Testing techniques to quantify drumlin height and volume; synthetic DEMs as a diagnostic tool

J. K. Hillier

Department of Geography, Loughborough University, LE11 3TU, UK.

M. J. Smith

School of Geography, Geology and the Environment, Kingston University, KT1 2EE, UK.

J. K. Hillier, Department of Geography, Loughborough University, LE11 3TU, UK. (j.hillier@lboro.ac.uk)

M. J. Smith, School of Geography, Earth and Environment, Kingston University, KT1 2EE, UK. (mike@hsm.org.uk)

Abstract. Glacial bedforms' heights, H, and volumes, V, likely preserve important information about the behaviour of former ice sheets. However, large systematic errors exist in the measurement of H and V. Three semi-automated methods to isolate drumlins from other components of the landscape (e.g., trees, hills) as portrayed by NEXTMap have recently been devised, however it is unclear which is most accurate. This paper undertakes the first quantitative comparison of such readily implementable methods, illustrating the use of statistically representative 'synthetic landscapes' as a diagnostic tool. From this analysis, guidelines for quantifying the 3D attributes of drumlins are proposed. Specifically, to avoid obtaining incorrect estimates caused by substantial systematic biases, interpreters should currently take three steps; declutter the DEM for estimating H but not for V, remove height data within the drumlin, then interpolate across the resultant hole to estimate a basal surface using Delaunay triangulation. Results are demonstrated through analysis of drumlins in an area in western Central Scotland. The guidance arguably represents the best current advice for subglacial bedforms in general, highlighting the need for more studies into the quality of mapped data using synthetic landscapes.

Key words: Subglacial, Drumlin, Bedform, DEM, Quantification

1. Introduction

Subglacial bedforms are a group of landforms created by ice-water-sediment interaction at the interface between glaciers and the terrain underneath [e.g., *Benn and Evans*, 2010]. 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*, 1979], (iii) ribbed moraine [*Hättestrand and Kleman*, 1999] and (iv) mega-scale glacial lineations (MSGL) [*Clark*, 1993].

The location and shape (e.g., length L, width W) of bedforms gives information about the kinematics and dynamics [e.g., *Prest and Grant*, 1968; *Ó Cofaigh et al.*, 2005; *Bradwell et al.*, 2008; *Hubbard et al.*, 2009], and possibly even mechanics [e.g., *Hindmarsh*, 1998; *Dunlop et al.*, 2008; *Chapwanya et al.*, 2011], of past ice flow. Subglacial bedform length, L, for instance, may be related to ice velocity [e.g., *Clark*, 1993; *Hart*, 1999; *Stokes and Clark*, 2002], and elongation ratio (i.e., L/W) may reflect the influence of bedrock on bedform genesis [e.g., *Phillips et al.*, 2010].

Height, H, is less often quantified and interpreted, but it has been used to distinguish ice flows of different ages [e.g., *Hättestrand et al.*, 2004] and its frequency distribution may be evidence that bedform growth is governed by processes or boundary conditions that are fundamentally stochastic (i.e., random in time) [*Hillier et al.*, 2013]. More directly, H has been analytically linked to the thermo-mechanical processes and physical properties (e.g., till rheology, ice velocity) of ice-sediment interaction within models of bedform creation; the 'till instability' model [*Hindmarsh*, 1998; *Fowler*, 2000] creates increasing relief mechanically, whilst *Hooke and Medford* [2013] employ a thermal driver to generate the feedbacks necessary for unstable growth. Qualitatively, but more broadly, and by analogy with other bedforms (e.g., fluvial) [e.g., *van der Mark et al.*, 2008; *Coleman and Nikora*, 2011], this work fits into a context of long-standing speculation that glacial bedform sizes vary with flow conditions and sediment flux [*Rose and Letzer*, 1977].

Volume, V, is much less used due to the historical lack of data at a sufficient spatial resolution [e.g., *Evans*, 1987]. Despite this, *Rose* [1989] was able to quantity volumes of sediment and, in combination with known dates, to estimate sediment flux, whilst *Shaw et al.* [1989] used volume to estimate potential meltwater quantity. Even so, it may be under-exploited. V is regarded as an important tool for understanding aeolian dunes where it is used to assess fluid flow (i.e., wind patterns), sediment behaviour (i.e., availability and flux) and wider flow regieme (i.e., climate) [see e.g., *Grohmann and Sawakuchi*, 2013]. So bedforms' heights and volumes likely preserve important information about different aspects of flow in former ice sheets.

This said, post-formational processes may have changed apparent sizes [e.g., *Hillier et al.*, 2013]; for example, post-glacial sedimentation potentially reduces estimates of volume [*Finlayson*, 2013] and height [e.g., *Boyce and Eyles*, 1991; *Smith et al.*, 2006; *Korkalainen et al.*, 2007; *Spagnolo et al.*, 2012]. Care should also be exercised when interpreting volume derived from surface expression alone as bedforms such as drumlins may have pre-existing (e.g., bedrock) cores, and thus a complex undulating base to the till layer [e.g., *Gluckert*, 1973; *Ó Cofaigh et al.*, 2002; *Stokes et al.*, 2011]. So, a drumlin's apparent height and volume may not all be due to till, and further information may be required when making some inferences about sub-glacial processes. Finally, significant random and, importantly, systematic errors exist in observations of H and V measured from digital elevation models (DEMs) [*Hillier and Smith*, 2012], hindering insights into physical processes that could be provided by subglacial bedforms.

Techniques to quantify the the height and volume of glacial bedforms must use DEM representations of landscapes. Predominantly, such DEMs contain not only the upper surfaces of these bedforms, but also anthropogenic 'clutter' and other signatures whose origin is not related to ice flow. A procedure for the isolation of height related to distinctly identifiable landforms, like subglacial bedforms, is 'regional-residual separation' (RRS) [e.g., *Wessel*, 1998; *Hillier and Watts*, 2004; *Hillier and Smith*, 2008, 2012; *Spagnolo et al.*, 2012]. This estimates the upper and basal surfaces of landforms, then, from the difference between these two (non-planar) surfaces, in conjunction with the outlines, H and V can be calculated. Currently for glacial bedforms these techniques contain manual and automated stages, and are generally termed 'semi-automated'. A number of different semi-automated methods have recently been used to isolate drumlins [*Smith et al.*, 2009; *Spagnolo et al.*, 2011; *Hillier and Smith*, 2012], however they make different methodological choices and it is unclear which is most accurate and least affected by systematic biases.

Hillier and Smith [2012] created 'synthetic landscapes', comprising idealised drumlins placed within a real DEM. These landscapes are statistically representative of the real landscape, at least with respect to the extraction of H and V using semi-automated methods. This allows an 'objective' (i.e., quantitative and reproducible) assessment of errors in measurement as retrieved values can be compared to the *a priori* known values. Synthetics thus enable an objective comparison of quantification techniques and whilst *Hillier and Smith* [2012] assessed the effect of altering one parameter in one method, they did not address three basic decisions that are made implicitly or explicitly whenever performing the RRS necessary to quantify the 3D properties of drumlins:

1. Is the removal of clutter in a DEM necessary when estimating H and V? If so, how should this be performed?

2. To accurately predict basal surfaces, should elevations within mapped outlines be discarded in a 'cookie cutter' [*Smith et al.*, 2009] style?

3. Which interpolation or extrapolation method best predicts bedforms' basal surfaces?

This work addresses these three questions, with the aim of providing practical recommendations for those mapping and quantifying drumlins. This is done through using the synthetic DEMs of *Hillier and Smith* [2012] to objectively assess the automated part of RSS procedures for quantifying H and V. RRS and procedures to be evaluated are explained in Section 2. The synthetic DEMs are of a study area in central Scotland, which is described in Section 3 along with a summary of how they were created, followed by a description of the research design used in this study to determine the most accurate quantification technique. Section 4 presents results, with discussion in Section 5 where recommendations about the best approach to use are made. Through using synthetic DEMs to conduct similar analyses it is hoped that observations clearly reflecting physical processes can be made for drumlins across the globe and insights gained into their genesis.

2. Regional Residual Separation

The computation of an underlying 'regional' surface, historically larger-scale and calculated first [see e.g., *Wren*, 1973; *Wessel*, 1998], is subtracted to leave a 'residual' layer. Typically, the residual layer is intended to represent height related to a physical process, such as glacial bedforms, and regional-residual separation (RRS) may be repeated to distinguish a number of layers [e.g., *Hillier and Watts*, 2005]. Calculation of H and V, where digital outlines are available for drumlins, requires definition of the upper and basal surfaces and the final quantification from the two surfaces; this study is concerned with the efficacy of the RRS techniques in defining the surfaces. In this work two RRS stages are involved: (1) removal of noise through filtering to 'declutter' the landscape estimating the upper surface, and (2) approximation of the basal surface of the drumlin. Decluttering is of interest across a range of disciplines [e.g., *Sithole and Vosselman*, 2004], and is itself an aspect of 'improving' raw digitial elevation data for morphological analysis [see e.g., *Milledge et al.*, 2009]. In previous studies decluttering has been neglected for simplicity [*Smith et al.*, 2009], or because it is found to introduce significant artefacts, but a choice about it is always made even if this is implicit. The following sections outline the methods for decluttering, methods for estimating basal surfaces, and an overview of the combinations selected for the three published methods applied to drumlins which, along with some other possibilities, are evaluated here.

2.1. Decluttering

Within a DEM, H can be described at any location (x, y) as the sum of n components (Eq. 1) [e.g., *Nettleton*, 1954; *Wren*, 1973; *Wessel*, 1998; *Hillier and Smith*, 2008]:

$$H_{DEM} = H_1 + H_2 + ...H_n \tag{1}$$

For the purpose of studying drumlins, the simplest approximation is the division of topography into three components: (1) 'noise' or surface clutter; these are small-scale height variations not genetically related to drumlin formation (e.g., trees, anthropogenic infrastructure); (2) 'drumlins' refers to the subglacial bedforms of interest and (3) 'hills' is shorthand for more regional topographic trends that are not drumlins. Heights may be then described by Eq. 2.

$$H_{DEM} = H_{noise} + H_{drumlins} + H_{hills} \tag{2}$$

Decluttering is filtering applied to a measured DEM to reveal the level of the 'bare earth' through the removal H_{noise}. In decluttering, the exact methods employed and definitions of which features should remain vary according to the data type being processed and proposed use of the output. Commonly, the terms digital surface model (DSM) and digital terrain model (DTM) are used to refer to DEMs before and after decluttering. A significant literature exists on decluttering including statistical, object-based, and multi-scale approaches [see e.g., *Sithole and Vosselman*, 2004; *Bartels and Hong*, 2010]. Extant methods are generally noted as less effective in hilly areas and there has been no explicit evaluation as to their performance for topographically subtle glacial landforms in this terrain.

Two main options are currently available to glacial geomorphologists interested in decluttering DEMs:

1. Do not declutter [e.g., *Livingstone et al.*, 2008; *Smith et al.*, 2009];

2. Use a simple filter [*Hillier and Smith*, 2012], which can be consistently applied to any DSM data;

In addition, for NEXTMap a proprietary decluttering algorithm [*Wang et al.*, 2001] has been applied to their DTM product [e.g., *Spagnolo et al.*, 2011]. Fig. 1 highlights the distortions caused to the height and width of drumlins by NEXTMap's decluttering

(Fig. 1b; grey line). Note that simple sliding window filters (e.g., median) can produce results similar to NEXTMap's DTM; a range of filter widths were tested, and the 110 m wide median that was found to minimise the average absolute vertical height difference between its output and the NEXTMap's DTM (dashed grey line).

2.2. Basal surface estimation

Height data within and around a drumlin may be used to calculate a basal surface, estimating the ground level were that feature not to exist. Early techniques to estimate regional surfaces underlying landforms of positive relief include frequency domain filters [e.g., *Watts and Daly*, 1981; *Cazenave et al.*, 1986] and statistics within a sliding window of stated width that moves across the DEM, specifically convolutions such as the mean [e.g., *Watts*, 1976] and Gaussian-weighted average [e.g., *McKenzie et al.*, 1980]. These were computationally efficient, and therefore possible, but in effect distributed height from the raised landforms rather than removing it detrimentally affecting the suitability of the output 'regional' surface as an estimator of an underlying basal surface [e.g., *Smith*, 1990; *Hillier and Watts*, 2004]. Windows returning the lowest [e.g., *Cobby et al.*, 2001; *Hillier et al.*, 2007], or the more statistically 'robust' [*Box*, 1953] median and mode [e.g., *Smith*, 1990; *Levitt and Sandwell*, 1996; *Crosby et al.*, 2006; *Kim and Wessel*, 2008] were therefore employed. The robust windowed filters operate better even where landforms are densely packed in space, effectively ignoring landforms as statistical outliers giving a basal surface much less biased by them [*Smith*, 1990]. The main limitation of these is that landforms vary in size, whilst a single window width must be selected. More sophisticated multi-scale techniques [e.g., *Sithole and Vosselman*, 2004], some of which automatically identify landforms [*Behn et al.*, 2004; *Hillier and Watts*, 2004; *Hillier*, 2008], now exist but have not yet been applied to drumlins. An overview of the possibilities can be gained through a combination of summaries in *Wessel* [1998], *Sithole and Vosselman* [2004], and *Hillier* [2011].

The techniques noted above require no manual intervention. Recent, 'cookie-cutter' style semi-automated techniques are distinguished by their use of manually digitised outlines, within which data are removed. This allows interpolation techniques such as bi-cubic splines [e.g., *Smith and Wessel*, 1990; *Smith et al.*, 2009] and Delaunay triangulation [e.g., *Watson*, 1982; *Shewchuk*, 1996; *Wessel and Smith*, 1998], which by definition fill holes in data, to also be used to estimate the basal surface within drumlins. These interpolations rely almost entirely upon data immediately outside each drumlin's outline, and thus are more prone to some errors than sliding window filters that estimate typical values for a regional trend from a wider spatial area (see e.g., Fig. 11b of *Smith et al.* [2009]).

2.3. Published Approaches

Three approaches have been published to perform regional-residual separations as the basis for isolating drumlins. Each makes different choices regarding the three main decisions in performing the RRS: decluttering, use of a cookie-cutter approach, and basal surface estimation.

Smith et al. [2009] do not declutter, they pioneer the cookie-cutter approach, and use a fully tensioned (i.e., $t = 1$) bi-cubic spline to estimate basal surfaces. *Spagnolo et al.* [2011] use NEXTMap's DTM implicitly accepting their decluttering algorithm, use the cookie-cutter, and use Delaunay triangulation to estimate basal surfaces. *Hillier and Smith* [2012] declutter with a 60 m wide sliding window median filter, do not use a cookie-cutter approach, and estimate basal surfaces with a 500 m wide median filter. NEXTMap's decluttering may be superior to simple methods [e.g., *Hillier and Smith*, 2012], but is imperfect (Fig. 1b) and due to its proprietary nature cannot be reproduced and tested. By not using manual mapping in the RRS, only using it to calculate H and V from the surfaces (Section 3.3.2), *Hillier and Smith* [2012] minimise sensitivity to the subjective mapping. This may be beneficial if mapping is uncertain, or detrimental because it does not fully exploit the information imparted by expert mappers (see Fig. 1b, or Fig. 4 of *Hillier and Smith* [2012]).

3. Methodology

3.1. Study Area

The 13 x 8 km study area (Fig. 2) is located in the western part of central Scotland and identical to that examined in a number of previous studies [*Smith et al.*, 2006, 2009; *Clark et al.*, 2009; *Evans*, 2012; *Hillier and Smith*, 2012]. It contains a variety of glacial landforms (see Fig. 5a of *Smith et al.* [2006]), 184 of which were interpreted as drumlins by *Smith et al.* [2009] . The landforms are Younger Dryas (YD) [ended ∼11.7 ka] and Last Glacial Maximum (LGM) [ended ∼14.5 ka] in age [*Smith et al.*, 2006; *Rose and Smith*, 2008]; ice flow trended approximately towards the East and South in both the LGM [*Sissons*, 1967] and the YD [*Rose*, 1987].

Drumlins range from prominent (i.e., ∼25% over 10 m tall) to subtle (i.e., 1-2 m) [*Hillier and Smith*, 2012], with broader-scale terrain ranging from hilly to flat in the lower Endrick and Blane valleys. Newer YD age landforms are sharply bounded, whilst LGM features have aprons of mass wastage deposits around their lower slopes [*Smith et al.*, 2006]. Non-glacial clutter, such as trees and anthropogenic infrastructure, vary in their width-scale and spatial density [e.g., *Smith et al.*, 2006]. All of these may impact upon drumlin mapping and measurement.

The drumlins shown (Fig. 2b) were digitised from the NEXTMap DSM and quantitatively compared to field mapping in *Smith et al.* [2006]. A combination of gradient, two orthogonal relief-shaded images, and local contrast visualisations, considered 'optimal' [*Smith and Clark*, 2005], were used in the digitisation to minimise bias in the orientations of the drumlins [*Smith and Wise*, 2007]. These mapped forms were used to create synthetic landscapes by *Hillier and Smith* [2012]. Note that alternative drumlin maps exist (e.g., Fig. 1a), but assessing potential errors in the manual mapping part of semi-automated procedures is beyond the scope of this study.

3.2. Synthetic DEMs

The synthetic DEMs of *Hillier and Smith* [2012] are used in this study. These are based upon mapping [*Smith et al.*, 2006] and manipulation of NEXTMap's DSM (e.g., Fig. 1). NEXTMap is a single-pass interferometric synthetic aperture radar (IfSAR) product presented at a 5 m spatial resolution, with a vertical accuracy estimated as 0.5–1 m [*Intermap*, 2004]. Consequently, the synthetic DEMs are gridded at 5 m.

Figure 3 illustrates the method used to create the synthetic DEMs; *Hillier and Smith* [2012] proposed a two stage method. In stage one the original drumlins are removed (Fig. 3a) and quantified (i.e., H,W,L) and in stage two, drumlins of these same known properties, H_{in} and V_{in}, are inserted into the synthetic DEM (Fig. 3b). The 10 DEMs used in this paper were created using Method 2 of *Hillier and Smith* [2012] because they are a close match to the real data (i.e., their volume errors are very close to those of the real landscape). There are 173 synthetic drumlins in each synthetic landscape; 1730 in total.

To allow a better appreciation of the strengths and potential weaknesses of the method that *Hillier and Smith* [2012] used to create synthetic DEMs, the remainder of this section briefly reviews the primary issues involved. Difficulties stem from not being able to *a priori* perfectly isolate drumlins in a complex landscape. In particular, (1) artefacts will remain after drumlins have been removed, and (2) their real sizes and shapes are not known. How exactly then is it possible to create a realistic, statistically representative synthetic DEM?

In order to address the artefacts issue *Hillier and Smith* [2012] utilised their qualitative observation that non-glacial features, which are causing the mapping problems, appear to be located randomly with respect to the drumlins. Drumlins in other spatial configurations that are randomly repositioned with respect to non-glacial features can then be imagined (and later checked) to have the same measurement error characteristics as the real landscape. This precludes systematic biases due to the synthetic drumlins being co-located with the artefacts, and allows multiple (e.g., 10) realisations so that random effects cancel in order to reveal a clear sense of the errors.

Unknown real sizes do not preclude some analyses; for instance, to evaluate errors for a drumlin of a particular size (e.g., H= 10 m, W= 200 m, L= 400 m) these may simply be inserted into a synthetic DEM. *Hillier and Smith* [2012] create more realistic synthetics by selecting a semi-automated technique to produce first estimates of the sizes. Clearly, the selection of one semi-automated method in creating the synthetics introduces the possibility of circularity; by putting in drumlins this method has found, the method may then be favoured. So, several steps were taken to ensure that no signature of procedures used, such as a preferred filter width (see Section 2), entered the synthetic DEMs. Firstly, W and L of the synthetic drumlins are those derived from the original manually determined outlines [*Smith et al.*, 2009] and are entirely uninfluenced by the automated aspects of techniques being assessed here. Secondly, an idealised 2D Gaussian shape was verified as representative then used for the synthetic drumlins. So, morphological signatures cannot be passed from the automated filter to the synthetic DEM. Thirdly, estimated $H_{drumlin}$ was only removed inside the outlines of *Smith et al.* [2009], which are not spatially coincident with the synthetic drumlins. This, to a large extent, preserves the roughness and frequency domain characteristics of the original landscape.

The ultimate test of the synthetics designed for the purposes of *Hillier and Smith* [2012] and this paper is that they are statistically representative of the real landscape, at least for the extraction of drumlins' H and V by semi-automated methods. Without *a priori* knowledge of the real drumlins' sizes their recovered sizes, H_r and V_r, provide the strongest test of the synthetic DEMs. Recovered size-frequency distributions for the real and synthetic drumlins match well. The simplest explanation for this is that the synthetic DEMs and drumlins well represent those digitised in *Smith et al.* [2006]. If there is an error in the synthetic DEMs that causes a systematic bias during semi-automated extraction, it must be cancelled by an equal and opposite effect to achieve the match.

3.3. Research Design

The aim of this paper is to establish how to achieve the most accurate results in respect of three aspects of the regional-residual separations that are the core of 3D quantifications of H and V for drumlins. To do this combinations of the choices in steps 1 and

2 of the following computational sequence were analysed. Manual mapping and other computational stages are taken as fixed. In each analysis, the sequence is:

1. Declutter;

2. Estimate basal surface, with or without an initial cookie-cutter step;

3. Quantify H and V values from each drumlin's outline, upper surface and lower surface (Sections 3.3.1, 3.3.2);

4. Estimate errors, ε_H and ε_V (Section 3.3.3).

Given that there are 3 decluttering options, an option to use the cookie-cutter approach, and at least 9 simple basal surface techniques (noted in Section 2.2) most of which have a variable parameter (e.g., window width), a large number of combinations are possible. Specifically, using only coarse steps for the variable parameters and selecting only 5 surface estimation methods (i.e., Table 1) 780 combinations exist, which would require the quantification of $\sim$1.35 million synthetic drumlins. A more efficient approach was therefore chosen.

Thirteen numerical 'experiments' ($E1$ to $E13$) (Table 1) were designed to explore the possibilities within 83 variations, computing errors for 143,590 synthetic drumlins. The single-valued unbiased metric, ε, necessary to compare errors between the variants is presented in Section 3.3.3, and the rationale for the experimental sequence is in Section 3.3.4.

3.3.1. Basal surface estimation

Introducing a cookie-cutter approach to regional-residual separation can cause computational problems if mapped drumlins are touching or in very close proximity [*Smith et al.*, 2009] (e.g., Fig. 4). Specifically, how is it possible to extrapolate a basal surface across a drumlin when no data remains on one side of it? To avoid such issues, a rectangular subset of data is selected for each synthetic drumlin **D**, starting from the complete DEM each time. This is performed for all analyses. Then, for cookie-cutter variants, data within drumlins (**D** and **d**) are deleted, but restored within a 20 m buffer zone **B2** [*Smith et al.*, 2009] outside the drumlin being quantified (Fig. 4). A buffer zone **B1** is a rectangular buffer set to be bigger than **B2**, and is always extended to half the width of any windowed filter used to avoid edge effects.

3.3.2. Measure of H and V used

A variety of possible ways exist to calculate H and V from any given outline and pair of upper and lower surfaces bounding $H_{drumlin}$ [e.g., *Hättestrand et al.*, 2004; *Spagnolo et al.*, 2011; *Hillier and Smith*, 2012]. Here, consistent with the process used to create the synthetic DEMs (Section 3.2), V is the volume between a drumlin's upper and basal surfaces within its outline, and H is the maximum vertical distance between the two surfaces [e.g., *Smith et al.*, 2009; *Hillier and Smith*, 2012]. Height could be corrected for the steepness of the underlying slope α as $H\cos(\alpha)$ [*Spagnolo et al.*, 2012], but the effect is relatively minor in this area [*Hillier and Smith*, 2012].

3.3.3. Error metric (ε)

Numerical searches for optimal parameters or sets of choices require a single metric for comparison. Errors for H and V, ε_H and ε_V, are treated in separate searches because, as discussed later, they behave differently. In this study 'optimal' is defined as most accurate, namely the RRS variant for which recovered values, H_r and V_r, are closest to input ones, H_{in} and V_{in}. No metric can ever uniquely claim to be 'best' [e.g., *Stein and Stein*, 1992; *Hillier and Watts*, 2005; *Goutorbe*, 2010], therefore those ones selected should reflect the aims of the study. The potential for ambiguity is illustrated by individual errors (i.e., V_r/V_{in}) plotted as a histogram for the 1730 synthetic drumlins assessed for the RRS method of *Smith et al.* [2009] and its equivalent including decluttering (Fig. 5b,d). The data contain some extreme outliers so measures that are statistically 'robust' [e.g., *Box*, 1953], namely insentitive to outliers, such as the Median Absolute Deviation (MAD) behave differently from standard error metrics such as root mean square error or standard deviation [e.g., *Stein and Stein*, 1992; *Fisher and Tate*, 2006]. Decluttering reduces the standard deviation (i.e., $\pm 2\sigma$) of the errors, but increases the MAD in line with a visually detectable deterioration in performance

This study's primary aims are firstly to select an RRS technique which extracts the majority of drumlins well, centring the modal peak of the V_r/V_{in} frequency distribution on 1.0, and secondly to give equal weight to both large and small drumlins even when the latter are much more common. The first is self explanatory. The second is particularly necessary since some filters involve a choice of scale, and could therefore potentially better select drumlins of a certain size. Since there are many small drumlins, an RRS variant dealing with these well at the expense of large ones could appear to be performing well whilst introducing systematic size-related distortions in recovered frequency distributions. Using a mean V_r/V_{in} close to 1.0 to select a best windowed filter, for instance, distorts the size-frequency distribution as the metric is dominated by the impact on the more numerous small drumlins (Fig. 6a). Since the shape of size-frequency distributions may be a key indicator of ice flow behaviour [*Hillier et al.*, 2013], such distortions are not desirable.

The metric used is based on ε_{ij}, which is error ε for drumlin i in synthetic DEM j. ε is the difference between a quantity of known value in the synthetic DEM, X_{in}, and the value it is recovered as, X_r, (e.g., Fig. 5) expressed as a fraction (Eq. 3).

$$\varepsilon = |1 - (X_r/X_{in})| \tag{3}$$

In equation 3 X is either height, H, or volume, V. This gives values near zero for recovered sizes X_r close to input 'correct' ones X_{in}. ε_{ij} values are then combined into a single error metric for each of height or volume, ε_H or ε_V. An arithmetic mean, $\bar{\varepsilon_{ij}}$, is simple but would not produce results consistent with this study's aims for reasons stated above (see Fig. 6a). So ε_H and ε_V were calculated through a three-stage process:

1. For each DEM, ε_{ij} were placed into bins of width 50 m according to that drumlin's length, L, and the median for each calculated. This gives size classes equal weight however many measurements populate the class, giving equal weight to small and large drumlins. Bins start from 0 m i.e., the first bin was L= 0-50 m;
2. Take the mean of these to create a single error value for each DEM;
3. Combine DEMs by taking the median of the values for the individual DEMs.

3.3.4. Experiments

The thirteen numerical experiments ($E1$ to $E13$) noted above were designed to more efficiently explore and compare the parameter space of the wide range of methods, options and parameters in the regional-residual separations. Each experiment deals with one choice combination, and allows the user-defined parameter to vary. The first method proposed for drumlins [i.e., *Smith et al.*, 2009] was used as a baseline [EB] for comparison,

$E1$ to $E12$ explore paths of variation away from it, and $E13$ is simply a verification that programming errors (e.g., truncating floating point variables) are small.

The first experiments, $E1$ to $E5$, explore RRS techniques to estimate basal surfaces, but without decluttering. The experiments compare a variety of simple windowed filters (i.e., mean, median, lowest) [e.g., *Cobby et al.*, 2001; *Hillier and Smith*, 2012] to interpolation schemes recently used upon drumlins (i.e., spline, triangulation) [e.g., *Smith et al.*, 2009; *Spagnolo et al.*, 2011]. $E6$ assesses the cookie-cutter approach by examining the most accurate method from $E1$ to $E5$ that can be implemented with or without it. The later experiments, $E7$ to $E9$, are to verify that the relative efficacy of the quantification techniques remained the same in conjunction with 'simple' decluttering using a 60 m wide median filter; this decluttering was seen to preform adequately under visual inspection [*Hillier and Smith*, 2012]. To put the efficacy of decluttering methods into context $E10$ represents 'perfect' decluttering, containing errors only due to the geometry of the underlying larger-scale trends. This is achieved by applying a method as in EB to a DEM containing no clutter, only $H_{drumlins}$ and H_{hills} estimated by a 500 m wide sliding median filter as is visually determined in *Hillier and Smith* [2012]. Finally $E11$ and $E12$ probe further into decluttering, ensuring that conclusions do not rest on the subjective choice of a filter width of 60 m.

Directly testing the decluttering of NEXTMap is not possible as the algorithm is proprietary and sufficient detail to reproduce the work is not given. Simply using the difference between NEXTMap's DTM and DSM is not appropriate because distortions caused by decluttering must spatially correlate with the synthetic drumlins. As discussed later (Section 5.1), however, strong constraints are possible by establishing an analogy with the closest simple decluttering technique.

4. Results

Results are presented in two distinct stages. In the first, the output of two RRS choice combinations are described in detail. The comparison is between the approach of *Smith et al.* [2009], which does not use decluttering, and an equivalent with it. This serves to illustrate the shape and character of the distribution of errors for individual drumlins, again demonstrating a need for the error metric ε. It also establishes that decluttering substantially affects results, and so must form part of the analysis in this paper. This was previously suggested, but not known. Lastly, it re-emphasises that commonly used [e.g., *Smith et al.*, 2009; *Clark*, 2010; *MacLachlan and Eyles*, 2013], if arguably sub-optimal [*Hillier et al.*, 2013], metrics of measured populations (e.g., the mean) can be substantively affected by choices made during 3D quantification. Then, in the second stage, results of the series of experiments ($E1$ to $E12$) (Table 1) to find the 'best' approach are reported, and lastly those for the best method described in detail.

Note that ε values closer to 0 reflect lower amounts of error, whilst ratios of recovered values to actual ones near 1 (e.g., Fig. 5) do the same. It is also necessary to distinguish metrics calculated for groups of mapped features as any geomorphologist might (e.g., mean H for 173 drumlins) and accuracy information for individual recoveries (e.g., bias and spread) made possible by the use of synthetics that give insight into these.

4.1. Detailed initial comparison

Figs. 5a and b show the H and V recoveries of individual drumlins for the method of *Smith et al.* [2009] (EB). In both, the modal peak indicates some tendency to recover sizes correctly, but with considerable scatter. The nature of this scatter is important in that input mean volume, $\bar{V}_{in}$, of 1.14×10^5 m^3 is recovered accurately using 173 drumlins at $1.09 \pm 0.06 \times 10^5$ m$^3 (2\sigma)$ since individual errors are random. In contrast, mean input H, $\bar{H}_{in}$, of 6.6 m is not recovered accurately at 11.7 ± 0.4 m 2σ (Fig. 7a) since individual errors are systematically and heavily positively skewed. ε_H is 0.863 and ε_V is 0.263. Figs. 5c and d show recoveries using a method identical apart from that it includes decluttering, which significantly affects the distributions of individual errors. Note that the spikes of values at zero, particularly present in small drumlins, are not artefacts of programming errors. They are due to the height attributed to small, thin drumlins placed in flat areas being removed as clutter by the 60 m wide median filter used. This width may not be optimal, but was found satisfactory in a visual investigation of filter widths and types in this study area by *Hillier and Smith* [2012]. Even this subjective approach, however, improves ε_H dramatically to 0.329. Mean height of the population of 173 drumlins is much better recovered at 6.8±0.2 m (2σ) (Fig. 7a) because individual H_r/H_{in} values are more symmetrically distributed around correct recovery at 1.0; i.e., errors become largely random (Fig. 5a,c). Decluttering, however, introduces a systematic bias into recovered mean V for 173 drumlins, underestimating it at $0.98 \pm 0.05 \times 10^5$ m$^3 (2\sigma)$, driven by small drumlins. This size-related effect, a bias driven by small drumlins (Fig. 5d), further clarifies why ε is used in order to maximise the number of H_r/H_{in} values near 1.0 whilst giving equal weight to each size class of drumlin.

4.2. Experiments

Fig. 8 shows results for a variety of methodological combinations investigated for their ability to recover sizes (H_r, V_r).

$E1$ varies spline tension, but does not improve results significantly, changing ε_H little from EB and with ε_V of 0.243 at best as compared to 0.263 of EB. The benefit of using a tensioned spline for calculating volumes, however, as documented by *Hillier and Smith* [2012], is captured. Delaunay triangulation in $E2$ reduces ε_V more, to 0.192, and it is comparable to the best methods for ε_H. Similar to tensioned splines, triangulation creates a surface linking heights immediately outside drumlins by a direct path. $E3$ and $E4$ reflect an artefact, discussed below. Thus initial indications are that, if interpolating, techniques approximating a highly tensioned rubber sheet stretching across the hole are best.

Still without decluttering, $E3$ to $E6$ explore the possibilities of sliding window filters. By considering heights from further outside a drumlin they may, for instance, deal better with trees on the boundary, a problem identified by *Smith et al.* [2009]. Windowed 'sliding median' filters, $E3$, which do not require heights within drumlins to be removed, attain a minimum ε_V of 0.415 at a best width of 420 m. This is a little better than similar filters based on the mean in $E4$. Where drumlins' widths are typically $\leq$300 m [Fig. 8 of *Hillier and Smith*, 2012] these are reasonable width scales across which to determine the basal surfaces, but Fig. 1b illustrates how errors may still result. In contrast, the filter estimating basal surfaces using the lowest value within a window, $E5$, performs very poorly.

In Fig. 8a, results for ε_H initially appear to contradict the picture of sliding window filters performing less well. Where the windowed filters are better than the interpolations (i.e., in $E3$ and $E4$), however, this is due to coherent errors: clutter, when present, artificially raises the estimated basal surface reducing the height overestimate caused by the clutter. Two errors performing in concert, however, are unlikely to do so reliably or give a more accurate estimate of drumlin morphology. So, sliding window filters are confirmed as typically performing less well.

Before discounting windowed filters, they were tested in conjunction with the cookie-cutter approach. $E6$ is a hybrid, where windowed median filters fill as much of the cut-out holes as their width permits, with the remainder filled using a fully tensioned spline. Even with additional complexity, this performs minimally better than the interpolations in $E1$ and $E2$. In short, extrapolation errors due to anomalous heights (e.g., trees) immediately outside drumlins' outlines are better mitigated by forcing a shorter path across the data gap by high spline tension or direct interpolation than attempting to statistically detect a regional trend in a variable landscape with a high spatial density of drumlins.

It is now necessary to verify that the observations above remain valid in conjunction with decluttering. Figs. 8c and d ($E7$ to $E9$) demonstrate that even simple methods to remove clutter improve ε_H dramatically for all variants on pure interpolation (spline and triangulation), but also removes the coherent height errors ($E9$ vs. $E4$ for $w \lesssim 100$ m). The interpolations perform a little better than the best sliding window filter, and are not dependant upon selection of a particular size scale, making them less subject to user-defined choices. As such, they are superior. Also note that decluttering degrades volume estimations (Table 1), consistent with the initial detailed comparison (Fig. 5b,d).

With the interpolations (i.e., spline and triangulation) identified as the two best alternatives, decluttering is objectively explored in Fig. 8e and f. The filter width used for simple decluttering is varied. Firstly, this confirms that overall, but by a small amount, the accuracy of triangulation exceeds a spline. The biggest difference is where triangulation better estimates volumes with no decluttering i.e., on NEXTMap's DSM. Note, however, that the difference is commonly within the random variability remaining in the experiment. The scatter of results from the 10 individual synthetic DEMs are shown (grey lines), illustrating this and justifying the use of multiple DEMs to stabilise results. Secondly, the results confirm that V is best estimated with minimal decluttering, whilst H needs it at approximately the value visually selected by *Hillier and Smith* [2012], also confirming the robustness of conclusions drawn from $E7$ to $E9$. The most accurate decluttering filter is not demonstrably different from the 60 m wide one proposed by *Hillier and Smith* [2012] and used in Fig. 8c and d. NEXTMap's DSM is the scenario with no decluttering, and NEXTMap's DTM is most closely matched by a 110 m wide median filter (Figs. 1 and 8).

4.3. Detailed results for the best method

From Fig. 8 the most robust way to quantify drumlins is using triangulation with decluttering when estimating H, but without for V: $E8$ and $E2$, respectively. For H, 60-100 m wide filters are of indistinguishable accuracy in this study area, so 60 m is taken as 'best' on the precautionary principle that smaller spatial filters cause least distortion and for ease of comparison with Fig. 8c. The purpose of Fig. 9 is to unpack the single ε values used to arrive at this conclusion, and to verify that minimising this quantity has achieved what was required in this study (Section 3.3.3).

Fig. 9 shows the associated recoveries for both individual drumlins and population metrics for the 'best' method. Errors for individual drumlins are still substantial but a comparison, for instance of Fig. 5b with Fig. 9c, demonstrates an improvement; the modal peak around correct recoveries is taller and more tightly constrained. Importantly, the modal peaks are centred on 1.0 so the tendency is to recover values correctly, large and small drumlins are extracted similarly, and errors are approximately symmetrical so that mean quantities are estimated acceptably. Systematic errors still exist in the estimated population parameters (i.e., mean H), but they are smaller and within the range of statistical variation, if only just (Fig. 9b,d). This is an achievement within a hilly, partially wooded area subject to significant anthropogenic influence.

5. Discussion

The results generated in this paper contribute directly to a discussion of how best to quantify drumlins, and potentially more generally glacial bedforms, as a precursor to using this information to understand the properties of flow in past ice sheets.

The numerical results show that the most robust way to quantify manually digitised drumlins in the presence of clutter is to remove data within the outline [*Smith et al.*, 2009], and then to use triangulation [e.g., *Spagnolo et al.*, 2011] to interpolate across the hole, decluttering the DSM [*Spagnolo et al.*, 2011; *Hillier and Smith*, 2012] when estimating H but not when estimating V. This is therefore the recommended general quantification protocol, distinct from the more specific parameters and techniques in the single best method.

It may seem counterintuitive that different approaches are needed for the connected properties H and V when the drumlin being measured and DEM remain the same. The explanation lies in systematic biases. Unlike random biases these are not alleviated at all by an increased number of observations in large datasets [e.g., *Clark et al.*, 2009]. For H the vast majority of clutter (e.g., trees) rise upwards from the surface of the solid Earth causing H to be systematically overestimated since only one object such as a tree is needed to distort the measurement. Also, trees' and drumlins' heights are of the same magnitude (e.g., 1b). So, for H, the need to remove clutter dominates, even if it is not completely removed (Fig. 9b). Volumes, however, are systematically underestimated after current, imperfect decluttering (Fig. 5d): Input $\bar{V}$ of 11.4×10^4 m^3 is recovered as $9.5 \pm 0.09 \times 10^4$ m$^3(2\sigma)$. This is because the volumes of clutter are typically substantially less than those of drumlins and height is pushed outside drumlins' outlines as decluttering smooths the topography, clearly seen where there is minimal visible clutter (Fig. 1).

In essence the cookie cutter approach succeeds because the information provided by manual digitisation is more powerful than simple statistical approaches using sliding window filters. A perfectly known outline more effectively prevents extraneous features contaminating the drumlin's basal surface. This highlights one of the assumptions of this analysis, that errors in mapped outlines are small. This is not necessarily the case (e.g., Fig. 1), but procedures are employed to maintain consistency and minimise bias [e.g., *Smith and Clark*, 2005; *Hughes et al.*, 2010].

Specifically regarding the best method, triangulation performs better than splines because it interpolates across the gap by the shortest paths, giving minimum weight to height anomalies immediately outside drumlins' outlines, to which interpolation is very sensitive (e.g., *Smith et al.* [2009] or Fig. 14 of *Hillier and Smith* [2012]). Untensioned splines follow the gradients immediately outside the boundary, are strongly influenced by anomalies such as trees, and so do not best estimate H and V. Tensioned splines mitigate this, preventing unrealistic deviations, and triangulation is in effect the limiting case of this tensioning. This said, the difference between triangulation and a fully-tensioned spline is small and only just distinguishable (Fig. 8e,f), so using either remains a valid option.

5.1. Scope of the guidelines

The results in this analysis are based on a single case study in Scotland. Are the guidelines formed from them generally applicable? Several lines of argument combine to suggest that they probably are. At least, in the absence of comparable studies on other areas or for other subglacial bedforms, they constitute a current best-assessment for subglacial bedforms in general.

Firstly, do they apply to DTM data sets created using NEXTMap's proprietary decluttering algorithm in the study area? This was assessed by applying the best method (Section 4.3) to real drumlins (i.e., Fig. 2). Values of H and V recovered from the real landscape (n = 178) for simple decluttering as in $E7$ to $E9$ and that of NEXTMap are similar, giving r^2 values of 0.76 and 0.97 respectively, both significant correlations ($p \ll 0.01$). Also, size histograms have closely matching forms, and standard metrics such as mean recovered volumes are close: $\bar{V_r} = 1.00 \pm 0.16 \times 10^5$ m$^3(1\sigma)$ for simple decluttering and $0.85 \pm 0.16 \times 10^5$ m$^3(1\sigma)$ for NEXTMap. This, insensitivity of results to filter widths in the range 50-100 m (Fig. 8e), and the ability of simple filters to closely replicate NEXTMap's DTM in a given locality (Fig. 1b), allow us to propose that the results of this study area are applicable to analyses based upon NEXTMap's multi-scale proprietary filter.

Secondly, is the case study area exactly representative, and if not does this alter the conclusions? This lowland area is considered to be a useful test site for its variety: the terrain ranges from hilly to flat, and non-glacial clutter varys in its width-scale and spatial density; drumlins are of two ages, differently affected by post-formational alteration, and are both topographically prominent and subtle; and although thinner (Fig. 6c) heights of the synthetic drumlins [*Smith et al.*, 2009; *Hillier and Smith*, 2012] closely resemble those of UK drumlins in general [*Clark et al.*, 2009; *Spagnolo et al.*, 2012] (Fig. 6b). Thus, it appears that both the guidelines and, more specifically, the best method may apply reasonably well to UK drumlins generally with the caveat that different optimal decluttering widths likely exist for focussed studies on different sub-regions.

Despite appearing representative, however, it is possible that the area is not so. For instance hillier, more challenging areas can be proposed (e.g., Wensleydale, UK [Fig. 11.15 of *Benn and Evans*, 2010]). The pertinent question is therefore whether or not this could alter the guidelines. Neither choices about the use of the cookie-cutter nor basal surface estimation method are senstive to decluttering (Section 4.2, Fig. 8a-d). So, in these respects it does not matter if clutter in the study area is exactly representative. With regard to the recommendations on decluttering, consider a locality with little clutter. Less severe decluttering measures could be used, which may distort the DTM less but a single tree would still produce a height overestimation, and currently available filters would cause some error in volume estimation. The same is true, but in the opposite sense, for areas overprinted by more clutter. So, the guidance given holds, up to the limiting case of no clutter or the design and verification of decluttering filters that cause minimal distortion. Hillier landscapes are more challenging for sliding window filters, but will least affect the semi-automated methods using interpolations as they rely only on heights at the manually identified outlines. So, the guidance also holds more generally in this respect.

Finally, note that the guidance applies to studies including large numbers of drumlins as well as detailed studies. The systematic biases, for instance in population means, do not reduce with large numbers of observations [e.g., *Spagnolo et al.*, 2012] like random errors do. The guidance mitigates but probably does not eliminate this issue. Fig. 9 illustrates that some level of systematic bias likely remains in most analyses of the 3D parameters of drumlins, even when an appropriate 3D quantification approach is used. Researchers should therefore remain aware of the possibility.

5.2. Future possibilities

The strength and weakness of the cookie-cutter type methods, found to be superior here, is their reliance on manually digitised outlines. Object, or vector, based approaches to automated landform delineation, suggested for drumlins by *Pike* [1995] may overcome this. *Irvin et al.* [1997] first attempted automated delineation, and uses of the multi-resolution segmentation algorithm of *Baatz and Schäpe* [2000] have been most successful [*Dragut and Eisank*, 2011; *Saha et al.*, 2011]. Through these, approaches based on other geomorphometric quantities such as curvature [e.g., *Rutzinger et al.*, 2012] or wavelets [*Kalbermatten et al.*, 2012; *Hillier*, 2008], or developments in related fields [e.g., *Wessel*, 1998; *Behn et al.*, 2004; *Hillier and Watts*, 2004], it may be possible to progress to fully automated techniques. An important caution is that even automated techniques ultimately rest on some level of subjective parameterisation, although methods are being developed to minimise this [e.g., *Anders and Seijmonsbergen*, 2011]. An alternative approach might be to manually map synthetic DEMs in order to investigate accuracies, recovery rates,

Experiment	Cut-out	Decluttering[1,2,4] Type	Decluttering Parameters	Basal surface estimation[2-4] Type	Basal surface estimation Parameters	Min. misfit ε_V (param.)[5]	ε_H (param.)
EB - Baseline	✔	✗		S	$t = 1$	0.263	0.863
No Decluttering							
*E*1	✔	✗		S	$t = 0.0 - 1.0$, $\Delta = 0.1$	0.243 (0.9)	0.825 (0.3)
*E*2	✔	✗		T		0.192	0.936
*E*3	✗	✗		Fm	$w = 20 - 500$, $\Delta = 80$	0.415 (420)	0.456 (20)
*E*4	✗	✗		Fb	$w = 20 - 500$, $\Delta = 80$	0.463 (340)	0.463 (20)
*E*5	✗	✗		Fl	$w = 20 - 500$, $\Delta = 80$	0.411 (20)	0.931 (20)
*E*6	✔	✗		Fm	$w = 20 - 500$, $\Delta = 80$	0.226 (20)	0.883 (20)
Decluttered							
*E*7	✔	Fm	$w = 60$	S	$t = 0.0 - 1.0$, $\Delta = 0.1$	0.279 (0.9)	0.329 (1.0)
*E*8	✔	Fm	$w = 60$	T		0.265	0.339
*E*9	✗	Fm	$w = 60$	Fm	$w = 20 - 500$, $\Delta = 80$	0.428 (420)	0.418 (260)
*E*10	✔	✗	No clutter in DEM	S	$t = 1.0$	0.102	0.082
Decluttering variants							
*E*11	✔	Fm	$w = 0 - 100$, $\Delta = 10$	S	$t = 1.0$	0.232 (20)	0.324 (80)
*E*12	✔	Fm	$w = 0 - 100$, $\Delta = 10$	T		0.182 (10)	0.310 (100)
Numerical error - $H_{drumlin}$ *only*							
*E*13	✔	n/a		S	$t = 1.0$	0.0019	0.0018

[1]Estimating upper surface of drumlin
[2]w = filter width in metres (m), and t is spline tension
[3]Interpolation methods: T = Triangulate; S = Spline
[4]Windowed filters: Fm = median; Fl = lowest; Fb = mean
[5]Value of parameter at which minimum misfit occurs is given in brackets

Table 1. Primary numerical experiments conducted (Fig. 8, Section 3.3.4).

distortions, and what is performed consistently. Perhaps this may even lead to agreement on the exact morphological definitions of particular bedforms. In general, numerous possibilities can be imagined for designing synthetic landscapes to assess methods for various glacial bedforms.

6. Conclusions

This work aims to provide practical recommendations for the mapping and quantification of drumlins; in particular, which semi-automated approach most robustly estimates heights and volumes, where 'robust' refers to insensitivity to outliers or biases. Using synthetic DEMs, it provides the first objective, reproducible, assessment of methods to optimise DEMs for the estimation of drumlins' heights and volumes.

A number of conclusions can be drawn from the analysis of semi-automated 3D quantification techniques applied to statistically representative synthetic landscapes:

1. Decluttering substantially affects measures of drumlin populations such as mean H and V for better and worse, respectively.

2. General guidelines to best quantify drumlins can be proposed. Specifically interpreters should i) declutter the DSM if estimating H but not V before ii) removing heights within the drumlin, then iii) interpolating to estimate a basal surface using Delaunay triangulation.

3. Researchers quantifying the 3D characteristics of drumlins should be aware of systematic biases, which will probably affect most analyses to some extent even when the best methods are used.

Whilst this study examines drumlins for a single study area in Scotland indications are that it is more widely applicable. At least, in the absence of studies on other areas or glacial bedforms, it constitutes a current best-assessment for glacial bedforms in general, albeit with the caveat that not all sources of error are accounted for here (e.g., mapping error). This analysis is also an example of the use of synthetic landscapes as a diagnostic tool in geomorphology for assessing otherwise intractable questions.

Acknowledgments. GMT [*Wessel and Smith*, 1998] was used for various calculations (e.g., `triangulate`, `surface`) and preparation of figures. The NEXTMap DEM data were supplied to MS. We thank Prof. S. Lane, Drs J. Hiemstra and I. Evans, and an anonymous reviewer for their comments that helped focus and improve this manuscript.

References

Anders, N. S., and A. C. Seijmonsbergen, Segmentation optimization and stratified object-based analysis for semi-automated geomorphological mapping, *Remote Sensing of Environment*, *115*(12), 2976–2985, 2011.

Baatz, M., and A. Schäpe, Multiresolution segmentation-an optimization approach for high quality multi-scale image segmentation, in *Angewandte Geographische Informationsverarbeitung*, edited by J. Strobl, T. Blaschke, and G. Griesebner, pp. 12–23, Wichmann-Verlag, Heidelberg, 2000.

Bartels, M., and W. Hong, Threshold-free object and ground point separation in LIDAR data, *Pattern Recognition Letters*, *31*(10), 1089–1099, doi: doi:10.1016/j.patrec.2010.03.007, 2010.

Behn, M. D., J. M. Sinton, and R. S. Deitrick, Effect of the Galapagos hotspot on seamount volcanism along the Galapagos Spreading Center, *Earth and Planetary Science Letters*, *217*, 331–347, 2004.

Benn, D. I., and D. J. A. Evans, *Glaciers and Glaciation*, 2nd ed., 802 pp., Hodder, Oxford, UK, 2010.

Boulton, G., The origin of glacially fluted surfaces - observations and theory, *J. Glaciology*, *17*(76), 287–309, 1976.

Box, G. E. P., Non-normality and Tests on Variances, *Biometrika*, *40*, 318–335, 1953.

Boyce, J., and N. Eyles, Drumlins carved by deforming till streams below the Laurentide ice sheet, *Geology*, *19*(8), 787–790, 1991.

Bradwell, R., M. S. Stoker, N. R. Golledge, C. Wilson, J. Meritt, D. Long, J. D. Everest, O. B. Hestvik, A. Stevenson, A. L. Hubbard, A. G. Finlayson, and H. E. Mather, The northern sector of the last British ice sheet: maximum extent and demise., *Earth Science Reviews*, *88*, 207–226, 2008.

Cazenave, A., K. Dominh, C. J. Allègre, and J. G. Marsh, Global Relationship Between Oceanic Geoid and Topography, *Journal of Geophysical Research*, *91*, 1986.

Chapwanya, M., C. D. Clark, and A. C. Fowler, Numerical computations of a theoretical model of ribbed moraine formation, *Earth Surf. Proc. Land.*, *36*, 1105–1112, 2011.

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

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

Clark, C. D., A. Hughes, S. L. Greenwood, M. Spagnolo, and F. S. Ng, Size and shape characteristics of drumlins, derived from a large sample, and associated scaling laws, *Quat. Sci. Rev.*, *28*(7-8), 677–692, doi:10.1016/j.quascirev.2008.08.035, 2009.

Cobby, D. M., D. C. Mason, and I. J. Davenport, Image processing of airborne scanning laser altimetery data for improved river flood modelling, *ISPRS Journal of Photogrammetry & Remote Sensing*, *56*, 121–138, 2001.

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

Crosby, A. G., D. McKenzie, and J. G. Sclater, The Relationship Between Depth, Age and Gravity in the Oceans, *Geophysical Journal International*, *166*, 553–573, 2006.

Dragut, L., and C. Eisank, Object representations at multiple scales from digital elevation models, *Geomorphology*, *129*, 183–189, 2011.

Dunlop, P., C. D. Clark, and R. C. A. Hindmarsh, Bed Ribbing Instability Explanation: Testing a numerical model of ribbed moraine formation arising from coupled flow of ice and subglacial sediment, *J. Geophys. Res.*, *113*(F3), doi:10.1029/2007JF000954, 2008.

Evans, I. S., A new approach to drumlin morphometry, in *Drumlin Symposium*, pp. 119–130, Balkema, Rotterdam, 1987.

Evans, I. S., Geomorphometry and landform mapping: What is a landform?, *Geomorphology*, *137*, 94–106, doi:10.1016/j.geomorph.2010.09.029, 2012.

Finlayson, A., Digitial surface models are not always representative of former glacier beds: Palaeoglaciological and geormophological implications, *Geomorphology*, *194*, 25–33, 2013.

Fisher, P., and N. Tate, Causes and consequences of error in digital elevation models, *Progress in Physical Geography*, *30*(4), 467—-489, 2006.

Fowler, A. C., An instability mechanism for drumlin formation, in *Deformation of Glacial Materials*, edited by A. J. Maltman, B. Hubbard, and M. J. Hambrey, geological ed., pp. 307–319, Geol. Soc. Publishing House, London, 2000.

Gluckert, G., *Two large drumlin fields in central Finland*, 37 pp., Societas Geographica Fenniae, Helsinki, 1973.

Goutorbe, B., Combining seismically derived temperature with heat flow and bathymetry to constrain the thermal structure of oceanic lithosphere, *Earth and Planetary Science Letters*, *295*, 390–400, 2010.

Grohmann, C. H., and A. O. Sawakuchi, Influence of cell size on volume calculation using digital terrain models: A case of coastal dune fields, *Geomorphology*, *180-181*, 130–136, doi:10.1016/j.geomorph.2012.09.012, 2013.

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

Hättestrand, C., and J. Kleman, Ribbed moraine formation, *Quat. Sci. Rev.*, *18*(1), 43–61, 1999.

Hättestrand, C., S. Gotz, J. O. Naslund, D. Fabel, and A. P. Stroeven, Drumlin formation time: Evidence from northern and central Sweden, *Geografiska Annaler Series A-Physical Geography*, *86A*(2), 155–167, 2004.

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

Hillier, J. K., Submarine Geomorphology : Quantitative Methods Illustrated with the Hawaiian Volcanoes, *Geomorphological Mapping: Methods and Applications*, *15*, 357–374, doi:10.1016/B978-0-444-53446-0.00012-4, 2011.

Hillier, J. K., and M. Smith, Residual relief separation: digital elevation model enhancement for geomorphological mapping, *Earth Surface Processes and Landforms*, *33*(14), 2266–2276, doi:10.1002/esp, 2008.

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

Hillier, J. K., and A. B. Watts, "Plate-like" subsidence of the East Pacific Rise - South Pacific Superswell system, *Journal of Geophysical Research*, *109*(B10102), 2004.

Hillier, J. K., and A. B. Watts, Relationship between depth and age in the North Pacific Ocean, *Journal of Geophysical Research*, *110*(B2), 1–22, doi:10.1029/2004JB003406, 2005.

Hillier, J. K., J. M. Bunbury, and A. Graham, Monuments on a migrating Nile, *Journal of Archaeological Science*, *34*(7), 1011–1015, doi:10.1016/j.jas.2006.09.011, 2007.

Hillier, J. K., M. J. Smith, C. D. Clark, C. R. Stokes, and M. Spagnolo, Subglacial bedforms reveal an exponential size-frequency distribution, *Geomorphology*, doi: 10.1016/j.geomorph.2013.02.017, 2013.

Hindmarsh, R. C. A., Drumlinization and drumlin-forming instabilities: viscous till mechanisms, *J. Glaciology*, *44*(147), 293–314, 1998.

Hooke, R., and A. Medford, Are drumlins a product of thermo-mechanical instability?, *Quaternary Res.*, doi:10.1016/j.yqres.2012.12.002, 2013.

Hubbard, A. L., T. Bradwell, N. Golledge, A. Hall, H. Patton, D. Sugden, R. Cooper, and M. S. Stoker, Dynamic cycles, ice streams and their impact on the extent, chronology and deglaciation of the British-Irish ice sheet, *Quaternary Science Reviews*, *28*(7-8), 758–776, doi:10.1016/j.quascirev.2008.12.026, 2009.

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

Intermap, Intermap product handbook and quickstart guide (v3.3), *Tech. rep.*, 2004.

Irvin, B., S. Ventura, and B. Slater, Fuzzy and isodata classification of landform elements from digital terrain data in Pleasant Valley, Wisconsin, *Geoderma*, *77*, 137–154, 1997.

Kalbermatten, M., D. van der Ville, P. Turberg, D. Tuia, and S. Joost, Multiscale analysis of geomorphological and geological features in high resolution digital elevation models using the wavelet transform, *Geomorphology*, *138*(1), 352–363, doi: 10.1016/j.geomorph.2011.09.023, 2012.

Kim, S., and P. Wessel, Directional median filtering for the regional-residual separation of bathymetry, *G3*, *9*(Q03005), 1–11, doi:10.1029/2007GC001850, 2008.

Korkalainen, T., A. Lauren, and T. Kokkonen, A GIS-based analysis of catchment properties within a drumlin field, *Boreal Environmental Research*, *12*, 489–500, 2007.

Levitt, D. A., and D. T. Sandwell, Modal Depth Anomalies from Multibeam Bathymetry: Is There a South Pacific Superswell?, *Earth and Planetary Science Letters*, *139*, 1–16, 1996.

Livingstone, S., C. Ó Cofaigh, and D. J. A. Evans, Glacial geomorphology of the central sector of the last British-Irish Ice Sheet, *Journal of Maps*, pp. 358–377, doi: 10.4113/jom.2008.1032, 2008.

MacLachlan, J. C., and C. Eyles, Quantitative geomorphological analysis of drumlins in the Peterborough drumlin field, Ontario, Canada, *Geografiska Annaler: Series A, Physical Geography*, *95*(2), 125–144, 2013.

McKenzie, D. P., A. B. Watts, B. Parsons, and M. Roufosse, Planform of Mantle Convection Beneath the Pacific Ocean, *Nature*, *288*, 442–446, 1980.

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

Milledge, D. G., S. Lane, and J. Warburton, The potential of digital filtering of generic topographic data for geomorphological research, *Earth Surface Processes and Landforms*, *34*(1), 63–74, doi:10.1002/esp.1691, 2009.

Nettleton, L. L., Regionals, Residuals, and Structures, *Geophysics*, *19*(1), 1–22, doi:10.1190/1.1437966, 1954.

Ó Cofaigh, C., C. Pudsey, J. A. Dowdeswell, and P. Morris, Evolution of subglacial bedforms along a paleo-ice stream, Antarctic Peninsula continental shelf, *Geophys. Res. Lett.*, *29*(8), doi:10.1029/2001GL014488, 2002.

Ó Cofaigh, C., J. A. Dowdeswell, C. Allen, J. F. Hiemstra, C. Pudsey, and J. Evans, Flow dynamics and till genesis associated with a marine-based Antarctic palaeo-ice stream, *Quaternary Science Reviews*, *24*(5-6), 709–740, doi:10.1016/j.quascirev.2004.10.006, 2005.

Phillips, E., J. D. Everest, and D. Diaz-Doce, Bedrock controls on subglacial landform distribution and geomorphological processes: Evidence from the Late Devensian Irish Sea Ice Stream, *Sedimentary Geology*, *232*(3-4), 98–118, doi:10.1016/j.sedgeo.2009.11.004, 2010.

Pike, R. J., Geomorphometry: progress, practice and prospect, *Zeitschrift fur Geomorphologie Suppl Bind*, *101*, 221–238, 1995.

Prest, V. K., and D. R. Grant, The Glacial Map of Canada, 1968.

Rose, J., Drumlins as part of a glacier bedform continuum, in *Drumlin Symposium*, edited by J. Menzies and J. Rose, pp. 103–116, Balkema, Rotterdam, 1987.

Rose, J., Glacier sediement patterns and stress transfer associated with the formation of superimposed flutes, *Sedimentary Geology*, *62*, 151–176, 1989.

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

Rose, J., and M. J. Smith, Glacial geomorphological maps of the Glasgow region, western central Scotland, *Journal of Maps*, *2008*, doi:doi:10.4113/jom.2008.1040, 2008.

Rutzinger, M., B. Hofle, and K. Kringer, Accuracy of automatically extracted geomorphological breaklines from airborne LiDAR curvature images, *Geografiska Annaler Series A-Physical Geography*, *94A*, 33–42, doi:10.1111/j.1468-0459.2012.00453.x, 2012.

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

Shaw, J., D. Kvill, and B. Rains, Drumlins and catastrophic subglacial floods, *Sedimentary Geology*, *62*(2), 177–202, 1989.

Shewchuk, J. R., Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator, in *First Workshop on Applied Computational Geometry (Philadelphia, PA)*, pp. 124–133, ACM, 1996.

Sissons, J. B., *The Evolution of Scotland's Scenery*, 259 pp., Oliver and Boyd, Edinburgh, 1967.

Sithole, G., and G. Vosselman, Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds, *ISPRS Journal of Photogrammetry & Remote Sensing*, *59*, 85—-101, 2004.

Smith, M. J., and C. D. Clark, Methods for the visualization of digital elevation models for landform mapping, *Earth Surface Processes and Landforms*, *30*(7), 885–900, doi: 10.1002/esp.1210, 2005.

Smith, M. J., and S. M. Wise, 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, 2007.

Smith, M. J., J. Rose, and S. Booth, 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, 2006.

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

Smith, W. H. F., Marine Geophysical Studies of Seamounts in the Pacific Ocean Basin, Ph.D. thesis, Columbia Univ., 1990.

Smith, W. H. F., and P. Wessel, Gridding With Continuous Curvature Splines in Tension, *Geophysics*, *55*, 293–305, 1990.

Spagnolo, M., C. D. Clark, P. Dunlop, and A. Hughes, The topography of drumlins; assessing their long profile shape, *Earth Surface Processes and Landforms*, *36*, 790–804, doi:10.1002/esp.2107, 2011.

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

Stein, C. A., and S. Stein, A Model for the Global Variations in Oceanic Depth and Heat Flow With Lithospheric Age, *Nature*, *359*, 123–129, 1992.

Stokes, C. R., and C. D. Clark, Are long subglacial bedforms indicative of fast ice flow?, *Boreas*, *31*(3), 239–249, 2002.

Stokes, C. R., M. Spagnolo, and C. D. Clark, The composition and internal structure of drumlins: complexity, commonality, and implications of a unifying theory of their formation, *Earth Sci. Rev.*, *107*(3-4), 398–422, doi:10.1016/j.earscirev.2011.05.001, 2011.

van der Mark, C. F., A. Blom, and S. J. M. H. Hulscher, Quantification of variability in bedform geometry, *J. Geophys. Res.*, *113*, F03,020, doi:10.1029/2007JF000940, 2008.

Wang, Y., B. Mercer, V. C. Tao, J. Sharma, and S. Crawford, Automatic generation of bald earth digital elevation models from digital surface models created using airborne IfSAR, in *CD-ROM Proceedings of the ASPRS Conference, April 23-27, St Louis, Missouri, USA*, 2001.

Watson, D. F., Acord: Automatic contouring of raw data, *Computers and Geosciences*, *8*, 97–101, 1982.

Watts, A. B., Gravity and Bathymetry in the Central Pacific Ocean, *Journal of Geophysical Research*, pp. 1533–1548, 1976.

Watts, A. B., and S. F. Daly, Long Wavelength Gravity and Topography Anomalies, *Annual Review of Earth and Planetary Sciences*, *9*, 415–448, 1981.

Wessel, P., An Empirical Method for Optimal Robust Regional-Residual Separation of Geophysical Data, *Mathematical Geology*, *30*, 391–408, 1998.

Wessel, P., and W. H. F. Smith, New, improved version of Generic Mapping Tools released, *Eos Transactions of the American Geophysical Union*, *79*, 579, 1998.

Wren, E. A., Trend surface analysis - a review, *Canadian Journal of Exploration Geophysics*, *9*, 39–45, 1973.

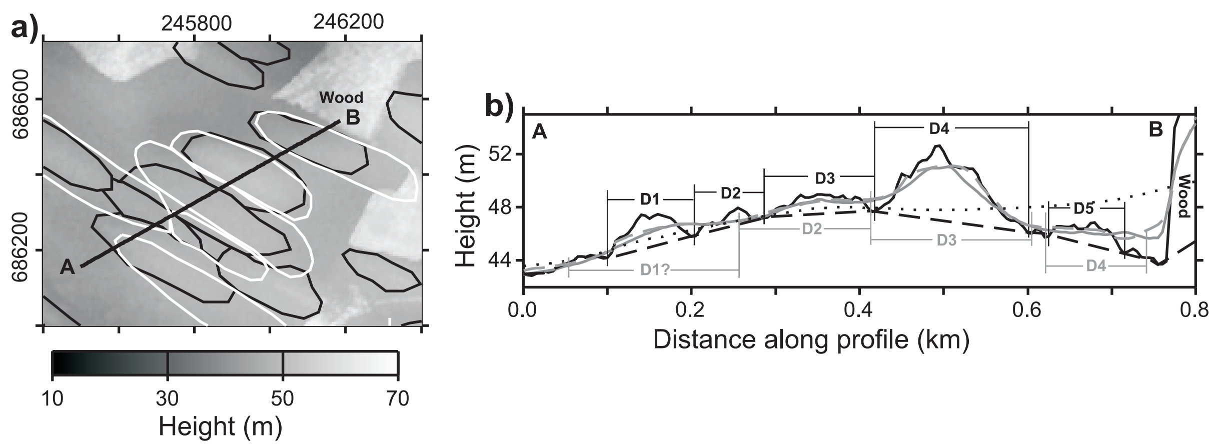

Figure 1. Effect of decluttering. a) Plan view of a sub-region of the study area locating the profile in b) (thick black line **A**-**B**). Underlying DEM is the DSM of NEXTMap. Black lines are drumlin outlines as mapped by *Smith et al.* [2006] (Fig. 2). White lines are outlines mapped by *Clark et al.* [2009], digitised from their Fig. 4, but only displayed where they are immediately proximal to the profile. b) Profile across the drumlins. Solid black line is the DSM, underlain by a manual determination of a basal surface based upon it (black dashed line). NEXTMap's DTM (grey line) is similar to the output of a best fitting (see text) 110 m wide sliding median filter (dashed grey line). Application of a 500 m wide median filter to the DSM is shown (dotted line) to illustrate errors that may occur for sliding window filters (e.g., in $E3$). Interpretations of drumlin locations from the profiles are denoted in the form 'D1' and are shaded black or grey to match the relevant elevation data.

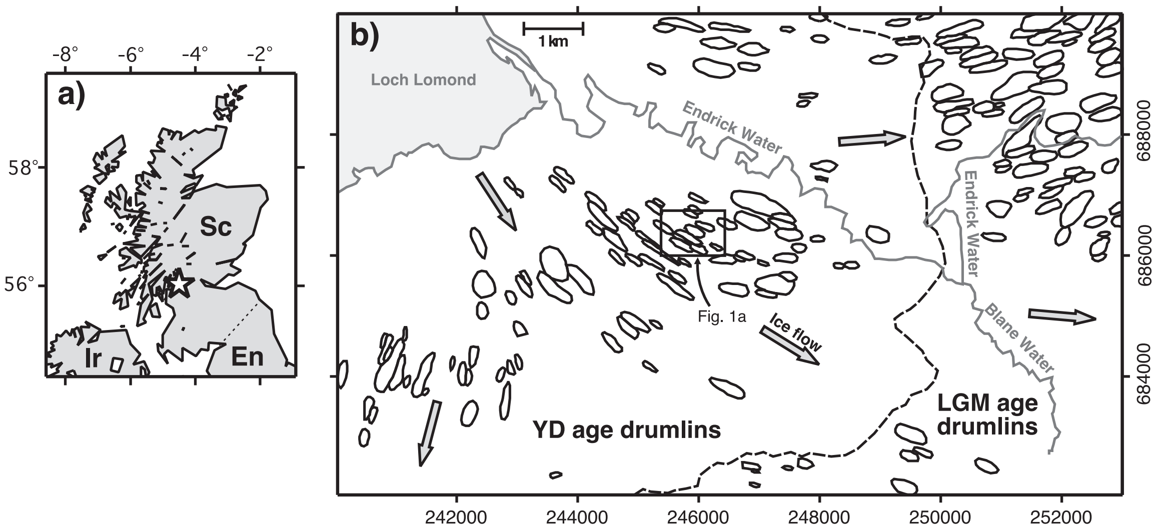

Figure 2. Location maps: a) The study area is located at the white star (4°28' W, 56°02' N). Countries are: England (En), Scotland (Sc), Ireland (Ir). Coastlines of both seas and major inland water bodies are shown. b) Study area, with main geomorphic features of interest highlighted; drumlins (black outlines), rivers (grey), water (grey shade). Dashed line separates drumlins of Younger Dryas (YD) and Last Glacial Maximum (LGM) ages [*Smith et al.*, 2006] to its west and east, respectively. Arrows indicate approximate ice flow trends in the LGM [*Sissons*, 1967] and YD [*Rose*, 1987]. Box indicates the extent of Fig. 1. Map coordinates are of the British National Grid.

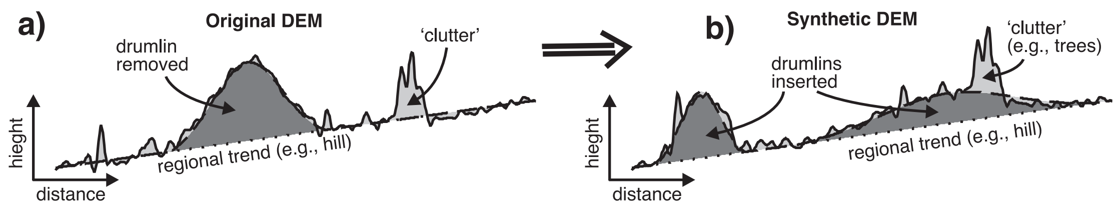

Figure 3. Idealised profiles to illustrate the process used to create synthetic DEMs [*Hillier and Smith*, 2012]. Glacial landscapes can contain three 'components': drumlins (dark grey shade), a large-scale regional slope, and non-glacial 'clutter' (light grey shade). a) Upper and lower surfaces of a drumlin are estimated to define it; dotted and dashed lines. It is then removed (height subtracted) from the measured DEM; solid line. After this, in b), Gaussian shaped drumlins, arbitrarily two in this example, are inserted (height added) to create the synthetic DEM. Critically, idealised drumlins are located randomly with respect to the causes of measurement error, noise and regional trends.

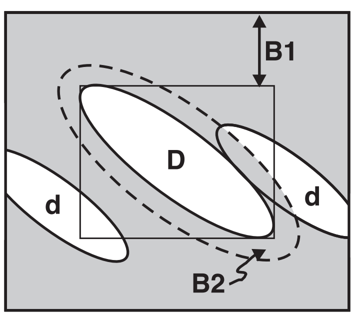

Figure 4. Spatial distribution of heights in DEM (grey) retained for basal surface estimation of a drumlin, **D**. Applies for all experiments, except those where all data were retained ($E3$ to $E5$). **d** are other nearby drumlins, whose exclusion or otherwise is significant for windowed filters ($E3$ to $E6$, $E9$). **B1** is buffer $\geq$ **B2**, extended to half the width of any windowed filter used to avoid edge effects. **B2** is a buffer to ensure that data completely encircles **D**, set at 20 m after *Smith et al.* [2009]. Further explanation of experiments in Section 3.3.4.

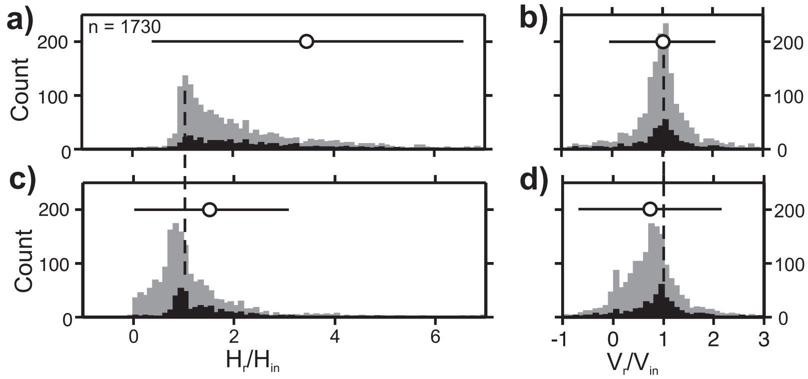

Figure 5. Recoveries of individual H and V. a) and b) method of *Smith et al.* [2009], i.e. without decluttering (EB). c) and d) equivalent method including simple decluttering [$E7$, $t = 1.0$]. Panels are histograms of recovered values, H_r or V_r, binned as a fraction of known values within the synthetic DEMs, H_{in} or V_{in}. Dashed lines indicate correct recovery; i.e., $V_r/V_{in} = 1$ at $V_r = V_{in}$. Circle is mean ratio, with bar of $\pm 2.96 MAD$ (95% of data) used to estimate $\pm 2\sigma$ as some extreme outliers exist. Grey bars are for all drumlins, and black bars are for only large ($L >$500 m) drumlins.

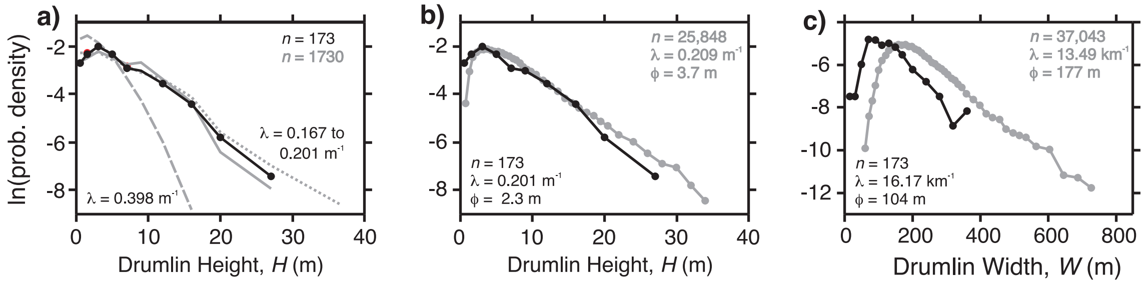

Figure 6. Empirical probability density functions for drumlin size, displayed on semi-log plots. a) H for synthetic drumlins as input (black) and recovered (grey), minimising ε for sliding median filters (solid grey, $E9$, $w = 260$) and interpolation using triangulation (dotted grey, $E8$), and selecting a mean V_r/V_{in} near 1.0 for median filters (dashed grey, $E9$, $w = 100$). The latter doubles the rate at which the prevalence of drumlins drops off with increasing size i.e., λ goes from ~ 0.2 to ~ 0.4. b) and c) Comparison between H and W for the study site (black) and the UK (grey) [*Clark et al.*, 2009; *Spagnolo et al.*, 2012], respectively. Dots are binned data; as input data in a) and so are not shown in there. Number of underlying data are indicated, in shades matching curves, on individual panels. See *Hillier et al.* [2013] for justification of, and calculation method for, semi-log plots and parameters λ (exponent) and ϕ (mode).

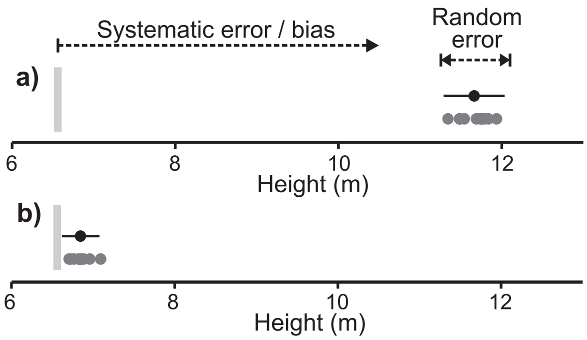

Figure 7. Reliability of recovering population parameter mean height, $\bar{H}_r$, for n = 173; a) without decluttering and b) with 'simple decluttering', 60 m wide median filter. Light grey bar is input height, compared to recoveries from the 10 synthetic DEMs (grey dots) whose mean and range ($\pm 2\sigma$) is displayed by the black dot and bar. Results from EB and $E7$, with $t = 1$ (Fig 8, Table 1).

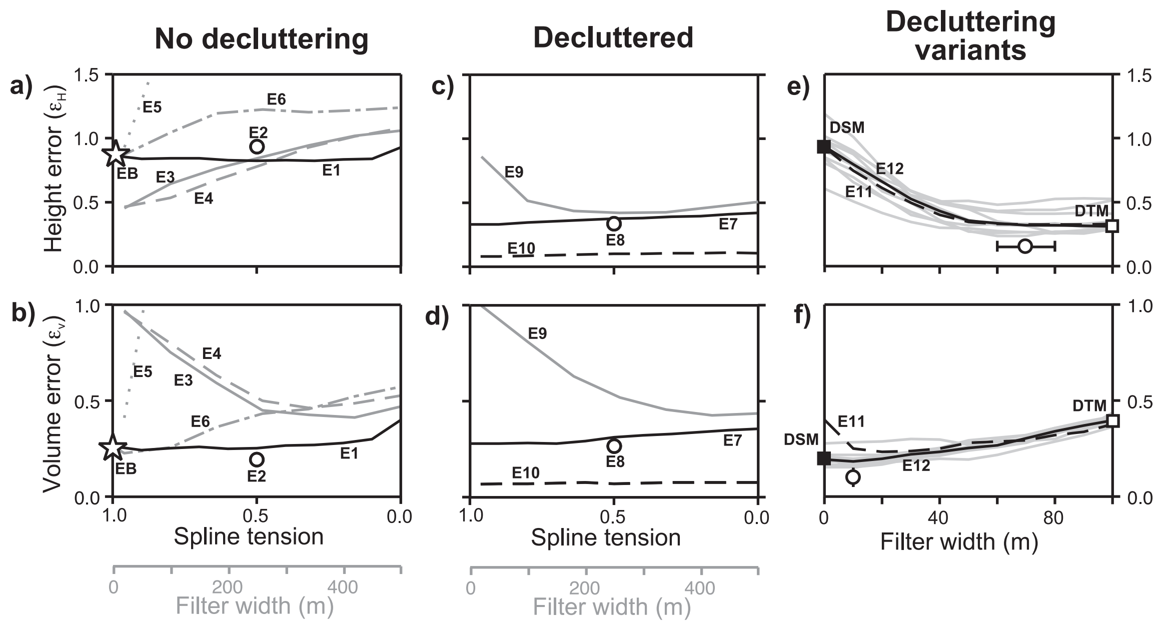

Figure 8. Results of error analysis from numerical experiments (Table 1). a) ε_H for height and b) ε_V for volume when a DSM is used without decluttering. Baseline for comparison is the method of *Smith et al.* [2009], *EB* (star). Grey lines are the errors for size estimates using sliding window filters with variable window widths (grey x-axis scale). Without cookie-cutter approach: $E3$ median (solid grey); $E4$ mean (dashed grey); $E5$ lowest (dotted grey). With cookie-cutter: $E6$ median (dot-dash grey). Black lines are scale-independent filters (black x-axis scale): bi-cubic spline is the solid line ($E1$); Delaunay triangulation is the white dot ($E2$) arbitrarily placed at $t = 0.5$. c) and d) are as a) and b), but decluttered with a 60 m median filter. e) and f) further investigate decluttering for more accurate techniques. Filter width is window size for decluttering using in conjunction with either a bi-cubic spline (black dashed line; $E11$) or triangulation (black line; $E12$). Grey lines are from each of the ten individual DEMs for $E12$, and white circles are median (±1 MAD) of these. Squares are for NEXTMap's DSM, and most analogous filter to its DTM (Fig. 1), respectively.

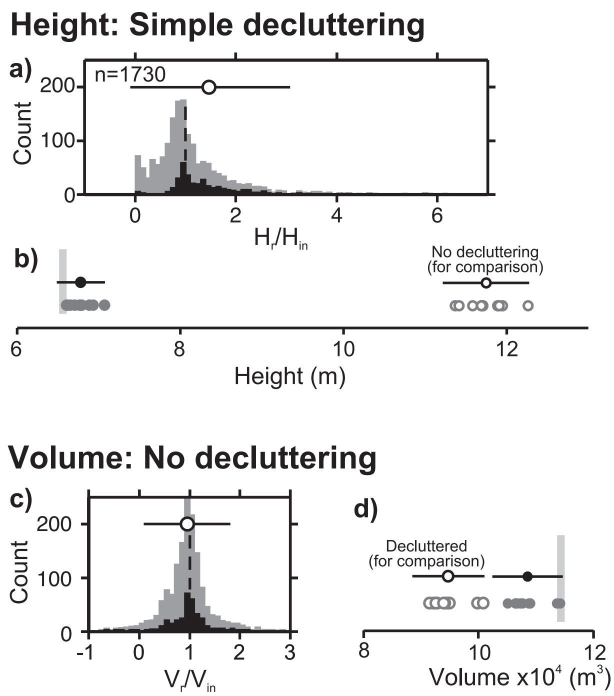

Figure 9. Recoveries of H and V for individual drumlins using the 'best' practical quantification method, $E8$ and $E2$ respectively: a cookie-cutter type approach, using triangulation, with and without decluttering using a 60 m wide median filter respectively. a) and c) are recoveries of individuals, with details as Fig. 5. b) and d) are recoveries of mean H and V respectively, with details as Fig. 7.