Subglacial bedforms reveal an exponential size-frequency distribution

J. K. Hillier[a,*], M. J. Smith[b], C. D. Clark[c], C. R. Stokes[d], M. Spagnolo[e]

[a] *Department of Geography, Loughborough University, LE11 3TU, UK.*
[b] *School of Geography, Geology and the Environment, Kingston University, KT1 2EE, UK.*
[c] *Department of Geography, University of Sheffield, S10 2TN, UK.*
[d] *Department of Geography, Durham University, DH1 3LE, UK.*
[e] *School of Geosciences, University of Aberdeen, AB24 3UF, UK.*

Abstract

Subglacial bedforms preserved in deglaciated landscapes record characteristics of past ice-sediment flow regimes, providing insight into subglacial processes and ice sheet dynamics. Individual forms vary considerably, but they can often be grouped into coherent fields, typically called flow-sets, that reflect discrete episodes of ice flow. Within these, bedform size-frequency distributions (predominantly height, width and length) are currently described by several statistics (e.g., mean, median, and standard deviation) that, arguably, do not best capture the defining characteristics of these populations. This paper seeks to create a better description based upon semi-log plots, which reveal that the frequency distributions of bedform dimensions (drumlin, mega-scale glacial lineation, and ribbed moraine) plot as straight lines above the mode (ϕ). This indicates, by definition, an exponential distribution, for which a simple and easily calculated, yet statistically rigorous, description is designed. Three descriptive parameters are proposed: gradient (λ; the exponent, characterising bedforms likely least affected by non-glacial factors), area-normalised y-intercept (β_0; quantifying spatial density), and the mode (ϕ). Below ϕ, small features are less prevalent due to i) measurement: data, sampling and mapping fidelity; ii) possible post-glacial degradation; or iii) genesis: not being created sub-glacially. This new description has the benefit of being insensitive to the impact of potentially unmapped or degraded smaller features and better captures properties relating to ice flow. Importantly, using λ, flow sets can now be more usefully compared with each other across all deglaciated regions and with the output of numerical ice sheet

*Corresponding author. Tel.: ++44 1509 223727; Fax: ++44 1509 223930
Email addresses: j.hillier@lboro.ac.uk (J. K. Hillier), mike@hsm.org.uk (M. J. Smith), C.Clark@Sheffield.ac.uk (C. D. Clark), C.R.Stokes@durham.ac.uk (C. R. Stokes), m.spagnolo@abdn.ac.uk (M. Spagnolo)

Preprint submitted to Geomorphology *February 19, 2013*

models. Applications may also exist for analogous fluvial and aeolian bedforms. Identifying the characteristic exponential and that it is typical of ‘emergent’ subglacial bedforms is a new and potentially powerful constraint on their genesis, perhaps indicating that ice-sediment interaction is fundamentally stochastic in nature.

Keywords: Subglacial; Bedform; Exponential; Stochastic; Flow-set; Fluvial.

---

1. Introduction

Subglacial bedforms are a group of landforms created at the interface between glaciers and the terrain underneath (e.g., Benn and Evans, 2010). Mainly comprised of glacial sediments (e.g., Stokes et al., 2011), they are often assigned to one of four categories based on their size and shape: (i) flutes (e.g., Boulton, 1976), (ii) drumlins (e.g., Menzies, 1979a), (iii) ribbed moraine (Hättestrand and Kleman, 1999) and (iv) mega-scale glacial lineations (MSGL) (Clark, 1993). Taken together, these range between 10^1 and 10^5 m long (Clark, 2010). Ribbed moraine form transverse to ice flow direction, whilst flutes, drumlins and MSGL form parallel to ice flow and are possibly a continuum of landforms (e.g., Aario, 1977; Rose, 1987) that are created by similar processes that operate under variable conditions. For example, it has been suggested that bedform length may be related to ice velocity (e.g., Clark, 1993; Hart, 1999; Stokes and Clark, 2002). Glacial bedforms are generally argued to be created directly by overriding ice flow (e.g., Benn et al., 2006; King et al., 2007; Clark, 2010; Ó Cofaigh et al., 2010), although an origin through sub-glacial floods has also been proposed (e.g., Shaw, 1983; Shaw et al., 2008). Due to their prevalence, drumlins have been most heavily studied, but even these remain enigmatic with their exact mode of formation still undetermined (e.g., Smalley and Unwin, 1968; Menzies, 1979b; Shaw, 1983; Boulton and Hindmarsh, 1987; Hindmarsh, 1998; Fowler, 2000; Clark, 2010).

The shapes of bedforms (e.g., height H, width W, length L, and orientation) preserve key information about the dynamics and mechanics of former ice sheets, an important guide as to how existing ice sheets will behave in the future. Observations are typically used descriptively to indicate properties such as ice extent or flow direction (e.g., Hollingsworth, 1931; Livingstone et al., 2008), for example to assess consistency with numerical ice sheet models (e.g., Evans et al., 2009), and only rarely to directly consider the mechanics of ice-sediment interaction and flow (Chorley, 1959; Morris and Morland, 1976; Smalley and Piotrowski, 1987; Smalley and Warburton, 1994).

Indeed, few theories of subglacial bedform genesis are yet to explicitly engage with empirical data on their shape and size. One that has made predictions of bedform dimensions is the instability theory (e.g., Hindmarsh, 1998; Fowler, 2000; Stokes et al., 2013), but it only considers them as quantitative constraints in the broadest sense, as an order of magnitude scale ground-truth (Dunlop et al., 2008; Chapwanya et al., 2011). A disconnect therefore exists between glacial geomorphology and glaciological modelling (e.g., Bingham et al., 2010).

As a step towards forming a link between the subglacial bedform record and the nature and mechanics of ice flow, this paper presents a descriptive development: a new statistical characterisation of bedform populations. The need for an improved description is two-fold. Firstly, population metrics should capture the signal of ice-sediment interaction, not artefacts of measurement or preservation. Secondly, individual population metrics should ideally capture key aspects of data allowing inter-comparison of data types, localities, and palaeo-environments. The origin and nature of potential artefacts and the implications of this for current metrics are considered in Section 2. It is demonstrated graphically in Section 3, using semi-log plots, that the size-frequency distributions of key properties (e.g. H, W, L, and L/W) of subglacial landforms are exponentially distributed above the mode. Following this, a simpler objective parameterisation of the data is created in Section 4 which consists of individual metrics better suited to isolating characteristics of bedform populations relating to ice flow. Then, by collating data sets for a variety of areas, Section 5 demonstrates the general applicability of the proposed description to subglacial bedforms. Finally, in Section 6, the selection of the exponential-based parameterisation is discussed and initial thoughts are offered on implications for the process of drumlin genesis.

2. Quantifying subglacial bedforms

Subglacial bedforms have been quantified in a variety of ways, both as individuals and populations (e.g., Gardiner, 1983; Smalley and Warburton, 1994). Individual forms vary considerably, even within a locality (e.g., Hollingsworth, 1931), so they are likely to best reflect flow regimes when grouped into spatially and temporally co-located flow-sets. Thus, quantifications for populations (e.g., Fig. 1a) are considered here, although error bars for parameters may be large enough to warrant particular attention for small populations (i.e., $n \lesssim 50$). For clarity we use the terms 'metric' or 'parameter' exclusively to refer to quantification statistics such as the mean or mode, as distinct from measurements to which they are applied such as L or 'aspect ratio' (i.e., L/W). Populations of observations from which the metrics are calculated are the final products of application of three compounding processes. Artefacts are due to i) measurement, ii) post-glacial preservation and iii) the process of glaciological interest i.e., bedform genesis itself. The artefacts must be accounted for to reveal information about ice-sediment interaction. In light of this, each metric has its strengths and weaknesses as a descriptor. Consequently, in attempting to faithfully capture process-related characteristics of the bedform populations it is necessary to choose metrics that will be minimally sensitive to systematic biases.

Measurement is the translation from the real, currently observable landscape to geometric quantities describing bedforms (e.g., H and L). In terms of size-frequency populations, this presents three specific issues concerning the efficacy of the measurements taken:

1. Effect of source data on mapping (e.g., Smith and Clark, 2005): Smith and Wise (2007) outline the primary controls on the 'detectability' of landforms mapped from satellite imagery or visualised digital elevation models (DEMs); namely solar elevation, solar azimuth and sensor spatial resolution. These factors resolve to sampling issues: there exists a population of phenomena from which our observational method necessarily involves the selection of a subset. Solar azimuth can, for instance, systematically reduce all L values. Perhaps the best understood sampling bias is sensor resolution; small landforms are not detectable in coarse, low resolution data. Resolution therefore may contribute towards the low number of small bedforms (e.g., Fig. 1a) by imposing a threshold below which sampling becomes more difficult. Spagnolo et al. (2012), for instance, note this with respect to H in previous databases (Francek, 1991; Wysota, 1994; Hättestrand et al., 2004), although inability to observe in no way precludes the landforms not being there in the first place. Without knowing the actual population or error associated with the sampling, true values for statistics derived from the whole population cannot be ascertained with certainty.
2. Quantification method: even for a given mapped outline and digital terrain model (DTM), a variety of algorithms exist to compute a bedform's properties (H, W, L, and volume V) (e.g., Spagnolo et al., 2010; Hillier and Smith, 2012). Values will vary, e.g. for H (Spagnolo et al., 2012), depending upon the method selected. Identical geometries, however, will be affected by the same proportion at all scales. Removing post-glacial clutter (e.g., trees) to

create a DTM will affect H for mapped forms (Hillier and Smith, 2012). This has not been systematically studied, but it seems probable that bedforms with small heights will more commonly be rendered unmappable.

3. Subjectivity of interpretation: Manual mapping of bedforms is subjective and reliant upon the expertise and experience of the mapper. Whilst the process is not objectively repeatable, procedures are employed to maintain consistency and minimise bias (e.g., Smith and Clark, 2005; Hughes et al., 2010). Interpretations may, perhaps inevitably, vary more towards both perceived limits of the size range of a bedform, creating the largest uncertainties there. This subjectivity may, in future, be alleviated by automated mapping (e.g., Hillier , 2008; Saha et al., 2011; Kalbermatten et al., 2012; Rutzinger et al., 2012), but most benefits depend upon agreement being reached on an exact formal definition of each bedform (e.g., Evans, 2012).

After measurement, post-glacial preservation rates also affect bedform populations. If the measurement issues could all be accounted for, it would be possible to interpret frequency information across the size spectrum in terms of physical processes. Even then, however, a low prevalence for palaeo-landforms does not necessarily mean they are not abundant in active environments. Relative abundances could still be an artefact of post-glacial degradation that varies with size, e.g. diffusive hillslope-type erosion (e.g., Putkonen and Swanson, 2003). The preservation of small features, flutes for instance, is thought to be low. Therefore, to best interpret mapped subglacial bedforms in terms of subglacial processes, it is likely important to use measures least affected by all the issues identified above. At the very least, doing this has no detrimental effects.

In terms of a size-frequency distribution, non-glacial distortions may be summarised as follows (also Fig. 1b). Artefacts affecting all sizes by a single factor do not change the distribution's shape, and are a minor issue. Most seriously, there is potentially significant undersampling of small features due to several limitations in source data, post-glacial erosion, perhaps the quantification method, and potentially the views of an interpreter when mapping landforms. This latter factor also introduces uncertainty into the upper end of the size distribution, potentially increasing or decreasing detections or introducing outliers by including genetically unrelated landforms. So, 'good' metrics will be insensitive to the potential absence of small landforms and either not be unduly influenced by outliers at the upper end of the size range or provide means to identify and exclude them. They should also not, if possible, depend on sample size or arbitrary choices. For utility, it is also desirable to have as succinct yet complete a description of the distribution as possible.

Currently both simple (H, W, and L) and derived morphometric measures such as 'elongation' (i.e., L/W) are collated for populations (e.g., Hoppe and Schytt, 1953; Boulton, 1976; Stokes and Clark, 2002; Dunlop and Clark, 2006; Clark et al., 2009; Smith et al., 2009; Phillips et al., 2010). Up to eight parameters or metrics (e.g., Clark et al., 2009) are used to describe each measure (e.g., Fig. 1a): minimum, maximum, mean, standard deviation, modal class, median, skewness and kurtosis. Whilst undoubtedly useful for initial assessment, the number and nature of these metrics is not necessarily ideal for describing populations. Problems include: (i) extreme values depend upon the number of observations (unless estimated using appropriate statistical techniques e.g., van der Mark et al. 2008), observational completeness, and distribution shape, (ii) modal class is dependent upon the selection of a bin width, and (iii) the use of all of four moments (i.e., mean, standard deviation, skew and kurtosis) to describe the shape of the distribution; comparisons between shapes are more straightforward for single characteristic shape parameters. Lastly, (iv) the mean is affected in the first order by the steepness and length of the right-hand tail (e.g., Fig. 1a), the location of the 'roll-over' at the mode, ϕ, and any outliers. Thus, this ensemble of metrics is somewhat unsatisfactory, primarily because smaller bedforms *may* be substantively under-represented (Fig. 1b), perhaps leaving larger bedforms best reflecting glacial processes (Fig 1b). A simpler description may be possible, however, whose parameters likely better relate to ice–sediment interaction and only requires the assumption that larger features are accurately observed. This would be a weaker requirement than that of accurate quantification at all sizes implicit in present analyses.

3. Graphical investigation

Appropriate parameterisation of a distribution requires knowledge of its form. Many univariate statistical distributions contain exponential or power-law elements (e.g., Leemis and McQueston, 2008). Exponential functions or forms plot as straight lines on semi-log plots, as do power law relationships on log–log ones. These plots are therefore useful in preliminary investigations of the characteristics of observed data. This section illustrates the utility of these approximations to subglacial bedform data through different plots of L for one data set relating to one type of bedform.

Through the production of a semi-log histogram (Fig. 1b) and plotting a linear fit through the data (see Section 4) above the modal 'roll over', it is possible to visually demonstrate that counts of drumlin lengths above the mode conform to an exponential distribution. Though being non-linear, Fig. 1c clearly demonstrates that no large part of the the distribution is power-law. Power-law segments in distributions are typical of fractals such as topography (e.g., Mandlebrot, 1983; Weissel et al., 1994; Cheng and Agterberg, 1996), natural phenomena (e.g., floods, earthquakes, and wildfires) (e.g., Main et al., 1999; Malamud et al., 2005; Kidson et al., 2006; Malamud and Turcotte, 2006), and linked to the notion of self-organising criticality in systems (e.g., Bak, 1996; Tebbens et al., 2001). Importantly, Haschenberger (1999) empirically relate the observed exponent of exponential distributions for fluvial bedforms to estimates of basal shear stress in that environment. Gradients of the fitted lines such as that in Fig. 1b may therefore not only capture an important property related to flow but also encapsulate it in a single value, facilitating easy intercomparison between data sets. Descriptively, e.g., in Fig. 1b, the exponential only applies to data above the mode. There are no grounds for plotting it at smaller sizes other than extrapolation. In the simplest possible model, continuing the trend may be seen as a continuation of the signature of a subglacial process, but there is no evidence here to support this. As noted in Section 2, the difference between data and extrapolation due to i) measurement: data, sampling and mapping fidelity, ii) possibly post-glacial preservation or iii) the roll-over being a signature of the processes of ice-sediment interaction resulting in smaller features not being created subglacially in the first place i.e., their genesis. Insufficient work has been published to make definitive, comprehensive comments upon which one dominates, but there are strong hints that commonly observed bedforms lack numerous smaller versions. For instance, in extension of the results of Smith and Wise (2007), Clark et al. (2009) suggest that a clear lower bound for W in UK drumlins is unlikely to be an artefact of imagery resolution, attributing it to glacial processes (i.e., smaller forms are less commonly created). For the smallest bedforms this is very probably true, and the exponential should certainly not be extended to the y-axis. Consider drumlins; size observations from recently deglaciated terrain (Johnson et al., 2010) conform with palaeo data, and very small drumlins are not reported. However, measurement and preservation issues seem to affect significant fractions of drumlins that are larger and yet below the mode (Smith et al., 2006). So, speculatively, the existence of a well-defined modal peak is a signature of physical processes. However, its location is not yet necessarily well determined, with the possibility that smaller features are not recorded. As such, non-glacial factors may have a large effect on measures such as the mean, particularly for mapping in areas where high-resolution DEMs are not available.

4. Objective parameterisation

Given that bedform size-frequency distributions appear well described by a right-hand exponential decay above the mode and a roll-over to low numbers below it (Fig. 1), a description using three parameters is proposed that is designed to best represent subglacial processes, facilitate comparison between regions and data sources, and whose computation is readily accessible to geomorphologists. The selected metrics to approximate the distributions, (see Fig. 2c), are:

1. Gradient (λ): magnitude of the gradient of the fitted line (e.g., Fig. 1b), which is the exponent of the decay (Eq. 2 in Appendix). This characterises the part of the distribution that is least likely affected by non-glacial factors. Larger bedforms will have greater endurance in the landscape and the observed frequency should be close to the expected frequency. No disproportionate weight in the fit is placed on the largest features whose interpretation may be uncertain (e.g., Fig. 2f), and features unrelated to the distribution can be identified and excluded.
2. Mode (ϕ): estimates the point at which bedforms are no longer representatively sampled, non-glacial factors become dominant, or ice-flow related behaviour somehow changes in its nature or effect. If many smaller features are missed, it will be influenced (e.g. Smith and Wise, 2007), but is much more robust than the mean or median.
3. Intercept (β_0): intercept of the exponential with the y-axis represents the spatial density of the landforms (i.e., number per unit area) in a way that is insensitive to the efficiency with which small ones are detected, unlike the mean (e.g., Smalley and Unwin, 1968; Miller, 1972; Menzies, 1979b). Whilst the area, A, of a bedform field remains inexactly defined, the use of this for subglacial bedforms is limited at present, but is a key parameter compared for seamount distributions (e.g., Jordan et al., 1983; Scheirer and Macdonald, 1995; Hillier and Watts, 2007) illustrating its potential.

It is anticipated that λ values, either individually or when plotted against each other for H, W or L (e.g., x–y or ternary diagrams), will be a powerful means of characterising landforms. This could, for example, contribute to the debate as to whether bedforms constitute a continuum (e.g., Rose and Letzer, 1977; Rose, 1987; Clark, 1993; Clark et al., 2009), with data points for localities for each bedform type either forming separate domains or a merging in a progression from one to the other. Using λ should make such analyses robust to the dataset or resolution used. Note also that λ will not vary with the size of the data set. ϕ is a natural measure of unimodal bedform distributions and is a useful metric whatever it is thought to represent. For instance, ϕ is a good indicator of the size at which imperfect detection arises perhaps due to data type where this dominates (e.g., Smith and Wise, 2007), and will reflect glacial processes where measurement is not an issue.

The question then is how to estimate values for these metrics. Various methods to estimate parameters of plotted data exist (e.g., Cornell and Speckman, 1967); fitting a line (e.g., by ordinary least squares – OLS) to counts from a selected portion of a histogram considered to be linear may be done for simplicity (e.g., Wessel, 1997), but is not optimal (e.g., Smith and Jordan, 1988; Solow et al., 2003; Bauke, 2007). OLS fits of power-laws to log–log frequency plots, for instance, are known commonly to introduce significant, systematic, unpredictable biases into estimates of gradient (e.g., Newman, 2005; Clauset et al., 2009). The results also depend on i) bin width and construction (e.g., Newman, 2005) and ii) range chosen. An insight into the limitations of applying OLS to plots such as Fig. 2 may be gained by considering that it fits to the x–y plot rather than the underlying data, and each point on the plot is assigned equal weight and accuracy despite containing a different number of data, although larger counts tend to be less variable. An objective, statistically valid method based upon the underlying data (i.e., not fitting a frequency plot) that minimises arbitrary choices is proposed to estimate λ, ϕ and β_0. The method of moments (e.g., Freund and Walople, 1980, p. 325) is used to estimate the mode using a Gamma distribution, then the gradient obtained through a maximum likelihood fit (e.g., Freund and Walople, 1980, p. 327) of an exponential distribution for data larger than ϕ. This may be performed without any specialist statistical software, requiring only the calculation of the mean and standard deviation (i in Appendix). Not only is this approach relatively straightforward, but with minor adaptation it allows parameter estimation from the published literature using data presented in histograms (iii in Appendix), although it is not able to recover information lost during binning. In fact, data digitised from published histograms were deliberately used in Fig. 2 to specifically illustrate this point.

The fitted lines (solid lines in Figs. 1 and 2) show the efficacy of the method, and whilst the Gamma distribution shown by dashed lines provides a poorer fit to the population, it is able to estimate ϕ particularly well. In Fig. 2a, ϕ is estimated as 424 m (3 s.f.), inside the range of the 393–441 of the modal bin of Clark et al. (2009). The same is true for W and H with 177 m inside 173-183 and 3.7 m inside 3.5–4.0, respectively (Fig. 2b,c). Note that no selection of a bin width is necessary. This approach of fitting a line to data larger than an objectively determined value for ϕ overcomes *ad hoc* criteria previously used to determine the range of data fitted (e.g., n in bin > 5; Rappaport et al., 1997). Furthermore the method outlined here avoids the systematic overestimation of λ that occurs when it is estimated by fitting a Gamma distribution (ii in Appendix). Once the exponential distribution is fitted, β_0 is calculated by simple geometry. A worked example detailing the procedure is provided in a Microsoft Excel spreadsheet as Supplementary Material accompanying this paper. In anticipation that readers may want to compare bedform populations, a method of statistically evaluating whether λ is significantly different for those populations is given in iv) in Appendix

5. Prevalence of the exponential tail

Fig. 2 demonstrates that a form of size-frequency distribution with roll-over and right-hand exponential tail is typical of subglacial bedforms and derived measurements. Data for Fig. 2 were selected to demonstrate this via a number of specific points. Fig. 2a,d depicts an exponential tail for large samples ($n > 10,000$) of the same measure, L, of a particular bedform (i.e., drumlins) in two discrete study areas. Evidence for the form is therefore not location dependent, and it may occur wherever bedforms do. Fig. 2a,e,g,h shows the output of at least four independent mappers demonstrating that the form is not a result of an individual's style or preference. Independent mapping of a sub-area of Fig. 2c is shown in Fig. 2e for the same measure of UK drumlins, H. So, the occurrence of the form over large areas is not purely the result of aggregation, but is directly related to and applicable to individual flow sets. Furthermore, Fig 2e illustrates the form's utility for even a relatively small sample ($n < 200$), although the error bars for descriptive parameters are larger; $\lambda = 0.209 \pm 0.003\ m^{-1}\ 2\sigma$ and $\lambda = 0.238 \pm 0.024\ 2\sigma$ for Fig. 2c and Fig. 2e, respectively, using the estimation method in Section 4. To be explicit, Fig. 2 demonstrates that the form applies to all glacial bedforms considered in this paper for which adequate data are available for assessment with drumlins in Fig. 2a–f, ribbed moraine in Fig. 2g and MSGL in Fig. 2h. Tentatively, this description may also apply to flutes (e.g., H), but evidence is limited (i.e., $n \sim 50$) (e.g., Hoppe and Schytt, 1953; Boulton, 1976) leading to much scatter in semi-log plots.

Despite the weight of evidence presented in Fig. 2, it is important to note that the fit to the right hand tail is neither perfect nor ubiquitous. Firstly, then, it is notable that the plots (Fig. 2) show some scatter for large bedforms, and exponentials fit imperfectly. Most data were, quite deliberately, digitised from published histograms, but its presence in Fig. 2h demonstrates that errors due to this re-use of data are not the main cause. Perhaps it originates from uncertainty in categorising and thus selecting larger forms. Secondly, some data sets show distinct deviations from a linear trend on semi-log plots. Height data from northern Sweden (Hättestrand et al., 2004; data pers. comm.), where 'crag-and-tail' bedrock-influenced drumlins dominate, show a distinct bend in their trend on a semi-log plot (Fig. 3a). Why? Few published bedform frequency plots exist to assess this. Elongation ratio (i.e., L/W) data digitised and re-plotted from Fig. 2 of Phillips et al. (2010) also shows trends of two distinct gradients. These data, however, come from neighbouring regions in which Phillips et al. (2010) consider the influence of bedrock in creating landforms. Fig. 3b shows a steep trend (blue line) in Zone 1 '*dominated by an extensive drumlin field*' and a shallower one in Zone 2 (red line) where landforms are of '*ice moulded bedrock*'. Speculatively, it seems possible that bedrock influence can create geometric extremes beyond those of till-dominated landforms consistent with local formational conditions. Thus, the Swedish data may be exhibiting the signature of bedrock influence. Note the contrasting studies in Fig. 2 (e.g., Clark et al., 2009; Spagnolo et al., 2012) pointedly seek to exclude bedrock influenced landforms from their drumlin catalogues, as do most studies (cf. Stokes et al., 2011). More generally, distinct trends will likely exist if a plot contains data aggregated from distinctly different flow regiemes, perhaps forms attributed to streaming ice (e.g., Fig. 2h) and 'typical' drumlins (e.g., Fig. 2d).

In summary, an exponential right-side tail is typical, and perhaps characteristic, of till-dominated 'emergent' subglacial bedforms (Clark, 2010), and the spread of observations is sufficient to suggest that this is a general characteristic. The consistency of form is remarkable considering the gradient and perhaps mode likely change with local conditions.

6. Discussion

6.1. Choice of parameterisation

This paper concerns the description of subglacial bedforms and, in particular, how to most usefully quantify populations within an area. The approach taken is that size-frequency distributions may contain information not best captured by currently used metrics. Primary difficulties for current statistical metrics in reflecting ice-sediment interaction are that many are not natural descriptors of heavily skewed distributions and are sensitive in the first order to imperfect detection rates for smaller features. Additional computational issues exist for some in that they are dependent upon sample size (e.g., min., max., range) or bin selection for aggregation (e.g., mode). Size-frequency data for measures of co-located subglacial bedforms display linearly on semi-log plots (Figs. 1 to 3). This demonstrates that they are commonly distributed exponentially, at least above their modal values. Noting this form creates the possibility to design a simpler description.

The simplest description would be an exponential probability density function. This has been used to characterise domains of submarine volcanoes (e.g., Jordan et al., 1983; Scheirer and Macdonald, 1995; Hillier and Watts, 2007) and fluvial scour depths (Haschenberger, 1999), but frequencies of these do not roll-over at small sizes. The single parameter, the exponent λ, could not capture this. In fluvial geomorphology a variety of two-parameter distributions (e.g., Gamma, Gaussian, Gumbel, Log-normal, and Weibull) have been evaluated for their potential to describe bedform size-frequency populations (Leemis and McQueston, 2008; van der Mark et al., 2008). The distributions approximate, with variable degrees of success, the *shape* of the size-frequency distributions of the populations including a roll-over. Whilst entirely statistically valid, their utility when applied to subglacial landforms may suffer as both their parameters are influenced by data across the whole size range. Ideally, for the purposes of description, the characteristics of the right-hand tail that likely best represent subglacial processes should not be influenced by potentially unmapped small features.

The method proposed here minimises the influence by fitting an exponential distribution, λ, to only data above the mode, ϕ: two shape parameters. Admittedly, λ and ϕ incompletely describe observations below the mode, giving only its starting point, but this is where observations are least securely related to glacial processes. λ represents a part of the distribution least likely affected by factors unrelated to ice-sediment interaction, and ϕ ensures that data selection for its calculation is objective. If features larger than the mode are unreliably detected it may not be entirely accurate, but biases due to this will be no worse than for other parameterisations. Synthetic landscapes (Hillier and Smith, 2012) may allow this to be quantified. β_0 is an additional scaling factor normalised for area to make it a useful measure i.e., of landform spatial density.

6.2. Utility of the parameterisation

Typically, in parameterising data, there is a trade-off between computational simplicity and ease (e.g., requirement for statistical software), and objectivity and rigour. This is optimised in the method suggested here as no subjective choices (e.g., bin width) exist: it fits underlying data, not a plot, and the whole calculation is possible without specialist software. It requires only the calculation of means and standard deviations (see i in Appendix or the accompanying worked example using Excel. These calculations may be biased by large, mis-identified outliers, but these are rare and the exponential form provides a mechanism for assessing if observations are consistent with the bulk of a population, leading to an iterative fitting solution if necessary. The method to estimate λ, ϕ and β_0 demonstrably (Fig. 2) works on both raw data and those already derived from published histograms.

The parameterisation proposed is entirely descriptive and non-genetic in that it is not necessarily related to any formational process: the description will be valid whether or not future work identifies it as a signature of any particular ice-sediment process. Its non-genetic nature is useful in a characterisation as it avoids tying it to process-related debates. It, and particularly λ, not only has the power to present a single, generally applicable measure of bedforms, but also apply it to a wide range of published size catalogues, mapped from data of various types and ages, allowing inter-comparison. For instance, flow sets can now be more usefully compared with each other across all deglaciated regions and with the output of numerical ice sheet models (e.g., flow velocity or basal shear stress). A method of determining whether λ is significantly different between flow sets is also given. With the same governing equations proposed to control the evolution of bedforms created by ice, water or wind (e.g., Fowler, 2002), and a similarity between glacial (e.g., Fig. 2) and fluvial size-frequency distributions (e.g., van der Mark et al., 2008), applications for the parameterisation may also exist for analogous fluvial and aeolian bedforms.

6.3. Bedform genesis

Perhaps the most exciting aspect of the work is the future potential to use the explanatory power of the exponential characterisation in terms of understanding physical processes that are operating. Some insights, however, are feasible now. The caveat is that caution is necessary as multiple processes or histories can lead to the same statistical distributions (e.g., Tuckwell, 1995; Beven, 2006; Newman, 2005).

Tentatively, it is possible to suggest that the similarity between distributions for different bedforms indicates some commonalities between processes creating them and progressions in the processes between the bedform types. In depth modelling of the underlying processes of bedform genesis is beyond the scope of this work, but the few indicators available suggest that λ may directly reflect aspects of physical processes. Specifically, Haschenberger (1999) empirically relate λ for fluvial scour depths to basal shear stress in that environment. Furthermore another simple form of size-frequency distribution, the power-law, has been interpreted and modelled in terms of process (e.g. Newman, 2005), for instance 'self-organised criticality' (e.g. Bak, 1996; Tebbens et al., 2001). 'Self-organised critically' involves a set of simple rules and randomness acting at multiple locations that combine to produce characteristic size-frequency distributions. Subglacial bedforms originating in the presence of random variations at multiple locations may have a similar ability to produce characteristic distributions. Indeed, fluvial bedforms with similar heavy-tailed size-frequency distributions (e.g., van der Mark et al., 2008; Singh et al., 2011) are considered to originate in random fluctuations in turbulent flow (e.g., Fredø se, 1996; McElroy and Mohrig, 2009; Coleman and Nikora, 2011) with H and L described there as 'stochastic variables' (van der Mark et al., 2008). Similarly, ice-sediment interaction may be fundamentally stochastic in nature i.e., bedform growth may be a process involving the convolution of randomness with simple rules about the rate of growth. This is consistent with geophysical studies that have revealed spatio-temporally variable bed conditions (Vaughan et al., 2003; Smith, 2006; Murray et al., 2008) and subglacial landforms (King et al., 2007; Smith and Murray, 2009) that evolve rapidly on sub-decadal timescales (Smith et al., 2007; King et al., 2009) under Antarctic ice streams. It is unclear, however, whether this variability at the bedform scale originates dominantly in the dynamics of ice-sediment-water interactions (e.g., water incursions or basal stick-slip events) or those between bedforms. This stochastic approach contrasts to a deterministic view whereby proto-bedforms of known size and shape always evolve similarly with time to a predictable final morphology; perhaps, each bedform's size may be individually limited by local physical conditions that vary in space such that an exponential distribution is created. It is not immediately clear, however, how neighbouring bedforms of dramatically different sizes, as commonly observed, originate in this theory, although it is likely that bedforms are 'born' at different times (cf. Smith et al., 2007), even within a single flow-set. So for this reason, and by a loose analogy with the processes creating exponential tails for fluvial bedform populations, we suspect that conditions that give rise to subglacial bedforms are fundamentally variable and stochastic.

Many possible processes can be conceived in which bedforms are created and destroyed under ice using randomness and growth with various rate characteristics. A limited number, however, will produce exponential size-frequency distributions. The observations are therefore a constraint on models of bedform genesis. For instance, can stochastic variability be incorporated into till instability theory of Hindmarsh (1998)? Considering the stability or otherwise of bedform populations with respect to time may also prove valuable. Are bedforms in steady state dynamic equilibrium? If so, λ values may relate to properties of ice flow, such as velocity. Alternatively, if size-frequency distributions continue to evolve with time, ϕ and λ might combine to provide some constraint upon both the rate and duration of bedform growth. Finally, note that for accurately measured and well-preserved size-frequency distributions, a different two-parameter distribution (see Section 6.1) may assist in providing further constraints by describing the roll-over as well as the exponential tail. So, we suggest that future progress will come through understanding the observed exponential in terms of the statistics and mechanics of ice flow.

7. Conclusions

This paper presents a simple yet robust descriptive parameterisation that can be used to summarise and compare populations of subglacial bedforms, e.g. in flow-sets. Whilst a variety of distributions have been used in other disciplines, an exponential characterisation is appropriate in this area and offers potential explanatory power in terms of the processes in operation. Through plotting observations of landform size, specifically for ribbed moraine, drumlins and MSGL, and for populations of different sizes, the following main conclusions may be drawn:

- Till-dominated subglacial bedform size-frequency distributions characteristically have an exponential right-hand tail.

- Semi-log plots are a useful tool with which to initially assess this since exponentials plot as straight lines.

- The distributions may be rigorously, objectively and practically approximated by using the method of moments and the Gamma distribution to estimate the mode ϕ, and then using a maximum likelihood method to estimate the exponent λ (i.e. gradient of the semi-log plot) for measurements larger than the mode.

- For observations *below* the mode, a combination of possible sampling error and probable absence means that there is some uncertainty here depending upon the data type used for mapping.

- λ is likely to reflect glacial processes significantly better than previously used metrics.

This description uses three parameters, rather than the selection of up to eight currently used. This simplicity makes it a preferable approach to developing understanding in unresolved areas such as the subglacial bedform continuum or spatial patterns of palaeo-flow. Future insights may come through the comparison of the spatial distribution of observed λ with the output of numerical ice sheet models, or through creating statistical models to link the mechanics of physical processes to observable characteristics of bedform populations. Indeed, it is consistent with the observed exponentially-tailed distribution that the growth and development of subglacial bedforms may be fundamentally stochastic in nature and involve the convolution of randomness with some, as yet unknown, simple rules about the rate of growth.

Acknowledgements

The authors wish to thank Clas Hättestrand for data. We are grateful to Prof. A. James and an anonymous reviewer for their constructive comments, and Prof. T. Oguchi for his help as editor.

Appendix

i) Parameterisation Method

The method proposed below is not the only possible solution (e.g., Fraile and Garcia-Ortega, 2005), but is objective, statistically valid, and easily implemented. Firstly, determine the range of the measured variable, x, that conforms to the exponential distribution: the linear part of the semi-log plot (Section 3). The mode is a visually reasonable, objective, estimate of the lower end of this range. Despite some previous practice to the contrary (e.g., Abers et al., 1988; Smith and Cann, 1992; Rappaport et al., 1997), all data of larger x are included here. To calculate the mode, based on the underlying data x_i where $i = 1 \ldots n$ and n is the number of individual observations, a Gamma distribution is used. A Gamma distribution is a two-parameter distribution (α,λ_{g}) which tends to an exponential at large x, but which rolls over to zero at small x, with a probability density function (pdf) (Tuckwell, 1995, , p. 62):

$$f(x) = \frac{\lambda_{\mathrm{g}}^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda_{\mathrm{g}} x}\ , x > 0\,;\lambda_{\mathrm{g}}, \alpha > 0 \tag{1}$$

Fig. 2 shows that the Gamma distribution approximates ϕ well. Maximum likelihood estimators (MLEs) of α and λ_g require numerical techniques, but may also be estimated by the method of moments as $\hat{\alpha} = (\bar{x}/s_x)^2$ and $\hat{\lambda_{\mathrm{g}}} = \bar{x}/(s_x)^2$ where $\bar{x}$ is the sample mean, and s_x is the sample standard deviation (Tuckwell, 1995, , p. 326). The mode of the Gamma distribution, ϕ, is then $(\hat{\alpha} - 1)/\hat{\lambda_{\mathrm{g}}}$. For length, L, of UK drumlins this is shown in red on Fig. 2a, as 424 m (3 s.f.) and is inside the range of the 393–441 modal bin of Clark et al. (2009).

Now, determine the gradient and intercept based upon data of size greater than ϕ. Data are fitted as a left-truncated exponential, which is equivalent to an exponential shifted by ϕ. Let $k_i = x_i - \phi$, and then for $k > 0$ the MLE estimator of the gradient of the exponential, $\hat{\lambda}$, is $\hat{\lambda} = 1/\bar{k}$ where $\bar{k}$ is the mean of the data (Tuckwell, 1995, , p. 329). This fully describes the pdf of the exponential distribution, which is defined by the following equation and has an area of 1 unit under its curve (Tuckwell, 1995, , p. 86 and 196).

$$f(x) = \lambda e^{-\lambda x} \tag{2}$$

Histograms and frequency plots are considered inferior to pdfs by many statisticians, but are common in the wider literature. So, how are the results related to the more familiar histogram (Figs. 1 and 2) deliberately used in this paper? The short answer is the line to be plotted on a histogram is given by the following equation

$$y = \hat{\lambda}x + \ln(n_\phi \hat{\lambda} w_\mathrm{b}) + \hat{\lambda}\phi \tag{3}$$

where $\hat{\lambda}$ and ϕ have been calculated as above, w_b is bin width, n_ϕ is the number of measurements greater than ϕ, and x and y are the variables relating to the axes of the semi-log plot.

To calculate the equation of the best-fit line for a frequency plot, firstly obtain the x-intercept x_0 through setting the exponential distribution equal to zero. Scaled up to an area of n_ϕ under its curve, the equation for the linear part of the histogram becomes $f(x) = n_\phi \lambda e^{-\lambda x}$. Taking logs and setting this to zero gives a horizontally shifted x-intercept of $\ln(n_\phi \hat{\lambda})/\hat{\lambda}$. This becomes $x_0 = [\ln(n_\phi \hat{\lambda} w_\mathrm{b})/\hat{\lambda}] + \phi$ when bin width w_b is used to multiply up for the conversion from count density (per unit x) to count within bins and the line is un-shifted and put back to its original location. With x_0 and λ estimated for the line, y_0, the y-intercept (i.e. y at $x = 0$) is by simple geometry $\hat{\lambda}x_0$. The equation of the line is therefore $y = \hat{\lambda}x + \hat{\lambda}x_0$ or $y = \hat{\lambda}x + \ln(n_\phi \hat{\lambda} w_\mathrm{b}) + \hat{\lambda}\phi$.

By plotting this equation it appears that the data are well-approximated (Fig. 2), and no arbitrary upper cut-off is required unless data clearly outlying from the distribution are known e.g. $L > 4$ km; an iterative technique may be used, perhaps excluding data by the probability that they could exist in the fitted distribution, calculated from the pdf of the exponential. The spatial density of landforms, β_0 is y_0/A, where A is the area of the study (km^2), although rigorous use of this will require work to define criteria by which to calculate A.

ii) Using Gamma distribution

Paola and Borgman (1991) estimated λ for fluvial bedforms by fitting a Gamma distribution. This assumed that fitted Gamma distributions become linear on semi-log plots after the mode; Fig 2b illustrates that this is not the case. The fitted distribution (dashed line) systematically increases with x and λ_g and so the gradient at large x is not reached on the plot; λ is systematically overestimated ($\lambda_\mathrm{g} = 0.00589$ whilst $\lambda = 0.00313$ for calculations as in Section (i) of this Appendix). Furthermore, this approach was not preferred since Fig. 2 a–c show that the gamma distribution (dashed line) fits more poorly than the exponential ($x > \phi$) and the fit gets poorer as α increases from 1 (exponential distribution): α_L =3.58, $\alpha_W = 6.68$, $\alpha_H = 2.11$. Namely, the extent of over-estimation depends upon the overall shape of the distribution, which is not desirable.

iii) Parameterising histogram data

Using, n underlying data points, x_i, the mean, $\bar{x}$, and standard deviation, s_x, of the sample are calculated using standard formulae. For n data with counts, c_j, of bins at x_j the related equations used are:

$$\bar{x} = \frac{1}{n} \sum c_j x_j \tag{4}$$

$$s_x = \sqrt{\frac{1}{n-1} \sum c_j \left(x_j - \bar{x}\right)^2} \tag{5}$$

iv) Comparing populations or sub-populations

Confidence intervals can be determined for $\hat{\lambda}$. Strictly, $\hat{\lambda}$ is distributed as $2n\lambda\bar{x} \sim \chi^2_{2n}$, assuming ϕ correctly delimits the linear portion of the size-frequency distribution. However, with large n, usually > 30, (i.e. using the central limit theorem) the sampling distribution of $\hat{\lambda}$ becomes approximately normal (Tuckwell, 1995, , p. 255–9). So, for large n an asymptotic unbiased approximation to the variance of a MLE estimate of a parameter may be determined using the Cramer-Rao lower bound (Tuckwell, 1995, , p. 313–4), giving

$$\lambda \sim N(\hat{\lambda}, \frac{\hat{\lambda}^2}{n_\phi}). \tag{6}$$

Stated more fully, the exponent of the observed sample (the gradient of the linear part of the semi-log plot) is distributed according to the normal distribution with a mean of $\hat{\lambda}$ and variance of $\hat{\lambda}^2/n_\phi$. The standard error of the sampling distribution of $\hat{\lambda}$ for an exponential distribution is

$$s \simeq \sqrt{\frac{\hat{\lambda}^2}{n_\phi}} \tag{7}$$

which, using standard tabulations for the normal distribution (Tuckwell, 1995, , p. 520), gives a 95% confidence interval of

$$\pm 1.96\sqrt{\frac{\hat{\lambda}^2}{n_\phi}}. \tag{8}$$

This then allows the difference between two sub-populations with estimated gradients to be assessed using a standard t-test. If independent random samples, of sizes $n_{1\phi}$ and $n_{2\phi}$ values above modes ϕ_1 and ϕ_2, are drawn from distributions $N(\lambda_1, \sigma_1^2)$ and $N(\lambda_2, \sigma_2^2)$, with standard deviations unknown *a priori*, $H_0 : \lambda_1 = \lambda_2$ can be tested using the test statistic

$$t_{n_{1\phi}+n_{2\phi}-2} = \frac{\hat{\lambda}_1 - \hat{\lambda}_2}{s_p\sqrt{\frac{1}{n_{1\phi}} + \frac{1}{n_{2\phi}}}} \tag{9}$$

where s_p is the pooled variance

$$s_p = \frac{(n_{1\phi} - 1)s_1^2 + (n_{2\phi} - 1)s_2^2}{n_{1\phi} + n_{2\phi} - 2} \tag{10}$$

within which s_1 and s_2 are estimated by Eq. 7 (Tuckwell, 1995, , p. 348). $t_{n_{1\phi}+n_{2\phi}-2}$ is the t statistic for $n_{1\phi} + n_{2\phi} - 2$ degrees of freedom, and can be compared to critical values obtained from standard tables or elsewhere. Note that this is a two-tailed test, so for 95% confidence the 0.025 tabulated value is the critical one. A z statistic may also be useful because samples are relatively large.

References

Aario, R., 1977. Classification and terminology of moranic landforms in Finland. Boreas 6, 77–100.

Abers, G.A., Parsons, B., Weissel, J. K., 1988. Seamount Abundances and Distributions in the southeast Pacific. Earth and Planetary Science Letters 87, 137–151.

Bak, P., 1996. How Nature Works: the Science of Self-Organized Criticality. Springer, New York.

Bauke, H., 2007. Parameter estimation for power-law distributions. Eur. Phys. J. B 58, 167–173.

Benn, D.I., Evans, D.J.A., 2010. Glaciers and Glaciation, 3rd Edition. Hodder, Oxford, UK.

Benn, D.I., Evans, D.J.A., Shaw, J., Munro-Stasiuk, M., 2006. Subglacial megafloods: outrageous hypothesis or just outrageous? / reply. In: Knight, P.G. (Ed.), Glacier Science and Environmental Change. Wiley-Blackwell, p. 42-60.

Beven, K., 2006. A manifesto for the equifinality thesis. J. Hydrology 320, 18–36.

Bingham, R.G., King, E.C., Smith, A. M., Pritchard, H.D., 2010. Glacial geomorphology: Towards a convergence of glaciology and geomorphology. Progress in Physical Geography 34, 327–355.

Boulton, G., 1976. The origin of glacially fluted surfaces - observations and theory. J. Glaciology 17, 287–309.

Boulton, G., Hindmarsh, R. C.A., 1987. Sediment deformatin beneath glaciers - rheology and geological consequences. Journal of Geophysical Research-Solid Earth and Planets 92 (B9), 9059–9082.

Chapwanya, M., Clark, C.D., Fowler, A.C., 2011. Numerical computations of a theoretical model of ribbed moraine formation. Earth Surface Processes and Landforms 36, 1105–1112.

Cheng, Q.M., Agterberg, F.P., 1996. Multi-fractal modelling and spatial statistics. Mathematical Geology 28, 1–16.

Chorley, R.J., 1959. The Shape of drumlins. J. Glaciology 3, 339–344.

Clark, C.D., 1993. Mega-scale glacial lineations and cross-cutting ice-flow landforms. Earth Surface Processes and Landforms 18, 1–29.

Clark, C.D., 2010. Emergent drumlins and their clones: from till dilatancy to flow instabilities. J. Glaciology 51, 1011–1025.

Clark, C.D., Hughes, A.L., Greenwood, S.L., Spagnolo, M., Ng, F.S., Apr. 2009. Size and shape characteristics of drumlins, derived from a large sample, and associated scaling laws. Quaternary Science Reviews 28, 677–692.

Clauset, A., Shalizi, C.R., Newman, M. E.J., 2009. Power-law distributions in empirical data. SIAM Review 51, 661–703.

Coleman, S.E., Nikora, V.I., 2011. Fluvial dunes: initiation, characterization, flow structure. Earth Surface Processes and Landforms 36, 39–57.

Cornell, R.G., Speckman, J.A., 1967. Estimation for a simple exponential model. Biometrics 23, 717–737.

Dunlop, P., Clark, C.D., 2006. The morphological characteristics of ribbed moraine. Quaternary Science Reviews 25, 1668–1691.

Dunlop, P., Clark, C.D., Hindmarsh, R. C.A., 2008. Bed Ribbing Instability Explanation: Testing a numerical model of ribbed moraine formation arising from coupled flow of ice and subglacial sediment. Journal of Geophysical Research 113 (F3).

Evans, D. J.A., Livingstone, S., Vieli, A., Ó Cofaigh, C., 2009. The palaeoglaciology of the central sector of the British and Irish Ice Sheet: reconciling glacial geomorphology and preliminary ice sheet modelling. Quat. Sci. Rev. 28, 739–757.

Evans, I.S., 2012. Geomorphometry and landform mapping: What is a landform? Geomorphology 137, 94–106.

Fowler, A.C., 2000. An instability mechanism for drumlin formation. In: Maltman, A.J., Hubbard, B., Hambrey, M.J. (Eds.), Deformation of Glacial Materials, geological Edition. Geol. Soc. Publishing House, London, pp. 307–319.

Fowler, A.C., 2002. Evolution equations for dunes and drumlins. Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales, Serie A. Mat. 96, 377–387.

Fraile, R., Garcia-Ortega, E., 2005. Fitting an exponential distribution. J. Applied Meteorology 44, 1620–1625.

Francek, M., 1991. A spatial perspecitve on the New York drumlin field. Physical Geography 12, 1-18.

Fredø se, J., 1996. The stability of a sandy river bed. In: Nakato, T., Ettema, R. (Eds.), Issues and Directions in Hydraulics. Balkema, Rotterdam, pp. 99–114.

Freund, F.E., Walople, R.E., 1980. Mathematical Statistics, 3rd Edition. Prentice Hall, London.

Gardiner, V., 1983. The relevance of geomorphometry to studies of Quaternary morphogenesis. In: Briggs, D.J., Waters, R.S. (Eds.), Studies in Quaternary Geomorphology. Geo Books, Norwhich, pp. 1–18.

Hart, J.K., 1999. Identifying fast ice flow from landform assemblages in the geological record: a discussion. Annals of Glaciology 28, 59–67.

Haschenberger, J., 1999. A probability model of scour and fill depths in gravel-bed channels. Water Resources Research 35, 2857–2869.

Hättestrand, C., Gotz, S., Naslund, J., Fabel, D., AP, S., 2004. Drumlin formation time: Evidence from northern and central Sweden. Geografiska Annaler Series A-Physical Geography 86A, 155–167.

Hättestrand, C., Kleman, J., 1999. Ribbed moraine formation. Quat. Sci. Rev. 18, 43–61.

Hillier, J.K., 2008. Seamount detection and isolation with a modified wavelet transform. Basin Research 20, 555–573.

Hillier, J.K., Smith, M.J., 2012. Testing 3D landform quantification methods with synthetic drumlins in a real DEM. Geomorphology 153, 61–73, doi:10.1016/j.geomorph.2012.02.009.

Hillier, J.K., Watts, A.B., 2007. Global distribution of seamounts from ship-track bathymetry data. Geophysical Research Letters 34, L113304, doi:10.1029/2007GL029874.

Hindmarsh, R. C.A., 1998. Drumlinization and drumlin-forming instabilities: viscous till mechanisms. J. Glaciology 44, 293–314.

Hollingsworth, S.E., 1931. The glaciation of western Edenside and adjoining areas and the drumlins of Edenside and the Solway basin. Quart. J. Geol. Soc. London 87, 281–359.

Hoppe, G., Schytt, V., 1953. Some observations on fluted moraine surfaces. Geografiska Annaler 2, 105–115.

Hughes, A., Clark, C.D., Jordan, C., 2010. Subgalcial beforms of the last British ice sheet. Journal of Maps, 543—-563.

Johnson, M.D., Schomacker, A., Benediktsson, I.O., Geiger, A.J., Ferguson, A., Ingolfsson, O., Sep. 2010. Active drumlin field revealed at the margin of Mulajokull, Iceland: A surge-type glacier. Geology 38, 943–946.

Jordan, T.H., Menard, H.W., Smith, D.K., 1983. Density and size distribution of seamounts in the Eastern Pacific inferred from wide-beam sounding data. Journal of Geophysical Research 88, 10508–10518.

Kalbermatten, M., van der Ville, D., Turberg, P., Tuia, D., Joost, S., 2012. Multiscale analysis of geomorphological and geological features in high resolution digital elevation models using the wavelet transform. Geomorphology 138, 352–363.

Kidson, R., Richards, K.S., Carling, P.A., 2006. Fractal analysis for natural hazards. Geol. Soc. Special Publication 261, 141–153.

King, E.C., Hindmarsh, R.C.A., Stokes, C.R., 2009. Formation of mega-scale glacial lineations observed beneath a west Antarctic ice stream. Nature Geoscience 2, 585–596.

King, E.C., Woodward, J., Smith, A.M., 2007. Seismic and radar observations of subglacial bed forms beneath the onset zone of Rutford Ice Stream Antarctica. J. Glaciology 53, 665–672.

Leemis, L.M., McQueston, J.T., 2008. Univariate distribution relationships. The American Statistician 62, 45–53.

Livingstone, S., Cofaigh, C., Evans, D. J. A., 2008. Glacial geomorphology of the central sector of the last British-Irish Ice Sheet. Journal of Maps, 358–377.

Main, I., Irving, D., Musson, R., Reading, A., 1999. Constraints on the frequency-magnitude relation and maximum magnitudes in the UK from observed seismicity and glacio-isostatic recovery rates. Geophys. J. Int. 137, 535–550.

Malamud, B.D., Millington, J. D.A., Perry, G. L.W., 2005. Characterizing wildfire regiemes in the United States. Proceedings of the National Academy of Sciences of the United States 102, 4694–4699.

Malamud, B.D., Turcotte, D.L., 2006. The applicability of power-law frequency statistics to floods. J. Hydrology 322, 168–180.

Mandlebrot, B.B., 1983. The Fractal Geometry of Nature. W. H. Freeman and Company, New York.

McElroy, B., Mohrig, D., 2009. Nature of deformation of sandy bed forms. J. Geophys. Res. 144, F00A04.

Menzies, J., 1979a. The mechanics of drumlin formation with particular reference to the change in pore-water content of the till. J. Glaciology 22, 373–383.

Menzies, J., 1979b. A review of the literature on the formation and location of drumlins. Earth Science Reviews 14, 315–59.

Miller, J.W., 1972. Variation in New York drumlins. Annals of the Association of American Geographers 62, 418–23.

Morris, E.M., Morland, L.W., 1976. A theoretical analysis of the formation of glacial flutes. J. Glaciology 17, 311–323.

Murray, T., Corr, H., Forieri, A., Smith, A.M., 2008. Contrasts in hydrology between regions of basal deformation and sliding beneath Rutherford ice stream, West Antarctica, mapped using radar and seismic data. Geophys. Res. Lett. 35, 10.1029/2006GL028207.

Newman, M. E.J., 2005. Power laws, Pareto distributions and Zipf's law. Contemporary Physics 46, 323–351.

Ó Cofaigh, C., Dowdeswell, J.A., King, E.C., Anderson, J.B., Clark, C.D., Evans, D. J.A., Evans, J., Hindmarsh, R. C.A., Larter, R.D., Stokes, C.R., 2010. Comment on Shaw J., Pugin, A. and Young, R. (2008): "A meltwater origin for Antarctic shelf bedforms with special attention to megalineations, Geomorphology 102, 364-375". Geomorphology 117, 195–198.

Paola, C., Borgman, L., 1991. Reconstructing random topography from preserved stratification. Sedimentology 38, 553–656.

Phillips, E., Everest, J.D., Diaz-Doce, D., Dec. 2010. Bedrock controls on subglacial landform distribution and geomorphological processes: Evidence from the Late Devensian Irish Sea Ice Stream. Sedimentary Geology 232, 98–118.

Putkonen, J., Swanson, T., 2003. Accuracy of cosmogenic ages for moraines. Quaternary Research 59, 255–261.

Rappaport, Y., Naar, D.F., Barton, C.C., Liu, Z.L., Hey, R.N., 1997. Morphology and Distribution of Seamounts Surrounding Easter Island. Journal of Geophysical Research 102, 24713–24728.

Rose, J., 1987. Drumlins as part of a glacier bedform continuum. In: Menzies, J., Rose, J. (Eds.), Drumlin Symposium. Balkema, Rotterdam, pp. 103–116.

Rose, J., Letzer, J. M., 1977. Superimposed drumlins. J. Glaciology 18, 471–480.

Rutzinger, M., Hofle, B., Kringer, K., 2012. Accuracy of automatically extracted geomorphological breaklines from airborne LiDAR curvature images. Geografiska Annaler Series A-Physical Geography 94A, 33–42.

Saha, K., Wells, N., Munro-Stasiuk, M., 2011. An object-orientated approach to landform mapping: A case study of drumlin. Computers and Geosciences 37, 1324–1336.

Scheirer, D.S., Macdonald, K. C., 1995. Near-axis seamounts on the flanks of the East Pacific Rise, 8N to 17N. Marine Geophysical Research 100, 2239–2259.

Shaw, J., 1983. Drumlin formation related to inverted melt-water erosional marks. J. GlaciologyJ. 29, 461–479.

Shaw, J., Pugin, A., Young, R.R., 2008. A meltwater origin for Antarctic shelf bedforms with special attention to megalineations. Geomorphology 102, 364–375.

Singh, A., Lanzoni, S., Wilcock, P. R., 2011. Multiscale statistical characterization of migrating bed forms in gravel and sand rivers. Water Resources Research 47, W12526.

Smalley, I., Unwin, D., 1968. The formation and shape of drumlins and their distribution and orientation in drumlin fields. J. Glaciology 7 (51), 377–390.

Smalley, I., Warburton, J., 1994. The shape of drumlins, their distribution in drumlin fields, and the nature of the sub-ice shaping forces. Sedimentary Geology 91, 241–252.

Smalley, I., Piotrowski, J.A., 1987. Critical strength / stress ratios at the ice-bed interface in the drumlin forming process: from ”dilatancy” to ”cross-over”. In: Menzies, J., Rose, J. (Eds.), Drumlin Symposium. Balkema, Rotterdam, pp. 81—-86.

Smith, A.M., 2006. Microearthquakes and subglacial conditions. Geophys. Res. Lett. 33, 10.1029/2006GL028207.

Smith, A.M., Murray, T., 2009. Bedform topography and basal conditions beneath a fast-flowing West Antarctic ice stream. Quat. Sci. Rev. 28, 584–596.

Smith, A.M., Murray, T., Nicholls, K.W., Makinson, K., Athalgerirdottir, G., Behar, A., Vaughan, D. G., 2007. Rapid erosion and drumlin formation observed beneath a fast-flowing Antarctic ice stream. Geology 35, 127–130.

Smith, D.K., Cann, J.R., 1992. The role of seamount volcanism in crustal construction at the Mid-Atlantic Ridge. Journal of Geophysical Research 97, 1645–1658.

Smith, D.K., Jordan, T.H., 1988. Seamount Statistics in the Pacific Ocean. Journal of Geophysical Research 93, 2899–2918.

Smith, M.J., Clark, C.D., 2005. Methods for the visualization of digital elevation models for landform mapping. Earth Surface Processes and Landforms 30, 885–900.

Smith, M.J., Rose, J., Booth, S., 2006. Geomorphological mapping of glacial landforms from remotely sensed data: an evaluation of the principal data sources and an assessment of their quality. Geomorphology 76, 148–165.

Smith, M.J., Rose, J., Gousie, M.B., 2009. The Cookie Cutter: A method for obtaining a quantitative 3D description of glacial bedforms. Geomorphology 108, 209–218.

Smith, M.J., Wise, S.M., 2007. Mapping glacial lineaments from satellite imagery: an assessment of the problems and development of best procedure. Int. J. Applied Earth Observation and Geoinformation 9, 65–78.

Solow, A.R., Costello, C.J., Ward, M., 2003. Testing the power law model for discrete size data. The American Naturalist 162, 685–689.

Spagnolo, M., Clark, C.D., Hughes, A.L., 2012. Drumlin relief. Geomorphology 153-154, 179–191.

Spagnolo, M., Clark, C.D., Hughes, A.L., Dunlop, P., Stokes, C.R., 2010. The planar shape of drumlins. Sedimentary Geology 232, 119–129.

Stokes, C.R., Clark, C.D., 2002. Are long subglacial bedforms indicative of fast ice flow? Boreas 31, 239–249.

Stokes, C.R., Clark, C.D., 2003. The Dubawnt palaeo-ice stream: evidence for dynamic ice sheet behaviour on the Canadian Shield and insights regarding the controls on ice stream location and vigour. Boreas 32, 263–279.

Stokes, C.R., Fowler, A.C., Clark, C.D., Hindmarsh, R. C.A., Spagnolo, M., 2013. The instability theory of drumlin formation and its explanation of their varied composition and internal structure. Quat. Sci. Rev. 62, 77–96.

Stokes, C.R., Spagnolo, M., Clark, C.D., 2011. The composition and internal structure of drumlins: complexity, commonality, and implications of a unifying theory of their formation. Earth Science Reviews, 10.1016/j.earscirev.2011.05.001.

Tebbens, S.F., Burroughs, S.M., Barton, C.C., Naar, D. F., 2001. Statistical self-similarity of hotspot seamount volumes modeled as self-similar criticality. Geophys. Res. Lett. 28, 2711–2714.

Tuckwell, H.C., 1995. Elementary Applications of Probability Theory, 2nd Edition. Chapman and Hall, London.

van der Mark, C.F., Blom, A., Hulscher, S. J. M.H., 2008. Quantification of variability in bedform geometry. J. Geophys. Res. 113, F03020, 10.1016/j.earscirev.2011.05.001.

Vaughan, D.G., Smith, A.M., Nath, P.C., Le Meur, E., 2003. Acoustic impedence and basal shear stress beneath four Antarctic ice streams. Annals of Glaciology 36, 225–232.

Weissel, J.K., Pratson, L.F., Malinverno, A., 1994. The length-scaling properties of topography. J. Geophys. Res. 99, 13997–14012.

Wessel, P., 1997. Sizes and ages of seamounts using remote sensing: implications for intraplate volcanism. Science 277, 802–805.

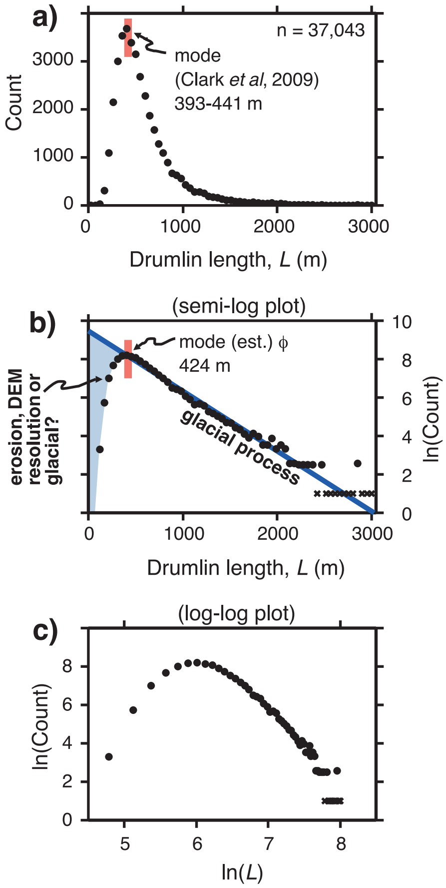
Wysota, W., 1994. Morphology, internal composition and origin of drumlins in the southeastern part of Chelmno-Dobrzy Lakeland, north Poland. Sedimentary Geology 91, 345–364.

Figure captions

Fig. 1:. Frequency plots of the lengths, L, of UK drumlins. Black dots are data digitised from Fig. 8 of Clark et al. (2009); bin width ~50 m. Larger drumlins ($L > \phi$) are, to a good first approximation, fit (see text) by a straight line in b), an exponential distribution. They are not power law, i.e. linear in c). Mode, ϕ, in b) estimated by fitting gamma distribution. Crosses indicate zero counts, placed at a nominal value of 1 in b) and c).

Fig. 2:. Semi-log plots for subglacial bedform properties (H, W, L and L/W) and types (drumlin, ribbed moraine, MSGL). Data (black dots) are exact (e, h) and digitised (a, b, c, d, f and g). Bin widths vary, and crosses indicate zero counts, placed at a nominal value of 1 if $n > 10,000$. Solid lines are the exponential distributions fitted to data above the mode ϕ. The exponent is the plotted gradient, λ. The red bars indicate ϕ, estimated from fitted gamma distributions, shown as dashed lines in a) to c) only. Hiller and Smith data are for 'best' isolation technique. Spagnolo et al. (2012) discard superimposed (i.e. cross-cutting) (e.g., Rose and Letzer, 1977) or slightly overlapping drumlins of Clark et al. (2009). MSGL are from Dubawant Lake ice stream flow-set (Stokes and Clark, 2003).

Fig. 3:. Size-frequency data possibly exhibiting the influence of bedrock. a) Swedish drumlin observations. H categorised discretely as 2 ,5, 7, 10, 20, 30 ... 80 m, so the number of drumlins per unit bin width (count density) is plotted. Lines fitted manually. b) Frequencies of drumlins (black dots) and streamlined bedrock forms (open circles) for L/W from neighbouring regions in Anglesey, UK. Lines fitted as in Section 4.



a)
n = 37,043
mode
(Clark et al, 2009)
393-441 m
Count
3000
2000
1000
0
0
1000
2000
3000
Drumlin length, L (m)
b)
(semi-log plot)
mode (est.) ϕ
424 m
erosion, DEM
resolution or
glacial?
glacial process
ln(Count)
10
8
6
4
2
0
0
1000
2000
3000
Drumlin length, L (m)
c)
(log-log plot)
ln(Count)
8
6
4
2
0
5
6
7
8
ln(L)

Figure 1:

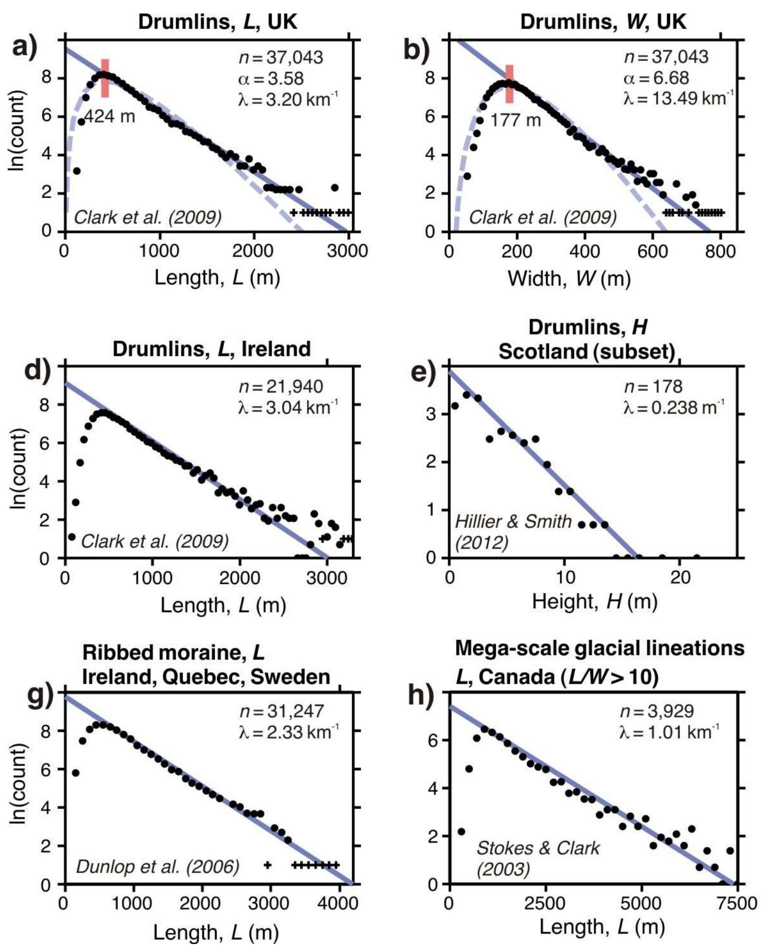

a)
Drumlins, L, UK
n = 37,043
α = 3.58
λ = 3.20 km-1
424 m
Clark et al. (2009)
Length, L (m)
ln(count)
b)
Drumlins, W, UK
n = 37,043
α = 6.68
λ = 13.49 km-1
177 m
Clark et al. (2009)
Width, W (m)
c)
Drumlins, H, UK
mode ϕ
n = 25,848
α = 2.11
λ = 0.209 m-1
gradient λ
3.7 m
Spagnolo et al. (2012)
Height, H (m)
d)
Drumlins, L, Ireland
n = 21,940
λ = 3.04 km-1
Clark et al. (2009)
Length, L (m)
e)
Drumlins, H
Scotland (subset)
n = 178
λ = 0.238 m-1
Hillier & Smith
(2012)
Height, H (m)
f)
Drum., aspect ratio, UK
n = 37,043
λ = 0.89
Min.
ratio 1:1
Clark et al.
(2009)
L/W
g)
Ribbed moraine, L
Ireland, Quebec, Sweden
n = 31,247
λ = 2.33 km-1
Dunlop et al. (2006)
Length, L (m)
h)
Mega-scale glacial lineations
L, Canada (L/W > 10)
n = 3,929
λ = 1.01 km-1
Stokes & Clark
(2003)
Length, L (m)

Figure 2:

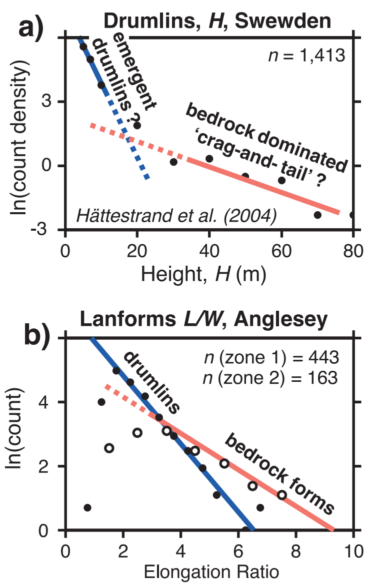

a)
Drumlins, H, Swewden
n = 1,413
emergent drumlins ?
bedrock dominated 'crag-and- tail' ?
Hättestrand et al. (2004)
ln(count density)
Height, H (m)
b)
Lanforms L/W, Anglesey
n (zone 1) = 443
n (zone 2) = 163
drumlins
bedrock forms
ln(count)
Elongation Ratio

Figure 3: