law) – however, the information-theoretic argument advanced by Kurgansky for its adoption is not without problems. The author is unaware of any physical reason why vortices should...
follow an exponential distribution. Exponential distributions arise in Poisson (random) processes – intervals between the decays of radioactive nuclei, or the length of telephone calls, for example. On the other hand, exponential distributions have been used empirically to describe populations of rain-cell sizes (indeed, for almost half a century, e.g. Dennis and Fernald, 1963; Sauvageot et al., 1999), although log-normal functions (e.g. Lopez, 1977) and power-law (Machado and Rossow, 1993) have also been applied successfully to statistics of cloud and rain systems. These analyses all present data in a preferred way (in the case of an exponential, plotting the logarithm of frequency against a linear size axis) and find a roughly straight line from which the authors conclude (prematurely, as in so many other fields – see Clauset et al., 2009) that the description is the appropriate one. In any case, the exponential is far from universally preferred in this field.

Log-normal distributions arise from the multiplication of random processes. Simply put, successive multiplications of a quantity correspond to successive additions of the logarithm of a quantity: thus via the Central Limit Theorem, the logarithms will tend to have a Gaussian (normal) distribution. Although the function is a reasonable description of some of the empirical data, it is far from obvious what properties might be multiplied to form a dust devil population – if anything, simple addition of vorticity elements seems a more likely physical mechanism, which would lead to a pure normal (i.e. not log-normal) population, which is not observed.

On the other hand, there is clear precedent in many natural systems for the emergence of power-law behavior, e.g. Newman (2005). He notes that power laws may emerge from a variety of processes, such as fragmentation (which leads to the power law in asteroid sizes, and thus of impact craters which we discuss elsewhere in this paper) and the Yule process (by which genera evolve stochastically, splitting into multiple genera or becoming distinct species). Of particular relevance to dust devils are two process types.

The first is self-organized criticality (SOC) e.g. Bak (1996). Measurements of the near-surface boundary layer during dust devil
activity show it to be formally unstable, and thus it seems likely to be generally poised on a threshold wherein the ongoing formation of convective plumes keeps pace with the buoyancy generated by strong heating. This scenario seems rather analogous to other critical phenomena that show power laws, such as avalanches in sand-piles, forest fires, and earthquakes (see also Turcotte, 1997). Machado and Rossow (1993) suggest tropical cloud clusters follow a $-2$ power law size distribution, and some recent work has noted that over two orders of magnitude, the power dissipation in cyclones also appears to follow a power law (Corral et al., 2010).

The second process type is of randomly-truncated exponential growth, as described by Reed and Hughes (2002). Specifically, processes that grow exponentially with time, but are terminated at random, will show power-law distributions. It is easy to see how exponential growth might appear – the convective heat flux ‘harvested’ by the dust devil will likely increase with the devil’s size as it draws in air over a larger area. Additionally, as suggested by Lorenz (2004) and elaborated by Furstenau (2006), the dust lofted in a dust devil provides local heating by absorption of sunlight in the column – a devil that becomes dusty may also become stronger. In areas which have available dust, it is easy to see, then, how the energy (or other properties) of a devil may increase exponentially. However, as a dust devil moves, it may encounter areas where friction is higher, or dust is not available, or the available superheated near-surface air has been removed by a gust or some other effect. Thus the devil may stochastically shrink or be starved altogether, leading to a power law as suggested for other analogous situations (Reed and Hughes, 2002).

Establishing detailed physically-based mechanistic models of dust devils and evaluating their statistical properties (and explicit numerical models such as Large Eddy Simulations (LES) may be of use here) will be left to future work, but these examples at least show that there are plausible means by which power-law behavior might emerge. Truncations to a power law are required, to avoid infinite populations or infinite integral properties: these can be specified in a number of ways, e.g. as simple limits (forming a trap-ezoidal area on a plot), or by introducing some other multiplicative function (while the spectrum of earthquake magnitudes follows a
power law overall, it is sometimes assumed to have an exponential tail to truncate the distribution at high values which approach the maximum energy storage in the lithosphere. In physical terms for dust devils, it is easy to imagine that these limits may be imposed by the character of the boundary layer. The height of the dust devil will likely be confined to the height of the atmospheric boundary layer, and the observed aspect ratios (height:diameter) of dust devils suggests this should translate into an upper diameter limit that is a factor of a few smaller than this. Similarly, the minimum size to the distribution may be the Obukhov length, as suggested by Kurgansky (2006).

6. Integrating across the population

One of the reasons why dust devils are of interest is their role in injecting radiatively-active dust into the martian atmosphere. The most widely-adopted approach in calculating this injection is to simplistically multiply the dust-lofting of a ‘typical’ dust devil by some number density of dust devils. As an example, Fisher et al. (2005) multiply a flux of $10^{-4}$ kg m$^{-2}$ s$^{-1}$ (a flux per unit area of the dust devil, from Ferri et al. (2003)) by the area of the smallest dust devil they resolve (thus the most common), and the total number density of devils. However, given a $-2$ power law of dust devils, doubling the size means there are four times fewer, but with four times the area. It follows, then, that this calculation underestimates the total dust devil area by a factor equal to the ratio between largest and smallest dust devil, i.e. a factor of at least $\approx 20$. The flux estimation could be an even more serious underestimate if there exists a positive correlation between dust density and/or vertical velocity and size – i.e. that larger devils have larger dust fluxes per unit area. Thus we argue that the skewed nature of the population demands care in calculating integrated properties.

These effects are illustrated in Fig. 6. A power-law ($-2$) population of 10,000 dust devils, with a minimum diameter of 2 m and a maximum diameter of 300 m is sampled, and various statistical properties are evaluated for the subsample. It is seen that the minimum of the population is discovered rather quickly. The maximum is not well-estimated and the mean shows significant (>30%) jumps even up to sample sizes of $\approx 100$. The sample-average values of mean-square and mean-cube diameters (divided by the square and cube of the population-mean, respectively) are shown. These properties are larger than the mean-based estimates by a factor of $\approx 6$ and $\approx 100$ respectively, and show significant jumps when large devils are encountered.
and so the mean shows stochastic jumps until the sample size grows to several hundred.

Properties of interest over the whole population are nonlinearly dependent on diameter – dust flux will likely be proportional to the 2nd or 3rd power of diameter. For a constant volume density of dust and fixed vertical velocity, the flux per devil will be proportional to the area and thus square of diameter; more probably there will be some dependence of vertical velocity and/or dust density on size, and thus the flux could be proportional to the cube or some higher power. When these nonlinear properties are integrated across the population, the result is very strongly affected by the largest members of the population, such that encountering a single large devil can increase the integral measure by an order of magnitude. It should be noted also that the population integrals are a substantial factor higher than those that would be estimated from the average values.

7. Requirements for future field observations

Preparations are underway at the time of writing for the November 2011 launch of the Mars Science Laboratory, a rover equipped with a suite of meteorological instrumentation suitable for characterizing several properties of dust devils (wind, temperature, pressure). Plans call for one or two hours of observations per day, at a cadence high enough to detect and measure dust devils. These observations may be augmented by event-triggered measurements. Operation for one Mars year should yield an optical survey comparable with that from the MER rovers (Greeley et al., 2006, 2010), but augmented by an in situ survey. We may hope that the good precedent set in detailed reporting of dust devil activity on Mars will be followed on this new mission.

However, as shown in this paper, it seems clear that the difficulty in comparing Mars and Earth populations arises from the poor quality of data on Earth, not on Mars. As has been discussed in this paper, and noted by Lorenz (2009), terrestrial ‘by eye’ surveys have relied on size binning that is too coarse to meaningfully discriminate candidate functions for the size distribution. An effort to acquire a suitable dataset would require good diameter determinations for many hundreds of dust devils, a demanding undertaking that would require careful field measurements over many weeks. In order to reliably detect convexity in the size distribution (and thus discriminate between exponential and power law populations), data over a factor of ~50 in size is desired, allowing the population of ~10 size bins of width ~/2 (since 2 ~ log 2(50) ~ 11). Given a ~2 power law and a count of 1 in the largest bin, the smallest bin will have a count of ~50^2 ~ 2500. The total number of dust devils in the survey would be therefore ~5000. This exceeds by a modest factor the largest surveys to date, although might be attainable with a modest area-time product if the minimum detectable size is sufficiently small (although it will be ultimately limited by the minimum in the population itself).

Timelapse imaging (e.g. Lorenz et al., 2010) offers the promise of providing the required data while exposing observers to the rigors of the field for much shorter periods. Initial experiments (Lorenz, 2011) have shown considerable promise.

Note that if a single dataset is to simultaneously detect the minimum and maximum in the population, then an even larger area-time product and number of events is needed. If one combines Mars lander and rover observations, which have detected devils with a diameter of ~5 m, with orbiter observations which show devils with diameters of 1–2 km, then the max/min diameter range is 200–400, and for the largest devil (in ~/2 bins) to be detected twice means 16–18 bins and thus some ~10^7 total events. Evidently, as befits a power-law phenomenon, a hierarchical observation approach combining several different scales may be the most efficient way to evaluate the population.

8. Conclusions and recommendations for future work

This paper has considered the distributions of a number of dust devil parameters, examining several datasets. Analytic functions that describe observed size distributions have been considered. Truncated power law and exponential distributions, and log-normal distributions, are found to be generally appropriate descriptions: data generally do not permit the determination of one of these functional descriptions as being (statistically) significantly better than any other. A few datasets seem better described by power laws, and power laws may be more convenient algebraically for predictive work, and appear to have a more robust theoretical foundation, but other functions cannot be excluded. We note that similar statistics, similarly described by a proliferation of analytic forms, have been encountered in terrestrial cloud and rain studies.

The observed populations of martian dust devils show larger sizes, but the extent to which this reflects a different mean size, a different minimum or maximum, is not yet clear. It is also not evident to which extent the published statistics reflect the underlying population, or is merely an observational artifact, reflecting higher detection efficiencies for small devilion Earth.

Some clear deficiencies in the existing data have been identified – in many cases, data are too coarsely-binned and/or sample sizes are too small to discriminate population functions or even to identify maximum and minimum sizes. It is recommended (following similar issues in the cratering community some 30 years ago) that observations in future surveys be tabulated explicitly, or at least in size bins small enough (2^−5) to permit their use. New terrestrial survey approaches using timelapse imaging (e.g. Lorenz et al., 2010; Lorenz, 2011) may offer a way forward.

Acknowledgment

Early parts of this work were supported by the NASA Applied Information Systems Research program.

References


Author's personal copy