GRANULAR PROPERTIES FROM DIGITAL IMAGES OF SEDIMENT: IMPLICATIONS FOR COASTAL SEDIMENT TRANSPORT MODELLING

Daniel Buscombe[1,3], Gerhard Masselink[1], and David M. Rubin[2]

Statistical techniques designed to obtain granular attributes from digital images of sediment are reviewed and evaluated. Such techniques promise to revolutionize field and laboratory studies into sediment transport and sedimentation dynamics, because of the non-intrusive, rapid and labour-saving nature of the method compared to traditional forms of particle size analysis. Methodological work undertaken thus far, in this relatively new field, is synthesized, as well as some ideas given on potential future developments and refinements. Finally, the potential utility of such methodological advances to the coastal science and engineering community is discussed. This includes an evaluation of the premises behind, and potential limitations of, 'digital grain-size analysis', as well as instances where high-resolution grain-size information could potentially inform a new generation of coastal sediment transport models, where granular attributes are included as free parameters rather than fixed at the boundary.

INTRODUCTION

Models for coastal sediment transport and morphological change/sedimentation require information on granular attributes such as mean/median size and sorting. Models for sediment transport can be highly sensitive to grain-size, so estimates need to be robust. This often means several samples taken from the field, and often repeatedly-so. Sophisticated models, especially for coarse or heterogeneous sedimentary environments, may need to account for the feedbacks induced by changing sediment beds to flows, and may need information on the entire grain-size distribution, shape, orientation of the sediments. Such information will come from field studies which document sedimentary change, and subsequently models which attempt to explain the physics of changes in granular characteristics from first principles.

However, there are several problems associated with field sediment sampling and subsequent grain-size analysis using traditional methods such as sieving and settling. These methods are labour-intensive and slow, since samples must be manually removed from the depositional environment, and brought back to the laboratory for detailed analyses. Not only is this laborious, making obtaining a grain-size distribution a lengthy process, by removing the sample

[1] School of Geography, University of Plymouth, 8 Kirkby Place, Plymouth, Devon, PL4 8AA, UK. g.masselink@plymouth.ac.uk
[2] United States Geological Survey Pacific Science Center, 400 Natural Bridges Dr., Santa Cruz, CA 95060, USA. drubin@usgs.gov
[3] present address: United States Geological Survey Pacific Science Center, 400 Natural Bridges Dr., Santa Cruz, CA 95060, USA. dbuscombe@usgs.gov

from the environment any information on the spatial configuration of those grains on the bed is lost, and by removing material in some sensitive systems, it could potentially alter subsequent system development. In addition, these techniques are not equivalent, because they measure slightly different things, and thus the computed grain-size distributions and sample statistics differ. The measurement techniques cannot straddle the size fractions from mud to cobbles. Settling and laser-diffraction, for example, are useful for mud, silts and sands, but not gravels or cobbles.

A single methodology is thus required which can estimate grain-size distributions remotely (i.e. non-intrusively), rapidly, and can be applicable across the grain-size fraction boundaries. Obtaining grain-size measures from digital images of sediment beds has thus long been an attractive option (e.g. Ibekken and Schleyer, 1986; Butler et al., 2001) since obtaining a grain-size distribution from an image should be less laborious and time-consuming than traditional methods. That is, if the process can be automated. Unfortunately, early attempts at obtaining grain-size information from photographic images were not automated, and thus were not an attractive alternative to traditional means. In addition, they were limited in scope to gravels and cobbles. The modern generation of image analysis methods are automated, thus fast and efficient. These algorithms (e.g. Butler et al., 2001; Sime and Ferguson, 2003; Graham et al., 2005) use sophisticated image segmentation and thresholding routines to give a robust estimate of grain-size distributions for coarse gravels by segmenting each individual grain, and return the axial and areal properties of each. However, problems remain with these methods, such as for use on sediment beds composed of grains difficult to segment. For example, a range of colours/mineralogies, or small gravel or sand, or where pebbles have inter-granular aberrations such as pock-marks, abrasion hollows and scratches.

In recent years, advances have been made in the automated and accurate quantification of granular attributes from the statistical analysis of digital images of seabed sediments. Rubin (2004) proposed a new method for estimation of grain-size based on the notion that the spatial autocorrelation of an image is sensitive to the relative size of objects in that image (*cf.* Preston and Davis, 1976; Lin, 1982). In other words, images of relatively large sediment have more neighbouring pixels with similar intensity values for a greater distance compared with images of smaller grains. Rubin (2004) suggested that, given careful calibration consisting of compiling correlograms for images containing known sediment sizes, reliable estimates of mean grain-size could be found by comparing the correlogram from a sample image with the calibration catalogue, through a simple least-squares problem.

Such digital photographic methods can now be used to measure grain-size information at a resolution comparable with measurements of hydraulic, hydrodynamic and morphological/topographical conditions, and without disrupting the sediment body by direct sampling. Advances in this area (called 'digital grain-size') are synthesized and evaluated here, followed by some ideas for further development in the future. Finally, a brief discussion of the potential implications of such rapid grain-size estimates may have on field and laboratory studies into coastal processes, and the development of coastal sediment transport models.

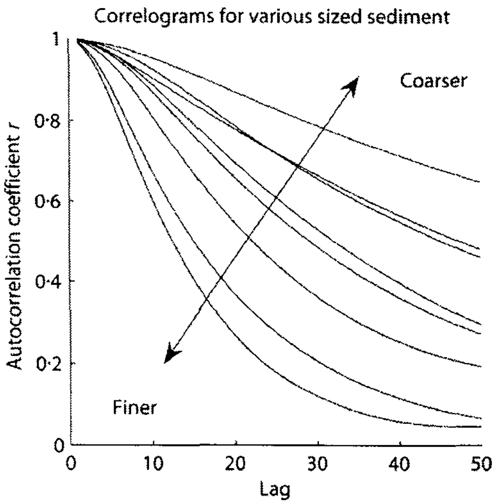

Figure 1. Images of finer grains have steeper correlograms than images of more coarse grains (from Buscombe and Masselink, 2008).

MEAN GRAIN-SIZE FROM DIGITAL IMAGES OF SEDIMENT

Following Rubin (2004), a sample image is subjected to a numerical technique which is sensitive to the statistical distribution of grain-sizes within that image, generating a numeric array which is the signature of the size information obtained within the image. Simple least-squares analysis is used to find the proportion of variance, at each position in the sample array, best explained by a calibration catalogue of several such numerical signatures associated with sediments of known grain-size. In other words, the analysis finds the numerical array which provides the collective 'best fit', or minimised residual, between the sample and the calibration catalogue. The array is indexed by the grain-size associated with each element within the calibration catalogue. The reader is referred to Rubin (2004), Rubin et al., (2007) and Buscombe and Masselink (2008) for a more detailed discussion on the principles involved, as well as methodological guidelines for the collection and analysis of calibration and sample images. There are a number of numerical routines which may be employed (Buscombe and Masselink, 2008), all of which have mathematical derivation from the most commonly applied to date, namely the autocorrelation approach of Rubin (2004).

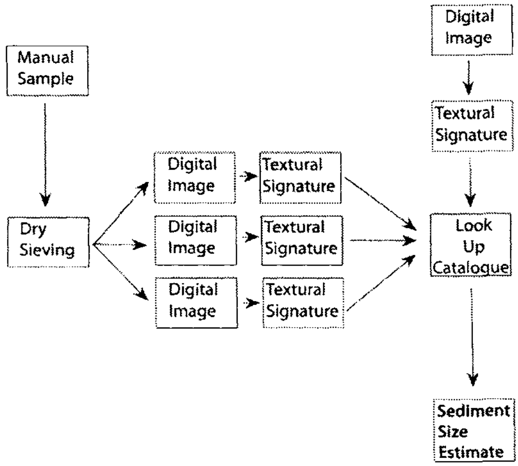

Figure 2. Schematic diagram of the stages involved in a 'digital grain-size' methodology (from Buscombe and Masselink, 2008).

The autocorrelation function, used as a measure of two-dimensional spatial independence, could be sensitive to the size of grains within images of sand, and thus, given careful calibration, could be used to derive a rapid, yet accurate, measure of sediment size. Positive spatial autocorrelation is the tendency for objects closer together to be more similar than objects further apart. For images of natural beds, pixels patches covering larger grains are more similar for a longer distance than pixel patches covering smaller grains (Fig. 1). Please refer to Rubin et al., 2007; Barnard et al., 2007; and Buscombe and Masselink (2008) for more details/guidelines on the general design (Fig. 2) of a project using digital grain-size techniques; both in terms of field/laboratory acquisition of images, and the optimum design of the calibration catalogue.

The general linear equations for the problem of solving for the proportions of calibrated sizes that collectively give the best fit to a given sample's numerical signature are given by $Ax = d$, where matrix A is the calibration catalogue composed of autocorrelation profiles to m lags, from n grain-size fractions; where array x is the grain-size solution; and where array d is the autocorrelation profile of the sample image (also to m lags). The vector solution x is one that minimises the sum of squared errors between A and d, or $(d-Ax)^T(d-Ax)$ where T denotes matrix transpose. Rubin (2004) showed that a matrix division of d into A, satisfying $\mathbf{min}(\|Ax-d\|)$ where $\|..\|$ denotes matrix norm (see Buscombe (2008)), does not satisfactorily resolve the grain-size solution because it may take on positive and negative values of grain-size, which is physically impossible. Alternatively, the least-squares result is commonly found by interpolation of d within A indexed by n, producing m values. The mean of the resulting vector has been shown to be a good estimate of mean grain-size (Rubin et al., 2007; Barnard et al., 2007; Buscombe and Masselink, 2008). Based on over 180 samples, for example, Buscombe and Masselink (2008) found mean grain-size of sieved and imaged beach gravels correspond to within between 84 and 92%. Barnard et al. (2007) found even better accuracy (~96%) on their method with over 200 Pacific beach sand samples.

There is a growing body of work which has begun to utilise these methods in experimental studies in fluvial and littoral settings, in both sandy (Gallagher et al., 2006; Rubin et al., 2006; Barnard et al., 2007; Mustain et al., 2007), and gravelly (Ruggiero et al., 2007; Warrick et al., 2007; Buscombe and Masselink, 2008; Austin and Buscombe, 2008) environments. Computations are so rapid that an estimate of grain-size from a sample image can be made in 'real-time' in the field using a hand-held computer and digital SLR camera (Rubin and Chezar, 2007).

RECENT ADVANCES IN DIGITAL GRAIN-SIZE

Grain-Size 'Distribution'

To broaden the applicability of these statistical automated techniques, they must accurately reproduce the entire grain-size distribution, or at least percentiles of that distribution, in addition to accurately quantifying mean/median sediment size. Only in doing so will physically meaningful and testable parameters be derived, which will drive progress in developing/testing models that incorporate time-updateable grain-size parameters in models for coastal sediment transport.

Whilst the interpolation method described above is reliable for ensemble statistics such as mean size, solving for every lag, *m*, rather than every size, *n*, creates additional problems when trying to solve for a grain-size distribution, and carrying out numerical error analyses. Rubin (2004) used an optimisation routine (Lawson and Hanson, 1974) to find the grain-size 'distribution', a vector solution which minimises $(d\text{-}Ax)^T(d\text{-}Ax)$ such that $x \geq 0$ (it is possible to further constrain the problem, such as an upper bound). The advantage is that physically unrealistic values are avoided, but it is computationally more involved, more unstable, and distributions thus obtained are dissimilar to natural size-distributions. The least squares method tends to produce a unimodal, unrealistically well-sorted distribution; and the least-squares with non-negativity method produces statistical artifacts such as a 'blocky', multimodal distribution (Rubin, 2004; Buscombe, 2008). While estimates of median size (D_{50}) may closely match the real sample distribution as derived through traditional means, derived measures such as sorting and skewness are often wildly inaccurate using these distribution estimation methods (Buscombe, 2008).

Buscombe (2008) outlined an alternative approach is to arrive at the vector solution x using a least squares approach, then to use x to compute a smooth probability density function using non-parametric kernel density estimation. There are several potential advantages to this technique, for example the kernel shape may be estimated directly from x so may be optimised to fit each sample. In addition, it may assume a given probability density function (e.g. normal or log-normal), and bounds may be easily specified, such as only allowing for positive grain-sizes. Importantly, the technique converts a solution array x indexed at m, to a solution of proportions which sums to unity whilst conserving probability mass. It therefore makes x amenable to measures of the extent to which the solution minimises the residual between d and A, which is a measure of the numerical 'fit' in the solution of the least-squares problem. This may be important for quality control purposes (i.e. the distributions with large residuals may be rejected). Measures of grain-size obtained from the imaging procedure correlate very well with grain-size measures derived from both the number-frequency and mass-frequency curve (Buscombe, 2008).

Digital Grain-Size in 2D

One problem with the autocorrelation function employed on images in the spatial domain is its sensitivity to the starting point and direction through which the function is applied. Because a mixture of grain-sizes and shapes are present in calibration images, differentiation of closely-sized fractions of material can be difficult, which often means that the calibration catalogue must be truncated at a relatively short lag, beyond which there is non-differentiation. Short correlograms threaten the statistical reliability of the computed mean grain-size and grain-size distribution. Buscombe and Masselink (2008) explored the use of spectral indices of spatial independence, and found that since image intensity at all scales and directions within the image can be mapped simultaneously, it is less sensitive to mixtures of shapes and sizes in calibration images (note that calibration catalogues must always contain a variety of shapes and sizes and colours, in order to characterize the sedimentary environment for repeat-surveys. Therefore the problem of non-differentiation in catalogues is unavoidable). It may be possible, they therefore concluded, to create calibrations which differentiate more distinctly over longer lags, using a spectral rather than spatial approach. Buscombe (2008) elaborated that this may allow for more calibration size-fractions, a better conditioned linear least-squares problem, greater statistical reliability (a larger sample size, both in terms of grain-size fractions and computed lags), and, potentially, better estimates of grain-size and moments of the distribution.

An independent numerical check can be made on the sensitivity/suitability of the calibration catalogue constructed for use in digital grain-size techniques. The sensitivity of a solution of a system of linear equations to errors in the sample data (d) is given by the matrix condition number $C = \|X\| \, \|X^{-1}\|$. The matrix is singular (it has either none or an infinite number of solutions) if $C = \infty$, and 'ill-conditioned' (small changes in d can lead to relatively large changes in x) if $C \gg 1$. Either of these situations must be avoided when calibration catalogues are constructed.

The calibration catalogue A may be expressed in an alternate form. Whereas A contains autocorrelation coefficients associated with each lag 1 to m, an alternative catalogue A_{mod} may be constructed to house the lags associated with a pre-determined vector, v, of autocorrelation coefficients: one for each grain-size (n) represented by the calibration. The relationship between catalogues A and A_{mod} could thus be expressed as $A_{mod} = A_{(m=v,n)}$. The relationship for a $n = 16$, $m = 50$ calibration catalogue may be seen in Fig. 3. Pre-determined v may be found as, for example, linearly-spaced values from the minimum value of r (autocorrelation) in A common to all n, to the maximum value of r in A common to all n. Trials have shown that the matrix condition numbers of the calibration catalogue expressed in the alternative form are consistently lower, indicating that the alternative form of calibration catalogue may be numerically more stable.

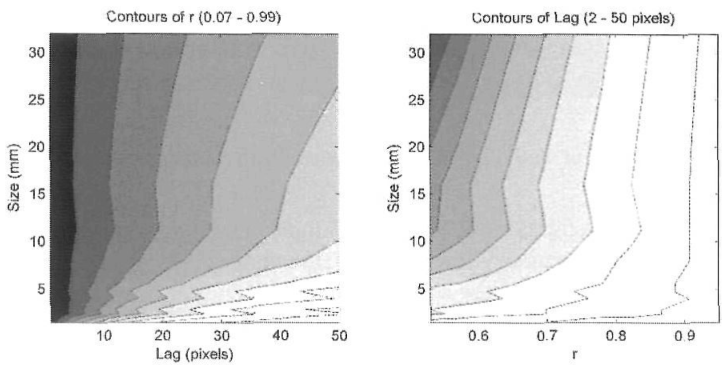

Figure 3. Two alternate forms of a calibration catalogue. Left: A (contouring values of autocorrelation, r), and right: A_{mod} (contouring values of lag associated with a specified vector of r, see text).

Following Preston and Davis (1976) and Lin (1982), who used two-dimensional Fourier transforms on characterising texture in binarised electron microscope images of sandstone thin sections, Buscombe (2008) presented a method by which the autocorrelation approach of Rubin (2004) could be extended into two dimensions. An image power spectrum maps intensity as a function of all frequencies and directions within the image. If intensity is measured along an arbitrary radius from the centre of the zero-shifted 2D power spectrum of an image (e.g. Fig. 4) to the edge, that is equivalent to a one-dimensional power spectrum of the image in that direction. A two-dimensional correlogram may then be obtained from the two-dimensional power spectrum under the Wiener-Khinchin theorem (the two-dimensional correlogram of a de-meaned image is thus the inverse Fourier transform of its two-dimensional power spectrum, normalise by its value at zero lag, which is the total power in the spectrum - see Buscombe (2008) for more details). Power spectral density and autocorrelation coefficient decreases exponentially as a function of wavelength, whereas autocorrelation decreases exponentially as a function of lag.

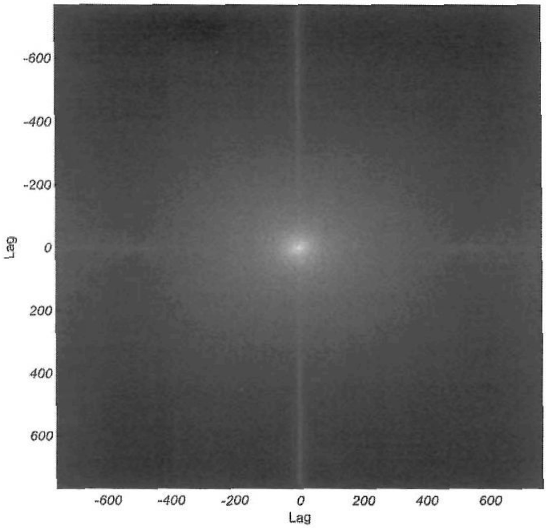

Figure 4. A digital image of sediment in the frequency domain (from Buscombe, 2008)

In a square image, elongation of this field of light in any direction indicates anisotropy in the spatial frequency distributions of image intensity. A perfectly circular closed contour would, on the other hand, indicate perfect isotropy in the features within the image, as would be achieved by uniformly-circular, non-overlapping grains. The closed contours are elliptical rather than circular which indicates that axial dimensions of grains are unlikely to be equal. A two-dimensional correlogram allows the construction of a calibration catalogue in the frequency domain. Potentially, this could be done for all orientations in the original image, although the amount of redundant information would render such an exercise intractable in any practical sense. However, it could be useful in revealing information on anisotropy in image intensity from the two or three dominant orientations of features within the image, which could yield information on the major and minor axes of grains within the image.

A method for constructing such calibration catalogues is outlined in Buscombe (2008). An ellipse may be fitted to each specified contour of autocorrelation, in order to obtain its minor and major axial length (in lags), and orientation. For example, Fig. 5 shows ellipses fitted to contours of $r = 0.1$–0.9. The major and minor axial lengths, in lags, for $r = 0.1$ are shown, as well as the principal axis orientation relative to N–S in the image. Cataloguing these coefficients as a function of both lag and s should enable the construction of two calibration catalogues: one for 'major', and one for 'minor' axial lengths (Fig. 6). Correlograms produced in this way are more distinct over longer lags, because the sensitive dependence of image intensity to start-point and direction is avoided. In addition, the computations are faster, and information in the whole image is utilised rather than in a section of it (Buscombe, 2008).

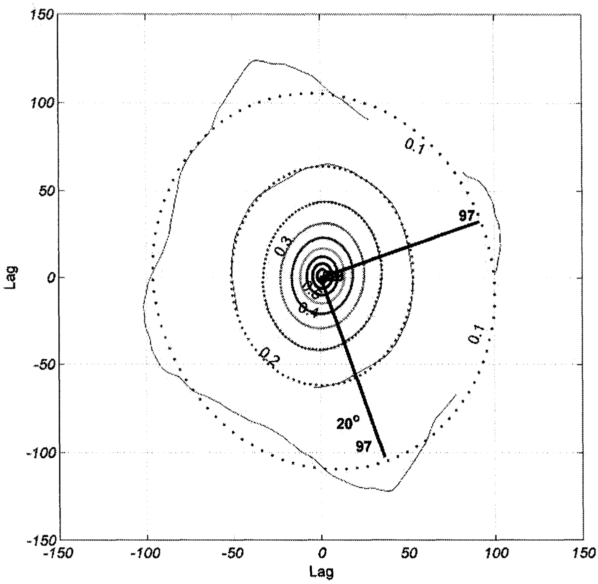

Figure 5. The correlogram of figure 4 (from Buscombe, 2008), contoured and ellipses fitted. The length (in lags) and orientation (in degrees, relative to the top of the image) of the correlogram surface shown for the outer contour (r = 0.1) have been computed using the method described in Buscombe (2008).

DISCUSSION

Implications for Coastal Sediment Transport Modelling and Monitoring

The availability of high-resolution grain-size distributions and derived parameters such as porosity in coastal field and laboratory data sets, may lead to the development of new class of sediment transport model. For example, one where the contribution of grain roughness to friction and energy dissipation is adequately parameterised; or where transport efficiency terms incorporate information on the micro-scale physics related to the relative abundance of coarse and fine particles at the instant of transport (better quantifying relative boundary layer protrusion, hiding and collision, etc). In field monitoring, high-resolution grain-size measurements of surficial sediments on beaches, inlets, spits, etc, may reveal structure not revealed by coarser sampling, but nevertheless have great relevance in the optimal characterization of the seabed, especially for modeling purposes. Figure 7, for example, shows the variability of surface texture (using samples taken every 1m^2) over a highly-2D gravel beach surface. Friction (bed roughness and infiltration) and particle mobility will be dependent on this spatial variability, even enough to induce spatial gradients in subsequent sedimentation.

High-resolution grain-size information may in future aid the development and calibration underwater instrumentation, especially for use in quantifying bedload transport (for example the use of hydrophones for gravel sediment transport). In addition, such information may be invaluable in interpreting underwater acoustic information, especially echoes from multibeam sonars, which are highly dependent on grain size distributions of bed, or indeed any instance where grain-size and sorting affects the acoustic properties of seabed sediments.

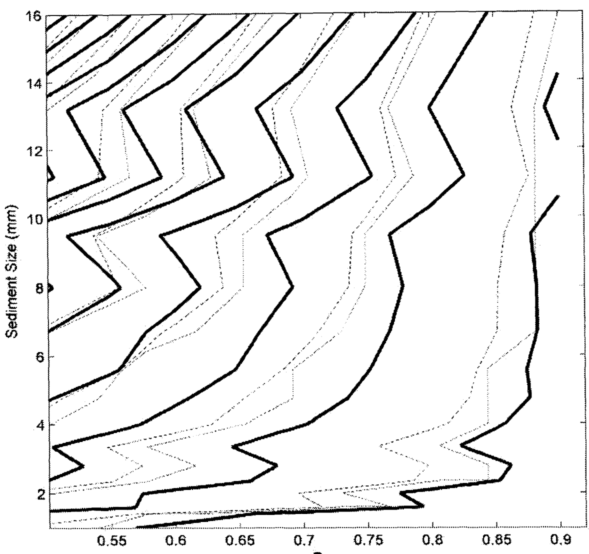

Figure 6. Calibration catalogues (in alternative form A_{mod}) for major (heavy line), mean (solid red line) and minor (dashed line) grain axes, derived from the 2D autocorrelogram method of Buscombe (2008).

Grain-Size Information from Digital Images: Evaluation

A number of challenges remain. For example, all methods thus far developed have needed to minimize or remove exterior light, which has limited digital grain-size techniques to be successfully applied on continuous streams of underwater video of the seabed composed of coarser grain-sizes (gravels/cobbles). Underwater images of sand beds can be taken in the very near-field with macro lenses, minimizing problems with lighting and turbid water, and already have been successfully used in the field (e.g. Mustain et al., 2007; Rubin et al., 2006, 2007). The challenge of statistically resolving (2D) particle shape, in an automated fashion, also remains unsolved. The techniques function only on the 2D image plane, so the full structure of the sediment bed is not quantified, and only the surface grains are sampled. Finally, the technique is only as good as the calibration catalogue it uses. This contribution contains a number of pointers as to the construction on an ideal calibration, and the reader is referred to Barnard et al., (2007) and Buscombe and Masselink (2008) for further examples and recommendations.

Particle-size analysis through traditional methods produces mass-frequency rather than number-frequency curves. The number of individual particles remains unknown, thus precluding the use of ordinary statistical measures of fit such as chi-squared, between measured grain-size distributions and parametric models. A related problem is that traditional measurement practices require the grain-size information is grouped, rather than continuous, which makes the numerical solution to some probability density functions difficult or impossible. Digital grain-size techniques share these (potential) limitations with traditional forms of particle size analysis.

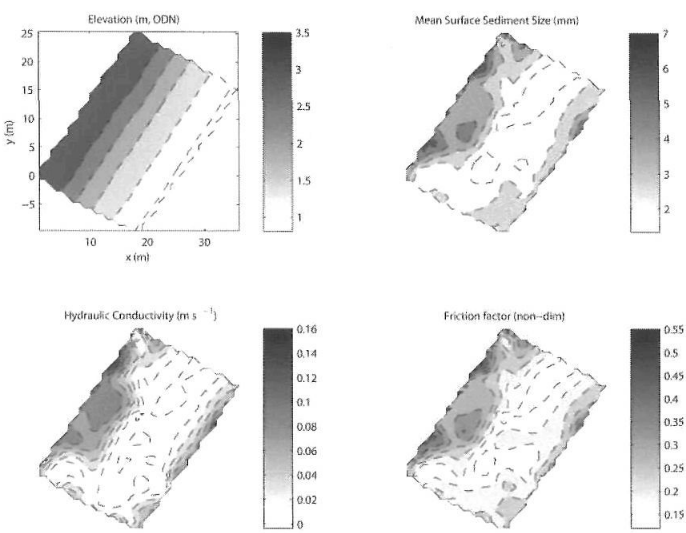

Figure 7. A 30 x 40 m area section of beach at Slapton Sands, UK. Top left: elevation (m, ODN). Top right: mean imaged surface sediment size (mm). Bottom right: friction factor (non-dimensional). Bottom left: hydraulic conductivity (cm s^{1}) (from Buscombe and Masselink, 2008).

CONCLUSION

Digital grain-size methods have undergone extension algorithmic development and validation in a relatively short space of time, and we predict that these statistical techniques will remain popular with researchers in disparate fields for rapid sampling of sediment beds, allowing larger areas to be covered quicker and easier/cheaper. The techniques have found utility in field studies of silt and sand environments (e.g. Mustain et al., 2007) through to gravels and cobbles (e.g. Ruggiero et al., 2007; Buscombe and Masselink, 2008). Recent methodological advances have made it possible to not only provide an accurate estimate of grain-size (Barnard et al., 2007), even in 'real-time' (Rubin and Chezar, 2007), but also reliable estimates of the distribution and parameters thereof (Buscombe, 2008). Work for the future includes validation of the 2D autocorrelation approach proposed by Buscombe (2008), which has the potential to provide more reliable estimates of sediment image correlograms, and also perhaps the construction of several calibration catalogues out of the same set of calibration images. This contribution has detailed a new approach to the construction of the calibration catalogue, on which the success of an individual application crucially depends. This approach allows more flexibility in the calibration stage, and also appears to be better equipped at providing a better posed set of simultaneous equations (linear problem) to be solved, quantified by matrix condition numbers.

ACKNOWLEDGMENTS

Thanks to the Challenger Society for Marine Science and the Plymouth Marine Science Education Fund for financial assistance.

REFERENCES

Austin, M.J., and D. Buscombe. 2008. Morphological change and sediment dynamics of the beach step on a macrotidal gravel beach, *Marine Geology* 249, 167-183.

Barnard, P.L., D.M. Rubin, J. Harney, and N. Mustain. 2007. Field test comparison of an autocorrelation technique for determining grain size using a digital 'beachball' camera versus traditional methods, *Sedimentary Geology* 201, 180-195.

Buscombe, D. 2008. Estimation of Grain Size Distributions and Associated Parameters from Digital Images of Sediment, *Sedimentary Geology*, 210, 1-10.

Buscombe, D., and G. Masselink. 2008. Grain Size Information from the Statistical Properties of Digital Images of Sediment, *Sedimentology*, in press (doi: 10.1111/j.1365-3091.2008.00977.x).

Butler, J.B., S.N. Lane, and J.H. Chandler. 2001. Automated extraction of grain-size data from gravel surfaces using digital image processing. *Journal of Hydraulic Research* 39 (5), 519–529.

Gallagher, E., J. MacMahan, P. Russell, G. Masselink and N. Auger. 2006. Grain-size from digital images and morphological variability, *30th International Conference Coastal Engineering*, San Diego, CA, USA.

Ibekken, H., and R. Schleyer. 1986. Photo-sieving: a method for grain size analysis of coarse-grained, unconsolidated bedding surfaces. *Earth Surface Processes and Landforms* 11, 59–77.

Lawson, C., and R. Hanson. 1974. *Solving Least Squares Problems*. Prentice-Hall.

Lin, C., 1982. Microgeometry I: autocorrelation and rock microstructure. *Mathematical Geology* 14, 343–360.

Mustain, N., G. Griggs, and P. Barnard. 2007. A rapid compatibility analysis of potential offshore sand sources for beaches of the Santa Barbara littoral cell. *Coastal Sediments 07*. Vol. 3. ASCE, pp. 2501–2514.

Preston, F., and J. Davis. 1976. Sedimentary porous materials as a realization of a stochastic process. In: Merriam, D.F. (Ed.), *Random Processes in Geology*, pp. 63–86.

Rubin, D.M. 2004. A simple autocorrelation algorithm for determining grain size from digital images of sediment, *Journal of Sedimentary Research* 74, 160-165.

Rubin, D.M., H. Chezar, J.N. Harney, D.J. Topping, T.S. Melis, and C.R. Sherwood. 2006. Underwater microscope for measuring spatial and temporal changes in bed-sediment grain size, *USGS Open File Report* 2006-1360.

Rubin, D.M., and H. Chezar, 2007. Instant Grainification: Real-Time Grain-Size Analysis from Digital Images in the Field, *Eos Transactions AGU, 88, Fall Meeting* Abstract H51I-0899.

Rubin, D.M., H. Chezar, J.N. Harney, D.J. Topping, T.S. Melis, and C.R. Sherwood. 2007. Underwater microscope for measuring spatial and temporal changes in bed-sediment grain-size, *Sedimentary Geology* 202, 402–408.

Ruggiero, P., P.N. Adams and J.A. Warrick. 2007. Mixed sediment beach processes: Kachemak Bay, Alaska, *Proceedings of Coastal Sediments 2007* (1), 463–476.

Warrick, J.A., D.M. Rubin, P. Ruggiero, A. Draut, and E. Todd. 2007. 'Cobble Cam': Grain-Size Analysis of Gravel Bars from Digital Photographs, *Eos Transactions AGU, 88, Fall Meeting* Abstract H51I-0900.