Experimental Censorship of Bed Load Particle Motions, and Bias Correction of the Associated Frequency Distributions

Francesco Ballio[1], Alessio Radice[1], Siobhan L. Fathel[2,3], and David Jon Furbish[2,3]

[1]Department of Civil and Environmental Engineering, Politecnico di Milano, Milano, Italy
[2]Department of Earth and Environmental Sciences, Vanderbilt University, Nashville, Tennessee, USA
[3]Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, Tennessee, USA

Key Points:

- Experimental censorship biases frequency distributions of particle hop length and travel time.
- Indirect censorship results from correlation between hop length and travel time.
- Censorship effects can be corrected up to a measuring window size.

Corresponding author: Francesco Ballio, francesco.ballio@polimi.it

Abstract

Knowledge of the statistical distributions of particle hop properties (distances, travel and rest times) enables a deeper understanding of the bed-load sediment transport. However, the measurement of particle hops is prone to censorship: Since many hops cross the boundaries of a spatial-temporal observation window, one knows that they exist but does not know how long they are. An option is to build particle hop samples considering only the hops that are completely observed and excluding (censoring) those which are observed only partially. Such a choice, however, biases the frequency distributions of the hop properties. Moreover, censorship acts in both space and time, and a hop that is censored in time will also not contribute to a sample of hop lengths, and *vice versa*. Time censorship similarly applies to particle rest times. This paper presents a theoretical formulation of censorship that leads to nonparametric bias corrections recovering estimates of values of the underlying distributions of hop distance, travel and rest time up to sampling window dimensions. We illustrate the occurrence and consequences of experimental censorship, and the benefit of applying the bias corrections, for both synthetic and laboratory samples of particle hops. The corrections reasonably recover the relative proportions of frequency distributions represented by the data up to the sampling dimensions, and improve the estimates of the first two moments of particle hop properties. Recommendations are given regarding how the size of an observation window may be chosen to reduce the bias to below some prescribed value, if the forms of the underlying distributions are known.

1 Introduction

Bed load particle motions can be described by various kinematic quantities. Four are particularly relevant to the description of bed load sediment transport and the behavior of tracer particles, namely, the instantaneous velocities and accelerations of the particles, and their hop distances and associated travel times (Einstein, 1950; Wilcock, 1997; Furbish et al., 2012, 2017; Campagnol et al., 2013, 2015; Ancey & Heyman, 2014; Heyman, 2014; Fathel, 2016; Fathel et al., 2015). In addition, particle rest times between motions are essential for understanding the residence time of particles on and within the streambed and the spreading behavior of tracer particles (Sayre & Hubbell, 1965; Bradley et al., 2010; Martin et al., 2012; Lajeunesse et al., 2013; Voepel et al., 2013; Iwasaki et al., 2017).

Probabilistic treatments of these quantities are based on the notion that they are represented by probability density functions whose forms and moments (mean, variance, etc.) are specific to the sediment properties and macroscopic flow conditions (Lajeunesse et al., 2010; Furbish et al., 2012, 2016; Houssais & Lajeunesse, 2012; Martin et al., 2012; Furbish & Schmeeckle, 2013; Fathel et al., 2015), and are aimed at determining these distributions and their parameters in relation to the sediment and flow characteristics. Indeed, this focus on the probability distributions of quantities involved in particle motions represents one aspect of a re-emerging interest in probabilistic formulations of transport, inasmuch as this topic figures prominently in describing rates of sediment transport and rates of dispersal of particles and particle-borne substances during transport. This interest stems from the recognition that particle motions are inherently stochastic, therefore the concepts and language of probability are particularly well suited to the problem of describing these motions — ideas that hark back to important early work, for example, that of Taylor (1921) concerning particle diffusion in turbulent flows, the pioneering work of Einstein (1937), who addressed bed load transport as a probability problem, and the contributions of Tsujimoto (1978) and Nakagawa and Tsujimoto (1980, 1984), who extended Einstein's work to an entrainment form of the Exner equation based on descriptions of particle hop distances.

With specific reference to the work presented here, the joint probability density function of particle hop distances measured start-to-stop L and associated travel times T, as well as the probability density function of rest times R, are central elements of formulations of the entrainment form of the flux and the Exner equation, where the advective part of the flux involves the product of the particle entrainment rate and the mean hop distance (Einstein, 1950; Wilcock, 1997; Parker et al., 2000; Seminara et al., 2002; Wong et al., 2007; Ganti et al., 2010; Furbish et al., 2012, 2016; Ballio et al., 2018b), and the diffusive part involves the variance of the hop distances (Furbish et al., 2012, 2017). In addition, distributions for L, T and R are a key part of descriptions of tracer particle motions (e.g., Sayre & Hubbell, 1965; Hassan et al., 1991; Ferguson et al., 2002; Bradley et al., 2010; Ganti et al., 2010; Furbish et al., 2012a; Martin et al., 2012; Lajeunesse et al., 2013; Voepel et al., 2013).

Interest for scholars in probability distributions for hop lengths, hop travel times and for the subsequent rest times justifies the focus of this paper, namely: How much do intrinsic limits imposed by experimental techniques to space and time dimensions of measuring windows affect the estimate of such distributions? As a consequence of the finite spatial (L_w) and temporal (T_w) sizes of the observation window, only a fraction of the observed hops are measured from start to stop, and only a fraction of the rest events are measured from stop to start; the remaining motion or rest events are only partially observed, so that we are aware of their existence but we cannot measure their full size. These events are typically discarded from the samples and, therefore, do not contribute to the experimental frequency distributions. Such procedure is hereafter referred as *censorship* (e.g. Fathel et al., 2015). As explained in detail in the next section, censorship does not uniformly act on events of varying dimension, because longer hops are more likely to be (spatially) censored then shorter ones, and rest events with higher duration are more likely to be (temporally) censored than those with a shorter duration. Non-uniform censorship thus biases the shape of the measured frequency distributions with respect to those of the original population, as well as their moments (Fathel et al., 2015; Fathel, 2016). Furthermore, all the events with (space or time) sizes larger than those of the observation window are evidently not measurable, and thus cannot be accounted for in the data samples where their frequency is zero; we will refer to this case as *truncation*.

While the concept of censorship of particle hops is, in the end, straightforward, its impact on measured statistics cannot be assessed *a priori*, since it depends on both operational parameters (first of all, the size of an observation window) and phenomenological properties (typical size of particle hops or rests) that, in turn, are linked to hydrodynamic and morphologic conditions. For example, the task of estimating the distribution of rest time and its moments may be particularly problematic if these times follow a power-law distribution whose moments are undefined or emerge only after long times in relation to particle burial and exhumation (e.g., Ferguson et al., 2002; Voepel et al., 2013; Iwasaki et al., 2017). The scientific literature on sediment transport mechanics offers limited examples of studies where the censorship problem has been recognized and addressed. Bialik & Karpinski (2018) demonstrated, on the basis of numerical simulations, that the exponents characterizing sediment dispersion change when windows of different size are used to simulate the process; they also stated that inadequate consideration of a window effect may lead to incorrect recognition of the diffusion regime. Ballio et al. (2018a) showed an example where the estimates of the mean hop length differed by 90% when experimental censorship was accounted for or not. Finally, a reanalysis of data by Fathel et al. (2015) will be presented later (section 5.2) to further demonstrate the quantitative impact of censorship on statistical moment estimates (with percent error around 25% and 60% for mean hop distances and associated variances, respectively).

The purpose of this paper is to present a generalization of the formulation of experimental censorship of hop distances provided by Fathel et al. (2015). In particular,

we examine the censorship of both hop distances L and travel times T, and we explicitly treat the censoring effects of the covariance between these quantities. In addition, we extend the analysis to censorship of rest times R. We then formulate a nonparametric bias correction that can be applied to experimental data to recover the forms of the underlying frequency distributions represented by the values up to the sampling dimensions L_w and T_w, and we present examples of censoring bias and the correction of this bias. Finally, we propose a quantification of how much first and second moments of key quantities can be expected to change as a result of censorship effects, based on inevitably required assumptions on the form of probability distributions.

The next section (Section 2) outlines in practical terms the source of experimental censorship of particle motions. The subsequent section (Section 3) presents a technical treatment of the probability of censorship, where we note that the formulation yields intuitively appealing results in limiting cases (e.g., when L_w or T_w becomes very large). Section 4 returns the treatment to practical terms of bias correction, and Section 5 describes specific examples. Section 6 quantifies sensitivity of the motion statistics to experimental bias and offers guidance for the choice of an appropriate window size. Finally, Section 7 provides a critical discussion of what we have learned.

2 The Source of Experimental Censorship

Let us consider a sediment transport process where particle motions are characterized by T and L with intervening rests events of duration R. For the sake of simplicity we only consider one-dimensional motion. The system is, therefore, fully described in a (t, x) plane, where t is time and x is the space coordinate. The particle entrainment events are assumed to be statistically uniform in space and time. However, we can only observe the process within a limited time-space window, namely [0,T_w] and [0,L_w]. We then assume that the window is large enough to contain a statistically representative number of motions and rest events.

Particle trajectories involve successions of motion and rest periods (Figure 1). Notice that trajectories are plotted in this figure over a (t, x) domain that is larger than the (T_w, L_w) observation window. Moreover, particle velocities are plotted as constant over individual motion events; this is a graphical simplification, as real velocities vary along the motion event, but such simplification does not affect the validity of the following modeling approach. It is apparent from Figure 1 that some events are fully captured from the observation window, while others are only partially captured as they are interrupted at least on one side by the limited time and/or space extension of the window. We may thus classify events as "complete" or "incomplete."

The crucial point is now the following: How should one construct a proper sample of events to represent the whole population? A first choice may be to consider only complete events (e.g., Fathel et al., 2016). This, however, produces a biased sample as longer (in time or space) events have a higher probability of being interrupted by the observation window relative to shorter events. In other words, the frequency distribution derived from a sample of complete events of the random variable is systematically smaller for large values of the random variable, and systematically larger for small values of the random variable, relative to the distribution of the true population. Alternatively, one could follow a different sampling strategy including all events that start within the window, independently of the fact that they are complete or incomplete. In this case, the mean value of hop distance and travel time could be computed dividing the total (i.e., for all particles) travelled distance and the total travel time by the number of entrainment events (e.g., Heyman et al., 2016; Ballio et al., 2018b). However, also in this case the frequency distribution would be biased: the real length or duration of the incomplete events is, of course, not known, as only a portion of their extent can be observed within the sampling window.

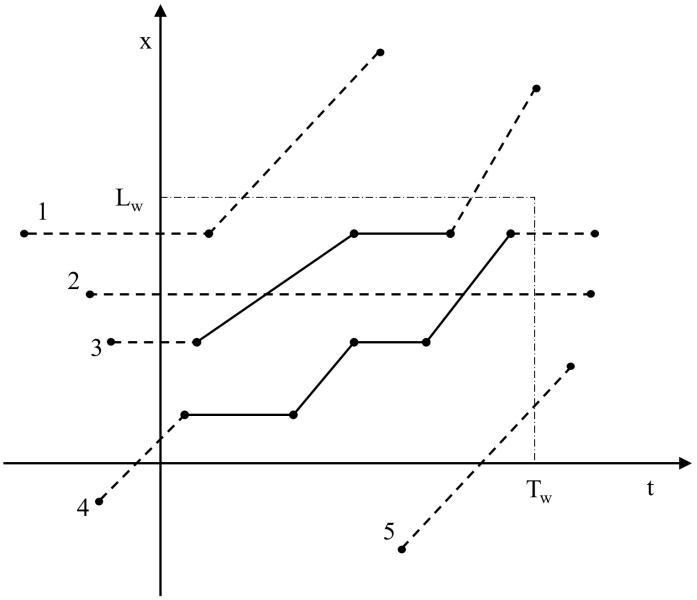


Figure 1. Definition diagram for particle trajectories within the unbounded time-position (t, x) domain sampled during a limited experimental window of duration T_w and length L_w (vertical and horizontal dash-dot lines). Dots represent particle entrainment and disentrainment events. Solid and dashed lines depict complete and incomplete motions or rests, where inclined lines are motions and horizontal lines are rests. From upper left to lower right: (1) motion starts within T_w but is then spatially censored upon leaving the window; (2) a particle is at rest for the entire duration of observation, therefore its rest time is truncated; (3) motion starts and stops within T_w (a complete hop) followed by a complete rest, then is spatially censored; (4) motion stops within T_w then involves two complete rests and two complete motions, while a final rest is temporally censored; (5) motion enters the sampling window but is spatially and temporally censored. These five example trajectories thus represent three complete hops and three complete rests, and four incomplete motions and four incomplete rests.

Let one first consider rest events that allow for a simplified treatment involving only the temporal dimension of the observation window. The sample contains N events represented by a cumulative distribution of values of R (Figure 2). This is an ideal distribution, representative of the whole population and containing all rest events within the window, even if their starting and/or ending instants are out of the window. The observable distribution differs from the ideal one for two reasons. First, values $R > T_w$ cannot be measured as all such events are necessarily incomplete (like for trajectory 2 in Figure 1); we define such exclusion as truncation and $N-N^*$ the number of truncated events, while N^* is the number of events with $R \leq T_w$. Second, a fraction of events with $R < T_w$ are incomplete (see, for example, the first and last rests of trajectories 3 and 4 in Figure 1, respectively), and this fraction progressively increases towards unity as $R \to T_w$; the residual cumulative distribution for complete events is plotted as a dashed line in Figure 2. We define such distortion as censorship, where the total number of complete events is N^c; the total effect of censorship is, therefore, as large as N^*-N^c. Considering now a subsample of events with specified duration, $R = R_1 < T_w$, we here show how the magnitude of censorship can be evaluated from observed values. Starting points of the events are uniformly distributed over the window. As a consequence, all the events starting within $0 < t < T_w - R_1$ are complete, while the remaining ones are interrupted.

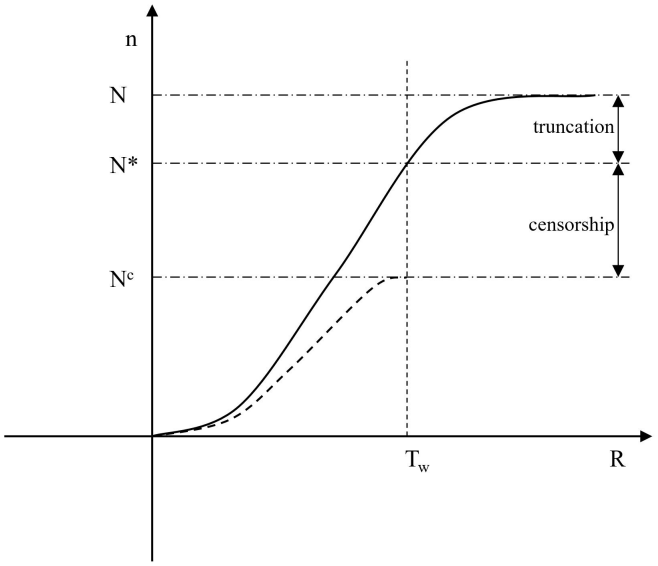


Figure 2. Plot of number of rest events versus duration of rests R showing expected effect of truncation and censorship on cumulative distributions of values of R. For a total sample containing N events and an observation window of duration T_w, N^* is the number of events with $R \leq T_w$ and N^c is the number of rests measured completely. The solid and dashed lines represent the true distribution and the truncated and censored distribution, respectively.

The fraction of complete events within the subsample is, therefore, $(T_w - R_1)/T_w = 1 - R_1/T_w$. The ratio R_1/T_w represents the effect of censorship for any subsample of duration R_1. This effect increases up to unity (total censorship) for $R_1 = T_w$, which is the limiting value distinguishing censorship from truncation. This effect explains the distortion of the measured distribution for complete events with respect to the true one in Figure 2. Moreover, the ability to quantify the bias offers in turn the possibility to correct it, as will be described in Section 4.

The previous discussion can be repeated for motion events. For these events, however, truncation and censorship can be generated by either of the two sizes of the observation window, T_w and L_w, or both (see again Figure 1). Correlation between T and L values makes things more complicated. The whole process is more rigorously formalized and extensively discussed in the next sections, with the aim of identifying proper corrections (where possible) for compensating censorship biasing effects on the frequency distributions of R, T and L. Nothing can be done for the truncation effects.

3 Probabilistic Formulation of Censorship

3.1 Initial Definitions

Let $f_{T,L}(T, L)$ [L^{-1} T^{-1}] denote the uncensored (unknown) joint probability density function of hop distances L and associated travel times T. Assuming only positive hop distances, then by definition,

$$\int_0^\infty \int_0^\infty f_{T,L}(T, L)\, \mathrm{d}T \mathrm{d}L = 1\,. \tag{1}$$

In general, L and T are correlated (e.g., Roseberry et al., 2012; Fathel et al., 2015). In turn, let $n_{T,L}(T, L)$ [L^{-1} T^{-1}] denote the number density such that, for a great num-

ber N of hops, $n_{T,L}(T,L)\mathrm{d}T\mathrm{d}L = Nf_{T,L}(T,L)\mathrm{d}T\mathrm{d}L$ is the expected number of hops within the small interval $\mathrm{d}T\mathrm{d}L$, that is, within the interval T to $T+\mathrm{d}T$ and L to $L+\mathrm{d}L$. Obviously,

$$\int_0^\infty \int_0^\infty n_{T,L}(T,L)\,\mathrm{d}T\mathrm{d}L = N\,. \tag{2}$$

Here, "a great number" refers to the statistical idea of large numbers, although in practice N refers to the number of all hops starting within the sampling window of length L_w during the sampling time T_w, both truncated ($L > L_w$ or $T > T_w$) and censored ($L \le L_w$ and $T \le T_w$).

By definition the marginal distribution $f_T(T)$ [T^{-1}] of travel times T is

$$f_T(T) = \int_0^\infty f_{T,L}(T,L)\,\mathrm{d}L\,, \tag{3}$$

and the marginal distribution $f_L(L)$ [L^{-1}] of hop distances L is

$$f_L(L) = \int_0^\infty f_{T,L}(T,L)\,\mathrm{d}T\,. \tag{4}$$

Similarly, the marginal number densities are

$$n_T(T) = \int_0^\infty n_{T,L}(T,L)\,\mathrm{d}L\,, \tag{5}$$

and

$$n_L(L) = \int_0^\infty n_{T,L}(T,L)\,\mathrm{d}T\,. \tag{6}$$

By definition, $f_T(T)$ and $f_L(L)$ integrate to one, and $n_T(T)$ and $n_L(L)$ integrate to N. The number of truncated hops is N-N^* with

$$N^* = N\int_0^{L_w}\int_0^{T_w} f_{T,L}(T,L)\,\mathrm{d}T\,\mathrm{d}L = \int_0^{L_w}\int_0^{T_w} n_{T,L}(T,L)\,\mathrm{d}T\,\mathrm{d}L\,. \tag{7}$$

The next task is to determine the number of censored hops N^*-N^c and, in turn, the number N^c of completed hops.

Let us now assume that the starting positions x_0 and starting times t_0 of the N motions are independent and uniformly distributed over the sampling window, that is,

$$f_{x_0,t_0}(x_0,t_0) = \frac{1}{L_w T_w}\,. \tag{8}$$

This assumption is a convenient starting point for our objective of illustrating the probabilistic elements of experimental censorship. Note, however, that this assumption may be incorrect for small sampling time T_w or observation length L_w, owing, for example, to turbulence structures or other factors whose scales are similar to or smaller than the sampling dimensions. Nonetheless, this assumption is reasonable for T_w and L_w larger than these scales. We comment further on this point in Section 7. In turn, let $H(u)$ denote the Heaviside step function defined by

$$H(u) = \begin{cases} 0 & \text{if } u < 0 \\ 1 & \text{if } u \ge 0 \end{cases}\,. \tag{9}$$

Independently of time, the probability p that a hop of length L will be spatially censored or truncated is

$$p = H(L_w - L)\frac{L}{L_w} + H(L - L_w)\,. \tag{10}$$

The first term on the right side of equation (10) pertains to censorship of motions with $L \le L_w$ and the second term on the right side pertains to truncation of motions with

$L > L_w$. The Heaviside function provides a "switch" such that the second term on the right side is zero for $L \leq L_w$ giving $p = L/L_w$, and the first term on the right side is zero for $L > L_w$ giving $p = 1$. Similarly, independently of space, the probability that a travel time T will be temporally censored or truncated is

$$q = H(T_w - T)\frac{T}{T_w} + H(T - T_w)\,. \tag{11}$$

The probability that a hop with length L and travel time T will be spatially censored or truncated *and* temporally censored or truncated is pq. The probability that such a hop will be spatially censored or truncated *or* temporally censored or truncated is $p+q-pq$. For a hop to be completed, it must be neither spatially nor temporally censored or truncated. Thus, the probability that a hop with length L and travel time T will be completed is $1 - [p + q - pq] = (1 - p)(1 - q)$.

We may now calculate the number N^c of completed hops. Namely,

$$N^c = N \int_0^\infty \int_0^\infty (1 - p)(1 - q) f_{T,L}(T, L)\,\mathrm{d}T\,\mathrm{d}L\,. \tag{12}$$

Using equations (10) and (11) this becomes

$$N^c = N \int_0^\infty \int_0^\infty \left[1 - H(L_w - L)\frac{L}{L_w} - H(L - L_w)\right]\left[1 - H(T_w - T)\frac{T}{T_w} - H(T - T_w)\right] \cdot f_{T,L}(T, L)\,\mathrm{d}T\,\mathrm{d}L\,. \tag{13}$$

Expanding the integrand in equation (13) and evaluating the integrals leads to (Appendix A)

$$\frac{N^c}{N} = \int_0^{L_w} \int_0^{T_w} (1 - L/L_w)(1 - T/T_w) f_{T,L}(T, L)\,\mathrm{d}T\,\mathrm{d}L\,. \tag{14}$$

Notice that in the limit of $T_w \to \infty$, equation (14) reduces to

$$\frac{N^c}{N} = \int_0^{L_w} (1 - L/L_w) \int_0^\infty f_{T,L}(T, L)\,\mathrm{d}T\,\mathrm{d}L\,. \tag{15}$$

Integrating with respect to T then gives

$$\frac{N^c}{N} = \int_0^{L_w} (1 - L/L_w) f_L(L)\,\mathrm{d}L\,. \tag{16}$$

We may similarly deduce that in the limit of $L_w \to \infty$,

$$\frac{N^c}{N} = \int_0^{T_w} (1 - T/T_w) f_T(T)\,\mathrm{d}T\,. \tag{17}$$

The N^c/N ratios given by equations (14), (16) and (17) become normalization factors in the formulation below.

3.2 Censorship of Hop Distances and Travel Times

The probability that a motion is within the small interval T to $T+\mathrm{d}T$ and within the small interval L to $L+\mathrm{d}L$ is $f_{T,L}(T, L)\mathrm{d}T\mathrm{d}L$, and the probability that a motion is within these intervals and is neither spatially nor temporally censored or truncated is $(1 - p)(1 - q)f_{T,L}(T, L)\mathrm{d}T\mathrm{d}L$. Thus, the (normalized) censored joint probability density function of T and L is

$$f^c_{T,L}(T, L) = \frac{(1 - p)(1 - q) f_{T,L}(T, L)}{\int_0^{L_w} \int_0^{T_w} (1 - L/L_w)(1 - T/T_w) f_{T,L}(T, L)\,\mathrm{d}T\,\mathrm{d}L} \qquad T \leq T_w, L \leq L_w\,, \tag{18}$$

where the normalization is provided by equation (14). This further allows us to simplify the notation to obtain:

$$f^c_{T,L}(T,L) = \frac{N}{N^c}(1-p)(1-q)f_{T,L}(T,L)\,. \tag{19}$$

Notice that, in the limits of $T_w \to \infty$ and $L_w \to \infty$ such that $p = q \to 0$, the censored distribution approaches the underlying true distribution, that is, $f^c_{T,L}(T,L) \to f_{T,L}(T,L)$. Finally, integrating equation (19) with respect to T and using the Heaviside functions to set the limits of integration yields the censored probability density function of hop distances, namely,

$$f^c_L(L) = \frac{N}{N^c}(1-L/L_w)\left[f_L(L) - \frac{1}{T_w}\int_0^{T_w} T f_{T,L}(T,L)\,\mathrm{d}T - \int_{T_w}^{\infty} f_{T,L}(T,L)\,\mathrm{d}T\right]. \tag{20}$$

In turn, integrating equation (19) with respect to L yields the censored probability density function of travel times, namely,

$$f^c_T(T) = \frac{N}{N^c}(1-T/T_w)\left[f_T(T) - \frac{1}{L_w}\int_0^{L_w} L f_{T,L}(T,L)\,\mathrm{d}L - \int_{L_w}^{\infty} f_{T,L}(T,L)\,\mathrm{d}L\right]. \tag{21}$$

Notice that, with $T_w \to \infty$, equation (20) reduces to

$$f^c_L(L) = \frac{N}{N^c}(1-L/L_w)f_L(L)\,, \tag{22}$$

and, with $L_w \to \infty$, equation (21) reduces to

$$f^c_T(T) = \frac{N}{N^c}(1-T/T_w)f_T(T)\,, \tag{23}$$

where the expression for N^c is specific to each case, that is, involving equation (16) or (17).

Let us now define the conditional probability density functions

$$f_{T|L}(T|L) = \frac{f_{T,L}(T,L)}{f_L(L)} \tag{24}$$

and

$$f_{L|T}(L|T) = \frac{f_{T,L}(T,L)}{f_T(T)}\,. \tag{25}$$

Rearranging these and substituting into equations (20) and (21) then leads to

$$f^c_L(L) = \frac{N}{N^c}(1-L/L_w)f_L(L)\left[1 - \frac{1}{T_w}\int_0^{T_w} T f_{T|L}(T|L)\,\mathrm{d}T - G(T_w,L)\right] \tag{26}$$

and

$$f^c_T(T) = \frac{N}{N^c}(1-T/T_w)f_T(T)\left[1 - \frac{1}{L_w}\int_0^{L_w} L f_{L|T}(L|T)\,\mathrm{d}L - G(T,L_w)\right], \tag{27}$$

where

$$G(T_w,L) = \int_{T_w}^{\infty} f_{T|L}(T|L)\,\mathrm{d}T \quad \text{and} \quad G(T,L_w) = \int_{L_w}^{\infty} f_{L|T}(L|T)\,\mathrm{d}L\,. \tag{28}$$

The integral in the bracketed part of equation (26) reflects that *indirect* censorship of hop distances occurs with finite sampling interval T_w due to the covariance between L and T. Similarly, the integral in the bracketed part of equation (27) reflects that indirect censorship of travel times occurs with finite window size L_w due to this covariance. The quantities $G(T_w,L)$ and $G(T,L_w)$ represent the effects of truncation. For example, $G(T_w,L)\mathrm{d}L$ represents the probability associated with the small interval L to $L+\mathrm{d}L$ that is "lost" from $f_L(L)\mathrm{d}L$ due to truncation associated with T_w. Similarly, $G(T,L_w)\mathrm{d}T$ represents the probability associated with the small interval T to $T+\mathrm{d}T$ that is lost from $f_T(T)\mathrm{d}T$ due to truncation associated with L_w. These quantities are unrecoverable (and cannot be estimated), as no information is available for truncated motions. We return to equations (26) and (27) in Section 4 describing a bias correction.

3.3 Censorship of Rest Times

Let N now denote a great number of disentrainment events during the sampling interval T_w. Rest times R are independent of hop distances and travel times. If $f_R(R)$ denotes the uncensored distribution of rest times, then the number of truncated rests is N-N^* with

$$N^* = N \int_0^{T_w} f_R(R)\,\mathrm{d}R\,. \tag{29}$$

Assuming that the instants at which particles come to rest are uniformly distributed over T_w, then the probability that a rest time will be censored is R/T_w. The number of completed rests is thus

$$N^c = N \int_0^{T_w} (1 - R/T_w) f_R(R)\,\mathrm{d}R\,. \tag{30}$$

In turn the censored distribution $f_R^c(R)$ of rest times R is

$$f_R^c(R) = \frac{N}{N_c}(1 - R/T_w) f_R(R)\,. \tag{31}$$

Note that, in the limit $T_w \to \infty$, $N^c \to N$ and $f_R^c(R) \to f_R(R)$.

4 Bias Correction

4.1 General Formulation

Let us rearrange equations (19), (26) and (27) to give

$$f_{T,L}(T,L) = \frac{N^c}{N} B_{T,L} f_{T,L}^c(T,L)\,, \tag{32}$$

$$f_L(L) = \frac{N^c}{N} B_L f_L^c(L)\,, \tag{33}$$

and

$$f_T(T) = \frac{N^c}{N} B_T f_T^c(T)\,, \tag{34}$$

where

$$B_{T,L} = \frac{1}{(1 - L/L_w)(1 - T/T_w)}\,, \tag{35}$$

$$B_L = \frac{1}{(1 - L/L_w)[1 - (1/T_w)\int_0^{T_w} T f_{T|L}(T|L)\mathrm{d}T - G(T_w, L)]}\,, \tag{36}$$

and

$$B_T = \frac{1}{(1 - T/T_w)[1 - (1/L_w)\int_0^{L_w} L f_{L|T}(L|T)\mathrm{d}L - G(T, L_w)]}\,. \tag{37}$$

The quantities $B_{T,L}$, B_T and B_L represent bias correction factors for the joint and marginal distributions of travel times and lengths. Namely, if the censored distributions $f_{T,L}^c(T,L)$, $f_T^c(T)$ and $f_L^c(L)$ can be estimated from data, and if $B_{T,L}$, B_T and B_L also can be determined, then these censored distributions can be used to estimate the underlying distributions $f_{T,L}(T,L)$, $f_L(L)$ and $f_T(T)$ up to $T = T_w$ and $L = L_w$. This is in principle straightforward for the bias correction $B_{T,L}$ in equation (35) involving the joint distribution $f_{T,L}(T,L)$ of equation (32). However, marginal distributions rather than the joint one are typically required; unfortunately, the conditional distributions $f_{T|L}(T|L)$ and $f_{L|T}(L|T)$ generally are not known, as these also are subject to experimental censorship, so that correction factors for the marginal distributions are also not known. One would thus need a closure model for their evaluation. In the following we suggest a formulation that gets closer to a bias correction, although it still cannot fully recover the underlying distributions, except in limiting cases.

4.2 Practical Formulation

Combining equations (32) and (35),

$$f_{T,L}(T,L) = \frac{N^c}{N} \frac{f^c_{T,L}(T,L)}{(1-L/L_w)(1-T/T_w)} \,. \tag{38}$$

Integrating this with respect to T then yields

$$f_L(L) = \frac{N^c}{N} \frac{1}{(1-L/L_w)} \left[\int_0^{T_w} \frac{f^c_{T,L}(T,L)}{1-T/T_w} \,\mathrm{d}T + \int_{T_w}^{\infty} \frac{f^c_{T,L}(T,L)}{1-T/T_w} \,\mathrm{d}T \right] . \tag{39}$$

Similarly, integrating equation (38) with respect to L yields

$$f_T(T) = \frac{N^c}{N} \frac{1}{(1-T/T_w)} \left[\int_0^{L_w} \frac{f^c_{T,L}(T,L)}{1-L/L_w} \,\mathrm{d}L + \int_{L_w}^{\infty} \frac{f^c_{T,L}(T,L)}{1-L/L_w} \,\mathrm{d}L \right] . \tag{40}$$

In equations (39) and (40), the first factor represents the correction for direct censorhip, while the term in parenthesis is the indirect contribution of correlation between hop distance and travel time. The first integral can be estimated, while the second one represent the effects of truncation and thus cannot be estimated. A correction of censorship relies on assuming that the contribution of truncation will be relatively small. In Section 5 we implement numerical approximations of equations (39) and (40) to estimate the densities $f_L(L)$ and $f_T(T)$ up to the dimensions L_w and T_w.

4.3 Limiting Cases

Consider the situation in which the sampling interval $T_w \to \infty$. In practice this means that T_w is much greater than the longest measured travel times (e.g., Fathel et al., 2015). In this case equation (36) reduces to

$$B_L = \frac{1}{1-L/L_w} \,, \tag{41}$$

which indicates that censorship only involves the finite widow size L_w. Conversely, in the situation where the window size L_w is much larger than the longest measured hop distance L, then equation (37) reduces to

$$B_T = \frac{1}{1-T/T_w} \,, \tag{42}$$

which indicates that censorship only involves the finite sampling time T_w. These limiting cases indicate that the indirect censorship of L and T in relation to their covariance becomes unimportant. For example, if T_w is larger than the longest measured travel times but L_w is short enough to produce censorship of hop distances, then the indirect censorship of L associated with finite T_w due to the covariance between L and T is negligible. Similarly, if L_w is larger than the longest measured hop distances but T_w is short enough to produce censorship of travel times, then any indirect censorship of T associated with finite L_w is negligible.

Observe that $Nf_L(L) = n_L(L)$, $N^c f^c_L(L) = n^c_L(L)$, $Nf_T(T) = n_T(T)$ and $N^c f^c_T(T) = n^c_T(T)$. We may thus rewrite equations (26) and (27) as

$$n_L(L) = B_L n^c_L(L) \tag{43}$$

and

$$n_T(T) = B_T n^c_T(T) \,. \tag{44}$$

These or their cumulative forms may be more suited for calculations with experimental measurements of hop distances and travel times.

4.4 Correction for Rest Times

With respect to rest times R and considering equation (31), the bias correction factor is

$$B_R = \frac{1}{1 - R/T_w} \,. \tag{45}$$

That is, censorship only involves the finite sampling time T_w. With this result in mind, let us note that the examples of censorship and bias correction presented in Section 5 are focused on hop distances and travel times, and do not include measurements of rest times and estimations of their distribution. Experimental measurements aimed at estimating the distribution of rest times and its moments are problematic, given that rest times may involve particle-bed exchanges with power-law behavior (e.g., Voepel et al., 2013). Nonetheless, the result embodied in equation (45) is correct for a distribution of rest times with finite moments, including power-law distributions whose moments emerge only after long times (see Section 6.2). We thus present this result as a theoretical conclusion, and now turn entirely to hop distances and travel times.

5 Applications

The theoretical formulations presented in the previous sections are now applied to demonstrate the validity of the method for correction of the censored frequency distributions. Because in a real experiment the unbiased sample is never known, we start here with a synthetic example to illustrate the occurrence and consequences of censorship, then present the bias correction. The true underlying distribution (that is known now) is the target of the correction, so this example provides a clear illustration of the fidelity and limitations of the correction. In the subsequent subsection we turn to an experimental data set, recently presented by Fathel et al. (2015). This example reinforces the points made concerning the synthetic data.

5.1 Synthetic Example

Particle hops are generated assuming uniform distributions for travel times and hop-averaged particle velocities, treating the two quantities as uncorrelated. Both travel times and hop distances take values within the interval $[0, 1]$. Using uniform distributions for times and velocities results in a distribution of hop distances $f_L = -\ln L$ for L within $[0, 1]$ (Appendix B). However, the shapes of the distributions have no physical basis and do not necessarily mimic experimental data. On the other hand, the example is particularly relevant for validating the bias correction as, in this ideal condition, any sample can be generated assuming that it is perfectly unbiased.

Three cases are considered, with different combinations of T_w and L_w (Table 1). In each case, the samples include 10,000 hops. Starting points (in both space and time)

Table 1. Properties of synthetic cases.

Case	T_w	L_w	N^c
S1	1	1	4,165
S2	0.8	10	3,988
S3	0.9	0.6	3,445

of hops are uniformly distributed in the observation window. The three cases imply different combinations of truncation and censorship. In case S1 truncation is absent, in case

S2 it is present only for travel times and in case S3 it is present for both travel times and hop distances.

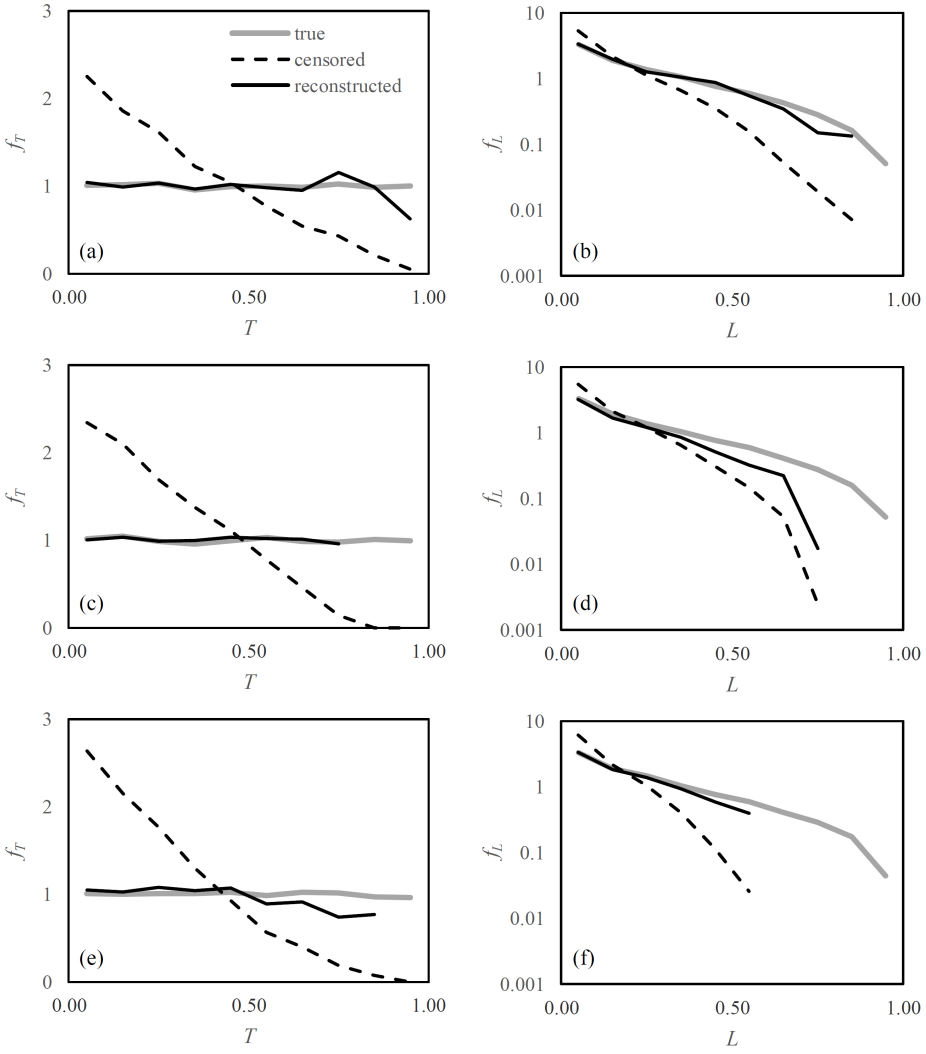


Figure 3. Bias corrections for the synthetic cases with (a), (b) for S1, (c), (d) for S2 and (e), (f) for S3. Plots show the underlying true distribution (gray line), the censored distribution of complete hops (dashed line) and the reconstructed distribution (solid black line).

This part of the analysis is supported by *Supplemental File 1* that contains the worksheets used for these synthetic cases. After generating the full samples, complete hops are first recognized based on their last points being within the observation window. Second, the number densities $n_{T,L}(T,L)$ and $n^c_{T,L}(T,L)$ are computed as the densities with reference to a matrix of 100 classes (10 for lengths by 10 for travel times). The sum of all the numbers in $n^c_{T,L}(T,L)$ equals the total number of complete hops N^c in Table 1 (from which it is also seen that an increasing impact of window dimensions is demonstrated by a decreasing number of observed complete hops). Third, two-dimensional frequency distributions and marginal distributions for travel times and hop distances are obtained considering both the true and the censored data. Fourth, $f_{T,L}(T,L)$ is reconstructed starting from the $f^c_{T,L}(T,L)$ using equation (38), and corresponding marginal

distributions are again obtained from equations (3) and (4). It is noted here that the correction is performed without any assumption on the shape of the underlying true distribution.

The bias corrections for these synthetic cases are presented in Figure 3. For each case we display the marginal probability distributions of travel times and hop distances. The plots include the true distribution for all hops starting within the observation windows, the censored distribution for complete hops and the estimate of the real distribution as reconstructed from the censored one.

For case S1 the bias correction is very satisfactory for both travel times and hop distances. However, the right tail of the distribution of L is missed, as visible in Figure 3(b). This is due to two effects. First, hops with a length of almost 1 are scarce since their presence in the true samples requires a combination of a travel time of almost 1 and a travel velocity of almost 1. In addition, for preserving these few hops in the censored distribution, the motion must start at the beginning of the time observation and at the beginning of the spatial observation. If any of these conditions is not satisfied, the hop is not completely measured and thus its effect on the distribution cannot be reconstructed. In the other cases (S2 and S3), the presence of truncation also affects the result because the censored and reconstructed frequency distributions have an upper bound that reflects the presence of truncation and prevents the right tail of the distribution to be observed and thus corrected. In this respect, case S2 also highlights the indirect truncation induced by correlation: hop distances are not truncated by L_w that is much larger than 1, but their distribution is still truncated because longer hops are also those with larger travel times. In cases S2 and S3 the performance of the correction decays as the bounds imposed by truncation are approached. However, the reconstructed frequency distribution is always better than the original censored one. As a quick indicator of the impact of bias correction, in Tables 2 and 3 we provide the mean and variance values of the quantities for all the distributions, with the following symbols: μ and σ^2 indicate means and variances, respectively; subscripts T and L denote the quantity under consideration; additional subscripts c and r (reconstructed) correspond to the censored and corrected distributions.

Table 2. Mean values (true, censored and corrected) of travel time and hop length for the synthetic cases.

Case	μ_T	μ_{Tc}	μ_{Tr}	μ_L	μ_{Lc}	μ_{Lr}
S1	0.45	0.29	0.44	0.24	0.13	0.22
S2	0.45	0.26	0.32	0.24	0.12	0.15
S3	0.45	0.25	0.36	0.24	0.10	0.15

5.2 Experimental Example

We consider the data of Fathel et al. (2015). Motions of 0.5 mm sand particles were recorded at 250 Hz over an observation window that was 7.5 cm in the stream-wise direction and 6 cm in the transverse direction. Duration of recording was 5 s. Several hops had a stream-wise distance equal to 0, corresponding to purely transverse particle motion, and were removed from the sample for consistency with the one-dimensional theoretical treatment introduced in this manuscript. The analysis of these experimental data was in the end based on 3,499 complete hops. The cumulative distribution of starting positions (not shown here) was represented well by a uniform distribution, as the sam-

Table 3. Variance values (true, censored and corrected) of travel time and hop length for the synthetic cases.

Case	σ_T^2	σ_{Tc}^2	σ_{Tr}^2	σ_L^2	σ_{Lc}^2	σ_{Lr}^2
S1	0.085	0.052	0.083	0.050	0.020	0.044
S2	0.085	0.038	0.068	0.049	0.019	0.030
S3	0.084	0.039	0.078	0.050	0.011	0.026

pling interval of five seconds was sufficient to cumulatively mask any small scale patchiness. However, the cumulative distribution of starting times could indicate a nonuniform distribution, likely due to fluctuations in entrainment associated with turbulence scales similar to the sampling interval (see Figure 1 in Fathel et al., 2015), where the imaging area is too small to sample all important turbulence scales (and entrainment) uniformly over time. Our assumption of equation (8) was thus not exactly met by the experimental data. The frequency distributions of measured travel times and hop distances for these hops (dashed lines in Figure 4) suggest that hop truncation was not significantly present in this experiment, since the duration of observation was 5 times the largest measured travel times and the length of the focus area was about twice the hop length for $f_L(L) \approx 0.01$.

Though not truncated, these data were affected by censorship and were thus corrected (Figure 4). This part of the analysis is documented in *Supplemental File 2* that

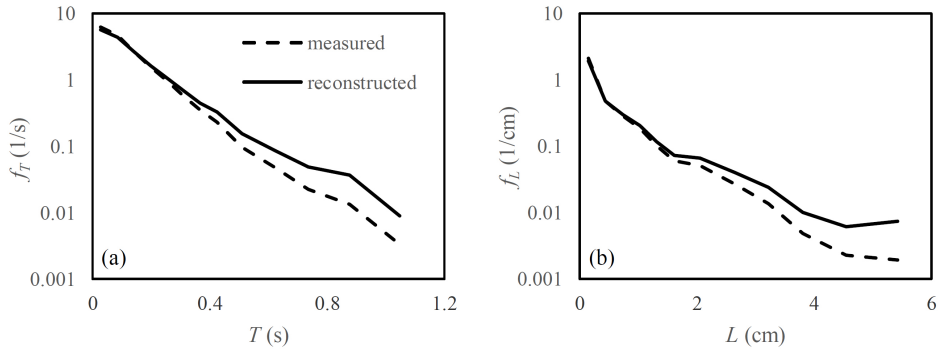


Figure 4. Experimentally measured (and thus censored, dashed lines) and reconstructed (solid lines) frequency distributions of (a) travel times T and (b) hop distances L for the data reported by Fathel et al. (2015).

contains the worksheet used for the computations. In this case, the correction could not be applied using equation (38) because incomplete hops were discarded and thus N was unknown, while $N^c = 3{,}499$. Therefore, the correction was applied to $n_{T,L}$ rather than $f_{T,L}$, using the following equation:

$$n_{T,L}(T,L) = \frac{n_{T,L}^c(T,L)}{(1 - L/L_w)(1 - T/T_w)}\,, \tag{46}$$

that results from the combination of equations (1), (2) and (38). Applying equation (46), a value of N is obtained as the sum of all the values in the $n_{T,L}$ matrix. Such a value

is a virtual number of hops that should have been measured if the sample were not censored, and is typically not an integer. A sample size of 3,908.6 corrected hops was determined starting from the 3,499 complete hops mentioned above. The mean and variance values of the properties are listed in Table 4, indicating that the correction increased

Table 4. Mean and variance values (censored and corrected) of travel time and hop length for the experimental example.

μ_{Tc}	μ_{Tr}	μ_{Lc}	μ_{Lr}	σ^2_{Tc}	σ^2_{Tr}	σ^2_{Lc}	σ^2_{Lr}
0.12	0.14	0.45	0.57	0.015	0.022	0.47	0.76

the mean values of travel times and hop distances by 16% and 27%, respectively.

To explore situations that may be more affected by truncation, we also analyzed these experimental data as if the observation window had been smaller than the one actually used. The considered window was 4 cm in the stream-wise direction and 2 s in time, and is depicted in Figure 5. The hops starting within this window were either truncated, censored or complete. We applied the bias correction given by equation (38) to the complete hops and compared the distributions with the best estimate of a true one, that is, the distribution obtained above and included in Figure 4. The results of this exercise (Figure 5) include the censored distribution for the reduced window, the corrected distribution for the same window and the best estimated distribution. The reduction of the sampling window implied an overestimation of the frequency of the shortest hops, which was present in both the censored and the corrected distributions. Conversely, frequencies in the body of the distributions were underestimated, with the correction partially fixing this mistake. The mean and variance values of travel time and hop length for the censored and corrected distributions are listed in Table 5, to be compared to those of Table 4. In this case, the correction was thus able to furnish distributions that, though still not fully similar to the true ones, were significantly better than the censored ones.

Table 5. Mean and variance values (censored and corrected) of travel time and hop length for the experimental example considering a sub-window to increase the impact of truncation.

μ_{Tc}	μ_{Tr}	μ_{Lc}	μ_{Lr}	σ^2_{Tc}	σ^2_{Tr}	σ^2_{Lc}	σ^2_{Lr}
0.08	0.11	0.28	0.45	0.009	0.017	0.21	0.51

6 Practical Guidance for the Choice of an Observation Window

In this section we exploit the concepts and relations developed so far to explore a crucial problem for the experimentalist: Given that, to some degree, censorship and truncation are unavoidable, how should the minimum size of the observation window (L_w, T_w) be chosen in order to keep the bias of the results to an acceptable level? We emphasize that there is, of course, no general answer to this question, as the biasing effects of censorship and truncation depend on the characteristics of the probability distribution of the quantity under investigation, as well as any covariance between L and T. However, with the aim of providing guidance on a possible strategy to address the problem, we present three (synthetic) examples; we simulate censoring and truncation effects of ob-

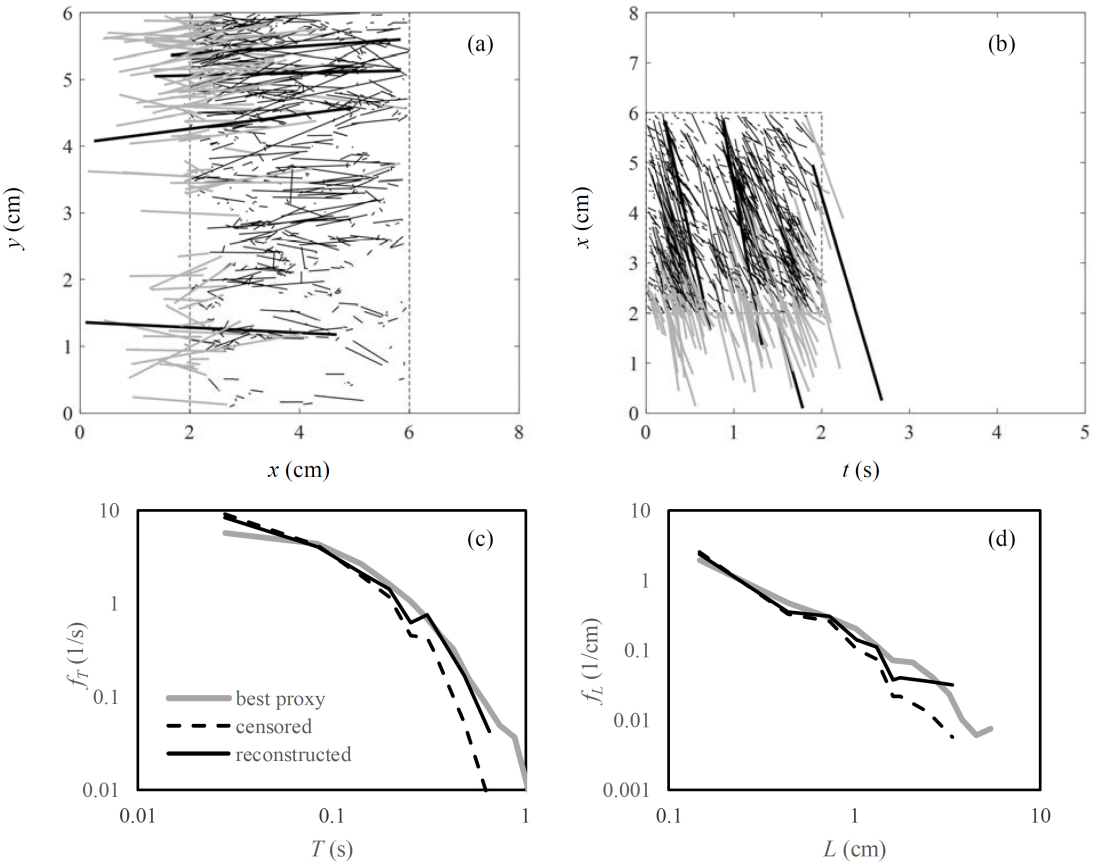


Figure 5. Sampling of the data reported by Fathel et al. (2015) with a smaller window. (a), (b) visualization of complete (black), censored (gray) and truncated (thick black) tracks (streamwise motion is towards smaller x values). (c), (d) censored (dashed lines) and reconstructed (black lines) distributions compared with the best estimates (gray lines) of the true distributions.

servation windows of different sizes and we show consequent bias effects on the mean and variance values.

6.1 Example 1: 1D Censorship with an Exponential Distribution of Travel Times

We consider here the case of an exponentially distributed random variable subject to censorship by only one dimension (length or duration) of an observation window. According to prior studies (e.g., Martin et al., 2012; Fathel et al., 2015; Furbish et al., 2016), an exponential distribution may fit samples for travel times T. If we further assume that the length L_w of the observation window is large enough to prevent any spatial truncation, we can analyze the effects of censorship and truncation as solely due to the time window. As discussed in section 4.3, in such a case the space-time correlation has no effect on censorship and, ideally, a perfect correction for the censorship bias is possible based on equation (42). In practice this is not true, as observed in section 5.1, because very few data (if any) are present for the largest values of the variable, close to the window dimension. However, as we are here working with analytical distributions, we can simulate the case of a population with a very large number of observations, so that the cen-

sored distribution is accurately represented along the whole observation window $0 \leq T < T_w$, and we can accurately reconstruct the true distribution within this window.

The true probability density function for travel times is $f_T(T) = (1/\mu_T)e^{-T/\mu_T}$, and the corresponding censored density function is $f_T^c(T;T_w) = N/N^c(1-T/T_w)(1/\mu_T)e^{-T/\mu_T}$. Note that the functional dependence of quantities on the size of the observation window is here explicitly indicated. Furthermore, since this example is addressed analytically, no synthetic data samples are used and, therefore, both N and N^c have no real meaning. An ideal ratio N/N^c is here used as a scaling factor to normalize all the distributions to unity.

The effect of the window size can be evaluated through the mean and variance calculated for the truncated distribution and for the censored distribution (a subscript t is used here to denote the moments for the truncated distribution):

$$\mu_{Tt}(T_w) = \int_0^{T_w} T f_T(T)\,\mathrm{d}T\,, \tag{47}$$

$$\mu_{Tc}(T_w) = \int_0^{T_w} T f_T^c(T;T_w)\,\mathrm{d}T\,, \tag{48}$$

$$\sigma_{Tt}^2(T_w) = \int_0^{T_w} T^2 f_T(T)\,\mathrm{d}T - \mu_{Tt}^2 \quad \text{and} \tag{49}$$

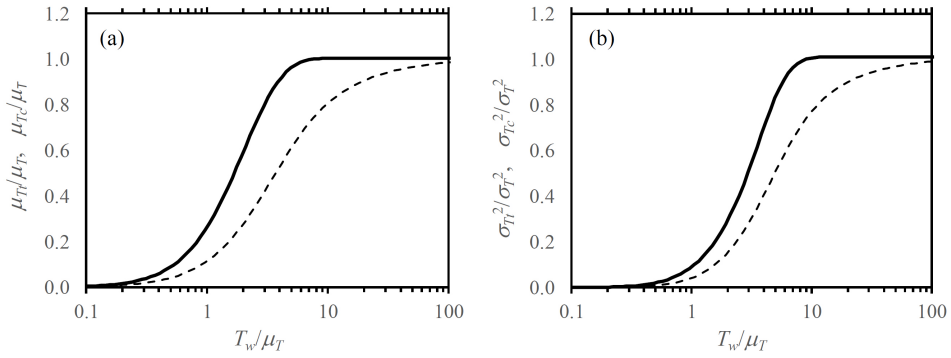
$$\sigma_{Tc}^2(T_w) = \int_0^{T_w} T^2 f_T^c(T;T_w)\,\mathrm{d}T - \mu_{Tc}^2\,, \tag{50}$$

All these moments are, necessarily, smaller than their true counterparts, μ_T and σ_T^2, the difference between them tending to zero for T_w tending to ∞. Figure 6 shows the result of the exercise. Note that, once normalized with the moments of the true exponential distribution, values are independent of the parameter μ_T of the distribution. Consider, for example, an observation window as large as six times the mean value of the distribution of time of motions ($T_w/\mu_T = 6$): for the truncated distribution, moments are estimated as $\mu_{Tt} = 0.98\mu_T$ and $\sigma_{Tt}^2 = 0.91\sigma_T^2$; the difference from the true values is due to the missing contribution to the truncated tail, $T > T_w$. Corresponding estimates for the censored moments are $\mu_{Tc} = 0.80\mu_T$ and $\sigma_{Tc}^2 = 0.59\sigma_T^2$; the increased bias is the effect of uncorrected censorship. From a different perspective, plots in Figure 6 provide guidance towards the minimum size of the observation window in order to achieve any given accuracy in the estimate of the mean and variance, the latter requiring larger observation windows for the same level of accuracy.

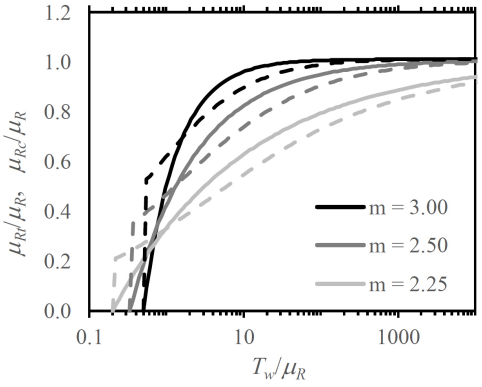
6.2 Example 2: 1D Censorship with a Power-Law Distribution of Rest Times

Example 2 explores the effects of censorship and truncation with a thick-tailed distribution. Rest time, R, is a suitable quantity that may present a heavy tail resulting from particle burial and reappearance (e.g., Ferguson et al., 2002; Voepel et al., 2013; Iwasaki et al., 2017); we here assume a power-law distribution for R. In this case the length of the observation window, L_w, has no effect on the variable and censorship or truncation effects are only due to the time window; as in example 1 we assume a perfect correction to the censorship bias to be achievable by equation (45).

The exact probability density function for rest times is $f_R(R) = (m-1)/R_{min}(R/R_{min})^{-m}$. The effect of the window size is evaluated, for different values of m, through the truncated mean calculated for the exact and for the censored distributions (expressions equivalent to those presented in example 1). We do not consider here the effects on variance, as the true distribution has no finite variance for $m \leq 3$. Results are plotted in Figure 7; note that, once normalized with the true mean, the values are independent of the parameter R_{min} of the distribution.



557 **Figure 6.** Ratios between the moments of truncated and censored distributions and those of
558 the true one for time of motion with an exponential distribution, as a function of the normalized
559 window size. Continuous lines: moments of the truncated distribution; dotted lines: moments of
560 the censored distribution. (a) mean and (b) variance of time of motion.



583 **Figure 7.** Ratios between the biased mean and its true counterpart for time of rest with a
584 power-law distribution and varying values of the exponent m, as a function of the normalized
585 window size. Continuous lines: moments of the truncated distributions; dotted lines: moments of
586 the censored distributions.

594 From a qualitative point of view, the results are similar to those for the exponen-
595 tial distribution of example 1: the bias due to truncation and censorship diminishes for
596 increasing window size. As expected, smaller values of m (corresponding to heavier tails)
597 increase the effect of both censorship and truncation: for example, for $m = 3.00$ an ob-
598 servation window as large as $T_w \approx 4.5\mu_R$ is sufficient to obtain $0.90\mu_R$ of the true mean
599 value, while for $m = 2.25$ a minimum window size of $T_w \approx 1,400\mu_R$ is required for
600 the same level of accuracy. Finally, the correction of censorship diminishes the bias in
601 the estimation of the mean, although the benefit is smaller than for the exponential dis-
602 tribution.

603 ### 6.3 Example 3: 2D Censorship with Exponential Distributions for Travel
604 Times and Velocities

605 In example 3 we explore the joint effect of time and space censorship and trunca-
606 tion on a two-dimensional thin-tailed distribution. We consider the case of hop lengths,

L, resulting from the product of uncorrelated exponential distributions of travel times and velocities (for the latter see, for example, Lajeunesse et al., 2010; Roseberry et al., 2012; Furbish et al., 2016). The resulting distribution for hop lengths is in this case a modified Bessel function of order zero (Appendix B).

The effect of the covariance between L and T on censorship and truncation precludes a simple analytical treatment. Therefore, we use a Monte Carlo approach as in section 5.1. The procedure is as follows. First, a sample of $N = 100,000$ values of hop length is generated as the product of uncorrelated values for travel time and velocity; this is the true sample. Second, for any given size of the observation window, hop values larger than the spatial extension of the window, $L > L_w$, are discarded, thus mimicking the effect of truncation. Third, starting points of the remaining hops are randomly located within the window considering uniform distributions along space and time. All the hops with a final time or space coordinate beyond the window boundaries are discarded, thus mimicking the effect of censorship. The remaining N^c values represent the censored sample, for which the distribution $n^c_{T,L}(T, L; T_w, L_w)$ is computed using a matrix of 10 by 10 bins and, correspondingly, marginal distributions $n^c_L(L; T_w, L_w)$ and $f^c_L(L; T_w, L_w)$ are also computed. Fourth, an uncensored distribution $n^r_{T,L}(T, L; T_w, L_w)$ is reconstructed by means of equation (46) and, correspondingly, marginal distributions $n^r_L(L; T_w, L_w)$ and $f^r_L(L; T_w, L_w)$ are also computed (the superscript r indicates indeed reconstructed distributions). Steps from 2 to 4 are repeated for a variety of combinations of window dimensions (L_w, T_w); for each of them, the mean and variance values are calculated from the truncated and censored marginal distributions, with equations equivalent to those presented in example 1.

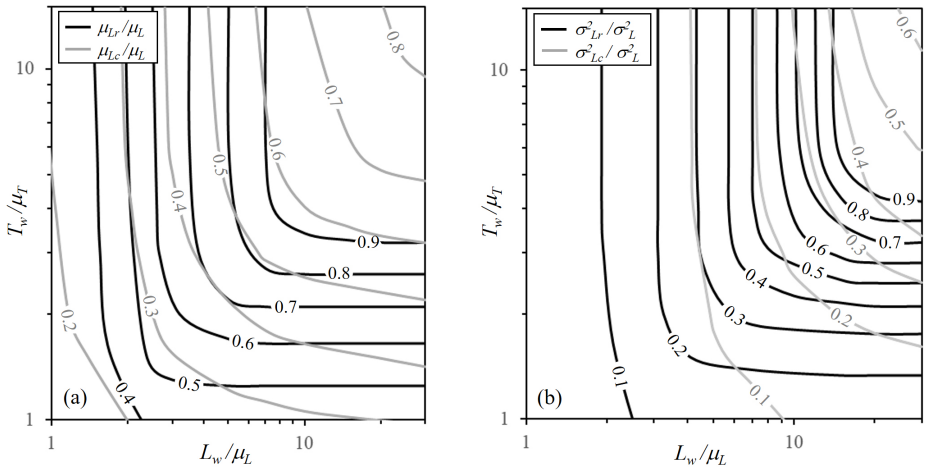


Figure 8. Contour lines of ratios between the biased moments and the true ones for hop length as a result of exponentially-distributed travel times and velocities. Black lines: reconstructed distribution; gray lines: censored distribution. (a) mean and (b) variance of hop length.

The results for the mean and variance values are shown in Figure 8 as contour lines of the ratios between the calculated values of the moments and their true values, over a $L_w/\mu_L, T_w/\mu_T$ plane (where μ_L and μ_T are the true means for hop lengths and travel times, respectively). The curves identify domains for the minimum sizes of the observation window that enable any given level of accuracy to be maintained in the evaluation of moments. For example, in order to achieve $0.90\mu_L$, a window larger than $L_w > 7\mu_L$ and $T_w > 3\mu_T$ is required for the mean of the reconstructed distribution, while $L_w >$

640 $15\mu_L$ and $T_w > 4.5\mu_T$ is required for the variance. The positive effect of correction of
641 censorship is evident for both moments.

642 7 Discussion

643 Censorship of particle motions can influence the measured proportions of all val-
644 ues of hop distances L and travel times T, while not involving just the truncation of the
645 longest motions. Our formulation of this idea hinges on the assumption of uniformly dis-
646 tributed starting positions x_0 and times t_0, in which case the probability of direct cen-
647 sorship of hop distances is equal to L/L_w and the probability of direct censorship of travel
648 times is equal to T/T_w. The pattern of entrainment positions (upon accumulating all
649 events over the duration of the imaging) must be uniform over L_w, masking any patch-
650 iness and intermittency of starting positions related to turbulence structures or other fac-
651 tors influencing entrainment at smaller scales. Similarly, the starting times should be ap-
652 proximately uniform over T_w. We emphasize that, even if one considers equilibrium sed-
653 iment transport, these represent ideal conditions, and are chosen as a convenient start-
654 ing point for our objective of illustrating the probabilistic elements of experimental cen-
655 sorship. If, for example, entrainment is approximately Poissonian (in space or time), then
656 the outcome is uniformity over these dimensions for sufficiently large N. However, cer-
657 tainly other factors might contribute to non-uniformity in starting positions and start-
658 ing times, for example, patchy surface-sediment texture or bedforms whose sizes are com-
659 parable to or smaller than the window dimensions. The condition of uniformity (or its
660 absence), therefore, must be evaluated for individual experiments. It may then be pos-
661 sible to incorporate any non-uniformity in evaluating and correcting the effects of cen-
662 sorship. For example, in the case of non-uniform starting times as in the data of Fathel
663 et al. (2015), one approach may be to bin the data within successive intervals for which
664 the starting times are approximately uniform, then weight each interval in the calcula-
665 tions according to the relative proportion of starting times in each. A similar approach
666 might be adopted for non-uniformity in starting positions.

667 Of particular importance is the theoretical and empirical demonstration of *indi-*
668 *rect* censorship related to the covariance between L and T. This means that, except for
669 the limiting cases in which either L_w is much larger than the largest measured hop dis-
670 tances, or T_w is much larger than the largest measured travel time, censorship of the marginal
671 distributions $f_L(L)$ and $f_T(T)$ generally should not be viewed separately. This result echoes
672 a similar admonition with respect to inferring the forms of these marginal distributions
673 in a manner that acknowledges the underlying mechanical basis of the covariance of these
674 quantities (Fathel et al., 2015; Furbish et al., 2016, 2017), an idea that merits further
675 examination as context for designing experiments to avoid effects of censorship.

676 The example involving synthetic data allows us to gain confidence in the formu-
677 lation of experimental censorship and the fidelity of the bias correction, given that the
678 underlying distributions are known and uniformity in starting positions and times are
679 specified. This example clearly illustrates the effect of truncation and the fact that the
680 truncated part of the underlying distribution cannot be recovered. It also illustrates the
681 occurrence of indirect censorship associated with the correlation between hop distances
682 and travel times, and the decrease in the fidelity of the bias correction as bounds imposed
683 by truncation are approached. Nonetheless, the reconstructed distributions are in all cases
684 better than the censored distributions. The example involving an experimental data set
685 highlights the limitations of not fully sampling all motions (while sampling just completed
686 motions), as well as the occurrence of indirect censorship as described in relation to the
687 synthetic data. Nonetheless, the bias correction suggests improvement in the reconstructed
688 distributions and, importantly, in the estimates of the mean values of the distributions,
689 which are larger than those reported by Fathel et al. (2015). Aside from bias correction
690 of these estimates, both examples highlight the importance of appropriately designing

experiments in order to mitigate effects of truncation that may come together with censorship (e.g., Radice et al., 2017).

We also emphasize that, depending on experimental objectives, required resolution, and constraints imposed by any experimental apparatus, censoring effects of window size sometimes cannot be avoided. The guidelines presented in Section 6 for choosing an observation window are based on results of censorship involving specified distributions with known parametric values. It was in general demonstrated that correcting the variance is (reasonably) much more challenging than correcting the mean, as well as that the bias correction is less effective for heavy-tailed quantities than for thin-tailed ones. Yet, in practice, these distributions and any covariance between L and T are not necessarily known *a priori*. One could start from some assumption on the shape of the distributions and related moments based on past studies, to provide some kind of practical (though imperfect) basis for choosing the appropriate measuring window. Furthermore, in practical terms, experiments likely require a certain amount of trial-and-error during design, for example, involving preliminary measurements and particle tracking to assess the likelihood and degree of censorship for given particle properties and planned flow conditions. This includes, for example, anticipating the likelihood of increasing effects of censorship for given sampling window dimensions in relation to increasing flow and near-bed fluid velocities, which typically induce longer hop distances and travel times (e.g., Campagnol, 2013; Fathel, 2016; Furbish et al., 2016).

Although we have not addressed here the question of assessing whether measured distributions represent time-invariant ensemble forms (Furbish et al., 2016), there is value in considering experimental measures of convergence to time-invariant forms with fixed moments. This may include, for example, convergence of moments to fixed values based on running averages (Anselmet et al., 1984), convergence of quartile-quartile (QQ) plots and probability-probability (PP) plots (Furbish et al., 2016), or comparison of the timescale of decay of autocorrelation functions relative to the sampling time (Fathel et al., 2015). Such measures, however, may be problematic with power-law distributions — which is likely with rest times R in the presence of particle-bed exchanges involving burial and exhumation (e.g., Ferguson et al., 2002; Voepel et al., 2013; Iwasaki et al., 2017). In order to fully examine rest times, we may need to redesign our experiments to involve long sampling times and a reassessment of needed resolution — in lieu of our current focus on relatively short hop distances and travel times and the resolution needed to characterize these motions, that is, frame rates of $\mathrm{O}(10^2)$ [s^{-1}] and spatial resolution of $\mathrm{O}(10^{-1})$ D [L] in current studies (Miao et al., 2018).

Although beyond the scope of this paper, an important objective is to use experimental data of hop distances, travel times and rest times (and other quantities such as particle velocities) to help identify the forms and parametric values of the underlying distributions of these random variables. However, we suggest that this is not just a statistical goodness-of-fit exercise, but rather, should be informed by mechanical considerations of these distributions (e.g., Furbish and Schmeeckle, 2013; Ancey and Heyman, 2014; Ancey et al., 2015; Furbish et al., 2016). Here we emphasize that the ideas outlined above concerning the bias correction for censorship are nonparametric techniques, whereas the fitting of data to an assumed underlying probability distribution with estimation of its parametric values represents a transition to parametric statistics. In this regard, the bias corrected data (up to the sampling window dimensions L_w and T_w) represent a starting point for selecting a possible underlying distributions based on inspection and standard statistical diagnostics, ideally in conjunction with mechanical considerations. Estimation of the distribution parameters may then follow, for example, maximum likelihood estimation or Bayesian analysis. Such methods use the data to aim at the best estimates of the parametric values of the assumed distribution. We emphasize, however, that Bayesian or similar analyses can only involve data up to L_w and T_w, and formally must be aimed at the truncated version of the assumed distribution (e.g., a truncated

exponential distribution or a truncated Pareto distribution). No information is available to constrain the truncated part of an assumed distribution. This also means that, strictly speaking, the parametric values of an underlying distribution that is truncated can never be known. Only the values of its *assumed* form can be estimated — which reinforces the value of connecting the choice of distribution with mechanical considerations. Our use of corrected data in the examples above to estimate parametric values is only possible because we have the luxury of specifying the underlying probability distributions and their parametric values *a priori* in order to illustrate the effects of censorship and truncation.

8 Conclusions

Experimental measurement of bed load particle motions can involve direct censorship of all hop distances L, travel times T and rest times R due to a finite sampling window length L_w and sampling time T_w, not just truncation of motions longer than the observation window. The likelihood of censorship increases with L, T and R for uniform starting and ending positions in the measuring window. For typical cases with finite covariance between L and T, censorship also acts indirectly; however, this does not happen for R as it is affected by time censorship only. As a consequence of direct and indirect experimental censorship, estimated frequency distributions of hop characteristics and their moments based only on completed hops are biased.

The novel procedure proposed in this work is able to reconstruct unbiased distributions if indirect censorship is absent; this is the case when either truncation effects along one of the two dimensions (space, time) are negligible or rest times are considered. For cases with significant spatial and temporal censorship and finite covariance between L and T, the procedure mitigates the effect of bias but does not fully eliminate it. In all such conditions the truncated part of the distributions (lengths larger than L_w and motion or rest times larger than T_w) is not reconstructed and remains unknown.

As truncation and, if the case, indirect censorship limit the possibility of obtaining complete and unbiased distributions and related moments, it would be desirable to set the size of the measuring windows in order to maintain such biases below a given level. This would require knowing the unbiased and untruncated distributions which are, in practice, normally not known. Therefore, only a recursive procedure can give an indication of the form of the distribution. As an alternative, the proposed procedure can provide estimates of the potential distortion as a function of the measuring window sizes once characteristics of the unknown probability distributions can be assumed.

The proposed approach to the problem relies on the sole hypothesis that entrainment and disentrainment events are uniformly distributed over the measurement window. Although such conditions are a reasonable approximation under equilibrium sediment transport, a systematic analysis with extension to nonuniform spatial and/or temporal event distributions would offer an improvement of the procedure.

A: Evaluating the Terms in (13)

Expanding the integrand in (13) and using the Heaviside functions to set the limits of integration leads to

$$\frac{N^c}{N} = \int_0^\infty \int_0^\infty f_{T,L}(T,L)\,\mathrm{d}T\,\mathrm{d}L$$

$$- \int_0^{L_w} \int_0^{T_w} \frac{T}{T_w} f_{T,L}(T,L)\,\mathrm{d}T\,\mathrm{d}L - \int_{L_w}^\infty \int_0^{T_w} \frac{T}{T_w} f_{T,L}(T,L)\,\mathrm{d}T\,\mathrm{d}L$$

$$- \int_{T_w}^\infty f_T(T)\,\mathrm{d}T$$

$$
\begin{aligned}
&- \int_0^{L_w} \int_0^{T_w} \frac{L}{L_w} f_{T,L}(T,L)\,\mathrm{d}T\,\mathrm{d}L - \int_0^{L_w} \int_{T_w}^{\infty} \frac{L}{L_w} f_{T,L}(T,L)\,\mathrm{d}T\,\mathrm{d}L \\
&\quad + \int_0^{L_w} \int_0^{T_w} \frac{L}{L_w}\frac{T}{T_w} f_{T,L}(T,L)\,\mathrm{d}T\,\mathrm{d}L \\
&\quad + \int_0^{L_w} \int_{T_w}^{\infty} \frac{L}{L_w} f_{T,L}(T,L)\,\mathrm{d}T\,\mathrm{d}L \\
&\quad - \int_{L_w}^{\infty} f_L(L)\,\mathrm{d}L \\
&\quad + \int_{L_w}^{\infty} \int_0^{T_w} \frac{T}{T_w} f_{T,L}(T,L)\,\mathrm{d}T\,\mathrm{d}L \\
&\quad + \int_{L_w}^{\infty} \int_{T_w}^{\infty} f_{T,L}(T,L)\,\mathrm{d}T\,\mathrm{d}L\,. \qquad \text{(A.1)}
\end{aligned}
$$

We now observe the following. The second term in the second line cancels the eighth line, and the second term in the fourth line cancels the sixth line. The first integral is equal to unity. The integral in the third line is the probability that a motion is greater than T_w. The integral in the seventh line is the probability that a motion is greater than L_w. The integral in the last line is the probability that a motion is greater than T_w *and* greater than L_w. Momentarily neglecting signs, the sum of the integrals in the third and seventh lines minus the integral in the last line must equal the probability that a motion is greater than T_w *or* greater than L_w. Thus, the integral in the first line minus this sum must equal the probability that a motion is less than T_w and less than L_w. That is, the sum of the integral in the first line, the third line, the seventh line and the last line is equal to

$$
\int_0^{L_w} \int_0^{T_w} f_{T,L}(T,L)\mathrm{d}T\mathrm{d}L\,. \qquad \text{(A.2)}
$$

Using these results, (A.1) reduces to

$$
\frac{N^c}{N} = \int_0^{L_w} \int_0^{T_w} \left(1 - \frac{L}{L_w} - \frac{T}{T_w} + \frac{L}{L_w}\frac{T}{T_w}\right) f_{T,L}(T,L)\mathrm{d}T\mathrm{d}L\,. \qquad \text{(A.3)}
$$

This is equivalent to (14) in the text.

B: The Distribution of a Product of Random Variables

Let X and Y be two random variables with the joint probability density function $f_{X,Y}(X,Y)$. The random variable $Z = XY$ is characterized by the following marginal density function (details not shown):

$$
f_Z(Z) = \int_{-\infty}^{\infty} \frac{1}{|w|} f_{X,Y}\left(w, \frac{Z}{w}\right) \mathrm{d}w\,. \qquad \text{(B.1)}
$$

If X and Y are two i.i.d. random variables with uniform marginal probability densities in the interval $[0,1]$, then (B.1) becomes

$$
f_Z(Z) = \int_{-\infty}^{\infty} \frac{1}{|w|} f_X(w) f_Y\left(\frac{Z}{w}\right) \mathrm{d}w\,, \qquad \text{(B.2)}
$$

with $f_X = 1$ for $0 \le w \le 1$ and $f_Y = 1$ for $0 \le Z \le w$. Therefore (B.2) becomes

$$
f_Z(Z) = \int_Z^1 \frac{1}{|w|}\,\mathrm{d}w = \begin{cases} -\ln Z & 0 \le Z \le 1 \\ 0 & \text{otherwise} \end{cases}\,. \qquad \text{(B.3)}
$$

If X and Y are two i.i.d. random variables with marginal exponential probability density functions in the interval $[0, \infty)$, then $f_X = \lambda e^{-\lambda w}$ for $0 \leq w < \infty$ and $f_Y = \lambda e^{-\lambda Z/w}$ for $0 \leq Z < \infty$ with $\lambda > 0$. Therefore (B.2) becomes

$$f_Z(Z) = \lambda^2 \int_0^\infty \frac{1}{w} e^{-\lambda(w+Z/w)} \, \mathrm{d}w = \begin{cases} 2\lambda^2 K_0(2\lambda\sqrt{Z}) & 0 \leq Z < \infty \\ 0 & Z < 0 \end{cases} , \qquad \text{(B.4)}$$

where K_0 is the modified Bessel function of order zero.

Acknowledgments

We acknowledge support by the Research Executive Agency through the 7th Framework Programme of the European Union (ITN-316546 'HYTECH' to FB), and the U.S. National Science Foundation (EAR-1226076 and EAR-1735992 to DJF). Monica Riva provided the proofs in Appendix B for probability distributions resulting from the product of two random variables. We appreciate discussions with Jonathan Gilligan regarding goodness-of-fit methods. John Buffington, Christophe Ancey, Raleigh Martin and two anonymous reviewers provided several cogent comments that enabled the manuscript to be considerably improved. The supplemental files may be accessed from the institutional site of the Politecnico di Milano at http://intranet.dica.polimi.it/people/radice-alessio/ and from a permanent repository at https://doi.org/10.5281/zenodo.1550750.

References

Ancey, C., & Heyman, J. (2014). A microstructural approach to bed load transport: mean behavior and fluctuations of particle transport rates. *Journal of Fluid Mechanics*, 744, 129–168. https://doi.org/10.1017/jfm.2014.74.

Ancey, C., Bohorquez, P., & Heyman, J. (2015). Stochastic interpretation of the advection-diffusion equation and its relevance to bed load transport. *Journal of Geophysical Research: Earth Surface*, 120, 2529–2551. https://doi.org/10.1002/2014JF003421.

Anselmet, Y., Gagne, Y., & Hopfinger, E. J. (1984). High-order velocity structure functions in turbulent shear flows. *Journal of Fluid Mechanics*, 140, 63–89. https://doi.org/10.1017/S0022112084

Ballio, F., Fathel, S. L., Furbish, D. J., & Radice, A. (2018a). On experimental censorship of particle hops in bed-load transport. *Proc. of River Flow 2018, Lyon, France.* https://doi.org/10.1051/e3sconf/20184005054.

Ballio, F., Pokrajac, D., Radice, A., & Hosseini Sadabadi, S. A. (2018b). Lagrangian and Eulerian description of bed load transport. *Journal of Geophysical Research: Earth Surface*, 123, 384–408. https://doi.org/10.1002/2016JF004087.

Bialik, R. J., & Karpinski, M., (2018). On the effect of the window size on the assessment of particle diffusion. *Journal of Hydraulic Research*, 56(4), 560–566. https://doi.org/10.1080/00221686.2

Bradley, D. N., Tucker, G. E., & Benson, D. A. (2010). Fractional dispersion in a sand bed river. *Journal of Geophysical Research: Earth Surface*, 115, F00A09, https://doi.org/10.1029/2009JF001

Campagnol, J. (2013). Characterization of bed load sediment transport at the grain scale. Ph.D. thesis, Politecnico di Milano, Milan, Italy.

Campagnol, J., Radice, A., Ballio, F., & Nikora, V. (2015). Particle motion and diffusion at weak bed load: Accounting for unsteadiness effects of entrainment and disentrainment. *Journal of Hydraulic Research*, 53(5), 633-648. https://doi.org/10.1080/00221686.2015.1085920.

Campagnol, J., Radice, A., Nokes, R., Bulankina, V., Lescova, A., & Ballio, F. (2013). Lagrangian analysis of bed-load sediment motion: Database contribution. *Journal of Hydraulic Research*, 51(5), 589–596. https://doi.org/10.1080/00221686.2013.812152.

Einstein, H. A. (1937). Bedload transport as a probability problem. Ph. D. thesis, Mitt. Versuchsanst. Wasserbau Eidg. Tech. Hochsch, Zürich, Switzerland.

Einstein, H. A. (1950). The bed-load function for sediment transportation in open channel flows. *Technical Bulletin 1026*, Soil Conservation Service, U. S. Department of Agriculture, Washington, D. C.

Fathel, S. L. (2016). Experimental analysis of bed load sediment motions using high-speed imagery in support of statistical mechanics theory. Ph.D. thesis, Vanderbilt University, Nashville, Tennessee.

Fathel, S.L., Furbish, D. J., & Schmeeckle, M. W. (2015). Experimental evidence of statistical ensemble behavior in bed load sediment transport. *Journal of Geophysical Research: Earth Surface*, 120(11), 2298–2317. https://doi.org/10.1002/2015JF003552.

Ferguson, R. I., Bloomer, D. J., Hoey, T. B., & Werritty, A. (2002). Mobility of river tracer pebbles over different timescales. *Water Resources Research*, 38(5), 31–39.

Furbish, D. J., Haff, P. K., Roseberry, J. C., & Schmeeckle, M. W. (2012). A probabilistic description of the bed load sediment flux: 1. Theory. *Journal of Geophysical Research: Earth Surface*, 117, F03031. https://doi.org/10.1029/2012JF002352.

Furbish, D.J., & Schmeeckle, M. W. (2013). A probabilistic derivation of the exponential-like distribution of bed load particle velocities. *Water Resources Research*, 49, 1537–1551. https://doi.org/10.1002/wrcr.20074.

Furbish, D. J., Fathel, S. L., & Schmeeckle, M. W. (2017). Particle motions and bed load theory: The entrainment forms of the flux and the Exner equation. In: Tsutsumi, D. and Laronne, J. B. (eds.), *Gravel-bed Rivers: Processes and Disasters*, Wiley-Blackwell, ISBN: 978-1-118-97140-6.

Furbish, D. J., Schmeeckle, M. W., Schumer, R., & Fathel, S. L. (2016). Probability distributions of bed-load particle velocities, accelerations, hop distances and travel times informed by Jaynes's principle of maximum entropy. *Journal of Geophysical Research: Earth Surface*, 121(7), 1373–1390. https://doi.org/10.1002/2016JF003833.

Ganti, V., Meerschaert, M. M., Foufoula-Georgiou, E., Viparelli, E., & Parker, G. (2010). Normal and anomalous diffusion of gravel tracer particles in rivers. *Journal of Geophysical Research: Earth Surface*, 115, F00A12. https://doi.org/10.1029/2008JF001222.

Hassan, M. A., Church, M., & Schick, A. P. (1991). Distance of movement of coarse particles in gravel bed streams. *Water Resources Research*, 27(4), 503–511. https://doi.org/10.1029/90WR02762.

Heyman, J. (2014). A study of the spatio-temporal behaviour of bed load transport rate fluctuations. Ph.D. thesis, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Heyman, J., Bohorquez, P., & Ancey, C. (2016). Entrainment, motion and deposition of coarse particles transported by water over a sloping mobile bed. *Journal of Geophysical Research: Earth Surface*, 121(10), 19311952, https://doi.org/10.1002/2015JF003672.

Houssais, M., & Lajeunesse, E. (2012). Bedload transport of a bimodal sediment bed. *Journal of Geophysical Research: Earth Surface*, 117, F04015. https://doi.org/10.1029/2012JF002490.

Iwasaki, T., Nelson, J., Shimizu, Y., & Parker, G. (2017). Numerical simulation of large-scale bedload particle tracer advection-dispersion in rivers with free bars. *Journal of Geophysical Research: Earth Surface*, 122(4), 847–874. https://doi.org/10.1002/2016JF003951.

Lajeunesse, E., Malverti, L., & Charru, F. (2010). Bed load transport in turbulent flow at the grain scale: Experiments and modeling. *Journal of Geophysical Research: Earth Surface*, 115, F04001. https://doi.org/10.1029/2009JF001628.

Lajeunesse, E., Devauchelle, O., Houssais, M., & Seizilles, G. (2013). Tracer dispersion in bedload transport. *Advances in Geosciences*, 37, 1–6. https://doi.org/10.5194/adgeo-37-1-2013.

Martin, R. L., Jerolmack, D. J., & Schumer, R. (2012). The physical basis for anomalous diffusion in bed load transport. *Journal of Geophysical Research: Earth Surface*, 117, F01018. https://doi.org/10.1029/2011JF002075.

Miao, W., Cao, L., Zhong, Q., & Li, D. (2018). Influence of the time interval on image-based measurement of bed-load transport. *Journal of Hydraulic Research*, 56(4), 567-575. https://doi.org/10.1080/00221686.2017.1399938.

Nakagawa, H., & T. Tsujimoto, T. (1980), Sand bed instability due to bed load motion, *Journal of the Hydraulics Division, American Society of Civil Engineers*, 106(12), 2029–2051.

Nakagawa, H., & Tsujimoto, T. (1984), Spectral analysis of sand bed instability, *Journal of Hydraulic Engineering*, 110(4), 467–483. https://doi.org/10.1061/(ASCE)0733-9429(1984)110:4(467).

Parker, G., Paola, C., & Leclair, S. (2000). Probabilistic Exner sediment continuity equation for mixtures with no active layer. *Journal of Hydraulic Engineering*, 126(11), 818–826. https://doi.org/10.1061/(ASCE)0733-9429(2000)126:11(818).

Radice, A., Sarkar, S., & Ballio, F. (2017). Image-based Lagrangian particle tracking in bed-load experiments. *Journal of Visualized Experiments*, 125, e55874. https://doi.org/10.3791/55874.

Roseberry, J. C., Schmeeckle, M. W., & Furbish, D. J. (2012). A probabilistic description of the bed load sediment flux: 2. Particle activity and motions. *Journal of Geophysical Research: Earth Surface*, 117, F03032. https://doi.org/10.1029/2012JF002353.

Sayre, W., and Hubbell, D. (1965). Transport and dispersion of labeled bed material, North Loup River, Nebraska. *U.S. Geological Survey Professional Paper*, 433-C. https://pubs.usgs.gov/pp/0433c/report.pdf.

Seminara, G., Solari, L., & Parker, G. (2002). Bed load at low Shields stress on arbitrarily sloping beds: Failure of the Bagnold hypothesis. *Water Resources Research*, 38(11), 1249. https://doi.org/10.1029/2001WR000681.

Taylor, G. I. (1921), Diffusion by continuous movements, *Proceedings of the London Mathematical Society*, 20, 196–212.

Tsujimoto, T. (1978), Probabilistic model of the process of bed load transport and its application to mobile-bed problems. Ph.D. thesis, Kyoto University, Kyoto, Japan.

Voepel, H., Schumer, R., & Hassan, M. A. (2013). Sediment residence time distributions: Theory and application from bed elevation measurements. *Journal of Geophysical Research: Earth Surface*, 118(4), 2557–2567. https://doi.org/10.1002/jgrf.20151.

Wilcock, P. R. (1997). Entrainment, displacement and transport of tracer gravels. *Earth Surface Processes and Landforms*, 22(12), 1125–1138.

Wong, M., Parker, G., DeVries, P., Brown, T. M., & Burges, S. J. (2007). Experiments on dispersion of tracer stones under lower-regime planebed equilibrium bed load transport. *Water Resources Research*, 43, W03440. https://doi.org/10.1029/2006WR005172.