Entrainment of coarse grains in turbulent flows: An extreme value theory approach

Manousos Valyrakis, 1 Panayiotis Diplas, 1 and Clint L. Dancey 1

[1] The occurrence of sufficiently energetic flow events characterized by impulses of varying magnitude is treated as a point process. It is hypothesized that the rare but extreme magnitude impulses are responsible for the removal of coarse grains from the bed matrix. This conjecture is investigated utilizing distributions from extreme value theory and a series of incipient motion experiments. The application of extreme value distributions is demonstrated for both the entire sets of impulses and the maxima above a sufficiently high impulse quantile. In particular, the Frechet distribution is associated with a power law relationship between the frequency of occurrence and magnitude of impulses. It provides a good fit to the flow impulses, having comparable performance to other distributions. Next, a more accurate modeling of the tail of the distribution of impulses is pursued, consistent with the observation that the majority of impulses above a critical value are directly linked to grain entrainments. The peaks over threshold method is implemented to extract conditional impulses in excess of a sufficiently high impulse level. The generalized Pareto distribution is fitted to the excess impulses, and parameters are estimated for various impulse thresholds and methods of estimation for all the experimental runs. Finally, the episodic character of individual grain mobilization is viewed as a survival process, interlinked to the extremal character of occurrence of impulses. The interarrival time of particle entrainment events is successfully modeled by the Weibull and exponential distributions, which belong to the family of extreme value distributions.


1. Introduction

[2] One of the fundamental objectives in Earth surface dynamics and engineering is to obtain a better understanding of the underlying dynamics of the interaction of turbulent flows and the bed surface that contains them, leading to the transport of coarse particles in fluvial, coastal, and aeolian environments. The precise identification of the critical flow conditions for the inception of sediment transport has many applications, ranging from the protection of hydraulic structures against scour to the assessment and regulation of flow conditions downstream of reservoirs to ecologically friendly stream restoration designs [e.g., Whiting, 2002; Lytle and Poff, 2004].

[3] The standard and widely employed method for the identification of incipient motion flow conditions is Shields' critical shear stress criterion. This is partly a reason why considerable effort has been spent to explain deviations from the Shields [1936] empirical diagram as well as to devise alternative plots for a variety of flow and sediment cases [Miller et al., 1977; Mantz, 1977; Yalin and Karahan, 1979; Bettes, 1984; Lavelle and Moffield, 1987; Wilcock and Mc Ardell, 1993; Buffington and Montgomery, 1997; Shvidchenko and Pender, 2000; Paphitis, 2001; Paphitis et al., 2002]. Since then, many researchers have adopted a deterministic perspective [White, 1940; Coleman, 1967; Wilberg and Smith, 1989; Ling, 1995; Dey, 1999; Dey and Debnath, 2000]. However, the comprehensive review of Buffington and Montgomery [1997] shows a scatter of field and laboratory threshold of motion results in excess of an order of magnitude. This, in addition to the subjectivity inherent in precisely defining a threshold for mobilization of sediment grains [e.g., Kramer, 1935; Papanicolaou et al., 2002], shows that a deterministic treatment of the turbulent flow processes leading to particle entrainment based on time- and usually space-averaged criteria does not suffice to accurately describe the phenomenon.

[4] In recognition of the variability of the hydrodynamic forces as well as of the local bed microtopography and grain heterogeneities, many researchers have supported a stochastic approach for the description of incipient movement [Einstein and El-Samni, 1949; Paintal, 1971; Cheng and Chiew, 1998; Papanicolaou et al., 2002; Wu and Yang, 2004; Hofland and Battjes, 2006]. A statistical description of the critical flow conditions by means of probability distributions is necessary because of the wide temporal and spatial variability of the parameters that control it, such as relative grain exposure or protrusion [Paintal, 1971; Fenton and Abbott, 1977; Hofland et al., 2005], friction angle [Kirchner et al., 1990], local grain geometry [Naden, 1987], and methods of estimation for all the experimental runs. Finally, the episodic character of individual grain mobilization is viewed as a survival process, interlinked to the extremal character of occurrence of impulses. The interarrival time of particle entrainment events is successfully modeled by the Weibull and exponential distributions, which belong to the family of extreme value distributions.

and bed surface packing conditions [Dancey et al., 2002, Papanicolaou et al., 2002]. Even for the simplified case of an individual particle resting on a fixed arrangement of similar grains, where the above parameters can be accurately specified, initiation of motion retains its probabilistic nature because of the variability of the near-bed turbulent stresses.

[5] The variability of the local grain configuration and bed surface geometry affects the features of the mechanisms that generate hydrodynamic forces on sediment particles, such as the turbulent flow structures [Schmeeckle et al., 2007] and bed pore pressure fluctuations [Hoiland and Battjes, 2006; Vollmer and Kleinhaus, 2007; Smart and Habersack, 2007]. These factors are essentially manifested in terms of fluctuating drag and lift forces, whose effect could generally be modeled with multivariate distributions. However, analysis of experimental results has shown that hydrodynamic lift has only higher-order effects and that high magnitude and sufficiently sustained drag forces are responsible for the case of particle entrainment by pure rolling [Heathershaw and Thorne, 1985; Valyrakis, 2011]. Then, neglecting such effects allows employing univariate distributions for probabilistic modeling of the phenomenon.

[6] The relevance of high-magnitude positive turbulence stress fluctuations in the vicinity of the boundary to the inception of particle entrainment was emphasized early in the literature [Einstein and El-Samni, 1949; Sutherland, 1967; Cheng and Clyde, 1972]. Recent detailed experiments and analyses have provided strong evidence for the significance of peak hydrodynamic forces for grain entrainment, particularly at low-mobility flow conditions [Hoiland et al., 2005; Schmeeckle et al., 2007; Vollmer and Kleinhaus, 2007; Gimenez-Curto and Corniero, 2009]. However, Diplas et al. [2008] demonstrated via carefully performed experiments that a rather small portion of the peak values results in particle dislodgement and, instead, proposed impulse, the product of force above a critical level and duration, as a more suitable criterion responsible for particle entrainment. Valyrakis et al. [2010] expanded and generalized the validity of the impulse criterion to a wide range of grain entrainment conditions by saltation and rolling. Analysis of experimental data indicates that turbulent force impulse values follow, to a good approximation, the lognormal probability density function [Celik et al., 2010]. However, their empirically derived critical impulse level, defined as the level above which the vast majority of impulses result in grain displacement, corresponds to the upper 3rd to 7th percentile of the entire distribution of impulses, for which the performance of the lognormal distribution is observed to decrease.

[7] Evidently of particular interest is the occurrence of relatively rare and extreme impulse events that are observed to dislodge a particle. In the following sections, impulse theory is reviewed (section 1), and the mobile particle flume experiments are described (section 3). The fundamental theorems and distributions of extreme value theory (EVT) are employed to develop an appropriate probabilistic framework for the stochastic analysis of impulses (section 2). Here the impulse events extracted from a series of experiments are shown to closely follow the Frechet distribution from the family of EVT distributions. This distribution is shown to be directly linked to a power law relation for the frequency and magnitude of occurrence of the impulses. The applicability of the peaks over threshold method is demonstrated for extracting the conditional exceedances of impulse data. For a sufficiently high impulse level the generalized Pareto distribution (GPD) is fitted to the distribution of the excess impulses. Finally, the response of an individual particle under different flow conditions is analyzed stochastically by means of reliability theory. The Weibull and exponential distributions are utilized to model the time between consecutive particle entrainments.

1.1. Impulse Theory

[8] Impulse is one of the fundamental physical quantities used to describe transfer of momentum. Bagnold [1973] was among the first researchers who employed the concept of “mean tangential thrust” to define the mean flow conditions required to sustain suspension of solids in the water column. Incipient displacement of a particle by rolling has been traditionally treated using a moments or torques balance [White, 1940; Coleman, 1967; Komar and Li, 1988; James, 1990; Ling, 1995], which also describes transfer of flow momentum to the particle. However, these static approaches refer to time-averaged quantities, thus being unable to incorporate the fluctuating character of turbulence. Recently, Diplas et al. [2008] introduced impulse \( I_t \) as the relevant criterion for the initiation of sediment motion. According to this concept, impulse is defined as the product of the hydrodynamic force \( F(t) \), with the duration \( T_i \), for which the critical resisting force \( F_{cr} \) is exceeded (Figure 1):

\[
I_t = \int_{t_i}^{t_i+T_i} F(t) dt, \quad F(t) > F_{cr}, \quad t_i < t < t_i + T_i.
\] (1)

The proposed criterion accounts for both the duration and the magnitude of flow events, introducing a dynamical perspective for the incipient entrainment of coarse particles. Valyrakis et al. [2010] provided a theoretical framework for the incipient saltation and rolling of individual particles by impulses of varying magnitude and duration, validated by bench top mobile particle experiments in air. They derived isomplitude curves corresponding to different particle responses ranging from incomplete movement (twitches) to energetic particle entrainment. In accordance with this theory, only impulses above a critical level \( I_{cr} \), which is a function of particle properties and local pocket geometry, are capable of the complete removal of a particle out of its resting position.

[9] An example illustrating the importance of duration in addition to the magnitude of a flow event is depicted in Figure 1 for the case of entrainment of a fully exposed grain by rolling (Figure 2). Two separate flow events with varying magnitude and duration and consequently different potential for momentum exchange impinge upon the particle under consideration. The instantaneous hydrodynamic force, parameterized with the square of the streamwise velocity, of the first flow event \( i \) peaks higher than the second flow event \( i + 1 \). However, the former is significantly more short-lived than the latter (\( T_{i+1} > T_i \)). According to equation (1), the integral of the hydrodynamic force over the
duration of the flow event (highlighted regions in Figure 1) is greater for the later impulse ($I_{i+1} > I_i$), leading to a more pronounced response of the particle [Valyrakis et al., 2010]. If this impulse exceeds the critical impulse level ($I_{i+1} > I_{cr}$), then the particle will be fully entrained (denoted by the vertical dotted line in Figure 1). Thus, before the probabilistic framework of impulse and grain entrainment is attempted, the definition of $u_{cr}$, used in extracting impulse events, as well as the theoretical impulse level for complete entrainment $I_{cr}$ has to be provided.

1.2. Detection of Impulse Events and Determination of $I_{cr}$

[10] As opposed to the traditional incipient motion identification techniques, the impulse concept provides an event-based approach, accounting for the dynamical characteristics of flow turbulence. In order to examine the statistical properties and distributions of impulse events and their effects on entrainment of coarse grains, the method employed for their identification must first be described. In the following, the applicability of the scheme proposed to extract impulse events and implementation of the theoretically derived critical impulse level for incipient rolling are critically reviewed.

[11] Typically, a simplified tetrahedral arrangement of well-packed spherical particles (Figure 2) is considered. Of interest is the response of the exposed particle, which is a function of the hydrodynamic and resisting forces, assumed to act through its center of gravity. Loss of initial stability may occur as a result of an impulse event imparting sufficient momentum to the particle. Over the duration of this event the sum of drag ($F_D$), lift ($F_L$), and buoyancy ($B_f$) force components along the direction of particle displacement, exceeds the corresponding component of particle’s weight $W$:

$$F_D \sin(\theta_0 - \alpha) + (F_L + B_f) \cos(\theta_0 - \alpha) \geq W \cos \theta_0,$$

where $\theta_0$ is the pivoting angle, formed between the horizontal and the lever arm ($L_{arm}$ in Figure 2) and $\alpha$ is the bed slope [Valyrakis et al., 2010]. Equation (2) describes the static equilibrium of forces or, equivalently, torques about the axis of rotation located at the origin of the polar coordinate system ($D'$ in Figure 2). Usually the effect of lift force for entrainment of completely exposed particles has been neglected without significant error [Schmeeckle and Nelson, 2003]. Inclusion of the hydrodynamic mass coefficient, $f_h = \left(\rho_s - \rho_f (1 - C_m)\right) / \left(\rho_s - \rho_f \right)$, where $C_m$ is the added mass coefficient (equal to 0.5 for water [Auton, 1988]), $\rho_f$ is the density of fluid, and $\rho_s$ is the particle’s density, increases the effect of the submerged particle’s weight, and equation (2) becomes

$$F_D \sin(\theta_0 - \alpha) \geq f_h W_s \cos \theta_0,$$

where $W_s = (\rho_s - \rho_f) V g$, is the submerged particle’s weight (assuming uniform flow), $V$ is the particle’s volume, and $g$ is the gravitational acceleration. Equations (2) and (3) may be solved for the critical drag force, considering the equal sign, to define the minimum level above which impulse events capable of dislodging a particle occur ($F_{cr}$ in Figure 1). For steady flows it is customary to parameterize the instantaneous drag force with the square of the streamwise local velocity component upstream of the exposed particle [e.g., Hofland and Battjes, 2006]. Then it
is convenient to define the critical flow conditions directly in terms of the square of the local flow velocity:

\[
\frac{u^2}{u^*} = \frac{2}{\rho_f C_D A} f_i W_i \frac{\cos \theta_0}{\sin (\theta_0 - \alpha)},
\]

(4)

where \( A \) is the particle’s projected area perpendicular to the flow direction and \( C_D \) is the drag coefficient, assumed here to be equal to 0.9. It can be shown that equation (4) is similar to the stability criterion suggested by Valyarakis et al. [2010] (if \( f_i = 1 \), by neglecting \( C_m \) at this stage) and identical to the critical level proposed by Celik et al. [2010] (after the appropriate algebraic and trigonometric manipulations are performed).

All of the detected flow impulses have the potential to initiate particle displacement. However, Valyarakis et al. [2010] predict complete removal of the exposed particle from its local configuration by rolling only when impulses exceed a theoretical critical impulse level, defined as the product of duration \( T_{roll} \) of the impulse event with the characteristic drag force \( F_D \) assumed constant over the duration of the flow event:

\[
I_{cr} = F_D T_{roll} = F_D \left[ \frac{L_{lam} (\frac{1.5}{\rho_f} + \rho_i C_m) V}{F_D \cos (\theta_0 - \alpha) + W_i \sin \theta_0} \right] \cdot \arcsin \left( \sqrt{2W \rho_f} \sqrt{F_D \cos (\theta_0 - \alpha) + W_i \sin \theta_0} \right),
\]

(5)

with

\[
\rho_i = (1 - \sin \theta_0) \left[ \frac{\cos \theta_0}{1 - \sin \theta_0} \sin \alpha + \left( \cos \alpha - \frac{\rho_f}{\rho_i} \right) \right]
\]

being a coefficient incorporating the effects of initial geometrical arrangement and the relative density of fluid and solid grain. Equation (5) is derived from the equation of motion of a rolling particle and has been validated for a range of particle arrangements via a series of laboratory experiments. This theoretical level corresponds to impulse events extracted using equation (2), neglecting hydrodynamic lift.

Application of equation (5) allows for the a priori determination of physically based impulse levels \( I_{cr} \), as opposed to their empirical estimation [Celik et al., 2010], which requires experimental identification of the impulses leading to entrainment. The two methods return equivalent results if the impulse values obtained by means of the former method are multiplied with an appropriate impulse coefficient \( C_f \). For the typically encountered cases of water flows transporting grains of specific density ranging from 2.1 to 2.6, \( f_i \) is close to 1.4, and \( C_f \) is determined to have a value of approximately 0.5 (by matching the theoretical and experimentally defined critical impulse levels).

Even though for an individual particle and local bed surface arrangement the critical conditions for entrainment can be deterministically defined, the randomness of turbulent flow forcing renders a statistical description of the critical flow conditions more meaningful. Development of a complete and reliable probabilistic theory for inception of grain entrainment requires consideration of the impulse theory together with appropriate statistical distributions that account for the intermittent character of the modeled phenomenon. In section 2 a stochastic framework for the accurate identification of the probability of entrainment of coarse grains for low-mobility flow conditions is considered.

2. Stochastic Modeling of Impulses

Many researchers have recently implemented and stressed the need for a stochastic approach to the initiation of sediment entrainment due to the action of near-bed turbulence [e.g., Dancey et al., 2002; Papanicolaou et al., 2002]. Central to such models is the assumption that the episodic removal of an individual particle from the bed surface is strongly linked to the occurrence of turbulent stresses exceeding a critical level (e.g., Figure 3a). Here, contrary to the past stochastic approaches, turbulence is treated as a discrete point process, where separate flow structures of varying magnitude and duration are modeled as impulse events \( I \) occurring at random instances in time (Figure 3b). Similarly, the sequence of conditional impulse exceedances \( (\xi = I - I_{thr}) \) above a certain threshold \( I_{thr} \) describes the point process of peak impulses. Here \( I_{thr} \) (not

![Figure 3. Representation of hydrodynamic forcing history on a coarse particle: (a) as a continuous time series of the drag force \( F_D \sim u^2 \) and (b) as a discrete point process of corresponding impulse events and impulse exceedances \( \xi \) (denoted by the thick vertical lines above \( I_{thr} \)) randomly occurring in time. The threshold impulse level \( I_{thr} \) (<\( I_{cr} \)) and impulse events associated with particle displacement (stars) are also shown.](Image)
to be confused with $I_{cr}$, which depends on the grain and local microtopography parameters (refers to the impulse level above which the tail of the distribution of impulses is defined (e.g., about 90% quantile of the distribution of flow impulse events; see section 4.3.1).

[16] The probability of particle entrainment $P$ may be approximated by the probability of occurrence of impulses in excess of the theoretically defined (e.g., equation (5)) critical level, $P_{E} = P(I_{i} > I_{cr})$. This concept is shown in Figure 4, where the very infrequent occurrence of such events, especially near threshold conditions, is evident. Since of interest are largely the extreme values of the distribution, its tail (region $I_{i} > I_{thr}$) may be modeled separately. Then the probability of particle entrainment may be found from the conditional probability that the critical impulse level is surpassed, $P(\xi_{i} > I_{cr} - I_{thr} | I_{i} > I_{thr})$. Thus, it is important to find statistical distributions that accurately model the magnitude and extreme character of impulses and their conditional exceedances for low-mobility flow conditions. For this purpose, distributions from EVT are considered to provide an appropriate statistical tool.

2.1. Extreme Value Modeling

[17] Mobile particle flume experiments discussed by Diplas et al. [2008] revealed the significance of high-magnitude impulses for grain entrainment. In their work it was first observed that only a few of the most extreme impulses, those that exceed an empirically defined critical level, result in particle entrainment. These peak impulses represented a small portion, about 4.4%, of the entire sample and belong to the upper tail of the distribution of impulses. Celik et al. [2010] proposed that impulses follow the log-normal distribution. A good fit is visually observed for the core of the distribution, in the 1–2.5 range of normalized impulses ($I = I/I_{mean}$, with $I_{mean}$ being the sample’s mean [Celik et al., 2010, Figure 9]). However, particularly for $I > 3$, the tail of the lognormal distribution falls quite faster than the distribution of the sample. The relatively high values of reported parameters such as the skewness and flatness support the observation that the impulse distribution has a heavy tail. It is also noted that the vast majority of impulses leading to particle displacement have values above an empirically defined critical level. Careful examination shows that for most experiments this level corresponds to $I = I_{cr}/I_{mean} > 2.5$ (e.g., using $I_{cr} = 0.0063$ and $I_{mean}$ values from Table 1). This implies that the lognormal may not be the most suitable distribution in the range of interest, which may also affect the accuracy of the probability of particle entrainment estimations. EVT provides a flexible stochastic framework with the potential to model impulse events more accurately because it has the ability to capture the extremal character of turbulence-particle interactions for near-threshold flow conditions.

2.2. Generalized Extreme Value Distribution (GEV)

[18] EVT provides representative distributions that model the stochastic character of extreme values from the sequence of impulses $I_{i}$ assumed to be independent and identically distributed (IID) [Gumbel, 1958]. The generalized extreme value (GEV) distribution unites the three types of extreme value distributions into a single family, allowing for a continuous range of possible shapes with a cumulative distribution function [e.g., Kotz and Nadarajah, 2000]:

$$F_{GEV}(x) = \begin{cases} \exp\left(-\frac{1 + \gamma(x - \mu) / \sigma}{\gamma}\right), & \gamma \neq 0, \\ \exp\left(-e^{-(x - \mu) / \sigma}\right), & \gamma = 0, \end{cases}$$

(6)

where $\gamma$ is the shape parameter (determines the type of the extreme value distributions), $\sigma$ is the scale parameter, and $\mu$ is the location parameter. For $\gamma = 0$, it corresponds to the Gumbel (type 1) distribution, for $\gamma > 0$, it corresponds to the Frechet (type 2) unbounded distribution, and for $\gamma < 0$ it is the Weibull (type 3) distribution with an upper bound. Application examples are given in sections 4.1 to 4.3 to illustrate the utility of these distributions.

2.3. Generalized Pareto Distribution (GPD)

[19] The generalized Pareto distribution is an additional family of EVT distributions. It is used to model the distribution of exceedances above a threshold and has been widely applied to a broad range of fields ranging from finance to environmental engineering and engineering reliability [Gumbel, 1954; Ashkar et al., 1991]. For the case of flow impulses, modeling the distribution of extreme values (maxima) separately is of particular interest, considering that common models may be biased in the right tail due to the relatively lower density of data.

<table>
<thead>
<tr>
<th>$I_{mean}$ (m s$^{-1}$)</th>
<th>$\tau$</th>
<th>$Re_{p}$ (m$^{2}$ s$^{-1}$)</th>
<th>$I_{thr}$</th>
<th>$f_{E}$ (events s$^{-1}$)</th>
<th>$f_{E}$ (entrainment s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1 0.248 0.007 424.18 0.0024 2.20 0.147</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2 0.243 0.007 412.75 0.0023 1.73 0.114</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E3 0.238 0.006 397.51 0.0021 1.31 0.051</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E4 0.230 0.006 384.81 0.0021 0.80 0.031</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E5 0.228 0.005 377.19 0.0022 0.33 0.012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E6 0.218 0.005 364.49 0.0019 0.47 0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Summary of Flow Characteristics for Mobile Particle Flume Experiments
According to the limit probability theory, the GPD is the appropriate distribution for exceedances \((I_i > I_{th})\), as it always fits asymptotically the tails of conditional distributions in excess of a sufficiently large threshold \((I_i > I_{th})\) [Pickands, 1975]. The GPD is a right-skewed distribution parameterized by shape \((\gamma_{GPD})\) and scale \((\sigma_{GPD})\) parameters, with probability density function

\[
f_{\text{GPD}}(I_i > x | I_i > I_{th}) = \frac{1}{\sigma_{GPD}} \cdot \left[1 + \frac{\gamma_{GPD}(x - I_{th})}{\sigma_{GPD}}\right]^{-\frac{1}{\gamma_{GPD}}}, \gamma_{GPD} \neq 0.
\]  

Equation (7) provides an accurate representation of the tail of the distribution provided that the exceedances are statistically independent and the selected threshold is sufficiently high. Similar to the GEV, GPD is classified to the Frechet (type 2) and the Weibull (type 3) distributions on the basis of the shape or tail index for \(\gamma_{GPD} > 0\) and \(\gamma_{GPD} < 0\), respectively. For \(\gamma_{GPD} = 0\), GPD becomes the two-parameter exponential distribution:

\[
f_{\text{GPD}}(I_i > x | I_i > I_{th}) = \exp\left[-\frac{(x - I_{th})}{\sigma_{GPD}}\right].
\]  

While the shape parameter for the GEV and GPD has identical meaning and value, the scale parameters are interlinked with the threshold according to \(\sigma_{GPD} = \sigma + \gamma_{x_{th}}\) [Coles, 2001]. The relation between the cumulative distribution functions of the two distributions is \(F_{\text{GPD}} = 1 + \ln(F_{\text{GEV}})\). GPD provides an adequate model, assuming that the threshold as well as the number of exceedances is sufficiently high, so that the asymptotic approximation of the distribution is not biased and is accurately estimated. The peaks over threshold (POT) method proposed by Davison and Smith [1990] is utilized to extract excess impulses above an appropriate threshold and fit the GPD model to the tail of impulses distribution.

Extreme impulses extracted using the block maxima method (where the time series is split into blocks from which the maximum value is obtained [e.g., Coles, 2001]) could also be modeled by attempting to fit them to the GEV distribution. However, such modeling is not directly applicable for the phenomenon under investigation because extreme impulses do not occur at regular, easy to identify intervals. To the contrary, GPD utilizes only the peak impulses in excess of a high, but below critical, impulse threshold \((I_{th} < I_{cr})\). This renders GPD ideal for modeling the tail of distribution of impulses for low-mobility conditions since for such flow conditions, \(I_{cr}\) is relatively large, allowing for a sufficiently high choice of \(I_{th}\) without biasing the distribution. The utility of GPD is demonstrated through the application of a threshold-excess method after the description of the experimental method and setup employed to obtain a series of sample impulse distributions.

### 3. Description of Setup and Experimental Process

Results from a series of incipient motion experiments (A. O. Celik, personal communication, 2009, see also Diplas et al. [2010]) were used to provide synchronous time series of the local flow velocity and particle position over a range of flow conditions. For completeness a summary of the experimental setup and conditions is provided below. Incipient motion experiments were performed to obtain coupled data for the entrainment of a fully exposed Teflon® (specific gravity of 2.3) spherical particle in water \((\rho_f = 1000 \text{ kg m}^{-3})\). The test section is located about 14.0 m downstream from the inlet of the 20.5 m long and 0.6 m wide flume to guarantee fully developed turbulent flow conditions. The sphere (12.7 mm diameter) rests on top of two layers of fully packed glass beads of the same size, forming a tetrahedral arrangement (Figure 5). The bed slope \(\alpha\) remains fixed at 0.25% for all of the conducted experiments. The series of conducted experiments refer to uniform and near-threshold to low-mobility conditions. For those flow conditions the use of data acquisition techniques that do not interfere with the flow renders possible the identification of the impulse events as well as entrainment instances, with greater accuracy [Diplas et al., 2010].

The motion of the mobile sphere is recorded via a particle tracking system composed of a photomultiplier tube (PMT) and a low-power (25–30 mW) He-Ne laser source. As seen in Figure 5, the He-Ne laser beam is aligned to partially target the test particle. Calibration of the setup showed that the angular dislodgement of the targeted particle is a linear function of the signal intensity of the PMT, which changes proportionally to the light received. A continuous series of entrainments is made possible because of a

![Figure 5](image-url)
restraining pin located 1.5 mm downstream of the mobile sphere (Figure 5), which limits the maximum dislodgement of the grain to the displaced position. The grain will not be able to sustain its new location for long and will eventually fall back to its initial position after the flow impulses are reduced below a certain level, without a need to interrupt the experiment to manually place the sphere back to its resting configuration.

[24] The time history of the streamwise velocity component one diameter upstream of the particle and along its centerline \( u(t) \) is obtained by means of laser Doppler velocimetry (4 W Argon ion LDV) at an average sampling frequency of about 350 Hz (Figure 5). These measurements are obtained simultaneously with the displacement signal, employing a multichannel signal processor. Utilizing equations (1) and (4), impulse events of instantaneous hydrodynamic forces exceeding a critical value can be extracted from the time series of \( F_D = f(u^+) \).

[25] A series of experiments (E1–E6) were carried out, during which coupled measurements of flow intensity and particle response were recorded for different low-mobility flow conditions. For each of the experimental runs the flow conditions were stabilized to achieve a constant rate of particle entrainment \( f_j \) over long durations (about 2 h). Impulses are extracted from the about 15 min long time series of the local flow to allow for their statistical representation. The main flow and grain response characteristics such as particle Reynolds number, \( Re_p \), are shown in Table 1 for each experimental run. All of these experiments refer to near-incipient motion conditions of about the same mean local velocity \( u_{\text{mean}} \), dimensionless bed shear stress \( \tau^* \), and turbulent intensity (equal to 0.27). Contrary to the aforementioned flow parameters, which remain relatively invariant, the rate of occurrence of impulses, \( f_i \), and \( f_j \), change more than an order of magnitude (Table 1). Thus, estimation of the mean rate of particle mobilization is less sensitive if based on \( f_j \) compared to using any of the above traditional flow parameters. In section 4 the relationship between the flow impulses and grain response is further explored under a probabilistic context.

4. Analysis and Results

4.1. Frequency-Magnitude Relationship

[26] Preliminary analysis of the impulse data obtained experimentally for near-threshold flow conditions showed that extreme events of both high magnitude and relatively low frequency of occurrence are linked to the instances of particle dislodgement. Thus, it is of interest to establish a relationship between the frequency of occurrence of extreme impulse events and their magnitude. To this purpose, the mean number of impulses per second \( f_i(l_i) \) above a certain level \( l_i \) is found for practically the whole range of magnitude of impulses for all experimental runs (E1–E6, corresponding to \( j = 1–6 \)). If the obtained pairs \((l_i, f_i)\) are plotted on a log-log scale, they are observed to closely follow an almost-straight line (Figure 6). This behavior is strongly indicative of power law dependence between the two variables:

\[
f_i(l_i) = b_j l_i^{-a_j},
\]

where \( a_j > 0 \) is the power law exponent and \( b_j \) is the base coefficient for a certain flow condition defined by index \( j \).

Figure 6. Variation of magnitude of impulses normalized with the critical impulse level \((l_i/l_{ci})\) with the mean frequency of their occurrence \( f_j \) (for run E3) and fitted power law relationship (the lognormal fit is also shown for comparison).

[27] Equation (9) is fitted to the sample pairs, for each experiment, to acquire the values \((a_j, b_j)\) that parameterize it (Table 2). For the range of flow conditions examined here the frequency of occurrence and magnitude of impulse events obey a power law relationship to a very good approximation, as confirmed by the high values of the coefficient of determination \((R^2 \sim 0.92–0.98\), Table 2). Here the scaling region, defining the range of applicability of power law, spans virtually the whole distribution of impulses. The value of the exponent remains almost constant \((a = 1.82 \pm 0.10)\) for the different experiments. The value of the coefficient \( b_j \) obtains higher values with increasing flow strength \((j \text{ from } 1 \text{ to } 6)\), implying that the peak impulses become more frequent. Comparison between the various flow conditions is facilitated by normalizing the impulse values with the critical impulse level, \( l_{ci} = 0.0063 \). Use of the normalized impulses changes the value of base coefficient to \( b_{n,j} = b_j l_{ci}^{-a_j} \), while the exponent remains the same (Table 2). The effect of decreasing flow conditions (from E6 to E1) on the relation between frequency and magnitude of normalized impulses is clearly demonstrated in Figure 7. Such a representation is of practical importance and predictive value since it directly provides the expected frequency of occurrence of impulses at the critical \((l_i/l_{ci})\) or multiples of it levels, which are of interest for particle entrainment.

[28] Power law models are attractive since they have the ability to describe a wide range of scale-invariant phenomena in Earth sciences \([\text{Schroeder, } 1991; \text{ Turcotte, } 1997]\),

Table 2. Summary of Power Law and Frechet Parameters Characterizing the Magnitude-Frequency Relation and Distribution of Impulses

<table>
<thead>
<tr>
<th>Run</th>
<th>( \alpha = \gamma )</th>
<th>( b )</th>
<th>( b_n )</th>
<th>( \sigma_n )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>1.73</td>
<td>0.000014</td>
<td>0.094</td>
<td>0.26</td>
<td>0.93</td>
</tr>
<tr>
<td>E2</td>
<td>1.75</td>
<td>0.000010</td>
<td>0.070</td>
<td>0.22</td>
<td>0.94</td>
</tr>
<tr>
<td>E3</td>
<td>1.79</td>
<td>0.000006</td>
<td>0.053</td>
<td>0.19</td>
<td>0.95</td>
</tr>
<tr>
<td>E4</td>
<td>1.69</td>
<td>0.000007</td>
<td>0.036</td>
<td>0.14</td>
<td>0.98</td>
</tr>
<tr>
<td>E5</td>
<td>1.92</td>
<td>0.000001</td>
<td>0.013</td>
<td>0.10</td>
<td>0.95</td>
</tr>
<tr>
<td>E6</td>
<td>1.83</td>
<td>0.000005</td>
<td>0.020</td>
<td>0.12</td>
<td>0.92</td>
</tr>
</tbody>
</table>
from the occurrence of rare natural hazards such as earthquakes [Bak and Tang, 1989; Rundle et al., 1997] to the self-similarity of channel networks and corresponding energy and mass distribution [Rodriguez-Iturbe et al., 1992; Rodriguez-Iturbe and Rinaldo, 1997]. Their superiority compared to more sophisticated models has been also demonstrated for the prediction of bed load transport rates [Barry et al., 2004]. Here the power law characterizes the momentum and energy contributed by flow structures toward particle entrainment.

4.2. Impulse Distribution

[29] In addition to their simplicity and effectiveness in expressing the frequency of high-magnitude impulses resulting in particle dislodgement, it may be shown that the proposed power law relation is statistically associated with the Frechet distribution [Leadbetter et al., 1983]. Assuming that the IID impulse events arrive according to a Poisson process with a mean frequency $f$, then the probability that no impulses of magnitude greater than $I_i$ occur in the unit time (s) is $F_j(I_i) = \exp[-f(I_i)]$ or, using equation (9),

$$F_j(I_i) = e^{-b_i^{1/\xi}}; \quad I_i > 0.$$  \hspace{1cm} (10)

which is the cumulative density function of the Frechet distribution as a special case of the GEV (equation (6)) with shape parameter $\gamma = a_i$ and scale parameter $\sigma = b_i^{1/\xi}$ (or $\sigma_n = b_n^{1/\xi}$ if $I_i$ is normalized with $I_{cr}$; Table 2).

[30] Equation (10) together with the lognormal distribution are plotted against the data sample for one of the experimental runs for comparison purposes (Figure 8). Generally, both distributions are seen to have a good overall fit to the sample impulses. However, the Frechet distribution has a relatively heavier tail behavior compared to other distributions such as the lognormal [Mitzenmacher, 2003]. This may easily be observed in the log-log representation of normalized magnitude and frequency of impulses (Figure 6), where the region of extreme impulses (normalized with $I_{cr}$) is emphasized. It is shown that the logarithmic distribution generally underestimates the frequency of occurrence of extreme impulses (Figure 6) or their probability (Figure 8) as opposed to the power law and corresponding Frechet distribution, respectively.

4.3. Distribution of Conditional Excess Impulses

[31] In sections 4.1 and 4.2 the extremal character of distribution of impulses has been demonstrated. Since out of the whole distribution it is rather the peak impulses that are linked to grain mobilization, separate modeling of the tail of the distribution is appropriate. Here, because of the heaviness of the tail, the POT method is employed to extract the conditional impulses $\xi_i$ in excess of a threshold level $I_{thr}$ and to fit the GPD. Since POT is threshold dependent, guidelines justifying the choice of $I_{thr}$ are provided. The GPD parameters are evaluated for the range of examined flow conditions, using different methods of estimation, and the model’s performance is accessed.

4.3.1. Application of POT

[32] For design applications, it is of interest to define the near-critical flow conditions for certain gravel properties and bed surface arrangements. The proximity to the critical flow conditions may be measured by the probability of grain entrainment $P_{Ei}$, which can be estimated through the conditional probability of impulse exceedances $P(\xi_i > I_{thr} | I > I_{thr})$.

[33] Evidently, the goodness of fit as well as appropriateness of the GPD model depends on the choice of threshold value. A very low threshold may utilize more data, but it may violate the assumption of the asymptotic nature of the model, biasing the distribution. On the contrary, a very high threshold will increase statistical noise because of the high variability of the extreme values, affecting the accuracy of estimated parameters. In the case when the sample size of conditional excess impulses is large (e.g., several hundreds of data points) for high-impulse quantiles, the sensitivity of the method to the threshold selection is not high. However, for flow conditions very close to critical, the sample of $\xi_i$ becomes relatively small, and the selection of an optimal threshold with which both statistical certainty and accurate parameter estimation are achieved requires further investigation.

[34] In practice, graphical diagnostic tools such as the mean excess over threshold plot are commonly employed in estimating a suitable $I_{thr}$ [e.g., Davison and Smith, 1990].
This graph depicts the pairs of the threshold impulse and corresponding mean excess over threshold function, \( \{I_{thr}, e_i(I_{thr})\} \), for a range of threshold values. The mean excess over threshold function \( e_i(I_{thr}) \) is defined as the ratio of the sum of impulses in excess of the threshold \( I_{thr} \) over the number of those exceedances \( n_{thr} \):

\[
e_i(I_{thr}) = \frac{1}{n_{thr}} \sum_{i=1}^{n_{thr}} (I_i - I_{thr}).
\]  

(11)

with \( I_i > I_{thr} \). Equation (11) provides an empirical estimation of the mean excess function \( E(I_i - I_{thr}|I_i > I_{thr}) \) of impulses. The mean excess over threshold function is plotted for a range of thresholds for run E1 (Figure 9a). The distribution of excess impulses follows a GPD above a threshold impulse, when the mean residual excess plot shows a line with approximately constant gradient [Davidson and Smith, 1990; Embrechts et al., 1997; Beirlant et al., 2004]. As an example, for experiment E1, this region corresponds to the range of 87.5%–97.5% quantiles of the impulse distribution. On the basis of the above observation and considering utilizing a relatively high number of data points (Figure 9b) for improved accuracy, a relatively low \( I_{thr} \), such as the 90% quantile of the distribution, may be chosen (\( I_{thr} = 0.005 \), vertical dashed line in Figures 9a and 9b). This impulse level provides an acceptable threshold for all the uniform flow conditions examined here. Alternatively, a practical, physically sound threshold may be predefined considering a value for the ratio \( I_{thr}/I_c (<1) \). Here for the selected threshold this ratio varied from about 0.6 to 0.8. For a lower threshold, e.g., corresponding to the 85% quantile, it may range from 0.53 to 0.7.

4.3.2. Estimation of GPD Parameters and Model Performance

[35] For the previously defined \( I_{thr} \) the basic properties of the GPD model are satisfied. Thus, the shape \( (\gamma_{GPD}) \) and scale \( (\sigma_{GPD}) \) model parameters are estimated for various choices of the threshold (corresponding to the 87.5% and 90% quantiles) and different methods of estimation.

[36] The estimation of GPD parameters may be performed using a variety of methods such as the maximum likelihood method (MLE) and the method of moments (MOM). The maximum likelihood method, as discussed by Embrechts et al. [1997], is employed to obtain parameter estimations and their standard errors (with 95% confidence interval), for the distribution of conditional excess impulses normalized with the mean of the distribution. Normalizing the impulse exceedances in this manner essentially removes the effect of increasing magnitude of excess impulses for increasing flow conditions, allowing evaluation of whether changes occur in the shape of the tail of the distribution. The method of moments [Hosking and Wallis, 1987] employs the mean \( (E(\xi)) \) and standard deviation \( (SD(\xi)) \) of impulse exceedances \( \xi_j \) to obtain the empirical estimates of \( \gamma_{GPD} \) and \( \sigma_{GPD} \), respectively:

\[
\gamma_{GPD} = \frac{1}{2} \left\{ 1 - \frac{E(\xi)/SD(\xi)}{E(\xi)/SD(\xi)} \right\}, \quad \sigma_{GPD} = \frac{1}{2} \left\{ \frac{1}{E(\xi)} - \frac{1}{E(\xi)/SD(\xi)} \right\}.
\]  

(12a) (12b)

[37] The summary of the estimated parameters is shown in Table 3. In particular, the variation of the parameters computed for different thresholds and method of estimation is shown to enable their comparison. The relative precision of the estimates, indicated by the standard errors, decreases as the sample size of normalized \( \xi \) is reduced (E1 to E6). Both the tail and scale parameters remain positive and decreasing for increasing flow conditions, allowing the conclusion that the relative changes are significant.

**Table 3. Summary of the GPD Shape \( (\gamma_{GPD}) \) and Scale \( (\sigma_{GPD}) \) Parameters, Estimated Using Maximum Likelihood Method (MLE) and Method of Moments (MOM) for All Experimental Runs**

<table>
<thead>
<tr>
<th>Run</th>
<th>( \gamma_{GPD} )</th>
<th>( \sigma_{GPD} )</th>
<th>( \gamma_{GPD} )</th>
<th>( \sigma_{GPD} )</th>
<th>( \gamma_{GPD} )</th>
<th>( \sigma_{GPD} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>0.09 (0.07)</td>
<td>1.20 (0.12)</td>
<td>0.05 (0.08)</td>
<td>1.30 (0.14)</td>
<td>0.04</td>
<td>1.31</td>
</tr>
<tr>
<td>E2</td>
<td>0.18 (0.08)</td>
<td>1.94 (0.10)</td>
<td>0.22 (0.09)</td>
<td>0.91 (0.11)</td>
<td>0.23</td>
<td>0.90</td>
</tr>
<tr>
<td>E3</td>
<td>0.13 (0.13)</td>
<td>1.02 (0.15)</td>
<td>0.10 (0.12)</td>
<td>1.29 (0.20)</td>
<td>0.07</td>
<td>1.33</td>
</tr>
<tr>
<td>E4</td>
<td>0.18 (0.14)</td>
<td>1.12 (0.19)</td>
<td>0.15 (0.15)</td>
<td>1.22 (0.23)</td>
<td>0.11</td>
<td>1.28</td>
</tr>
<tr>
<td>E5</td>
<td>0.13 (0.15)</td>
<td>0.86 (0.17)</td>
<td>0.15 (0.18)</td>
<td>0.86 (0.20)</td>
<td>0.10</td>
<td>0.89</td>
</tr>
<tr>
<td>E6</td>
<td>-0.03 (0.23)</td>
<td>1.11 (0.31)</td>
<td>-0.18 (0.22)</td>
<td>1.35 (0.39)</td>
<td>-0.08</td>
<td>1.24</td>
</tr>
</tbody>
</table>

*Standard errors are shown in parentheses.*
relatively constant with a mean of $\gamma_{\text{GPD}} = 0.14(\pm 0.12)$ and $\sigma_{\text{GPD}} = 1.08(\pm 0.16)$ for experiments E1–E5 (Table 3). For the experiment closest to the critical flow conditions (E6) the shape estimate becomes negative, indicating a possible change in the form of the distribution (from unbounded to a distribution with an upper bound). However, the uncertainty for such an observation is relatively high, considering that the standard error is greater than the estimated value and the confidence intervals span above zero. This probably occurs because of the relatively small sample of impulse exceedances (only 29 for $I_{\text{thr}} = 90\%$ quantile). The within–sampling error invariance of the model parameters with threshold selection further justifies the appropriateness of the GPD for modeling impulses above the chosen threshold.

[38] The probability density function (equation (7)) and cumulative density function predicted by the GPD model are plotted against the empirical observations in Figures 10a and 10b for run E1. Despite the relative uncertainty in the estimation of the tail parameter, the GPD model provides an excellent fit to the tail of impulse events distribution, as assessed visually. The fitted distributions for all of the experiments are shown collectively in Figure 11, represented in a log-log scale, to emphasize the region of greater normalized $\xi$. In agreement with the previous observations and within statistical uncertainty the shape of the distributions remains invariant for the different examined flows. This implies lack of any significant trend for the shape parameter for uniform flow conditions of about the same turbulence intensity levels but different Reynolds numbers.

4.4. Reliability of Grain Dislodgement

[39] In addition to the rate of occurrence of impulses in a flow, the interarrival times between the instances of entrainment may be further statistically studied under an EVT context. Specifically, considering the episodic nature of the phenomenon, the time to grain entrainment for low-mobility flow conditions is analyzed here utilizing reliability (or survival) theory. Under this framework, the complete entrainment of individual coarse particles may be viewed as a stochastic process, with a certain reliability or probability of survival of entrainment events, for a specific time interval and flow conditions. Similar concepts have been employed by Ancey et al. [2008], who considered the flux of coarse grains as a birth-death process, as well as Tucker and Bradley [2010], who investigated the probability of grain entrainment while moving along a transport path.

4.4.1. Empirical Estimation

[40] Consider a surface particle resting in its local configuration. Survival of the particle past time $t_i$ is defined as the probability that the particle remains in its position after time $t_i$, without being entrained: $S(t_i) = P(T_c > t_i)$, with $T_c > 0$ the random variable representing the time to entrainment. According to the multiplication rule for joint events, this probability may also be expressed as $S(t_i) = P(T_c > t_i | T_c > t_i)P(T_c > t_i)$, where the probability of entrainment occurring at least after $t_i$ equals the probability of surviving past the time $t_{i-1}$: $P(T_c \geq t_i | T_c > t_i)P(T_c > t_i) = S(t_{i-1})$. By recursive application of these formulas, it is possible to express the survival from entrainment past time $t_i$ in terms of all the conditional probabilities for entrainment in times before $t_i$, leading to the product-limit formula: $S(t_i) = \prod_{j=1}^{i} P(T_c > t_j | T_c > t_j)$. \hspace{1cm} (13)
[41] For a particular flow condition, equation (13) implies that the survival function is a decreasing function of the time for grain entrainment.

[42] Equivalent to equation (13) is the Kaplan-Meier estimator [Kaplan and Meier, 1958], which is widely used for the nonparametric empirical estimation of the survival function. Utilizing the time series of particle dislodgement obtained from the He-Ne laser for different flow conditions, the interarrival times for each entrainment event can be defined as the time intervals between the instances of deposition and entrainment of the mobile grain. The order statistics for the interarrival times \( t_i, i = 1, \ldots, m \) for a total of \( m \) complete entrainment events may be obtained for each experiment. This is equivalent to having a population of \( m \) different particles, which get entrained at time \( t_i \), when a particular low-mobility flow condition is imposed. For such flows it is safe to assume statistical independence between different grain mobilization events. Defining \( n_i \) as the number of particles that have “survived” entrainment just before time \( t_i \) and \( k_i \) as the number of dislodgements (or “deaths”) occurring at time \( t_i \), the Kaplan-Meier estimation of the survival function is expressed as

\[
S(t) = \prod_{j=1}^{i} \frac{n_j - k_j}{n_j}, \quad i = 1, 2, \ldots, m. \tag{14}
\]

where \( t \) belongs to the duration interval over which the particle is expected to dislodge \( (t_i < t < t_{i+1}) \). Then the survival function \( S(t) \) indicates the percentage of grains that have “survived” entrainment by time \( t \). As expected, for a fixed level of \( S \), the higher the flow conditions, the lower the estimated time to displacement is (e.g., compare \( S(t) = 0.5 \) decreasing from E4 to E1; Figure 12).

[43] Estimates of the survivability of a grain to entrainment or of its reliability to not get entrained until past a certain time instance \( t \) are provided by equation (14) and plotted for various flow conditions in Figure 12. Usually, the right-hand tail of the survival function becomes unreliable when the number of grains remaining at risk for entrainment becomes small (e.g., less than 10 may be used as a rule of thumb). Consequently, it is not statistically meaningful to evaluate the survivability for experiments with a very low number of entrainments (E5 and E6). Similarly, it may be observed that longer sampling times may be required for accurate assessment of \( S(t) \) at very low mobility conditions and in accordance to the findings of Singh et al. [2009] and Bunte and Abt [2005], who observed the dependence of the bed load transport on the sampling interval in an experimental flume and real rivers, respectively. A useful parameter signifying the reliability of grain entrainment is the mean time between entrainments, which is the inverse of the mean frequency of complete grain entrainments (Table 4).

### 4.4.2. Parametric Modeling

[44] A number of parametric models for survival times have been proposed in the literature. Typically, distributions such as the exponential have been used to characterize the interarrival of rainfall events in hydrology [e.g., Adams and Papa, 2000]. Herein the application of Weibull and exponential distributions is demonstrated in modeling the extremal character of interevent times of particle entrainment.

First, the two-parameter Weibull survivor function is defined as

\[
S(t) = P(T \geq t) = \exp\left[-(\lambda W t)^{\gamma W}\right], \tag{15}
\]

the positive parameters \( \lambda W \) and \( \gamma W \) denote the scale and shape parameters of the Weibull model, respectively. The mean time to full grain dislodgement as predicted by the Weibull model may be analytically estimated as

\[
E_W = \lambda W^{-1} \Gamma\left(1 + \frac{1}{\gamma W}\right), \tag{16}
\]

where \( \Gamma(\cdot) \) is the Euler Gamma function, defined as \( \Gamma(x) = \int_0^\infty v^{x-1}e^{-v}dv \). In addition to the survival function, the hazard rate function signifying the instantaneous particle entrainment rate may be of practical utility (e.g., section 5.3). Here the hazard rate \( h(t) \) may be defined as the conditional probability that a particle is entrained at time \( t \). It may be calculated as the number of entrainments per unit time until time \( t \) divided by the number of particles that have survived entrainment:

\[
h(t) = \frac{d}{dt}S(t) = \lambda W\gamma W(\lambda W t)^{\gamma W-1}. \tag{17}
\]

#### Table 4. Summary of Empirical Model Estimators and Weibull and Exponential Model Parameters and Mean Time to Displacement (\( E_W \) and E), Characterizing the Survival Function for Various Flow Conditions (E1-E4)

<table>
<thead>
<tr>
<th>MTBE*</th>
<th>( \lambda_W )</th>
<th>( \gamma_W )</th>
<th>( E_W )</th>
<th>( R^2 )</th>
<th>( \lambda_{exp} )</th>
<th>( E_x )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>8.7</td>
<td>0.103</td>
<td>1.124</td>
<td>9.3</td>
<td>1.00</td>
<td>0.136</td>
<td>7.4</td>
</tr>
<tr>
<td>E2</td>
<td>10.5</td>
<td>0.075</td>
<td>1.175</td>
<td>12.6</td>
<td>1.00</td>
<td>0.114</td>
<td>8.8</td>
</tr>
<tr>
<td>E3</td>
<td>29.1</td>
<td>0.035</td>
<td>1.060</td>
<td>28.1</td>
<td>0.99</td>
<td>0.042</td>
<td>23.6</td>
</tr>
<tr>
<td>E4</td>
<td>45.1</td>
<td>0.023</td>
<td>0.960</td>
<td>43.8</td>
<td>0.98</td>
<td>0.020</td>
<td>50.6</td>
</tr>
</tbody>
</table>

*aMean time between entrainments.
The shape parameter defines the behavior of the hazard function, which is predicted to be monotonically increasing for $\gamma_W < 1$, decreasing for $\gamma_W < 1$, and constant otherwise. Best fit values for the Weibull model parameters are provided in Table 4. The coefficients of determination ($R^2 \sim 0.99$) indicate an excellent model fit. The mean estimated time to entrainment compares well with the empirical estimation. It is observed that the shape parameter for the different flow conditions is very close to 1, which within statistical uncertainty implies a constant hazard rate. This, in turn, confirms that entrainment of individual grains in low flow rates is a Poissonian-like process without memory, contrary to the case of higher flows when grains are entrained collectively [Ancey et al., 2008].

Even though the Weibull model exhibits flexibility, which may be useful for higher-flow conditions as well, for near-threshold conditions it is more realistic and physically sound to assume that the survival times are sampled from a population with a constant hazard rate. To this purpose the one-parameter exponential distribution, a special case of the Weibull model with $\gamma_W = 1$, may be used. The survival function and density function of the exponential distribution are

\[
S(t) = P(T \geq t) = e^{-\lambda_{exp}t},
\]

\[
f(t) = \lambda_{exp}e^{-\lambda_{exp}t},
\]

with $\lambda_{exp}$ being the time-independent hazard rate of the exponential model. The simplicity of the model is further showcased considering that the expected time to entrainment $E_{tr}$ is the reciprocal of the hazard rate. The performance of the exponential distribution is very good as assessed by the high values of the coefficient of determination ($R^2 \sim 1$). Practically, the curves predicted by the two models almost collapse on the same curve, providing a good fit to the empirical estimation, as may be assessed visually (Figure 12). Overall, the mean time to entrainment and hazard rate of the exponential model provide easy to calculate and efficient tools for characterizing the frequency of grain response to uniform and relatively low mobility flow conditions.

5. Discussion

5.1. Probability of Entrainment for an Individual Particle

In the previous sections EVT models were fitted to the distribution of impulses and their tail (sections 4.1, 4.2, and 4.3 respectively). It was observed that any change in the Reynolds number is reflected mainly in the alteration of the scale parameter (in the case the fit is performed for $I_\xi$ or $I_\xi$ or, equivalently, of $I_{mean}$ (when impulses or their exceedances are normalized with $I_{mean}$). Herein the ability of the suggested models to predict the probability of entrainment for individual particles is evaluated assuming that impulses above the defined critical level are responsible for their mobilization.

For the case of impulses that closely follow the Frechet distribution this probability is straightforwardly obtained from the probability of exceeding the theoretically derived critical level, $P = P(I > I_{cr})$. Considering that the GPD model provides the conditional impulse exceedances above impulse threshold, the probability of grain entrainment is given by $P = P(I > I_{cr} > I_{th})$, with

\[
P(I_{cr} > I_{th}) = P(I > I_{th})P\left(\xi > I_{cr} - I_{th}\right|I > I_{th}),
\]

where the probability of exceedance at the threshold value ($P(I > I_{th})$) is estimated as the ratio of the number of impulses exceeding the threshold impulse over the impulses of the entire data set (for instance, $P(I > I_{th}) = 0.10$ for $I_{th}$ corresponding to the 90% quantile).

The probability of particle entrainment $P_E$ may be approximated employing various methods. For instance, Ancey et al. [2008] used for $P_E$ the ratio $\tau_e/\left(\tau_e + \tau_w\right)$, where $\tau_e$ is the mean duration of entrainment and $\tau_w$ is the average waiting time for entrainment to occur. This definition works well for the case when grains may freely dislodge downstream. For the case when the downstream motion of the particle is restrained (Figure 5), the same rationale may be implemented by approximating $\tau_e$ as the average time the particle remains displaced. Then $P_E$ may be plotted against each model distribution $P$, as shown in Figure 13a.

In addition to the previous method, $P_E$ may be estimated by the relative frequency of impulses resulting in grain mobilization or, equivalently, the ratio of mean frequency of impulse events $f_I$ to the average rate of particle displacement $f_E$ [Celik et al., 2010]. The results obtained via this method are plotted in Figure 13b. This definition of $P_E$ accounts for all impulse events or, equivalently, the flow events for which $u^2 > u_{cr}^2$, along with their observed impact on mobilizing the particle (from twitches to complete dislodgements). However, there exists some uncertainty associated with estimating the above frequencies ($f_I$, $f_E$). First, it is assumed that all particle mobilization events (even small displacements) are accounted for, while many of them may not be discernable because of being hidden by the small-scale, high-frequency fluctuations inherent in the displacement signal. Second, the estimation of $f_I$ or mean number of impulses per unit time is quite sensitive to the selection of $u_{cr}$ [Celik et al., 2010], implying that the estimation of $P_E$ will vary depending on whether very small (and possibly ineffective) impulses are accounted for or not.

On the basis of the previous observations the exceedance probability $P(I_{cr} > I_{th})$ may be more appropriately estimated as the ratio $f(I_{cr})/f_C$, with $f_C$ being the rate of occurrence of complete particle entrainments and $f(I_{cr})$ being the mean frequency of impulses above critical $(I_{th})$. Using the theoretical prediction of $u_{cr}$ (derived from equation (2)), the result of $P_E$ is plotted against the various model predictions ($P$) in Figure 13c. Essentially, the latter estimation method of $P_E$ considers the effectiveness of the extreme impulse events (above a critical level $I_{cr}$), avoiding some of the uncertainties associated with the previous method.

Careful observation of Figures 13a, 13b, and 13c reveals the same trend for each model distribution, independent of the various methods of estimation of $P_E$ presented above. The overall performance of the Frechet and GPD distributions is very satisfactory and at least comparable to the lognormal distribution. Particularly, it is seen that the GPD has a high predictive ability when the model
assumptions are satisfied (e.g., E1 and E4). For the case of relatively small data samples of impulse exceedances (E5 and E6) the uncertainty of statistical estimation increases, implying that the closer to threshold the flow conditions are, the longer the required flow sampling time needs to be.

On the contrary, the Frechet distribution has a consistent and superior performance throughout the range of examined flow conditions. The predictions from this distribution fall very close to the line of perfect agreement (\( P_E = P_{E,2} \)), returning a smaller error compared to the other models (Figure 13c). As opposed to the GPD, it can model the whole distribution of impulses.

5.2. Extension of the Utility of the Power Law Relation to Variable Grain and Flow Characteristics

The connection of the proposed extreme value distributions to a power law relation for the magnitude and frequency of impulses renders it a tool of great utility in characterizing the impact of a particular flow on individual particles. Specifically, two particle arrangements may have varying critical impulse levels (\( I_1 \) and \( I_2 \), with ratio \( m = I_2/I_1 \)) because of a number of factors such as different geometrical characteristics of the local bed configuration and/or the grains composing it (see equation (5)). By applying equation (9) a relation between the power law coefficients (\( b_1 \) and \( b_2 \)) describing the two critical flow conditions and the ratio of the critical levels may be derived. For approximately constant exponent \( a \) (which is demonstrated to be valid here for a relatively small range of flow conditions), the ratio of rate of occurrence of two different impulses is

\[
\frac{f_1(I_1)}{f_2(I_2)} = \frac{b_1}{b_2 m^{-a}}.
\]

Different threshold conditions are defined in terms of the same probability of entrainment (\( P_E \)), which is reduced to \( f_1(I_1)/f_2(I_2) = 1 \) if the efficiency of transfer of flow momentum to the particle is assumed to be constant.

On the basis of the above assumptions and using equation (20), the relationship between the two threshold flow conditions becomes \( b_2 = b_1 m^{a} \).

Similarly, for a given flow condition (or, equivalently, fixed values of the power law parameters), flow events of \( m \) times the magnitude of a reference impulse \( I_{ref} \) appear \( m^{a} \) times less frequently:

\[
\frac{f(I_{ref})}{f(mI_{ref})} = \frac{1}{m^{-a}}.
\]

For example, considering that the range of flow conditions for the performed experiments is characterized by a relatively invariant negative exponent, \( a = 1.82 \), flow events of magnitude \( 2I_{ref} \) have a rate of occurrence decreased by 3.5 times (for the same flow condition). This means that grains of the same diameter and about 52% greater density or, equivalently, the same density and 2.1-fold greater diameter
are entrained downstream at a rate 3.5 times smaller compared to the case of the reference grain arrangement at the flow conditions examined here. In addition, when the bed surface is characterized by a range of particle sizes, the relative contribution of each size fraction to the total bed load transport may be estimated by measuring the average distance over which each particle will dislodge when an impulse above \( I_c \) is applied to it (e.g., using particle tracking and synchronized flow measurement techniques).

\[ \text{[56] Grass [1970] viewed incipient motion as a stochastic process and used the distributions of the applied hydrodynamic as well as resisting critical stresses to account for their variability. He suggested that the threshold conditions and, generally, any level of grain movement could be modeled by the overlap between the distributions of instantaneous bed shear stresses and the critical shear stresses required for the inception of motion of the bed surface grains. Following the same reasoning, the frequency distributions of flow impulses \( I_i \) and critical impulses \( I_c \) may be employed to characterize the flow-forcing and motion-resisting conditions (because of heterogeneities in grain and local microtopography parameters). Then the overlap between the two distributions of impulse (Figures 14a and 14b) may provide an alternative means of accurately accessing marginal bed load transport rates. The above approach demonstrates conceptually how the utility of an impulse model may be extended to characterize threshold conditions for a general bed configuration and variable grain parameters.}

5.3. Implications for Bed Load Transport

\[ \text{[57] At near-threshold flow conditions and relatively low transport rates, grain entrainment becomes a highly intermittent process [Singh et al., 2009], which has been experimentally shown to affect mean bed load transport rates [Radice and Ballio, 2008]. The wide spatial and temporal fluctuations of bed load transport rates that have been observed in both laboratory experiments [e.g., Kuhnle and Southard, 1988; Ancey et al., 2008] and field studies [e.g., Drake et al., 1988] demonstrate the appropriateness of treating bed load movement as a probabilistic process. Einstein [1937] was among the first to recognize the stochastic nature of grain entrainment and develop distribution functions for the number of grains passing through a cross section. Recent probabilistic bed load transport studies model the intermittently occurring grain entrainment and disentrainment as a Markov process [e.g., Lisle et al., 1998; Papanicolaou et al., 2002; Ancey et al., 2008; Turowski, 2009]. Particle instability may be triggered when the instantaneous hydrodynamic forces exerted to it exceed the resisting frictional and gravitational forces (\( u > u_c \)).}

\[ \text{[58] Here, under a similar context, flow events or impulses of sufficient magnitude, \( I_i > I_c \), may impart enough momentum for particle removal to a downstream location. As shown in section 5.1, the distribution of impulse exceedances modeled by means of extreme value distributions, \( P( I_i > I_c ) \), are directly linked to \( P_E \) and consequently to the time average frequency of particle entrainment (\( 1/n_1 \)). In addition, the frequency of impulse exceedances above critical (as predicted by equation (9)) exhibits an almost-linear relation (\( R^2 = 0.93 \)) to the average frequency of entrainment (modeled by the exponential distribution of time to entrainment, equation (18b))). Those observations strengthen the significance of impulse on particle entrainment.}

\[ \text{[59] Herein, for low-mobility flow conditions the interarrival times between entrainments are modeled by an exponential distribution. This result offers an experimental validation of Einstein’s [1937] assumption of exponentially distributed rest durations when no collective transport occurs. Furthermore, assuming an exponential distribution for the waiting time between entrainments, the rate of transport may be shown to follow a Poisson distribution [Ancey et al., 2008; Turowski, 2010]. The Poisson distribution for bed load is found to provide a good fit to high-resolution field data when the transport is not dominated by bed form motion [Turowski, 2010].}

\[ \text{[60] In agreement with these studies, it is shown that at relatively low transport rates, particle entrainment (with \( T_c \) following the exponential distribution) is a process without memory, implying that \( P_E \) should remain constant over time (for the same flow conditions and bed geometry). As the transport rate increases, particles may be set into motion from bed material already entrained, which is not captured from a Poissonian representation of bed load [Ancey et al., 2008]. For such cases, it may be more appropriate to implement the Weibull distribution (equation (15)) to model the distribution of \( T_c \) because of its flexibility to account for time dependence.}
6. Conclusions

[61] The extremal character and episodic nature of the occurrence of high-magnitude impulse events and associated time to entrainments are considered here by employing stochastic measures and distributions from the extreme value theory for low-mobility flow conditions. The probability of particle entrainment is approximated by the probability of impulse exceedances above a theoretically defined critical level. Impulses and conditional impulse exceedances are treated as random occurrences of flow events of different magnitudes and durations.

[62] It is demonstrated that the distribution of impulses closely follows a Frechet distribution that is associated with a power law relation for the frequency and magnitude of impulses. The exponent of this relation did not show any significant trend for the range of examined flow conditions. The increase in flow rates was mainly demonstrated by an increase of the base coefficient. Such a description offers a useful tool for the prediction of particle entrainment for particular flow conditions.

[63] Additionally, the generalized extreme value distribution is shown to be an acceptable model for the tail of the distribution of impulses. The peaks over threshold method is implemented to extract the conditional excess impulses above a certain threshold. Guidelines for appropriate selection of the impulse threshold are provided, and the methods’ sensitivity to this threshold is also assessed. Different methods are employed for the estimation of the model parameters. The robustness of the method is indicated by the satisfactory fit of the generalized Pareto distribution to the sample of conditional excess impulse data.

[64] The overall performance of the distributions is at least comparable to or better than the lognormal distribution, as assessed by direct comparison of the predicted and observed probabilities of entrainment for different flow conditions. In direct analogy to the statistical concept of Grass [1970], an extension of the utility of the proposed power law relation is offered by expressing the distribution of forces driving and resisting grain mobilization in terms of impulses rather than shear stresses.

[65] Further, the grain response is statistically described by employing concepts from reliability theory to model the time to full grain entrainment. The exponential distribution is a useful model providing mean time to entrainment and hazard rates for dislodgement, which efficiently characterize the intermittent nature of the phenomenon for low flow rates. The goodness of fit of the exponential model to the empirical distribution provides an experimental validation of the assumption employed by a number of bed load transport models.

[66] In addition to providing good statistical approximations to impulses and time to occurrence of grain entrainment, EVT models provide enhanced understanding and simulation abilities, which are required for the development of predictive equations for sediment entrainment.

Notation

\( \alpha \) bed slope.
\( \Gamma(\cdot) \) Euler Gamma function.
\( \gamma \) shape parameter of the GEV distribution (equation (6)).
\( \gamma_{\text{GPD}} \) shape parameter of the GPD (equations (7) and (8)).
\( \gamma_{\text{W}} \) shape parameter of the Weibull model (equation (15)).
\( \theta_0 \) pivoting angle.
\( \lambda_{\text{W}} \) scale parameter of the Weibull model (equation (15)).
\( \lambda_{\exp} \) hazard rate of the exponential model (equation (17)).
\( \mu \) location parameter of the GEV distribution (equation (6)).
\( \xi_j \) conditional exceedance of impulse \( I_i \) above a threshold level.
\( \rho_f \) density of the fluid.
\( \rho_p \) density of the particle.
\( \rho_s \) coefficient including effects of local grain arrangement and relative density of fluid and solid grain.
\( \sigma \) scale parameter of the GEV distribution (equation (6)).
\( \sigma_{\text{GPD}} \) scale parameter of the GPD (equations (7) and (8)).
\( \tau^* \) dimensionless bed shear stress.
\( a_j \) power law exponent (equation (9)) for flow conditions defined by index \( j \).
\( b_j \) base coefficient (equation (9)) for flow conditions defined by index \( j \).
\( B_f \) buoyancy force.
\( C_D \) drag coefficient.
\( C_f \) impulse coefficient.
\( C_m \) added mass coefficient.
\( E_w \) Weibull mean time to particle entrainment (equation (16)).
\( E_h \) exponential mean time to particle entrainment (\( = 1/\lambda_{\exp} \)).
\( e_x(I_{\text{thr}}) \) function of mean excess impulses over a threshold (\( I_{\text{thr}} \)).
\( F(t) \) total hydrodynamic force.
\( F_{\text{GEV}} \) cumulative distribution function of the GEV distribution (equation (6)).
\( F_{\text{GPD}} \) cumulative distribution function of the GPD.
\( F_{\text{cr}} \) critical force level.
\( F_D \) hydrodynamic drag force.
\( F_L \) hydrodynamic lift force.
\( F_m \) mean force level.
\( f_c \) mean rate of complete particle entrainments.
\( f_e \) mean rate of particle entrainments.
\( f_{\text{GPD}} \) probability density function of the GPD (equations (7) and (8)).
\( f_i \) mean rate of occurrence of impulses.
\( f_h \) hydrodynamic mass coefficient.
\( f(I_i) \) frequency of impulses in excess of \( I_i \) for flow conditions defined by index \( j \).
\( g \) gravitational acceleration.
\( h(t) \) hazard rate function (equation (17)).
\( I \) impulse normalized with mean of distribution sample (\( I_{\text{mean}} \)).
\( I_i \) impulse event \( i \) (equation (1)).
\( I_{\text{cr}} \) critical impulse level (equation (5)).
\( I_{\text{thr}} \) threshold impulse level (defining the tail of distribution of impulses).
Acknowledgments. The support of the National Science Foundation (grants EAR-0439663, EAR-0738759, and CBET-1033196) and the Army Research Office is gratefully acknowledged. The authors also thank A. Ozan Celik for providing access to experimental data used in this research.

References


Kirchner, J. W., W. E. Dietrich, F. Iseya, and H. Ikeda (1990), The variability of critical shear stress, friction angle, and grain protrusion in water-worked sediments, Sedimentology, 37, 647–672.


