CALCULATION OF EXPLICIT PROBABILITY OF ENTRAINMENT BASED ON INERTIAL ACCELERATION MEASUREMENTS

Georgios Maniatis [1], Trevor B. Hoey [2], Marwan A. Hassan [3], Joseph Sventek [4], Rebecca Hodge [5], Tim Drysdale [6], Manousos Valyrakis [7]

ABSTRACT

A new method for the approximation of the explicit probability of entrainment for individual coarse particles is presented. The method is based on the derivation of inertial acceleration measurements, space-state approximation of the dynamics close to entrainment and the probabilistic approximation of the threshold inertial acceleration that causes incipient motion. Results from flume experiments with a custom-made Inertial Measurement Unit enclosed in an idealized spherical enclosure, under varied flow conditions (achieved via slope change) and two different arrangements (saddle and grain-top positions) are presented to demonstrate the application of the method. The analysis supports the modification of the existing flow velocity based entrainment criteria so they respect the particle-frame realisation of forces during incipient motion.

Keywords: grain entrainment, inertial sensors, incipient motion

[1]School of Geographical and Earth Sciences and School of Computing Sciences, University of Glasgow. UK. E-mail: g.maniatis.1@research.gla.ac.uk
[2]School of Geographical and Earth Sciences. University of Glasgow. UK
[3]Department of Geography. University of British Columbia. Canada
[4]Department of Computing and Information Science. University of Oregon. US
[5]Department of Geography. Durham University. UK
[6]Department of Engineering and Innovation. Open University, Milton Keynes. UK
[7]School of Engineering. University of Glasgow. UK

INTRODUCTION

Water-related erosion and sedimentation processes interact with the biosphere at a global scale (Walling 2009) and affect many human activities (Dotterweich 2008). One of the most important aspects of sediment transport, that is also one of the most complex and difficult problems in this field, is determination of the dynamics that drive the initiation of sediment movement which reflect the combined effects of fluid turbulence, grain arrangement and local topographic variability (Grass 1970; Buffington and Montgomery 1997; Dey 2014). In spite of the large number of studies since the classic work of Shields (Buffington and Montgomery 1997; Shields 1936), aspects of the problem of sediment entrainment remain unsolved which partly-explains the absence of a widely accepted model for the prediction of sediment transport in streams (Merritt et al. 2003).

Although it is generally accepted that sediment entrainment and hence material transport rates increase in a non-linear manner as the flow rate increases, a unique, explicit parameter capable of characterising the threshold conditions for incipient motion does not yet exist (Garcia 2008; Buffington and Montgomery 1997). Furthermore, the variability of sediment transport and the plethora of related parameters (e.g. the wide range of particle sizes, bed surface structure, hiding and exposure, the complex history of the channel bed) lead to bedload transport models and corresponding entrainment criteria that are valid only within specific conditions (Habersack and Kreisler 2013).

Accurate modelling of bedload transport processes is complicated by this transport taking place across a range of temporal and spatial scales (from the grain to catchment scale, and from rapid single-grain movements to annual bed displacements). These scale ranges lead to two fundamentally different descriptions of sediment transport: the Eulerian deterministic approach formalized by Shields (Shields 1936), mainly applicable at reach to catchment scales, and the Lagrangian stochastic step-length model introduced by Einstein (Einstein 1937), which is by definition relevant to the grain scale. The specification of an appropriate dynamic field for these approaches (Eulerian or Lagrangian) concerns the frame of reference

for the water-flow generated dynamics: in Shields' model individual particles move under a time-averaged mean bed shear stress, while in Einstein's model grain movement is consequent on local turbulent stresses (Papanicolaou et al. 2002). Formal ways to link the two frames of reference in the context of sediment transport are logical next steps in improving our understanding of transport processes.

The most widely used criterion for incipient motion is Shields' critical shear stress (τ_c). τ_c is the bed shear stress produced by the water flow (if uniform flow is assumed this is approximated as a channel slope-depth product) that is capable of mobilising each specific sediment size class (which, for grain sizes yielding particle Reynolds' Numbers > 70, is correlated with the median diameter of the sediment).

Since Shields' (1936) work, a series of empirical values have been suggested to account for a range of factors including the relative depth of the flow, grain shape and protrusion (Ashida and Michiue 1971; Fenton and Abbott 1977; Shvidchenko and Pender 2000). In parallel, Shields' criterion has been extensively criticized for its ambiguity and limited applicability (Church et al. 1998; Buffington and Montgomery 1997; Parker et al. 2003; Bunte et al. 2013) and the validity of a single criterion or even the existence of measurable critical threshold conditions have been questioned (Einstein 1950; Lavelle and Mofjeld 1987). Parallel work has associated the effects of bed micro-topography (Kirchner et al. 1990; Buffington et al. 1992; Prancevic and Lamb 2015), the near bed flow turbulence (Nelson et al. 1995; Papanicolaou et al. 2002) and impulsive (Diplas et al. 2008; Valyrakis et al. 2010) or energetic (Valyrakis et al. 2013) flow events on incipient motion. The combined result of these phenomena cannot be accounted for within a deterministic time-averaged mean stress calculation, although such approaches can still yield useful results.

The core problem with the inclusion of all the above phenomena in any analysis of entrainment is that each of them is difficult to measure or quantify. Advances in monitoring techniques have improved the accuracy of measurements of grain scale near bed forces (Papanicolaou et al. 1999; Schmeeckle et al. 2007) as well as enabling monitoring of impulse

events and their energy potential (Valyrakis et al. 2013). These laboratory measurements reveal great variability of flow dynamics at micro-scale which, combined with the random character of the micro-topography, justifies the treatment of incipient motion as an inherently stochastic processes (comparable to Einstein's description). A summative review of recent studies that define and explore the concept of "pickup probability" or "entrainment probability" as attributed to conditions related to both flow turbulence and sediment arrangement is presented in Marion and Tregnaghi (2013).

Marion and Tregnaghi (2013) show how all the stochastic studies of entrainment reutilize and extend the conceptual framework introduced by Grass (1970). Grass suggested calculating the probability of entrainment as a joint probability derived from Probability Density Functions (PDF hereafter) of critical shear stresses (connected to the resistance to entrainment of the sediment grains) and of the distribution of hydrodynamic forces (derived from near bed flow velocities). The probability of entrainment (P_E) is calculated as the exceedance of a random near bed velocity (U_f) represented by a cumulative distribution F_{U_f} having a PDF of f_{U_f}, as:

$$P_E = P(U_f > u_f = u_g) = \int_{u_g}^{\infty} f_{U_f}(u)du = 1 - F_{U_f}(u_g) \tag{1}$$

where u_g represents the threshold velocity for grain entrainment. Note that the form of equation (1) is general and also applies to definitions of entrainment in terms of other relevant variables such as shear stress or turbulent kinetic energy.

A development of the stochastic description of sediment transport is that the process has been described using a range of mathematical approaches including state-space descriptions (e.g. Markov chains, Tsai and Lai (2014)) and inference techniques in both adaptive neuro-fuzzy (Valyrakis et al. 2011) and classical Bayesian (Schmelter and Stevens 2012) contexts.

These approaches rely on robust calculation of probabilities such as the probability of entrainment, since they can be utilized either as real transition probabilities for the state-space derivations or as training functions and priors for the inference systems.

An aspect of the entrainment problem that has not been extensively investigated is that the existing criteria for sediment entrainment are essentially implicit in the sense that they are based on near particle flow features (e.g. flow turbulence) rather than characteristics of each individual particle and its local arrangement. This problem has often been identified (Cao et al. 2006), but has only recently been formally treated by measuring the dynamics that occur in the inertial frame of the particles close to the threshold of entrainment. It is now technically possible to measure inertial dynamics at scales appropriate for gravel sized sediment, since the miniaturisation of sensing equipment has made the concept of a "smart pebble" (a small, free-moving multi-sensor capable of measuring inertial dynamics such as acceleration and angular velocity) feasible (Akeila et al. 2010; Šolc et al. 2012; Frank 2014).

Maniatis et al. (2013) have shown how this technology can be optimized for natural fluvial environments, demonstrating the capability of the sensor to capture accurate, representative and robust dynamical information over a broad range of imposed forces. However, interpretation of the inertial data in a theoretical framework for incipient motion (Frank 2014) has so far been restricted to the utilisation of Shields' conceptual model.

Following from these theoretical and technical developments, the contributions of this paper are to provide:

- an evaluation of the mobile sensor presented by Maniatis et al. (2013) in entrainment threshold experiments. These results provide supporting evidence towards the formation of an explicit entrainment criterion that has the potential to be utilised across the range of natural river flow regimes.
- description of the derived time-series with dynamic linear models in order to make space state approximations for a representative underlying entrainment process. This approximation is performed by the application of a simplified Kalman filter.

- illustration of attribution of categorical variables to the approximated states, and calculation of the probability of entrainment as a function of inertial acceleration using logistic regression analysis. This result connects directly the inertial dynamics of individual particles to the more relevant probabilistic mathematical context for the description of incipient motion.
- finally we introduce a metric to evaluate the performance of the probabilistic criteria that are relevant to grain incipient motion: the overlapping coefficient (OVL) (Weitzman 1970). The derivation of the OVL requires the numerical approximation of the PDF of the recorded measurements (for pre- and after entrainment conditions) which is achieved non-parametrically using Kernel Density Estimates (KDE).

METHODS

Flume experiments

Initial laboratory experiments used a prototype sensor designed specifically for flume deployment. This prototype consists of a wireless mote platform deployed with a 3-axis accelerometer with a measurement range of $\pm 4g$ precisely located at the cntre of mass of the particle. The electronics were enclosed in a spherical enclosure of $111mm$ diameter and the overall assembly weighed $1.43kg$ (Maniatis et al. 2013), giving a density of $2383kg.m^{-3}$, which is within the range of natural materials. Higher density can be achieved by adding removable weights to voids designed within the case (Maniatis et al. 2013). The prototype was tested in a series of experiments in the $6m \times 0.6m$ $(L \times W)$ recirculation flume in the Mountain Channel Hydraulics Experimental Laboratory (MCHEL), University of British Columbia. The scope of the experiments was to make a first evaluation of the $3D$ inertial acceleration measurements from the prototype sensor under varying flow and slope conditions.

We constructed an idealized bed of hemispheres (Figure 1) of the same diameter as the the sensor ($111mm$) using the rapid prototyping technique described in Maniatis et al. (2013). The result was a bed topography with saddle and grain-top positions that approximated the model described in Kirchner et al. (1990). In each experiment, the sensor was placed in a saddle position (position A in Figure 1) and the flow initiated from zero with a steadily increasing rate of $0.014l.s^{-2}$ up to a maximum rate of $6.25l.s^{-1}$ which was reached after 446 seconds. Upstream of the bed of hemispheres, large sediment particles generated fully turbulent flows which approximated uniform conditions (equal water depth along the length of the flume) over the full range of discharges used. Flow velocities where monitored using a *Vectrino II* Acoustic Doppler Profiler (Craig et al. 2011) for the discharges where flow depths were sufficient to permit the placement of the probe to be far enough from above the bed to derive measurements. The experiments were designed to observe the following series of processes:

- Entrainment from the saddle position and subsequent deposition in the grain top position (position B in Figure 1). As well as providing information on movement from saddle positions, this step removes bias from the placement of the sensor, since the deposition in the second grain top position can be treated as natural with minor random variations in position and pebble orientation.
- A second entrainment of the instrumented particle which transports the sensor out of the monitoring area.

At each of four different flume slopes, 0.026, 0.037, 0.044 and 0.057, nine repeat experiments were undertaken giving 36 experiments in total. The inertial acceleration of the sensor was monitored at a frequency of $4Hz$.

Experimental limitations

Our sensor has idealised physical characteristics compared to a natural coarse particle, and the sensor was only subjected to shallow flows (up to $0.16m$) and velocities up to $0.37m.s^{-1}$ (Figure 2), such that, except at the highest flows at the lowest slope (0.026), flow depth was less than sensor diameter. When using the data from these experiments the assumptions and the errors inherent to the sensing process need consideration.

The sensor has been developed for natural environments and the testing of relevant signal transmission and wireless communication technologies was crucial for the evaluation for the prototype. Hence for this prototype a sampling rate of $4Hz$ was used, the maximum sampling rate at which acceleration data could be reliably transmitted over the underwater radio.

The optimal sampling frequency for entrainment of large grains in turbulent flows is less than the frequency for characterising flow turbulence due to the lower sensitivity of particle movement to micro-turbulence. However, the system must record particle dynamics at sampling rates that reflect the physical meaning of the derived inertial measurements. Although previous experiments with artificial pebbles have used very high sampling frequencies for both intertial (e.g. Šolc et al. (2012)) and turbulence (e.g. Valyrakis et al. (2013)) mea-

surements, over-sampling involves redundant data storage and may lead to artefacts in the data. The optimal sampling frequency for particle motion can be defined with reference to the velocity of the particle. Assuming maximum displacement velocities of the order of 100 diameters per second, as demonstrated in experiments for particles of smaller diameter (Ancey et al. 2002), we suggest a target sampling frequency of about $> 50Hz$.

To enhance the analysis by isolating the impact of the flow, we report the absolute total acceleration after gravity compensation (CA) which is the total acceleration minus the acceleration due to gravity (Equation 2):

$$A = \sqrt{a_x^2 + a_y^2 + a_z^2} \quad , \quad CA = A - \mid g \mid \tag{2}$$

Where A is the magnitude of total acceleration, a_x, a_y, a_z are the accelerations recorded along the x, y and z axes respectively, CA is the acceleration norm after gravity compensation and g is the acceleration norm due to gravity ($9.81m.s^{-2}$ or $1g$). From this point, the term absolute acceleration (acceleration norm) refers to the absolute total inertial acceleration after gravity compensation (CA) as described in Equation 2.

More accurate compensation for gravity is possible, with the monitoring of rotation and the removal of the gravitational effect from the axis parallel to the gravity vector (Nagrath et al. 2008). This form of compensation was not possible using our current sensor which did not contain a gyroscope and magnetometer, but its absence does not affect the calculation of the absolute compensated acceleration value. The addition of a gyroscope would also enable comparison with evidence for directional entrainment from saddle positions (Chin and Chiew 1993).

Another important implication of not measuring angular velocities is the inability to integrate the accelerations in order to derive velocities, which has two aspects. Firstly, although the noise threshold was identified during (manufacturer proposed) calibration and subtracted

from the measurement, inaccuracies remain and are highly relevant to the sensor's response to gravity (Woodman 2007). Secondly, in the field of electrical engineering the error accumulation in MEMS based IMUs is one of the most intensively researched problems (Zekavat and Buehrer 2011). It is known that, without a restriction of the degrees of freedom of the motion, the error propagation during integration makes the velocity (and the displacement results after two integrations) unusable. To access the"velocity response" of the sensor the measurement of angular velocity is necessary and we address this in a subsequent paper.

Finally a combination of restricted sampling frequency and absolute gravity compensation leads to a masking of the pre-entrainment conditions during the statistical treatment of the signal. More specifically entrainments and pre-entrainment motions occur in the same time-window when the signals from individual experiments are synchronized (Figure 3). This is an artefact of the data-processing in order to increase the confidence on the magnitude of the recording dynamics. An other type of analysis that includes advanced filtering of the individual signals (see Section 2) of higher frequency is needed to estimate accurately the fluctuation of pre-entrainment dynamics.

CA and the fundamental forces in a fluid flow

Gravity compensation is important since, for the inertial frame of the sensor, gravity is fictitious force. After removing the fictitious forces from the accelerometer measurements we are left with the linear coordinate acceleration, which is the acceleration that mobilises the sensor relative to the bed (or the Eulerian frame of the flow if an explicit description is required). As a result the CA represents the magnitude of the resultant force, the (3D-tensor) force generated on the centre of mass of the particle when the force balance is disturbed.

The above can simplify significantly the mathematisation of particle entrainment. Using recent definitions of the force balance on a spherical particle that is exposed to a fluid flow (Bialik et al. 2012), the resultant force is given by:

$$\Sigma F = F_D + F_L + F_M + W_{sub} \tag{3}$$

Where F_D and F_L are drag and lift generated forces, respectively, F_M is the added mass force and W_{sub} is the submerged weight of the particle. For a parametrisation of these forces see Bialik et al. 2012. Interestingly, the CA acceleration parameterises directly the left part of Equation 3. $\Sigma F = ma$, where m is the mass of the particle and a is the acceleration tensor applied on the centre of mass of the particle. If an ideal accelerometer (without noise) was placed precisely on the centre of the mass of the exposed particle then we could write $CA =\mid a \mid$ since all the forces of are non-fictitious. Unfortunately real accelerometers are not ideal and this is why, in this work, we choose to treat the acceleration measurements statistically.

Categorisation and summary of Total Acceleration (CA)

Pre-entrainment position grouping

The experiments produced two different modes of movement. For the two lowest slopes (0.026 and 0.037) initial entrainment from the saddle position was followed by settling in a grain top position, where the sensor remained until entrained for a second time. On the contrary, for the two higher slopes (0.044 and 0.057), although entrainment from the saddle position was clearly recorded, the sensor did not remain stationary in the grain-top position for sufficient time prior to its second, grain-top, entrainment to allow isolation of pre-entrainment conditions.

For the following analysis entrainments from the saddle position for the low slopes are omitted and data are grouped in two limiting cases:

- high resistance to entrainment (low slope, grain-top position), entrainment from the grain-top position for the lowest two slopes;
- low resistance to entrainment (high slope, saddle position), entrainment from the saddle position for the two higher slopes.

This grouping avoids inconsistent comparisons and allows investigation of a wider range of pre-entrainment dynamics. Note that for the lower slopes, entrainment from the saddle position was identified from the data as an orientation change on the acceleration vector (a_x, a_y, a_z). In the total acceleration signals reported here orientation changes are masked. A representative signal at the entrainment point from an individual experimental run for one slope (0.026) and grain top position is shown in Figure 3a. Note that in this individual signal a pre-entrainment wobbling is also recorded c.10sec before the entrainment event.

Derivation of aggregated time series for each slope

The acceleration time series from each individual run were synchronised with the corresponding flow rate curve (Figure 3b). Instead of approximating the underlying process for individual runs, the individual acceleration signals for each slope have been aggregated.

For a time domain $t_0, .., t_n$ the acceleration signal is defined by a series of $A_0, .., A_n$ absolute acceleration values. If k is the number of experiments for one slope the summed acceleration signal is given by the set:

$$A = \{\{A_1, ..., A_k\}_{t_0}, \quad ..., \quad \{A_1, ..., A_k\}_{t_n}\} \tag{4}$$

The above formulation states the full range of absolute accelerations recorded in each of k repeat experiments ($k = 9$ in this case) for each time point, hence discharge value (Figure 3c). The analysis of the aggregated signal has two benefits compared to individual signal analysis:

- given that our analysis is purely statistical all the approximations are performed with a larger input sample of accelerations which increases statistical confidence and thus the significance of the results; and,
- the resulting individual time series for each slope is more representative of a raw signal derived in a natural environment, extending the application range of the presented method.

Analysis of Absolute Acceleration (CA) : Statistical Techniques

Dynamic Linear Model Filtering

The aim here is to approximate the underlying dynamical process for each slope by analysing the combined acceleration signal. Space-state estimation techniques for time series analysis are commonly used (Box et al. 2013). Here we follow (Zhang and Li 1996) and use a recursive algorithm for space state estimation to enhance numerical stability and the square root version of the Kalman filter (Kalman 1960). We used these algorithms as coded for the R-statistical software by (Petris 2010).

The summed time series is approximated with a first order polynomial model of the form:

$$\begin{cases} y_t = \theta_t + \upsilon_t, \ \upsilon_t \sim N(0, V) \\ \theta_t = \theta_{t-1} + \omega_t, \omega_t \sim N(0, \Omega) \end{cases} \tag{5}$$

where y_t is the vector of observed absolute accelerations after gravity compensation (CA). θ_t is a vector which represents the underlying process (the state) of the system. The observation vector is related to the process vector with the addition of Gaussian noise with variance V ($\sigma^2 = 0.2$). Similarly, the process vector, Ω, is defined as its preceding value with additional covariance ($\sigma^2 = 0.3$). The estimation of these variances is based on the calibration of the sensor under zero (non-gravitational) acceleration conditions. The recursive algorithm of Zhang and Li (1996) is then applied to compute the filtering distribution and its variance for each one of the observations. Finally, the variances are used to calculate probability intervals as shown in Figure 3c.

Note that the combination of Equation 4 with the linear Kalman filtering of the model of Equation 5 should not be confused with the Ensemble averaging presented in other works (eg. Fathel et al. 2015). Here we only group the derived sychornised accelerations (CA) and this aggregation does not represent an Expected Value or any other function.

The Kalman filter only approximates the process given the overall time series and the

relative variances we introduced from calibration. As a result the difference between the individual entrainment of Figure 3a and the approximated entrainment of Figure 3c is explained by the fact that more entrainments happened at a later time in the aggregated data. Characteristically for the slope discussed in Figure 3 (0.026) the individual entrainments occurred in a range between between 20 and 48mg of CA as shown below (Figure 7a).

Logistic Regression Calculation of the Probability of Entrainment

The probability of entrainment is assessed using the derived signals for each individual run, to which a binary categorical variable was attributed with pre-entrainment and post-entrainment states (0 and 1, respectively). This allows the recorded accelerations to be grouped according to grain condition and for calculation of separate probability densities for each condition (Figure 4a). Note that as explained previously at the lower slopes, the measured entrainments are from grain-top positions following initial displacement from saddle positions, whereas at higher slopes entrainment was from the saddle positions (Figure 6). The point of entrainment is shown in the derived signals as a sudden drop in the inertial acceleration, followed by high acceleration values due to impacts of the sensor with the flume bed down-stream of the monitoring area (Figure 3c). In two cases of ambiguity (for slope 0.57) the time of entrainment was verified from video recordings.

To calculate the probability of entrainment we fitted a binomial model using logistic regression (Hosmer et al. 2000) between the binary variable that corresponds to entrainment and the accelerations derived after the space-state filtering of the time-series. As a generalized linear model, logistic regression for binomial data is expressed with the linear predictor function:

$$logit = \left(E\left[\frac{Y_i}{n_i} \mid X_i \right] \right) = logit(p_i) = ln\left(\frac{p_i}{1-p_i} \right) = \beta X_i \tag{6}$$

where Y_i is the dependent binary variable with:

$$Y_i \sim B_{in}(n_i, p_i) \quad \text{for} \quad i = 1, ..., n \tag{7}$$

X_i being the vector of predictors and β being the vector of regression coefficients (Kay and Little 1987).

In this case the model was reduced to one predictor variable, equal to the filtered acceleration values. As an optimisation process to estimate the probabilities p_i and the regression coefficients β, we implemented the Maximum Likelihood Estimation in the R-statistical package for the default function *glm* (Faraway 2005). The values fitted by the above process are an explicit calculation of the probability of entrainment as a function of the recorded inertial acceleration. The threshold of entrainment is determined by the acceleration corresponding to 0.5 probability as shown in Figure 4b. The determination of the entrainment threshold as a probability of 0.5 is consistent with other applications of probabilistic entrainment criteria that utilize near bed turbulence measurements (Papanicolaou et al. 2002). The acceleration values corresponding to probabilities < 0.5 represent dynamic conditions that act in favour of the resistance of the particle to entrainment, whereas values corresponding to probabilities ≥ 0.5 represent dynamic conditions where the potential for entrainment is enhanced.

A statistical framework for the evaluation of incipient motion criteria

Rationale

The probabilistic derivations for incipient motion differ in terms of the physical definition of the conditional probability that defines the threshold of motion. In the initial framework of Grass (1970) the conditional probability is expressed as a function of Shields' shear stress, while in recent derivations the same probability is physically determined by point (eg. Papanicolaou et al. (2002)) or streamwise (eg. Bottacin-Busolin et al. (2008)) flow velocities (Equation 1). Furthermore, here we propose a new physical definition based on the inertial dynamics of the target particle.

However, one observation is relevant to all the studies, including this one: the definition

of a non-abrupt threshold implies that the the measured physical instance (shear stress, flow velocity or acceleration) is defined by two $PDFs$. One of these defines pre-entrainment conditions (mainly representing the resistance to movement of the particle) and one defining the post-entrainment conditions (mainly recording the mobilisation of the particle). This is a representation in an inertial frame of reference of the idea introduced by Grass (1970) in the form of overlapping PDFs, and it implies that the critical point for entrainment lies within the area of overlap of these two distributions. This has been demonstrated in much later work, regardless of the physical definition of this probability that was used in each case.

In a probabilistic context, the degree of overlap between the pre- and post-entrainment $PDFs$ defines the domain of the critical point, hence the exactness of the entrainment criterion. A large overlap of the two distributions suggests a large domain for the threshold, and is thus a less well-defined criterion. A smaller overlap shrinks the domain for the threshold and the derived criterion is better defined. As a result any improvement in the definition of incipient motion thresholds can be quantified by the degree of overlap of the pre- and post-entrainment distributions.

A formal measure for the overlap between two $PDFs$ is the Overlapping Coefficient (OVL) initially proposed by Weitzman (1970). OVL has been used since to quantify the degree of overlap for a range of distributions, from theoretical normal distributions (Inman et al. 1989) to empirical density functions (Schmid and Schmidt 2006; Clemons and Bradley 2000) which are directly relevant to the analysis presented in this work. Since the Kernel Density Estimation (KDE) of the $PDFs$ is an important step of the analysis, we discuss this before the definition of OVL.

A note on the non-parametric estimation of $PDFs$

The KDE is an established technique for the approximation of $PDFs$ of random variables, when no hypothesis can be established for the underlying distribution (non-parametric). Full description of this technique is outside of the scope of this work, however it is necessary to introduce it from the point of application for the coherence of the presented analysis.

We performed the approximations using the default routines implemented in the R-statistical software which are based on the Fast Fourier transform of the kernel estimator introduced by (Rosenblatt 1956). The basic algorithm was derived by Silverman (Silverman 1982; Silverman 1986) for Gaussian kernels, which is also the type of kernel we chose for KDE in this paper (the default in R-statistical software).

The KDE, like all smoothing techniques, requires the selection of bandwidth. Numerous automatic bandwidth selectors have been devised (see Heidenreich et al. (2013, Sheather et al. (2004)) however they do not all perform equally well (Park and Turlach 1992).

To highlight this effect, we use two bandwidth estimators:

- for display purposes and to derive simple inferences about the data (Figure 4) we use Silverman's rule of thumb (Silverman 1986) which tends to over-smooth the data 5.
- for more accurate calculations, such as to calculate the OVL coefficient, we use the data-based method proposed in Sheather and Jones (1991), which for the variability in our data gives more representative approximations (Figure 5).

Both of these methods are options of the default library of the R-statistical software with Silverman's rule of thumb being the default method.

The Overlapping coefficient (OVL)

After the approximation of the $PDFs$ the Overlapping Coefficient is calculated as:

$$OVL = \int_{R_n} min[f_1(x), f_2(x)]dx \tag{8}$$

where $f_1(x)$ and $f_2(x)$ are two overlapping PDFs and R_n is the n-dimensional space of real numbers.

The OVL coefficient is always in the range $[0, 1]$ and complete overlap between $f_1(x)$ and $f_2(x)$ has $OVL = 1$, while complete distinction gives $OVL = 0$ (Clemons and Bradley 2000). In the context of evaluating entrainment criteria we are looking for OVL closer to 0 which

would suggest smaller overlap between the pre- and post-entrainment $PDFs$.

More specifically if we accept that the threshold of entrainment exists in the region where pre-and post entrainment distributions overlap, then the exactness (and the significance) of the threshold is related to how different the two distributions are. If the distributions were normal and had the same variance the difference of the distributions would be approximated by the separation of the means. The OVL coefficient quantifies this difference for empirically approximated distributions.

A smaller OVL means that the entrainment threshold is better defined. More precisely an $X\%$ reduction of the OVL coefficient represents the maximum % reduction of the variance of the approximated threshold.

RESULTS

Absolute Inertial Acceleration (CA) thresholds

The methods described above for the acceleration analysis were applied to data for all the slopes. Figure 6 shows the filtered acceleration signals and the fitted probabilities of entrainment: these results are summarised in Table 1.

As slope increases the discharge at which entrainment occurs is reduced (Figure 6). However, inertial accelerations recorded by the sensor show a more complex pattern (Figure 6b). For the lower slopes with final entrainment from grain top positions, there is considerable overlap between accelerations at entrainment with the higher (0.037) slope having the highest accelerations. The steeper slopes, with entrainment from saddle positions, also show considerable overlap but accelerations are significantly lower than for the lower slopes.

As a statistical evaluation for the derived binomial models (Figure 6b), the p-values for the significance of the coefficients of the independent variable (acceleration, Table1) are given. Another relevant metric is the Walden test which is used for the evaluation of single predictor models, but also to evaluate competing models with different numbers of predictors. The p–value of the Walden test for the four fitted probability models was < 1 x10^{-20} which

increases our confidence regarding the significance of the derived models (Montgomery and Runger 2010).

Measured inertial accelerations at the point of entrainment were of the order $50mg$ (Figures 3 and 6). The acceleration of mean velocity as a result of the steady increase in discharge through each experiment is four orders of magnitude lower than these inertial accelerations at c.$2x10^{-2}mg$, justifying the assumption of gradually varied flow that has no direct influence on entrainment forces.

Comparison of Total Inertial Acceleration $PDFs$ with Velocity $PDFs$

Figure 7 shows how all the derived signals are synchronised over the same time domain. For the two lower slopes (0.026 and 0.037) the approximated acceleration thresholds (44 and $51mg$, respectively) were projected back to Threshold Discharges (6.15 and $4.3l.s^{-1}$, respectively). The latter were used to separate the recorded flow velocities (Figure 2) to pre- and post-entrainment distributions the $PDFs$ of which were approximated by KDE and the bandwidth selection technique of Sheather and Jones (1991). Finally the OVL coefficient was calculated for both the velocity and total acceleration $PDFs$

The results in Figure 7 show that for both of the slopes the OVL coefficient for the acceleration $PDFs$ (0.36 for $slope = 0.026$ and 0.21 for $slope = 0.037$) is smaller than the $OVLs$ for the velocity $PDFs$ (0.44 for $slope = 0.026$, and 0.33 for $slope = 0.037$).

DISCUSSION

Evaluation of applied techniques

Filters based on a Dynamic Linear Model have many advantages over traditional time series regression analysis as they can be applied without the associated assumptions of stationarity. Another important advantage is that the filtered signal corresponds to exactly the same time domain as the unfiltered series (which is not the case when some other techniques, e.g. moving average, are applied). The latter point becomes crucial since the entire time domain along with the space state characterisation of the process make the attribution of categorical variables to each one of these states both feasible and conceptually consistent. Similarly, logistic regression is a versatile technique that can be applied without the strict assumptions of linear regression and becomes very useful when categorical characterisation of states is necessary (e.g. Entrainment-No Entrainment).

Our results suggest that current technology (inertial-sensors) permits the modification of equation 1 to a form of:

$$P_E = P(A_f > a_f = a_g) = \int_{a_g}^{\infty} f_{A_f}(a)da = 1 - F_{A_f}(a_g) \tag{9}$$

where A_f is a random inertial acceleration variable for an individual pebble, represented by a cumulative distribution F_{A_f} with a probability density function of f_{A_f}, and a_g is the threshold acceleration for grain entrainment as approximated statistically in the current work. This derivation has the potential to enhance the accurate determination of P_E as it utilises the explicit dynamics of the particles being entrained instead of using implicit flow-related metrics.

Here a clarification is necessary; the fact that the above criterion is explicit does not mean that we treat the entrainment process in a non-stochastic framework. This observation is highly relevant to the use of the proposed criterion and methods under different hydrody-

namic conditions (eg. uniform vs non-uniform flow). There is a range of hydrodynamics that can produce the same threshold of inertial acceleration (or more specifically the range of inertial accelerations where entertainment can occur). However, this range of hydrodynamics corresponds to a smaller range of inertial dynamics. This is reflected in the definition of inertial acceleration (Section 2), and results in the approximated acceleration thresholds varying in a range of only $19mg$ regardless of the distinctively different force-balance conditions (slopes and initial placement). Consequently data collected under different hydrodynamic regimes will improve the determination of the inertial threshold and connect it with previous results.

The technique that we propose for the comparison of this inertial acceleration-based threshold with a flow velocity based prediction (OVL coefficient), suggests that the overlap of pre- and post-entrainment distribution is reduced by c.10%. Moreover, the reduction is greater for the higher slope where the variability of the hydrodynamics is greater as demonstrated from the $PDFs$ of Figure 7. Based on this result, it is possible to form the hypothesis that, for medium - large grain scales, the inertia of the particles exerts a more significant control on their motion than flow generated forces (Bathurst 1985). It also important that other geomorphological characteristics can be described by overlapping distributions of dynamics (eg. Ze'ev and Schumm (1984)), which extends the applicability of the OVL coefficient beyond the detection of incipient motion.

As a result, the study of inertial dynamics of the sediments has the potential to improve prediction across the modes of sediment transport. Formalisation of statistical definitions of entrainment can lead to further improvements to the conceptual model introduced by Grass (1970) since new technologies enable dynamical measurements at high frequency and accuracy.

Further study of the proposed criterion under varied conditions is required and is likely to reveal a range of types of behaviour dependent on the same issues which lead to variability in the definition of Shields' criterion (Buffington and Montgomery 1997). However, the range

of behaviour may be better constrained as actual forces are directly measured rather than being inferentially related to measurable parameters such as grain shape and protrusion. The approach therefore has the potential to lead to a general inertial force -based equivalent of Shields' diagram which will not be restricted by the assumption of uniform flow (or any other flow characterisation) and will have broad applicability.

Future work

A new prototype sensor (of diameter $< 80mm$) is under development, instrumented with all the sensors required for full determination of inertial dynamics (accelerometer, gyroscope and a magnetometer which contributes an extra constant reference axis). Reducing the size of the overall unit is crucial for increasing the range of pebble sizes and shapes which we can be tested, either by reducing the diameter of spherical pebbles, or with pebbles with non-unity aspect ratios with one dimension smaller than $111mm$ (the new sensor will be able to be housed in non-spherical casings, which extends its generality). This sensor will be capable of higher frequency (up to c. 100Hz) sampling allowing pre-entrainment and motion dynamics to be recorded coherently in $3D$ space.

Finally field deployment of the sensor will contribute to a better description of all the stages of sediment transport (Entrainment -Translation- Deposition). Currently the technology permits the construction of robust enclosures that, in terms of physical characteristics, are mainly relevant to the sediment sizes found in upland streams, debris flows and some gravel beaches.

CONCLUSION

We provide a new method to approximate the probability of entrainment for individual coarse particles based on inertial acceleration measurements (Maniatis et al. 2013). This became possible after prototyping a purpose specific I.M.U. The key steps of the method are:

1. Recording of inertial dynamics at an appropriate high frequency (inertial acceleration): in this initial study, data were recorded at 4Hz for practical reasons and this has been shown to be adequate for laboratory conditions. For field deployment, recording frequencies of over 50Hz are required, although over-sampling should be avoided to ensure both efficiency and reliable interpretation of results.
2. Bayesian filtering of the derived signals (Kalman Filter): we conducted 9 replicates of each entrainment experiment. Rather than analyse each separately, combining these into one synchronised data set allows robust interpretation and specification of statistical uncertainty in the results using an appropriate process model (Equation 5). This model illustrates both the inter-experiment variability and the trends in the data (Figures 3 and 6).
3. Categorical characterisation of the filtered signals for pre- and post-entrainment conditions: the individual data sets (e.g. Figure 3a) show pre-entrainment vibration increasing through time, entrainment, and post-entrainment oscillations. Categorising these data into pre- and post-entrainment accelerations allows the probability of entrainment to be considered as a function of increasing flow intensity (approximated here by discharge).
4. Logistic approximation of the relationship: After analysing inertial accelerations pre- and post-entrainment, logistic regression provides a way of statistically expressing a gradual increase in the probability of entrainment with increasing accelerations. This also provides confidence intervals (Figures 4 and 6) which clarify the differences in entrainment processes between different grain positions (saddle vs. grain top) and

which also show considerable overlaps between data obtained at different slopes but for the same grain positions. These differences suggest that secondary effects, such as the orientation of initial grain movement, may be significant even for spherical grains.

5. Calculation of the probability of entrainment as a function of inertial acceleration from the conditional threshold probability ($p_i = 0.5$): the notion of entrainment as a gradational increase in probability of movement as flow intensity increases is well-established (Grass 1970), but has been difficult to quantify for natural conditions. Extension of our approach to natural particle shapes and positions will help to address this data requirement.

The results support the implementation of explicit dynamical metrics with reference to the inertial frame of the particle under entrainment. Further research is needed to expand this type of entrainment criterion to a range of particle sizes and dynamical schemes.

ACKNOWLEDGEMENTS

The authors want to thank Shawn Chartrand and Matteo Saletti for assisting the preparation and the execution of the experiments as well the discussions on the meaning of the derived data. Rick Ketler offered valuable technical support during the experimental evaluation (University of British Columbia). The Editor and the Associate editor of the journal, as well as two reviewers, provided feedback on an earlier version that contributed significantly to the manuscript.

APPENDIX I. REFERENCES

Akeila, E., Salcic, Z., and Swain, A. (2010). "Smart pebble for monitoring riverbed sediment transport." *Sensors Journal, IEEE*, 10(11), 1705–1717.

Ancey, C., Bigillon, F., Frey, P., Lanier, J., and Ducret, R. (2002). "Saltating motion of a bead in a rapid water stream." *Physical review E*, 66(3), 036306.

Ashida, K. and Michiue, M. (1971). "An investigation of river bed degradation downstream of a dam." *in Proceedings of 14th Int. Association of Hydraulic Research Congress*, Vol. 3, Wallingford, U.K., 247–255.

Bathurst, J. C. (1985). "Flow resistance estimation in mountain rivers." *Journal of Hydraulic Engineering*, 111(4), 625–643.

Bialik, R., Nikora, V., and Rowiński, P. (2012). "3d lagrangian modelling of saltating particles diffusion in turbulent water flow." *Acta Geophysica*, 60(6), 1639–1660.

Bottacin-Busolin, A., Tait, S. J., Marion, A., Chegini, A., and Tregnaghi, M. (2008). "Probabilistic description of grain resistance from simultaneous flow field and grain motion measurements." *Water resources research*, 44(9).

Box, G. E., Jenkins, G. M., and Reinsel, G. C. (2013). *Time series analysis: forecasting and control*. John Wiley & Sons.

Buffington, J. M., Dietrich, W. E., and Kirchner, J. W. (1992). "Friction angle measurements on a naturally formed gravel streambed: Implications for critical boundary shear stress." *Water Resources Research*, 28(2), 411–425.

Buffington, J. M. and Montgomery, D. R. (1997). "A systematic analysis of eight decades of incipient motion studies, with special reference to gravel-bedded rivers." *Water Resources Research*, 33(8), 1993–2029.

Bunte, K., Abt, S. R., Swingle, K. W., Cenderelli, D. A., and Schneider, J. M. (2013). "Critical shields values in coarse-bedded steep streams." *Water Resources Research*, 49(11), 7427–7447.

Cao, Z., Pender, G., and Meng, J. (2006). "Explicit formulation of the Shields diagram for

599 incipient motion of sediment." *Journal of Hydraulic Engineering*, 132(10), 1097–1099.
600 Chin, C. and Chiew, Y. (1993). "Effect of bed surface structure on spherical particle stabil-
601 ity." *Journal of waterway, port, coastal, and ocean engineering*, 119(3), 231–242.
602 Church, M., Hassan, M. A., and Wolcott, J. F. (1998). "Stabilizing self-organized structures
603 in gravel-bed stream channels: Field and experimental observations." *Water Resources*
604 *Research*, 34(11), 3169–3179.
605 Clemons, T. E. and Bradley, E. L. (2000). "A nonparametric measure of the overlapping
606 coefficient." *Computational statistics & data analysis*, 34(1), 51–61.
607 Craig, R. G., Loadman, C., Clement, B., Rusello, P. J., and Siegel, E. (2011). "Characteriza-
608 tion and testing of a new bistatic profiling acoustic doppler velocimeter: The vectrino-ii."
609 *Current, Waves and Turbulence Measurements (CWTM), 2011 IEEE/OES 10th*, IEEE,
610 246–252.
611 Dey, S. (2014). "Sediment threshold." *Fluvial Hydrodynamics*, Springer, 189–259.
612 Diplas, P., Dancey, C. L., Celik, A. O., Valyrakis, M., Greer, K., and Akar, T. (2008). "The
613 role of impulse on the initiation of particle movement under turbulent flow conditions."
614 *Science*, 322(5902), 717–720.
615 Dotterweich, M. (2008). "The history of soil erosion and fluvial deposits in small catch-
616 ments of central europe: deciphering the long-term interaction between humans and the
617 environmenta review." *Geomorphology*, 101(1), 192–208.
618 Einstein, H. (1937). "Bedload transport as a probability problem." *Sedimentation (reprinted*
619 *in 1972). Water Resources Publications, Colorado*, 105–108.
620 Einstein, H. A. (1950). *The bed-load function for sediment transportation in open channel*
621 *flows*. Number 1026. US Department of Agriculture.
622 Faraway, J. J. (2005). *Extending the linear model with R: generalized linear, mixed effects*
623 *and nonparametric regression models*. CRC press.
624 Fathel, S. L., Furbish, D. J., and Schmeeckle, M. W. (2015). "Experimental evidence of
625 statistical ensemble behavior in bed load sediment transport." *Journal of Geophysical*

626 *Research: Earth Surface*, 120(11), 2298–2317.

627 Fenton, J. and Abbott, J. (1977). "Initial movement of grains on a stream bed: The effect of
628 relative protrusion." *Proceedings of the Royal Society of London A: Mathematical, Physical*
629 *and Engineering Sciences*, Vol. 352, The Royal Society, 523–537.

630 Frank, D. P. (2014). "Incipient motion within the wave bottom boundary layer." Ph.D.
631 thesis, University of New Hampshire, University of New Hampshire.

632 Garcia, M. H. (2008). "Sediment transport and morphodynamics." *ASCE Manuel No. 110,*
633 *Sedimentation Engineering*, 2008, 21–163.

634 Grass, A. J. (1970). "Initial instability of fine bed sand." *Journal of the hydraulics division*,
635 96(3), 619–632.

636 Habersack, H. and Kreisler, A. (2013). "Sediment transport processes." *Dating Torrential*
637 *Processes on Fans and Cones*, Springer, 51–73.

638 Heidenreich, N.-B., Schindler, A., and Sperlich, S. (2013). "Bandwidth selection for kernel
639 density estimation: a review of fully automatic selectors." *AStA Advances in Statistical*
640 *Analysis*, 97(4), 403–433.

641 Hosmer, D. W., Lemeshow, S., and Sturdivant, R. X. (2000). *Introduction to the logistic*
642 *regression model*. Wiley Online Library.

643 Inman, H. F., Bradley, J., and Edwin, L. (1989). "The overlapping coefficient as a measure
644 of agreement between probability distributions and point estimation of the overlap of two
645 normal densities." *Communications in Statistics-Theory and Methods*, 18(10), 3851–3874.

646 Kalman, R. E. (1960). "A new approach to linear filtering and prediction problems." *Journal*
647 *of Fluids Engineering*, 82(1), 35–45.

648 Kay, R. and Little, S. (1987). "Transformations of the explanatory variables in the logistic
649 regression model for binary data." *Biometrika*, 74(3), 495–501.

650 Kirchner, J. W., Dietrich, W. E., Iseya, F., and Ikeda, H. (1990). "The variability of critical
651 shear stress, friction angle, and grain protrusion in water-worked sediments." *Sedimentol-*
652 *ogy*, 37(4), 647–672.

Lavelle, J. W. and Mofjeld, H. O. (1987). "Do critical stresses for incipient motion and erosion really exist?." *Journal of Hydraulic Engineering*, 113(3), 370–385.

Maniatis, G., Hoey, T., and Sventek, J. (2013). "Sensor enclosures: Example application and implications for data coherence." *Journal of Sensor and Actuator Networks*, 2(4), 761–779.

Marion, A. and Tregnaghi, M. (2013). "A new theoretical framework to model incipient motion of sediment grains and implications for the use of modern experimental techniques." *Experimental and Computational Solutions of Hydraulic Problems*, Springer, 85–100.

Merritt, W. S., Letcher, R. A., and Jakeman, A. J. (2003). "A review of erosion and sediment transport models." *Environmental Modelling & Software*, 18(8), 761–799.

Montgomery, D. C. and Runger, G. C. (2010). *Applied statistics and probability for engineers*. John Wiley & Sons.

Nagrath, I., Nagrath, I., and Gopal, M. (2008). *Textbook Of Control Systems Engineering (Vtu)*. New Age International.

Nelson, J. M., Shreve, R. L., McLean, S. R., and Drake, T. G. (1995). "Role of near-bed turbulence structure in bed load transport and bed form mechanics." *Water Resources Research*, 31(8), 2071–2086.

Papanicolaou, A., Diplas, P., Evaggelopoulos, N., and Fotopoulos, S. (2002). "Stochastic incipient motion criterion for spheres under various bed packing conditions." *Journal of Hydraulic Engineering*, 128(4), 369–380.

Papanicolaou, A. N., Diplas, P., Balakrishnan, M., and Dancey, C. L. (1999). "Computer vision technique for tracking bed load movement." *Journal of Computing in civil engineering*, 13(2), 71–79.

Park, B. and Turlach, B. (1992). "Practical performance of several data driven bandwidth selectors." *Report no.*, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Parker, G., Seminara, G., and Solari, L. (2003). "Bed load at low shields stress on arbitrarily sloping beds: Alternative entrainment formulation." *Water resources research*, 39(7).

680 Petris, G. (2010). "An R package for dynamic linear models." *Journal of Statistical Software*,
681 36(12), 1–16.
682 Prancevic, J. P. and Lamb, M. P. (2015). "Particle friction angles in steep mountain chan-
683 nels." *Journal of Geophysical Research: Earth Surface*, 120(2), 242–259.
684 Rosenblatt, M. (1956). "Remarks on some nonparametric estimates of a density function."
685 *The Annals of Mathematical Statistics*, 27(3), 832–837.
686 Schmeeckle, M. W., Nelson, J. M., and Shreve, R. L. (2007). "Forces on stationary particles in
687 near-bed turbulent flows." *Journal of Geophysical Research: Earth Surface (2003–2012)*,
688 112(F2).
689 Schmelter, M. L. and Stevens, D. K. (2012). "Traditional and bayesian statistical models in
690 fluvial sediment transport." *Journal of Hydraulic Engineering*, 139(3), 336–340.
691 Schmid, F. and Schmidt, A. (2006). "Nonparametric estimation of the coefficient of overlap-
692 ping?theory and empirical application." *Computational statistics & data analysis*, 50(6),
693 1583–1596.
694 Sheather, S. J. et al. (2004). "Density estimation." *Statistical Science*, 19(4), 588–597.
695 Sheather, S. J. and Jones, M. C. (1991). "A reliable data-based bandwidth selection method
696 for kernel density estimation." *Journal of the Royal Statistical Society. Series B (Method-*
697 *ological)*, 683–690.
698 Shields, A. (1936). "Application of similarity principles and turbulence research to bed-load
699 movement, translated from german.
700 Shvidchenko, A. B. and Pender, G. (2000). "Flume study of the effect of relative depth on the
701 incipient motion of coarse uniform sediments." *Water Resources Research*, 36(2), 619–628.
702 Silverman, B. (1982). "Algorithm as 176: Kernel density estimation using the fast fourier
703 transform." *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 31(1),
704 93–99.
705 Silverman, B. W. (1986). *Density estimation for statistics and data analysis*, Vol. 26. CRC
706 press.

Šolc, T., Stefanovska, A., Hoey, T., and Mikoš, M. (2012). "Application of an instrumented tracer in an abrasion mill for rock abrasion studies." *Strojniški vestnik-Journal of Mechanical Engineering*, 58(4), 263–270.

Tsai, C. W. and Lai, K.-C. (2014). "Three-state continuous-time markov chain model for mixed-size sediment particle transport." *Journal of Hydraulic Engineering*, 140(9).

Valyrakis, M., Diplas, P., and Dancey, C. L. (2011). "Prediction of coarse particle movement with adaptive neuro-fuzzy inference systems." *Hydrological Processes*, 25(22), 3513–3524.

Valyrakis, M., Diplas, P., and Dancey, C. L. (2013). "Entrainment of coarse particles in turbulent flows: An energy approach." *Journal of Geophysical Research: Earth Surface*, 118(1), 42–53.

Valyrakis, M., Diplas, P., Dancey, C. L., Greer, K., and Celik, A. O. (2010). "Role of instantaneous force magnitude and duration on particle entrainment." *Journal of Geophysical Research: Earth Surface (2003–2012)*, 115(F2).

Walling, D. E. (2009). *The impact of global change on erosion and sediment transport by rivers: current progress and future challenges*. Unesco, Paris.

Weitzman, M. S. (1970). *Measures of overlap of income distributions of white and Negro families in the United States*, Vol. 22. US Bureau of the Census.

Woodman, O. J. (2007). "An introduction to inertial navigation." *University of Cambridge, Computer Laboratory, Tech. Rep. UCAMCL-TR-696*, 14, 15.

Ze'ev, B. and Schumm, S. A. (1984). "Gradational thresholds and landform singularity: significance for quaternary studies." *Quaternary Research*, 21(3), 267–274.

Zekavat, R. and Buehrer, R. M. (2011). *Handbook of position location: Theory, practice and advances*, Vol. 27. John Wiley & Sons.

Zhang, Y. and Li, X. (1996). "Fixed-interval smoothing algorithm based on singular value decomposition." *Control Applications, 1996., Proceedings of the 1996 IEEE International Conference on*, IEEE, 916–921.

APPENDIX II. NOTATION

The following symbols are used in this paper:

PDF = Probability Density Function

KDE = Kernel Density Estimates

OVL = Overlapping coefficient

P_E = probability of Entrainment

U_f = random near bed velocities (L/T)

F_{U_f} = cumulative distribution of U_f

f_{U_f} = Probability density function of F_{U_f}

A = Magnitude of total acceleration (mg);

g = Acceleration due to gravity ($1g$)

$\mid CA \mid$ = Absolute acceleration with gravity compensation (mg)

a_x, a_y, a_z = Accelerations on x, y and z axes respectively (mg);

y_t = Observation vector (Kalman Filter Definition)

θ_t = Unobserved vector (Kalman Filter Definition)

V, Ω = Observation and Evolution (process) covariances (Kalman Filter Definition)

p_i = Probabilities

Y_i = Binary variable (Logistic Regression)

X_i = Vector of predictors (Logistic Regression)

β = Vector of regression coefficients (Logistic Regression)

A_f = random total inertial acceleration (L/T^2)

F_{A_f} = cumulative distribution of A_f

f_{A_f} = probability density function of F_{A_f}

List of Tables

TABLE 1. Summary Results

Slope	Threshold acceleration (mg)	Threshold Discharge ($l.s^{-1}$)	Type of position pre-entrainment	Significance of logistic regression model for the calculated probability
0.026	44	6.15	Grain–top	< 2 x10^{-16}
0.037	51	4.3	Grain–top	< 2x10^{-16}
0.044	32	3.21	Saddle	< 2 x10^{-16}
0.057	25	2.2	Saddle	< 2x10^{-16}

List of Figures

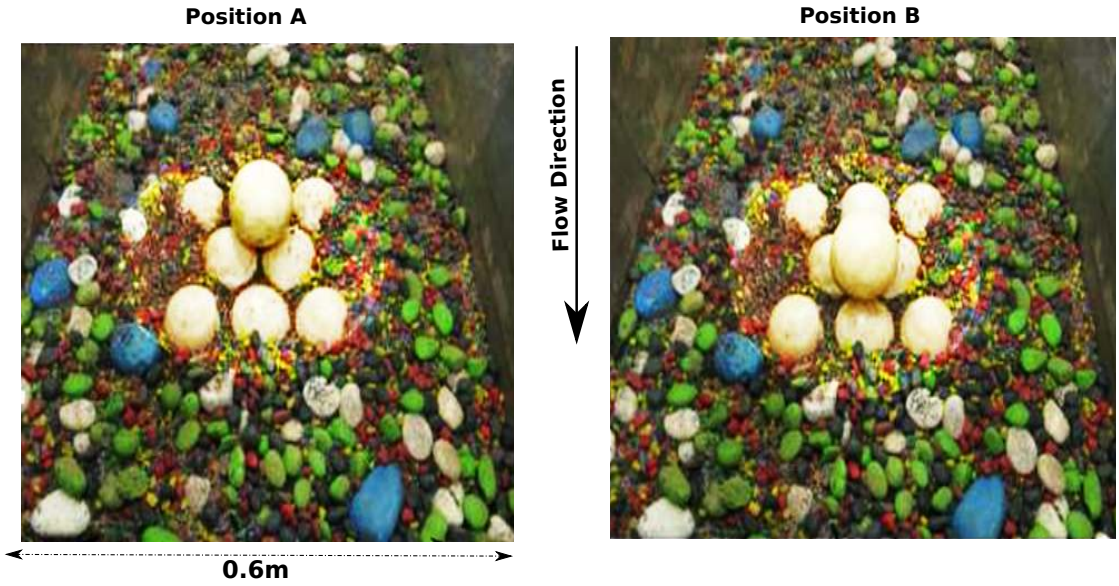

FIG. 1. Description of the experimental setting

The setting represents the theoretical model described in Kirchner et al. (1990) with two characteristic entrainment positions (saddle (A) and grain-top (B)). The physical differences of the sensor compared to natural sediment are its shape and its density. A sphere was used to enable robust determination of sensor dynamics in the inertial frame (requiring accurate definition of the center of mass and locations of points of contact) during calibration.

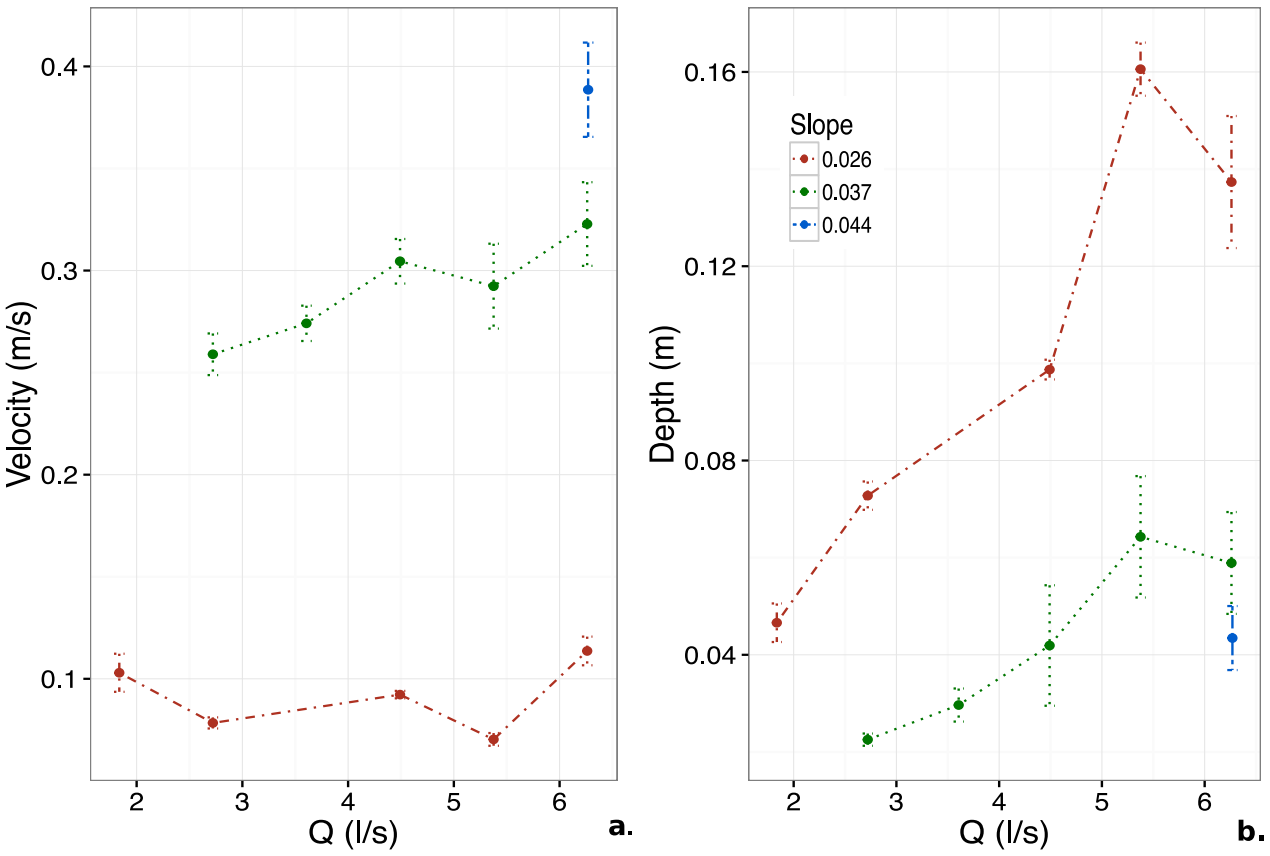

FIG. 2. Description of hydraulic conditions

Near bed flow velocities (a), measured c.0.5m upstream of the bed of hemispheres. Figure 2b. shows water depths at the measurement points on the bed of hemispheres in the absence of the sensor particle. This was possible for the lower two slopes (0.026 and 0.037) at a range of discharges. At slope = 0.044 velocities could only be measured for the highest discharge ($6.25l.s^{-1}$) and depths were too low at slope =0.057 for any velocity measurements. In all cases the lower end of the probe was placed at a distance of 15mm from the bed. The calculated Froude number ($F = u/\sqrt{gd}$) was subcritical, in the range 0.57 to 0.71. The low depth:sediment diameter ratios mean that it was not possible to use a uniform flow approximation of the Shields stress (τ) for cross comparison with the inertial accelerations (Shields 1936). These low ratios also account for some variability in the responses of depth and velocity to increasing discharge as larger roughness elements were progressively drowned out.

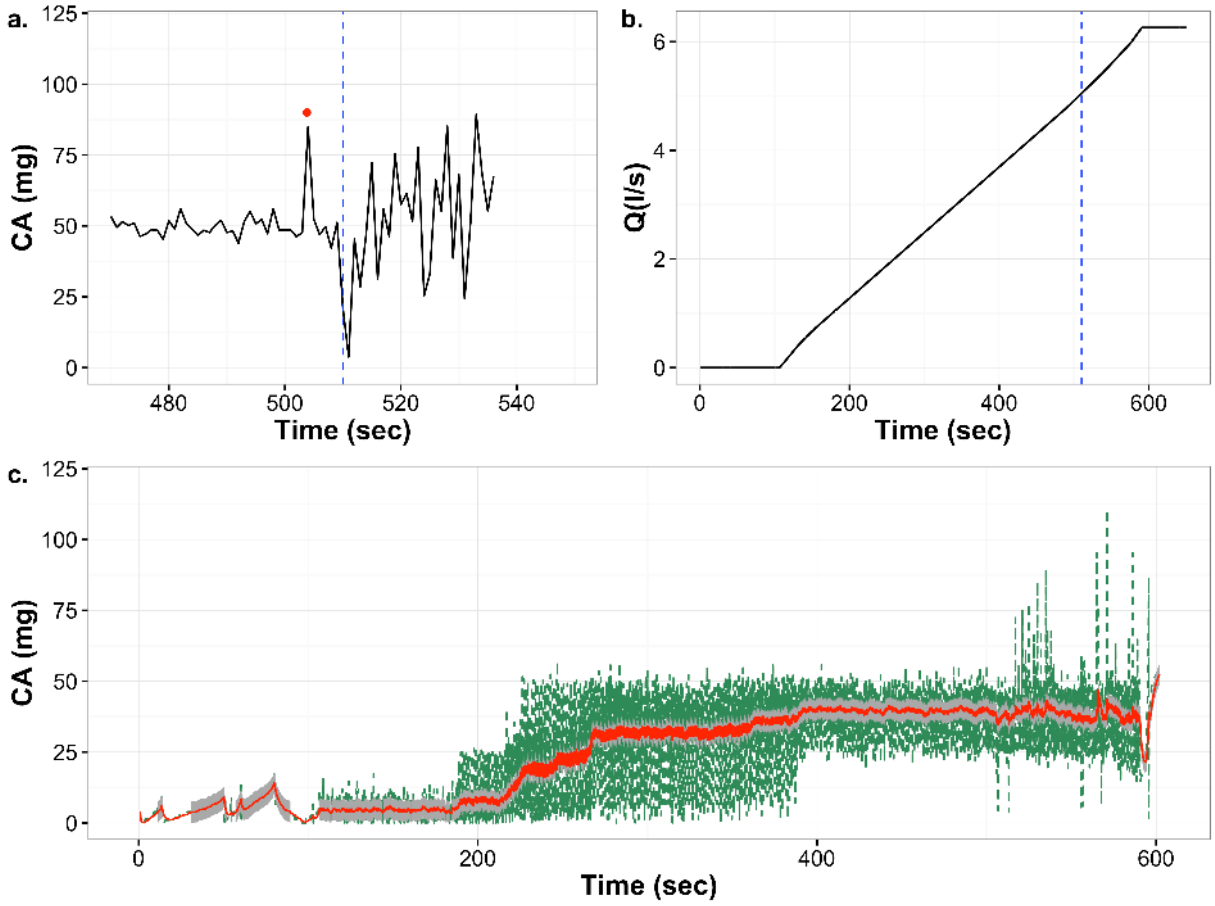

FIG. 3. Definition of the underlying entrainment process

Figure 3a shows the total absolute acceleration close to the point of entrainment from one run at slope = 0.026. The sampling frequency is $4Hz$ and the point of entrainment ($50mg$ at $t = 507sec$, blue-dotted line) is shown as a sudden reduction of the acceleration (dislodging) followed by variable smaller accelerations due to subsequent vibrations. The red dot indicates a distinct pre-entrainment vibration c.$10sec$ before the entrainment event for this experiment. **Figure 3b** shows the corresponding flow increase (steady increase rate of $0.014l.s^{-2}$ up to a maximum rate of $6.25l.s^{-1}$ for all experiments) and the entrainment point from Figure 3a expressed as discharge ($5.17l.s^{-1}$ at $t = 507sec$, blue line). **Figure 3c** shows the summary signal derived by the process described in Equations 4 and 5 for all nine replicates with $slope = 0.026$. The red line is the underlying process as approximated after the application of the Kalman filter (Eq. 5). The grey band shows the process noise which is modelled as a Gaussian distribution with $\sigma^2 = 0.3$.

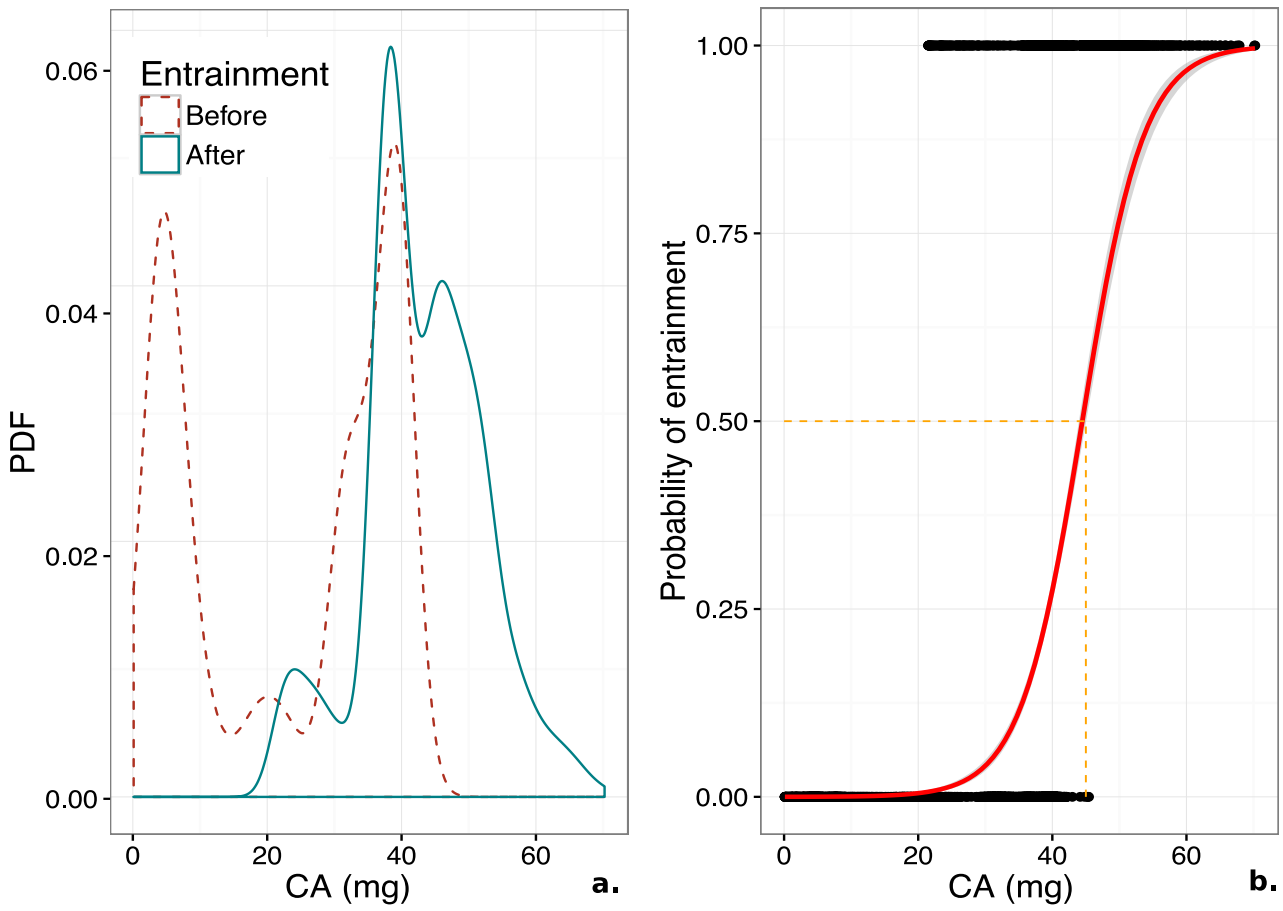

FIG. 4. Calculation of the probability of entrainment

Figure 4a shows the classification of acceleration according to the pre- and post-entrainment conditions ($slope = 0.026$). The probability density of the post-entrainment accelerations shows that the grain is subject to greater forces than pre-entrainment which is consistent with the experimental procedure. **Figure 4b** shows the calculation of the probability of entrainment after the application of the logistic regression model (Equations 5 and 6). The orange-dotted line indicates the acceleration threshold of entrainment ($44mg$ at 0.5 probability). The acceleration value corresponding to 0.5 probability is interpreted as the acceleration where the dynamics start to act in favour of entrainment. The grey band indicates the 95% confidence bands of the logistic regression model.

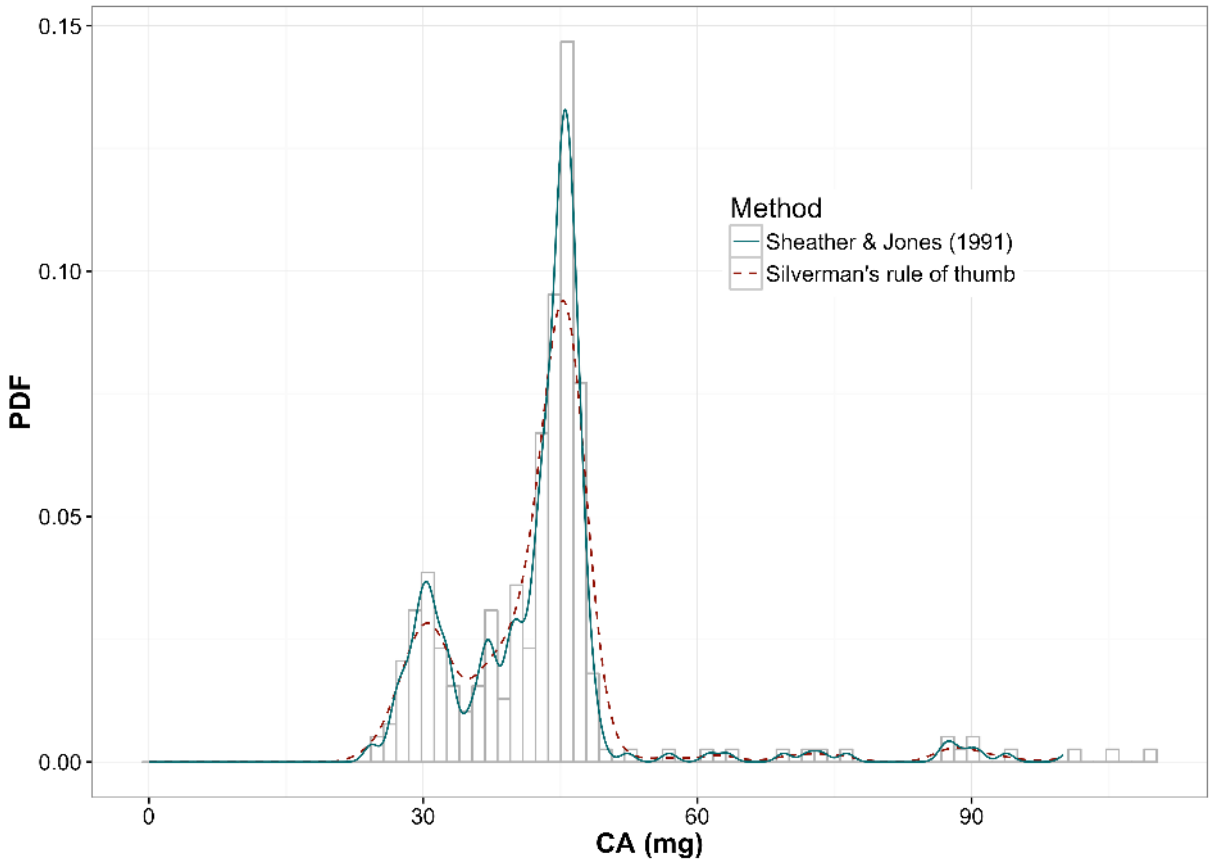

FIG. 5. KDE approximation for inertial acceleration

Figure 5 shows two KDE approximations for the absolute accelerations recorded after the entrainment point for $slope = 0.037$ (histogram). The technique proposed by Silverman (1986) is used for display purposes (Figure 4) while for the calculation of the Overlapping coefficient (Figure 7) the $PDFs$ are approximated using the bandwidth selection method proposed in Sheather and Jones (1991).

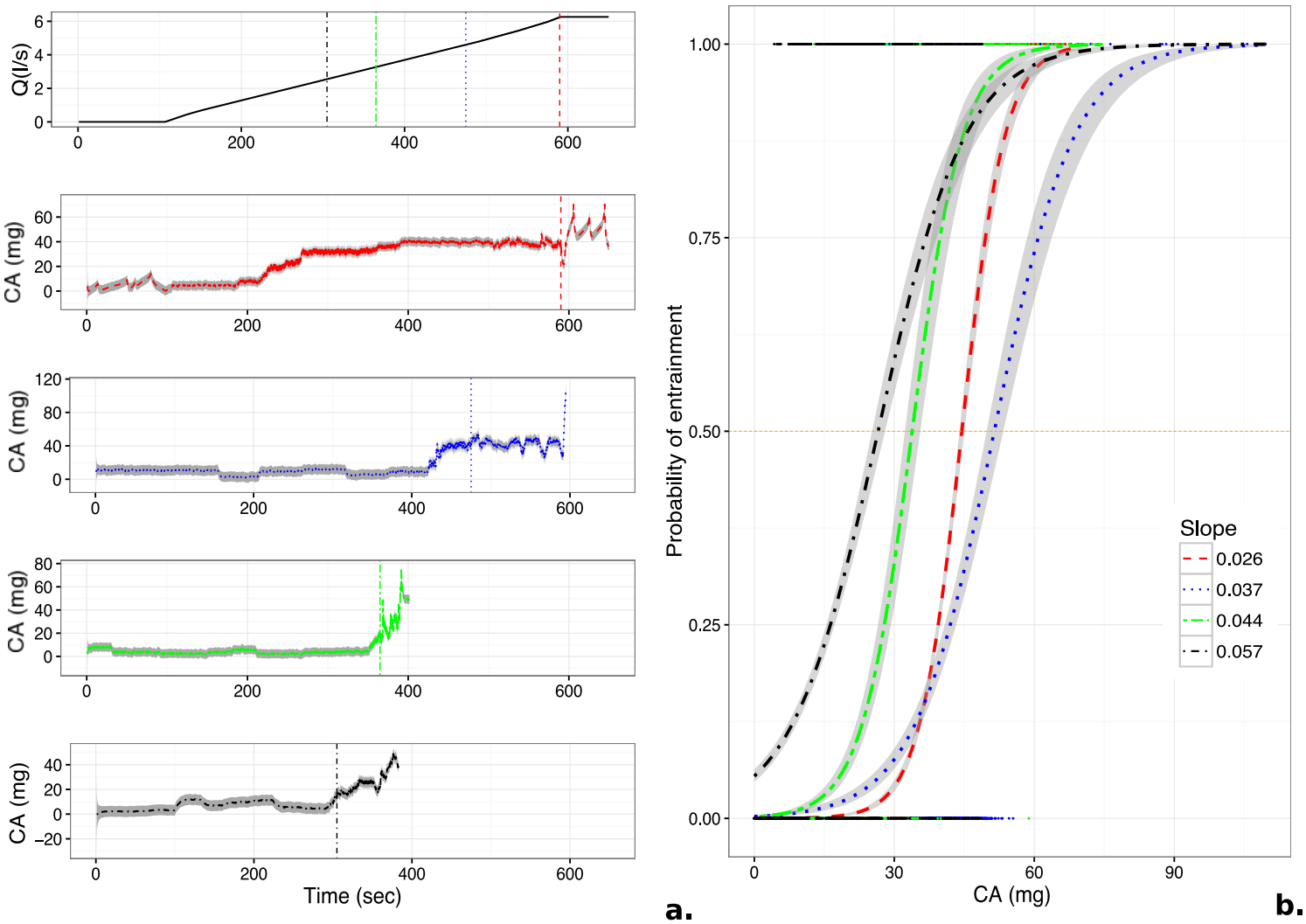

FIG. 6. Summary results

Figure 6a total inertial acceleration and the thresholds of entrainment (vertical lines). The same thresholds are synchronised with the corresponding flow increase which is the same for all experiments (top diagram). Slope changes are colour coded as in Figure 6b. **Figure 6b** shows the calculation of the probability of entrainment for all slopes by logistic regression. Grey bands indicate the 95% confidence bands.

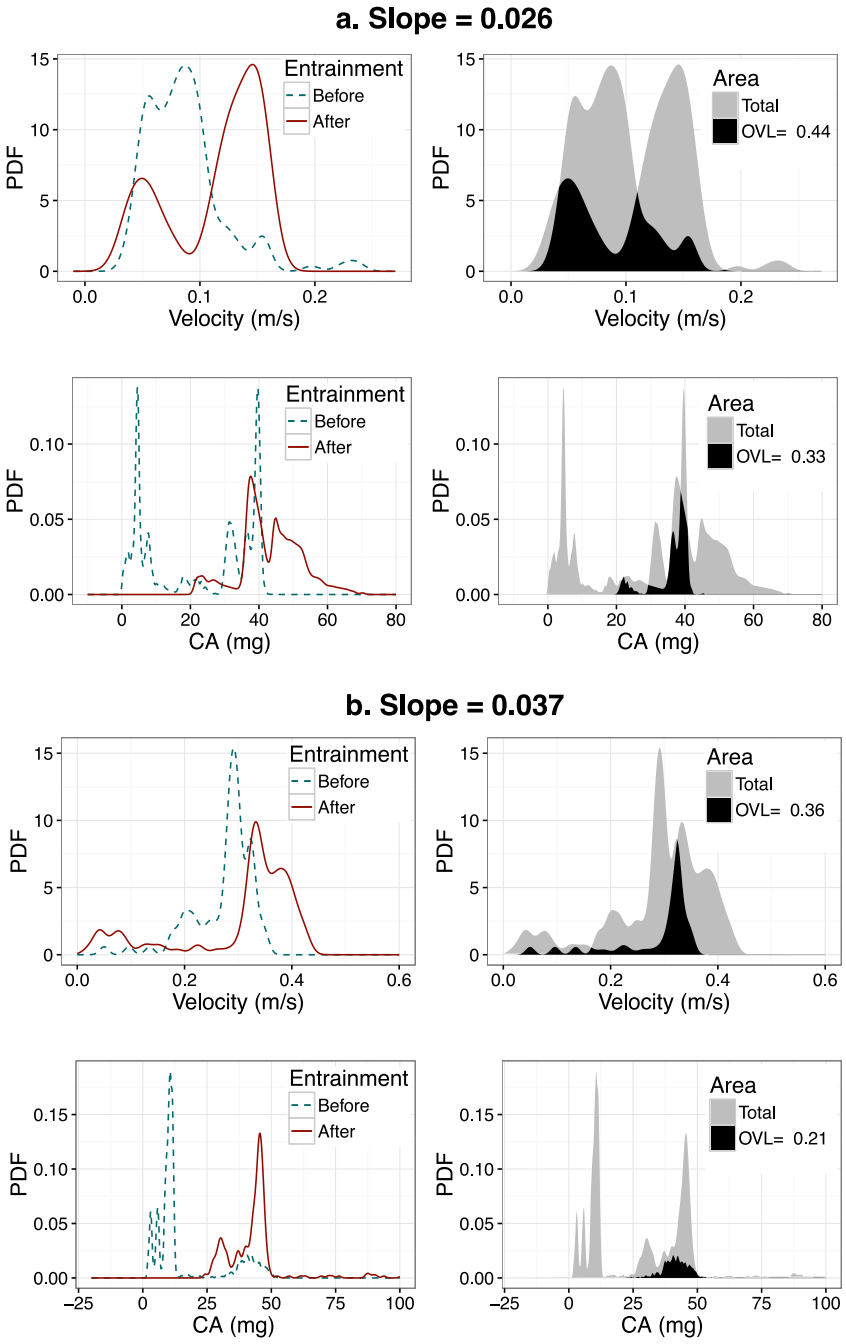

FIG. 7. Comparison of velocity and acceleration

Figures 7a and **b** show pre- and post-entrainment $PDFs$ for two slopes (0.026 and 0.037). Plots on the left side are smoothed PDFs, an the right hand plots show the areas of overlap used to calculate OVL values. The calculation of the OVL coefficient suggests that an incipient motion criterion based on inertial acceleration has the potential two improve prediction as the overlap is reduced at an order of 10% in both cases.