Windblown sand saltation: A statistical approach to fluid threshold shear velocity

Windblown Sand Saltation: a statistical approach to threshold shear velocity

Lorenzo Raffaele[a,d,*], Luca Bruno[b,d], Franco Pellerey[c], Luigi Preziosi[c,d]

[a] *Politecnico di Torino, Department of Structural, Geotechnical and Building Engineering, Corso Duca degli Abruzzi 24, I-10129, Torino, Italy*
[b] *Politecnico di Torino, Department of Architecture and Design, Viale Mattioli 39, I-10125, Torino, Italy*
[c] *Politecnico di Torino, Department of Mathematical Sciences, Corso Duca degli Abruzzi 24, I-10129, Torino, Italy*
[d] *Windblown Sand Modeling and Mitigation joint research group*

Abstract

The reliable prediction in probabilistic terms of the aeolian events magnitude is a key element for human activities in arid regions. Threshold shear velocity is in turn a key component of such a prediction. It suffers the effects of a number of uncertainties, such as the ones related to the physical phenomena, measurement procedures, and modelling. Semi empirical models are often fitted to a small amount of data, while recent probabilistic models needs the probability distribution of a number of random variables. Triggered by this motivation, this paper proposes a purely statistical approach to threshold shear velocity for sands, treated as a single comprehensive random variable. The data ensemble is defined collecting a huge number of studies available in literature. Estimates of probability density functions of threshold shear velocity for given sand classes are obtained. The obtained statistical moments are critically compared to some deterministic semi empirical models refitted to the same collected data. The proposed statistical approach allows to obtain high order statistics useful for practical purposes.

Keywords: windblown sand, saltation, threshold shear velocity, uncertainty, statistics

1. Introduction

Aeolian sand transport is a complex process that is induced by the interaction between subfields such as wind, air suspended particles and bed-particles. It contributes to soil erosion and landform evolution. Understanding and modeling its features is of fundamental interest in many research fields. Beside the importance of windblown sand and dust to the Earth sciences (Kok et al., 2012), from the engineering perspective, simulating windblown sand phenomena is relevant because of the interaction with a number of human activities and related infrastructures in arid environments (e.g. Middleton and Sternberg, 2013; Rizvi, 1989; Alghamdi and Al-Kahtani, 2005; Zhang et al., 2007; Cheng and Xue, 2014). In the infrastructure design perspective and within a probabilistic approach to design, the engineer is interested in relating a sand erosion or transport condition to a given, preferably low enough, probability of exceedance.

Among the transport mechanisms responsible of sand transport, saltation largely prevails in term of sand mass. The evaluation of the involved sand flux is usually given in term of sand transport rate by several laws, revised e.g. in Dong et al. (2003); Kok et al. (2012); Sherman and Li (2012). Most of such laws require the definition and evaluation of the fluid or static threshold, i.e. the value of the wind shear stress at which saltation is initiated. Usually, such a threshold is given in terms of fluid threshold shear velocity u_{*t}. In turn, such a threshold value depends on a number

*Corresponding author. Tel: (+39) 011.090.4870. Fax: (+39) 011.090.4999.
Email address: `lorenzo.raffaele@polito.it` (Lorenzo Raffaele)
URL: `http://www.polito.it/wsmm` (Lorenzo Raffaele)

of parameters belonging to both the wind and sand subfields.
Several authors have investigated such dependencies and proposed fluid threshold models, many of them reported e.g. in the overviews by Shao (2008); Pye and Tsoar (2009); Merrison (2012); Kok et al. (2012).
Systematic *experimental studies* addressed to u_{*t} versus the grain diameter d were carried out by e.g. Bagnold (1937), Chepil (1945), Zingg (1953), Fletcher (1976), Iversen et al. (1976). These measurements ground the consolidated literature data base. They are reported in Figure 1. A significant scatter among data can be observed notably at low values of the particle diameter. However, two general trend can be observed, divided by a local minimum at about 75-100μm (Kok et al., 2012).
A number of *deterministic models* of the threshold shear velocity have been proposed in literature so far. They can be categorized in two classes with respect to both modelling scale and goal. *Microscopic models* discuss the equilibrium of the moments of the forces acting on the single particle resting on a bed of other particles (Shao, 2008). They aim at pointing out the physical phenomena underlying each force and at modelling it. In a general framework, entraining aerodynamic forces (drag and lift ones) induce saltation, while stabilizing forces (gravitational and the interparticle ones) counteract them (Greeley and Iversen, 1985; Shao and Lu, 2000). On one hand, the effective gravitational force including buoyancy, and the drag force correspond to well known phenomena and their modelling is widely accepted, see e.g. Greeley and Iversen (1985) and the cited reviews. On the other hand, the same does not hold for the other forces: the resultant lift force results form the Saffman one (Saffman, 1965) and the lift induced by vortical structures; the overall interparticle force results from several kinds of forces, including van der Waals forces, water adsorption forces and electrostatic forces. Although interparticle forces are expected to scale with the soil particle size (e.g. Shao and Lu, 2000), their modelling for aspherical and rough sand and dust remains poorly understood (Kok et al., 2012). In particular, such forces depend upon a number of parameters such as surface cleanliness, surface roughness at micro/nano meter scale, air and grain humidity, mineralogy and surface contaminants affecting hydrophilicity (Merrison, 2012). Semi-empirical *macroscopic models* aim at approximating the threshold shear velocity trend versus the particle diameter. Some of them are compared to the experimental data in Figure 1(a). Because of the above modelling difficulties, they do not analytically include the contribution of lift and interparticle forces while they explicitly retain the gravitational and drag ones. Any other contribution is accounted for in a semi empirical approach by introducing one or more free parameter(s), and the value of the latter obtained by fitting experimental data. The pioneering Bagnold (1941) model involves a single dimensionless constant A_B, i.e. independent from the grain diameter or, in other terms, from Reynolds number: a monotonic increasing trend of $u_{*t}(d)$ results. The model by Iversen and White (1982) defines the same parameter $A(\mathrm{Re}_{*t})$ as a piece-wise empirical function of the friction Reynolds number Re_{*t} to mimic the effects of lift and interparticle forces: the resulting $u_{*t}(d)$ law is no longer monotonic and qualitatively reflect the trend of the experimental data. The model by Shao and Lu (2000) is more compact than the previous one. It neglects the Re_{*t} dependency, and at the same time generalizes the Bagnold one by introducing a novel correction term to account for the interparticle forces. A second dimensional constant free parameter γ [N/m] is included in the correction term. More recently, McKenna (2003) have considered the effect of soil moisture on the interparticle cohesive force by defining $\gamma(\Delta P, d)$ as a function of the capillary-suction pressure deficit and of the grain diameter. Other laws of u_{*t} have been proposed for natural surfaces: they account for the effects of soil texture, soil moisture, salt concentration, surface crust, vegetation and/or pebbles on the surface. The review of such models is out of scope of the present paper: interested readers can refer to Shao (2008); Webb and Strong (2011).
The *probabilistic modelling approach* is a promising alternative to the deterministic one, having in mind that the modelling difficulties outlined above are mainly due to the uncertainties which affect the sand-acting forces (Merrison, 2012). We suggest to ascribe such uncertainties to distinct comprehensive sources of uncertainty, that are:

- randomness of the *grain features*. Among these features, *grain size distribution* is traditionally recognized in literature as an important sand feature affecting u_{*t} (e.g. Edwards and Namikas, 2015, and included references), beside the mean diameter. In fact, smaller particles interspersed among the large particles provide additional cohesive forces in natural sands, resulting in higher threshold conditions (Roney and White, 2004). The early studies on $u_{*,t}$ (e.g. Bagnold, 1937) usually assume nominally uniform sand, but this restriction clearly does not hold in a probabilistic framework. Others random/uncontrolled features are raised in literature, such as grain shape, surface microstructure (e.g. Duan et al., 2013), grain position relative to the other bed particles (Phillips, 1980), grain mineralogy and surface cleanliness (Merrison, 2012);

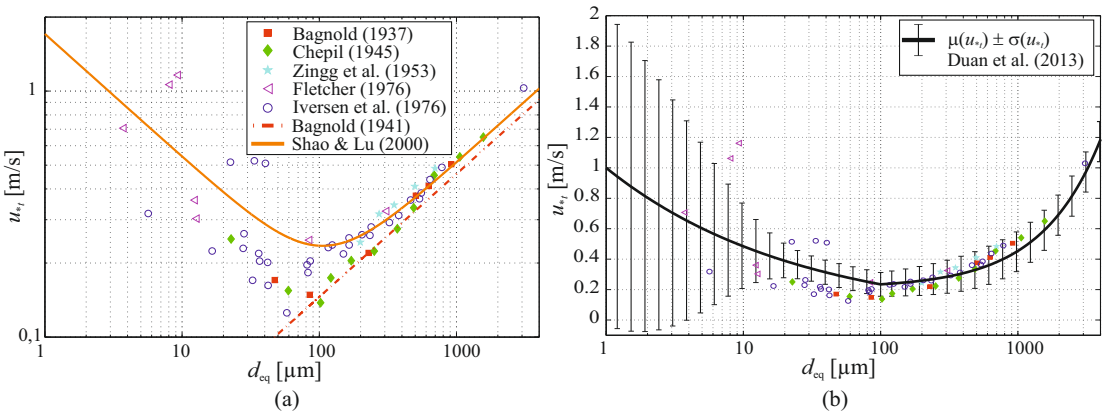

Figure 1: Threshold friction velocity: experimental data (symbols) compared with semi-empirical deterministic models (a, redrawn after Kok et al., 2012), and the probabilistic model by Duan et al. (2013) (b)

- the inborn variability of the *environmental conditions* even in wind tunnel facilities, e.g. air temperature and air relative humidity (e.g. Greeley and Iversen, 1985; Jones et al., 2002);

- *epistemic uncertainties* due to modelling, measurement procedures and techniques adopted to evaluate the bulk granulometry (Blott and Pye, 2006; Zhang et al., 2014) and/or the threshold shear velocity (Barchyn and Hugenholtz, 2011).

The smaller the grain size, the major the role of the interparticle forces, the higher the expected effect of the above uncertainties on the threshold shear velocity. Having this qualitative dependency in mind, Shao (2008) conjectured that while it is meaningful defining a threshold shear velocity as a single value for sand-sized particles, it is not meaningful to do the same for dust particles. This conjecture seems to be confirmed by the scatter of the experimental data at low values of d even for a common nominal setup condition (Figure 1a). Zimon (1982) first suggested to treat cohesive forces acting upon dust particles as random variables (r.v.s). He argued from experimental data that the cohesive forces probability distribution can be approximated by a lognormal one. Following Zimon's findings, Shao (2008) assumed that also the threshold shear velocity is log-normally distributed. Such an assumption looks questionable from an analytical point of view even by assuming the cohesive force the sole random variable among the grain acting forces: in fact, u_{*t} does not result from a simple rescaling of the cohesive force.
Fueled by these problem features, Duan et al. (2013) have recently proposed a probabilistic model for threshold shear velocity. The study is grounded on a microscopic model, where the drag force, the electrostatic force, the gravity force and the cohesion force are described as functions of four microscopic r.v.s owing to the random nature of the microstructure of soil surface, of the particle shape and of positions in the bed irregular particle. The threshold shear velocity is then expressed as a function of these random quantities, some of them independent, some dependent, and its Probability Density Function (PDF) then evaluated through a statistical estimation of the distributions of the predictors. Subsequently, the mean value and standard deviation of the threshold shear velocity are fitted as functions of d. On one hand, the innovative model by Duan et al. (2013) has the merit to tackle for the first time the statistical characterization of the threshold shear velocity. On the other hand, the obtained results (Figure 1b) are not not entirely convincing. First, at very low values of d, mean value minus standard deviation $\mu(u_{*t}) - \sigma(u_{*t})$ is negative, while $u_{*t} \in \mathbb{R}^+$. Second, the standard deviation $\sigma(u_{*t})$ is monotonically increasing for $d \geq 100\ \mu$m and asymptotically tends to 0.132, while the scatter of experimental data clearly decreases for increasing d. Third, the mean $\mu(u_{*t})$ is a linear function of d for $d > 100\ \mu$m, while its deterministic counterpart, i.e. the nominal values obtained by semi-empirical macroscopic models, is not. Finally, the study of Duan et al. (2013) does not evaluate high order statistical moments of the threshold shear velocity, i.e. skewness. In our opinion, such critical features can be ascribed to both modelling and technical difficulties. Among the former ones, the challenging task in writing a microscopic model inclusive of all the r.v.s affecting the sand grain acting forces. Among the second ones, the difficulties in obtaining probability

distribution for each microscopic r.v. from measurements and in handling mutually dependent r.v.s.

According to the authors, three main questions rise from the state of art briefly reviewed above: i. Is the deterministic approach able to face to the sources of uncertainties introduced above? ii. Is a statistical approach to the threshold shear velocity required only for dust particles or for sand-sized particles too? iii. How to overcome the difficulties encountered by probabilistic mechanical models in handling a number of microscopic r.v.s?

The present study aims at contributing in shedding some light on such issues. The deterministic approach is critically reconsidered in the light of a huge collection of experimental measurements. Then, a purely statistical approach to threshold shear velocity is proposed, where the effects of all kinds of uncertainty sources are comprehensively included and merged. Finally, the two approaches are compared.

The paper develops accordingly to the above objectives through the following sections. In Section 2 the collected measurements and the resulting ensemble of selected data are described. In Section 3 some semi-empirical macroscopic models are refitted to the ensemble by means of non-linear regression. In Section 4 the statistical description of the threshold shear velocity is given by referring to both analytical distributions (Subsect. 4.1) and the non-parametric one (Subsect. 4.2). The deterministic and statistical approach are critically compared in Section 5, while conclusions and research perspectives are outlined in Section 6.

2. Data collection and ensemble setting

The data already collected in Figure 1 are complemented by additional experimental measures collected from review papers (Kok et al., 2012; Edwards and Namikas, 2015) and studies addressed to the evaluation of sand transport rate for single particle diameters. Table 1 summarizes in chronological order the considered studies, while the complete ensemble of retained sand experimental measurements of u_{*t} is plotted in Figure 2(a) versus d.

All studies test nominally dry granular matters. Except for Fletcher (1976) and Iversen et al. (1976), granular matter

Table 1: Collected setups: reference paper, number of samples, reference diameter

	#	d [mm]
Bagnold (1937)	6	$0.05 \leq d \leq 0.92$
Chepil (1945)	11	$0.02 \leq d \leq 1.57$
Kawamura (1951)	2	0.25, 0.31
Zingg (1953)	5	$0.20 \leq d \leq 0.72$
Chepil (1959)	5	$0.20 \leq d \leq 0.72$
Belly (1964)	1	0.44
Kadib (1964)	1	0.15
Lyles and Krauss (1971)	3	$0.24 \leq d \leq 0.72$
Fletcher (1976)	7	$0.01 \leq d \leq 0.31$
Iversen et al. (1976)	33	$0.01 \leq d \leq 3.09$
Logie (1981)	4	$0.15 \leq d \leq 0.43$
Logie (1982)	1	0.24
Horikawa et al. (1983)	1	0.28
McKenna Neuman and Nickling (1989)	3	$0.19 \leq d \leq 0.51$
Nalpanis et al. (1993)	2	0.12, 0.19
Nicking and McKenna Neuman (1997)	1	0.20
Dong et al. (2002)	9	$0.13 \leq d \leq 0.90$
Dong et al. (2003)	9	$0.13 \leq d \leq 0.90$
Cornelis and Gabriels (2004)	3	$0.16 \leq d \leq 0.36$
McKenna Neuman (2004)	1	0.27
Roney and White (2004)	12	$0.31 \leq d \leq 0.39$

is sand and/or dust. For each considered study, the number # of the tested samples is given: an overall collection of

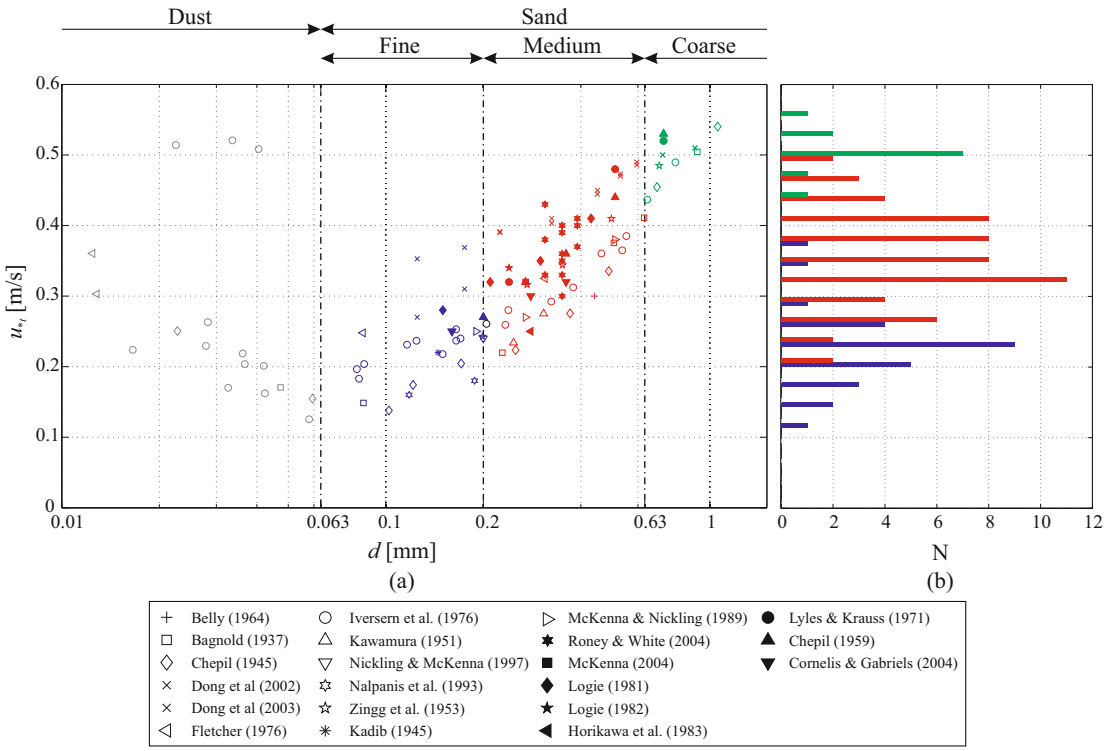

Figure 2: Threshold shear velocity measurements collected in literature (a), and their histograms for each sub-ensemble (b): fine (blue), medium (red), coarse (green) sands

120 setups follows. For each setup, the cited papers provide the grain mean, or median, diameter. In order to account for the effect of different density of the grain constitutive materials, the equivalent particle diameter defined by Chepil (1951); Kok et al. (2012) is evaluated. In Table 1 and in the following the equivalent reference diameter is noted as d for the sake of conciseness. In the rest of the paper, d is treated as a deterministic quantity.
A significant dispersion of the data can be easily observed in Figure 2(a), notably for fine and medium sands. In other terms, u_{*t} takes different values at the same d. Such a feature suggests the ensembles are potentially constituted by heterogeneous setups or, in other terms, setup parameters other than d are conjectured to affect the quantity of interest. As anticipated in Sect. 1, several uncertain/uncontrolled parameters can be detailed for the considered studies.

- In the selected setups the *grain size distribution* is often qualitatively described, e.g. "as uniform as possible" in Bagnold (1937), "very well and poorly sorted" in Belly (1964), "naturally graded" in Kawamura (1951). Such a qualitative description is usually complemented by the nominal size-range of grains (e.g. Bagnold, 1937; Dong et al., 2003), while in some papers the cumulative grain size distribution is plotted (e.g. Belly, 1964; Nalpanis et al., 1993; Kawamura, 1951; Nicking and McKenna Neuman, 1997; Roney and White, 2004). Recently, Edwards and Namikas (2015) have made an effort to evaluate a measure of the diameter variability by evaluating the sorting coefficient for a number of studies: in spite of some difficulties in obtaining such a measure from nominal size-range, it is worth recalling that non negligible variability (e.g. sorting$\approx$ 0.05, coefficient of variation $c.o.v. \approx 0.12$ in Chepil (1959)) results also from sieving addressed to obtain sands as uniform as possible. Even greater variability characterizes natural sands (e.g. sorting$\approx$ 0.65, $c.o.v. \approx 0.35$ in Kawamura (1951)). Other randomness of the grain features (e.g. grain shape, surface microstructure, grain position relative to the other bed particles, grain mineralogy) are not specified in the collected studies.

- *Air humidity* during wind tunnel tests is given and systematically addressed only by Kadib (1964) to our best knowledge.

- Analogously, u_{*t} *measurements* and *post processing techniques* are heterogeneous among the studies (Blott and Pye, 2006; Zhang et al., 2014), Roney and White (2004) prove their effects on fluid threshold shear velocity by adopting three different techniques.

- The *quantitative definition of the fluid threshold shear velocity* is not commonly adopted in all the studies, only Lyles and Krauss (1971) provide several u_{*t} values from visual observations depending on the kind of grain motion.

In short, the experimental data ensemble is naturally and inevitably affected by a huge number of uncertainties, belonging to both physical setup and epistemic uncertainty.

The present paper is devoted to the characterization of threshold shear velocity of sand only. Hence, setups adopting dust, i.e. having $d < 0.063$ mm according to ISO14688-1:2002, are first discarded (empty light grey markers in Figure 2-a). An overall sand ensemble having # = 97 results. A deterministic dependency on d is clearly confirmed along the ensemble. Hence, on one hand, the definition of sub-ensemble is therefore advisable. On the other hand, the limited number of realization in the ensemble does not allow to define a huge number of sub-ensembles. Having these issues in mind, we gather realizations in three sand classes according to the common practice in aeolian geomorphology and referring to ISO14688-1:2002:

- Fine sand ($0.063 < d \leq 0.2$ mm), # = 27;

- Medium sand ($0.2 < d \leq 0.63$ mm), # = 58;

- Coarse sand ($0.63 < d \leq 1.2$ mm), # = 12.

It is worth noting that "very coarse" sand as defined by Friedman and Sanders (1978) is not included in the coarse sub-ensemble because scarceness of available experimental data. The histograms for each sub-ensemble are plotted in Figure 2(b). The adequateness of each sub-ensemble cardinality in providing accurate statistics will be carefully checked in the study.

3. Deterministic approach: non-linear regression

Prior to the statistical analysis of the sub-ensembles above, non-linear regression is applied to the whole collected data in order to refit some of the semi-empirical macroscopic models available in literature. The refitting objective is twofold: on the one hand, the field of application is limited to sands, i.e. on an entrainment physics simpler than the one governing dusts; on the other hand, model parameters are fitted to a number of data higher than the one originally adopted by the authors of the models. Bagnold (1941) (Eq.1) and Shao and Lu (2000) (Eq.2) models are selected because of their compactness, i.e. their dependence from a small number of empirical parameters (A_b, A_s and γ). The two semi-empirical models are

$$u_{*t} = A_b \sqrt{\frac{\rho_p - \rho_a}{\rho_a} g d}, \tag{1}$$

$$u_{*t} = A_s \sqrt{\frac{\rho_p - \rho_a}{\rho_a} g d + \frac{\gamma}{\rho_a d}}. \tag{2}$$

where ρ_p and ρ_a are particle and air density, respectively, and g is gravitational acceleration. Beside the single-valued estimates of a goodness of fit, for each model the prediction Confidence Intervals (CIs) are evaluated at 5th and 95th percentiles, i.e. the interval within which the true value is expected to lie. Figure 3 compares the refitted laws to the original ones, while the corresponding model parameters are summarized in Table 2. The following remarks can be outlined:

- generally speaking, the refitted laws predict higher values of u_{*t} for given d. It is worth pointing out that the ensemble includes a number of poorly sorted and natural sands, while the ensemble originally adopted by Bagnold (1941) and Shao and Lu (2000) were limited to sand as uniform as possible. Hence, interspersed small particles provide additional cohesive forces also for medium and coarse natural sands (Roney and White, 2004);

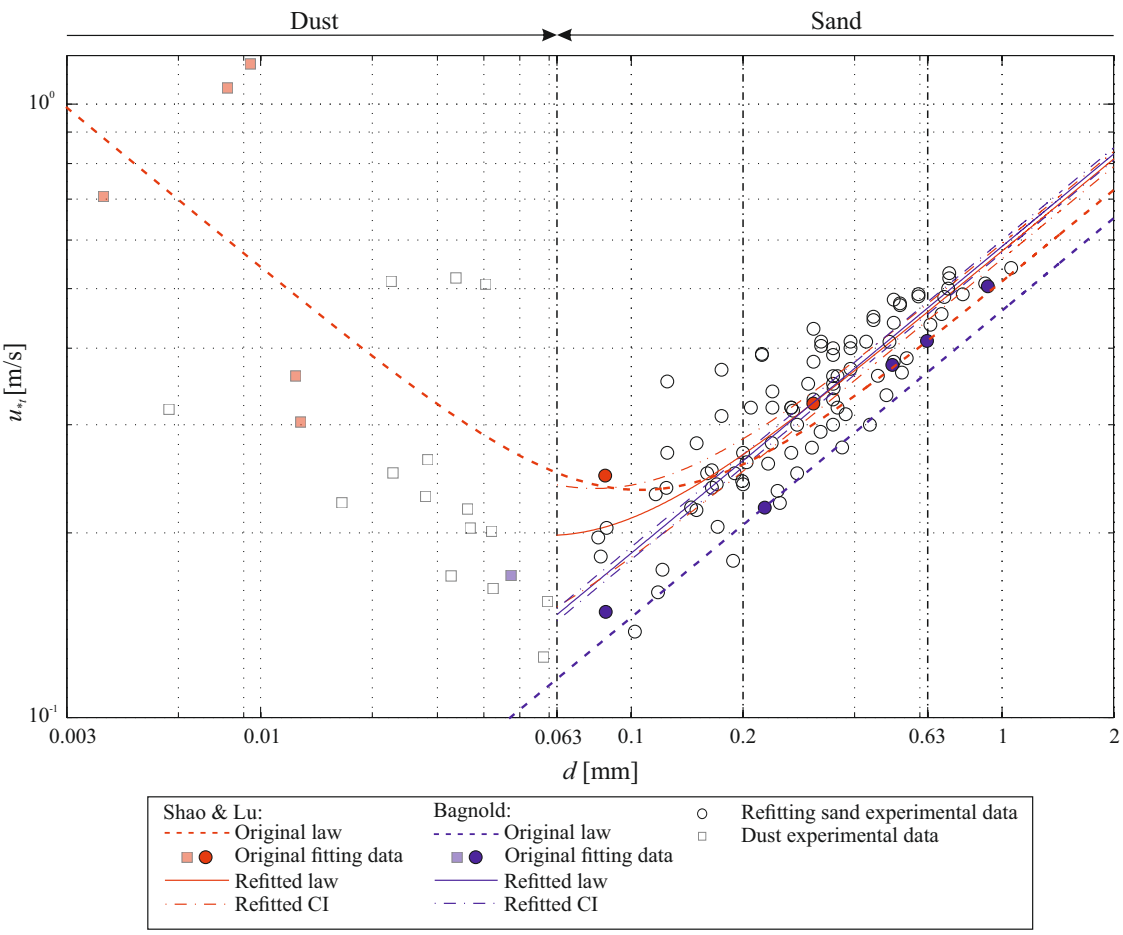

Figure 3: Non-linear regression and CIs for Bagnold (1941) and Shao and Lu (2000)

Table 2: Original and refitted model parameters

	Original parameters		Refitted parameters	
	Bagnold (1941)	Shao and Lu (2000)	Bagnold (1941)	Shao and Lu (2000)
A [-]	0.100	0.111	0.127	0.125
γ [N/m]	-	2.9×10^{-4}	-	9.15×10^{-5}
R^2	-	-	0.76	0.77

- both laws pretty agree for medium and coarse sands ($d > 0.2$ mm), i.e. they share both the asymptotic trend due to the common dependency of u_{*t} on $\sqrt{d}$, and the intercept, i.e. $A_b = 0.127 \approx A_s = 0.125$. This finding is consistent with the spirit of the Shao's model, whose corrective term $\gamma/\rho_a d$ is conceived to modify Banold's model at low d only;

- as regards Shao and Lu (2000) model, the refitted law predicts lower u_{*t} values for small d than the original one, because fitting is restricted to sands and exclude dusts. In other words, the refitted Shao's law mimics herein only the sand physics, and its trend at low d is not driven by the dust physics, and notably by the very high values $u_{*t} \approx 0.5$ m/s provided by Iversen et al. (1976) at $d = 0.023, 0.034, 0.041$ mm and $u_{*t} > 1$ m/s provided

by Fletcher (1976) at $d = 0.008, 0.009$ mm. A lower value of γ for the refitted law follows;

- CI of the Bagnold fitting is quite narrow (being the model easily reducible to a linear regression model). On the contrary, CI for Shao's model becomes wider as d decreases, because of the statistical uncertainty on the parameter γ in the term $\gamma/(\rho_a d)$, which has its main effects for small values of d;
- for both fittings $R^2 \approx 0.76$. This value, although satisfying, highlights a shortcoming of the deterministic approach: both laws cross over different kind of sands (fine, medium, coarse), while distinct regimes can be expected in each sand class, notably as regards data dispersion and skewness.

4. Statistical approach

In the following a statistical approach is proposed, having in mind the shortcomings of the deterministic approach, and the perspective practical needs in design infrastructures in arid environments. In fact, engineers are interested in evaluating low percentiles of u_{*t}, i.e. values having not-exceedance low probability, which reflect in high percentiles of the transport rate, i.e. values having an exceedance low probability.

Within such an approach, each source of uncertainty and related microscopic parent r.v.s are not described in statistical terms, because of the lack of data. Conversely, the threshold shear velocity is adopted as a single comprehensive r.v. On the one hand, its variability comprehensively includes and reflects the effects of all the parent r.v.s. On the other hand, the effects of a given single parent r.v. cannot be isolated.

The statistics of u_{*t} are obtained for each sub-ensemble resulting from the sand grading as illustrated in Sect. 2.

4.1. Analytical distributions fitting

Since the probability distributions of u_{*t} for the three size ranges are a priori unknown, we first aim at assessing if a parametric distribution can be adopted to describe the threshold shear velocity for each kind of sand. The data in each sub-ensemble are fitted to some guess reference distributions by means of the maximum likelihood estimation method. The considered distributions are normal, lognormal, and Weibull. Figure 4(a), (b), (c) collects the empirical and fitted Cumulative Distribution Functions (CDF). In order to assess the goodness of the fit, two approaches are used.

First, we employ the Anderson and Darling (1952) empirical distribution test because of the high weight placed on observations in the tails of distribution. The null hypothesis is never rejected, being the resulting p-values always greater than 0.1 for all tested distributions (with a range from 0.11 to 0.43 for fine sand, always greater than 0.5 for medium sand, and with range a from 0.22 to 0.89 for coarse sand). In particular, for medium sand, the normal distribution obtains the largest p-value ($p \approx 0.88$), while for coarse sand, the Weibull distribution is highly scored ($p \approx 0.89$).

However, having in mind the high levels of probability of errors of the second kind in goodness of fit tests like Anderson and Darling (1952) dealing with small samples, and the fact that we are interested in evaluating extreme percentiles of u_{*t} (notably, the low ones), a second analysis based on the so-called quantile $q-q$ plots is adopted as an exploratory visual aid to assess the local goodness of fitting of the reference distributions (Figure 4-d, e, f). The $q-q$ plots exam reveals that parametric distributions generally fail in describing experimental data. Only the goodness of normal fit is confirmed for medium sand ($\mu(u_{*t}) = 0.355$ mm, $\sigma(u_{*t}) = 0.068$ mm) also close to the tails, while a significant departure of the Weibull quantiles from bisector is observed at the lower tail for coarse sand.

4.2. Non-parametric distribution fitting

Since most of the parametric distributions do not seem able to correctly fit data in the tails, we decide to adopt a non-parametric density estimation based on kernel methods.

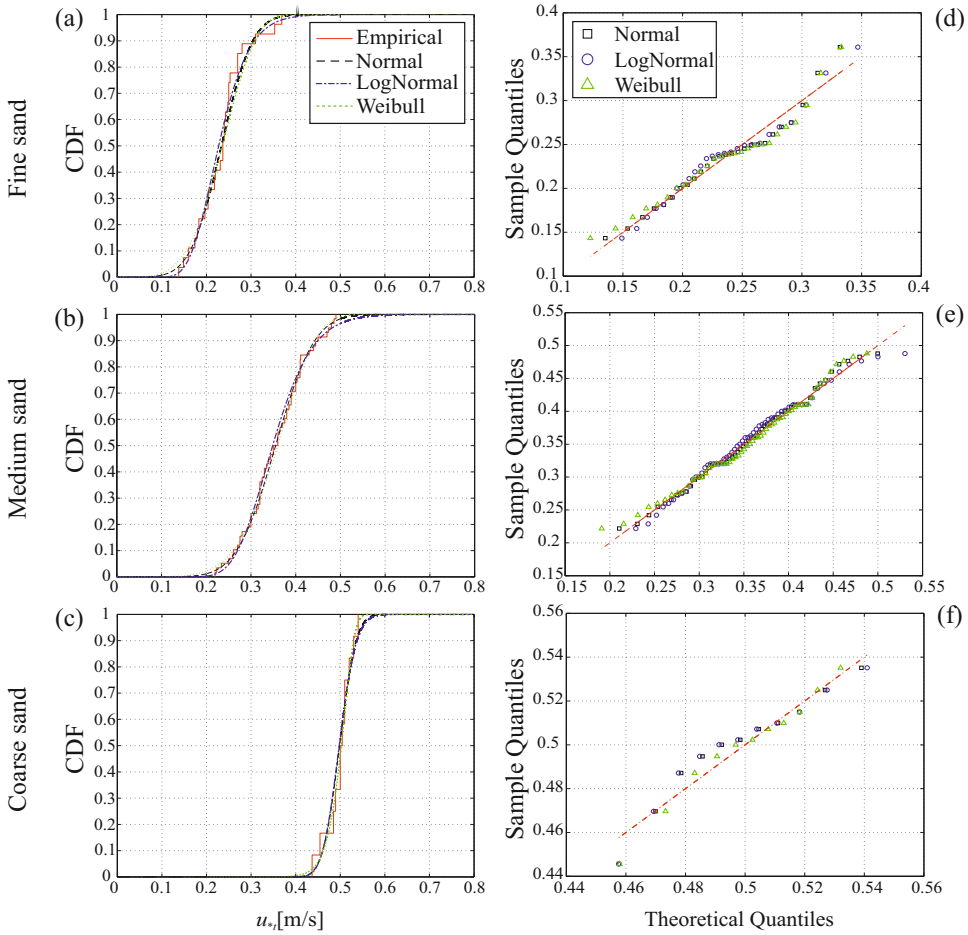

Figure 4: Analytical distribution results: cumulative distribution functions (a), (b), (c), and $q-q$ plots (d), (e), (f) for fine, medium and coarse sand

4.2.1. *Adopted methods*

The PDF kernel estimate method is based on representation of the density as a weighted sum of the kind

$$\hat{f}_h(x) = \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right),$$

where the x_i's are the observed values, the kernel K is a suitable unimodal probability density symmetric about zero, while h is a suitable smoothing parameter known as the bandwidth of the estimate. While the choice of the kernel K seems not to affect too much a proper non-parametric estimate $\hat{f}$ of the density f (commonly, Gaussian kernels are applied), the specific choice of the bandwidth h controlling the smoothness of the resulting density curve is extremely important, since the bias and the variance of the estimator $\hat{f}$ strongly depend on h in a non-linear relation. In fact, the Mean Integrated Square Error ($MISE$, Eq. 3)

$$MISE(h) = E\int (\hat{f}_h(x) - f(x))^2 dx, \tag{3}$$

which provides a measure of the difference between the estimate density function and the true density, can be expressed as

$$MISE(h) = \int Bias^2(\hat{f}_h(y))dy + \int Var(\hat{f}_h(y))dy, \tag{4}$$

where $Bias(\hat{f}_h(y)) = c_1 h^2 + o(h^2)$ and $Var(\hat{f}_h(y)) = \frac{c_2}{h} + o(\frac{1}{h})$, being c_1 and c_2 values depending on K and on the true density f (see Sheather, 2004, for details).
Since the $MISE$ is not mathematically tractable, common methods for bandwidth selection employ the Asymptotic Mean Integrated Square Error ($AMISE$), whose minimum as a function of h can be less hardly evaluated, as well as a variety of alternative automatic, data-based methods. Among all, the most common are Plug-In (PI) and Least Squares Cross-Validation (LSCV) bandwidth estimate methods. Both these two techniques provide good performances, but while LSCV often shows tendency to undersmooth, PI tends to oversmooth in the case of densities with high fluctuations (see, again Sheather, 2004), as our raw data suggest. Because of this reason, and since oversmoothing provides a prudential approach in estimation of extreme quantiles (i.e., tends to estimate larger interquantiles differences), the PI method is adopted. The Matlab ©Kernel Smoothing Toolbox developed by Koláček and Zelinka (2012) (see Horová et al., 2012, for details) is used to numerically evaluate the values of h through the PI method and Gaussian kernels.

4.2.2. Results

As anticipated, the convergence of the first two statistical moments is checked for increasing ensemble cardinality $\# = n$, i.e. number of data included in each sub-ensemble. The weighted residual of the generic parameter φ for growing cardinality of a sub-ensemble is defined as $\varphi_{res,n} = |(\varphi_n - \varphi_{n-1})/\varphi_n|$, and averaged over 20.000 random permutations of the order of the data. Residual convergence versus n is plotted on loglog scale in Figure 5. The rate

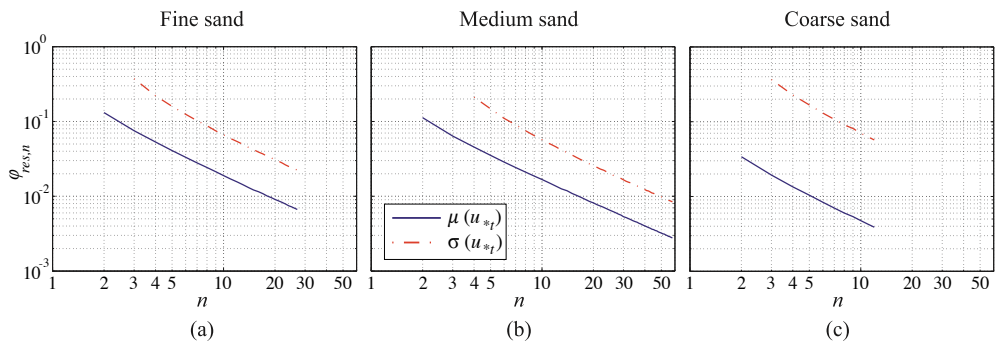

Figure 5: Convergence of the the mean $\mu(u_{*t})$ and standard deviation $\sigma(u_{*t})$ for each sub-ensemble (fine, medium and coarse sand)

of convergence is the same for both moments and for every sand class. Hence, the key element in convergence is the cardinality of each sub-ensemble. The complete set of collected events allows one to reach a threshold of about $3 \times 10^{-3} \leq \mu_{res,\#} \leq 7 \times 10^{-3}$ for the mean values, and of $8 \times 10^{-3} \leq \sigma_{res,\#} \leq 6 \times 10^{-2}$ for the standard deviation. The obtained final residual error is acceptable from a practical engineering point of view, except for the one of the standard deviation for coarse sand (Figure 5-c). In spite of such an encouraging convergence, the fitting of high order moments and extreme percentiles would benefits of higher cardinality of the sub-ensembles. We encourage further independent experimental measurements to enrich the ensembles.

The non-parametric PDFs estimated from the complete sub-ensembles are shown in Figure 6(a), (b), (c), for fine, medium and coarse sands, respectively. Statistical moments, such as mean values, standard deviations and skewness sk, coefficient of variation ($c.o.v.$), 1st and 3rd quartiles, and 5th and 95th percentiles are obtained from the fitted PDFs. They are summarized in Table 3. Remarkably, the non parametric estimate gives results coherent with the goodness of fit assessment for the analytical distributions: u_{*t} for medium sand is confirmed to be pretty normally distributed (Figure 6-b), with analogous mean value and standard deviation, and with a very low value of skewness ($sk = 0.033$). Conversely, the PDFs for fine and coarse sands are confirmed to be far from gaussian, positively and negatively skewed, respectively. It is worth pointing out that, even if no constraints are a priori imposed on the support of the non parametric PDFs, also the low percentiles are positive, i.e. $p_5(u_{*t}) > 0$ for fine sand too. The monotonic growth of both $c.o.v.$ and skewness for decreasing reference diameter d properly reflects the expected growing role played by interparticle forces and related uncertainties. In particular, the coefficient of variation for fine and medium sands is about $c.o.v. \approx 0.25$: even if this is a relatively moderate $c.o.v.$ value with respect to other environmental r.v.s (e.g. turbulent wind velocity), it implies the 5th percentile is about 0.6 times the mean value.

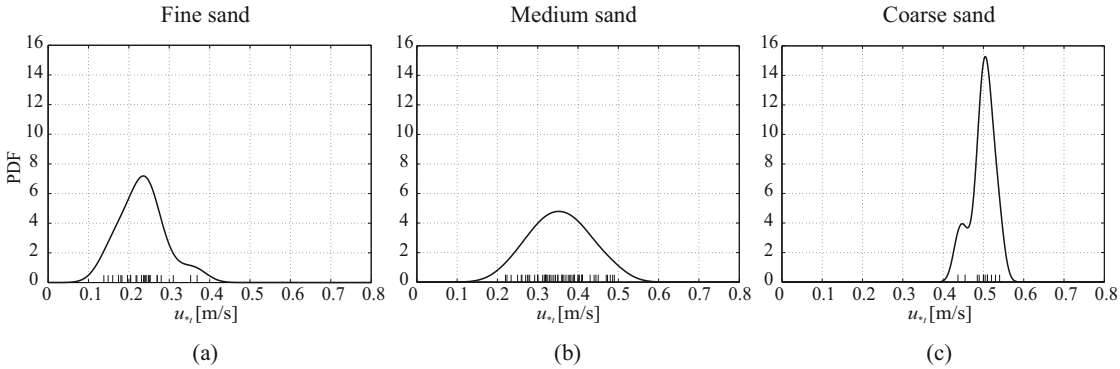

Figure 6: Non-parametric PDFs of the threshold shear velocity for fine, medium and coarse sands. Lower vertical bars stand for sand sample measurements

Table 3: Statistics of threshold shear velocity from non-parametric distributions

		Fine sand	Medium sand	Coarse sand
$\mu(u_{*t})$	$[m/s]$	0.234	0.355	0.498
$\sigma(u_{*t})$	$[m/s]$	0.062	0.080	0.032
$sk(u_{*t})$	[-]	0.360	0.033	-0.539
$c.o.v.$	[-]	0.266	0.224	0.063
$p_5(u_{*t})$	$[m/s]$	0.136	0.225	0.437
$Q_1(u_{*t})$	$[m/s]$	0.192	0.299	0.482
$Q_3(u_{*t})$	$[m/s]$	0.269	0.411	0.520
$p_{95}(u_{*t})$	$[m/s]$	0.351	0.488	0.544

5. Comparison between deterministic and statistical approach

Finally, the main findings of the proposed statistical approach are critically compared to the results of the deterministic approach. In Figure 7(a) the mean values $\mu(u_{*t})$ obtained from the non-parametric distribution for the three sand classes are superimposed to the refitted deterministic semi-empirical expressions of the nominal value of u_{*t}. It is

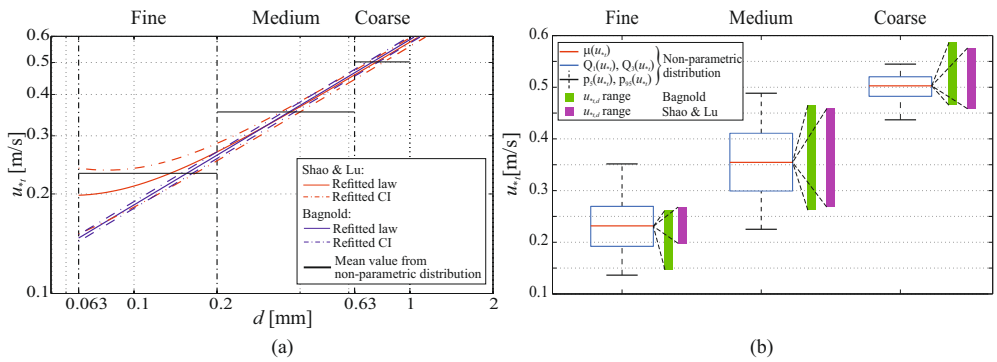

Figure 7: Comparison between statistical non-parametric distributions and semi-empirical deterministic models: mean values $\mu(u_{*t})$ versus refitted Bagnold (1941) and Shao and Lu (2000) models (a), boxplots versus deterministic ranges (b)

clear that statistics over data for a given sand class, i.e. a given d range, implies the mean value has a step-wise trend versus the reference diameter d. In other words, the statistical approach apparently involves loosing information with respect to the continuous deterministic laws $u_{*t}(d)$, if reference is made to the mean value only. In fact, this apparent under-sampling, is largely compensated by high-order statistics, that substantially enrich the description of u_{*t} for each sand class. In order to testify this feature, the box plots corresponding to the sand classes are plotted in Figure 7(b) together with the corresponding deterministic range of the nominal values of the threshold shear velocity $u_{*t,d}$ predicted through Bagnold (1941) and Shao and Lu (2000) refitted models. From both Figures the role of skewness is clearly depicted: the mean value of u_{*t} as a r.v. is very close to the nominal value deterministically evaluated at the mid-range diameter d_m only for null skewness (medium sand); otherwise, $\mu(u_{*t}) \geq u_{*t}(d_m)$ for $sk > 0$ (fine sand) and viceversa (coarse sand). Finally, the statistical approach allows to associate a given probability of exceedance to any value of the threshold shear velocity, while the nominal value from a deterministic law does not.

6. Conclusions

The present study critically compares deterministic and statistical approaches to threshold shear velocity on the basis of the collection of a huge amount of experimental measurements collected ad hoc from literature. Since the description of each random variable affecting u_{*t} is hard to be practically tractable, each source of uncertainty is merged within the single and comprehensive random variable u_{*t}.

Deterministic approach is updated thanks to the amount of collected data: in spite of a satisfying fitting of the $u_{*t}(d)$ nominal law, the lack of information about u_{*t} variability remains a shortcoming of the approach.

The proposed statistical approach allows to enrich the threshold shear velocity description providing measures of its variance and high order statistics, notably extreme percentiles. From a practical perspective in a number of application fields, the proper definition of u_{*t} values associated to given non exceedance probabilities allows to predict aeolian events and in turn to assess the performances of mitigation measures in probabilistic terms. Moreover, statistics are obtained over broad sub-ensembles defined w.r.t. sand classes, rather than on narrow d intervals. From a practical perspective, this allows to apply the statistical approach to mesoscale problems (e.g. infrastructures crossing several landforms along their path), where a single local reference sand diameter (e.g. at a dune toe or crest) is no longer relevant.

In the light of the obtained results, we suggest two research perspectives. First, a high cardinality of the dataset allows the full convergence of the statistical estimates: hence, the authors hope that further independent experimental studies will enrich the data ensemble. Second, the uncertainty propagation from the threshold shear velocity to the sand transport rate would worth to be further investigated.

Acknowledgments

The study has been developed in the framework of the Windblown Sand Modeling and Mitigation joint research, development and consulting group established between Politecnico di Torino and Optiflow Company. The authors wish to thank Davide Fransos, member of the WSMM group, for the helpful discussions about the topics of the paper. The research that lead to the present paper was partially supported by a grant of the group GNFM of INdAM.

References

Alghamdi, A.A., Al-Kahtani, N.S., 2005. Sand Control Measures and Sand Drift Fences. J. Perform. Constr. Facil. 19, 295–299. doi:10.1061/(ASCE)0887-3828(2005)19:4(295).

Anderson, T., Darling, D., 1952. Asymptotic theory of certain goodness of fit criteria based on stochastic processes. Ann. Math. Stat. 23, 193–212.

Bagnold, R., 1941. The Physics of Blown Sand and Desert Dunes. Methuen. doi:10.1007/978-94-009-5682-7.

Bagnold, R.A., 1937. The transport of sand by wind. Geogr. J. 89, 409–38. doi:10.2307/1786411.

Barchyn, T.E., Hugenholtz, C.H., 2011. Comparison of four methods to calculate aeolian sediment transport threshold from field data: Implications for transport prediction and discussion of method evolution. Geomorphology 129, 190–203. doi:10.1016/j.geomorph.2011.01.022.

Belly, P.Y., 1964. Sand movement by wind. Technical Report 1. U.S. Army - Corps of Engineers. US Army Coastal Eng. Res. Center, Washington, DC.

Blott, S., Pye, K., 2006. Particle size distribution analysis of sand-sized particles by laser diffraction: an experimental investigation of instrument sensitivity and the effects of particle shape. Sedimentology 53, 671–685. doi:10.1111/j.1365-3091.2006.00786.x.

Cheng, J., Xue, C., 2014. The sand-damageprevention engineering system for the railway in the desert region of the qinghai-tibet plateau. J. Wind Eng. Ind. Aerodyn. 125, 30–37. doi:10.1016/j.jweia.2013.11.016.

Chepil, W., 1945. Dynamics of wind erosion: Ii. initiation of soil movement. Soil Sci. 60, 397–411. doi:10.1097/00010694-194511000-00005.

Chepil, W., 1959. Equilibrium of soil grains at threshold of movement by wind. Soil Sci. Soc. Proc. 23, 422–428.

Chepil, W.S., 1951. Properties of soil which influence wind erosion: Iv. state of dry aggregate structure. Soil Sci. 72, 387–401.

Cornelis, W., Gabriels, D., 2004. A simple model for the prediction of the deflation threshold shear velocity of dry loose particles. Sedimentology 51, 39–51.

Dong, Z., Liu, X., Wang, H., Wang, X., 2003. Aeolian sand transport: a wind tunnel model. Sediment. Geol. 161, 71–83. doi:10.1016/S0037-0738(02)00396-2.

Dong, Z., Liu, X., Wang, H., Zhao, A., Wang, X., 2002. The flux profile of a blowing sand cloud: a wind tunnel investigation. Geomorphology 49, 219–230. doi:10.1016/S0169-555X(02)00170-8.

Duan, S.Z., Cheng, N., Xie, L., 2013. A new statistical model for threshold friction velocity of sand particle motion. Catena 104, 32–38. doi:10.1016/j.catena.2012.04.009.

Edwards, B., Namikas, S., 2015. Characterizing the sediment bed in terms of resistance to motion: Toward an improved model of saltation thresholds for aeolian transport. Aeolian Res. 19, 123–128. doi:10.1016/j.aeolia.2015.10.004.

Fletcher, B., 1976. Incipient motion of granular materials. J. Phys. D Appl. Phys. 9, 2471–8. doi:10.1088/0022-3727/9/17/007.

Friedman, G., Sanders, F., 1978. Principles of sedimentology. Wiley, New York.

Greeley, R., Iversen, J.D., 1985. Wind as a Geological Process on Earth, Mars, Venus, and Titan. Cambridge University Press, New York. doi:10.1017/CBO9780511573071.

Horikawa, K., Hotta, S., Kubota, S., Katori, S., 1983. On the sand transport rate by wind on a beach. Coast. Eng. Japan 26, 101–120.

Horová, I., Koláček, J., Zelinka, J., 2012. Kernel Smoothing in Matlab: Theory and Practice of Kernel Smoothing.

ISO14688-1:2002, . Geotechnical investigation and testing - identification and classification of soil - part 1: Identification and description.

Iversen, J.D., Pollack, J.B., Greeley, R., White, B.R., 1976. Saltation threshold on mars: The effect of interparticle force, surface roughness, and low atmospheric density. Icarus 29, 381–93. doi:10.1016/0019-1035(76)90140-8.

Iversen, J.D., White, B.R., 1982. Saltation threshold on earth, mars and venus. Sedimentology 29, 111–9. doi:10.1111/j.1365-3091.1982.tb01713.x.

Jones, R., Pollock, H., Cleaver, J., Hodges, C., 2002. Adhesion forces between glass and silicon surfaces in air studied by afm: effects of relative humidity, particle size, roughness and surface treatment. Langmuir 18, 8045–8055. doi:10.1021/la0259196.

Kadib, A.L., 1964. Sand Transport by Wind - Addendum II to Sand movement by Wind. Technical Report 1. U.S. Army - Corps of Engineers. US Army Coastal Eng. Res. Center, Washington, DC.

Kawamura, R., 1951. Study of sand movement by wind. Technical Report. Translated (1965) as University of California Hydraulics Engineering Laboratory Report HEL 2-8. Berkley.

Kok, J.F., Parteli, E.J.R., Michaels, T.I., Karam, D.B., 2012. The physics of wind-blown sand and dust. Rep. Prog. Phys. doi:10.1088/0034-4885/75/10/106901.

Koláček, J., Zelinka, J., 2012. MATLAB toolbox. URL: http://www.math.muni.cz/english/science-and-research/developed-software/232-matlab-toolbox.html.

Logie, M., 1981. Wind tunnel experiments on dune sands. Earth Surf. Proc. Land. 6, 365–374.

Logie, M., 1982. Influence of roughness elements and soil moisture on the resistance of sand to wind erosion. Catena Suppl. 1, 161–174.

Lyles, L., Krauss, R., 1971. Threshold velocities and initial particle motion as influenced by air turbulence. Trans. ASAE , 14563–14566.

McKenna, N.C., 2003. Effects of temperature and humidity upon the entrainment of sedimentary perticles by wind. Boundary Layer Meteorol. 108, 61–89. doi:10.1023/A:1023035201953.

McKenna Neuman, C., 2004. Effects of temperature and humidity upon the transport of sedimentary particles by wind. Sedimentology 51, 1–17.

McKenna Neuman, C., Nickling, W., 1989. A theoretical and wind tunnel investigation of the effect of capillary water on the entrainment of sediment by wind. Can. J. Soil Sci. 69, 79–96.

Merrison, J., 2012. Sand transport, erosion and granular electrification. Aeolian Res. 4, 1–16. doi:10.1016/j.aeolia.2011.12.003.

Middleton, N., Sternberg, T., 2013. Climate hazards in drylands: A review. Earth-Sci. Rev. 126, 48–57. doi:10.1016/j.earscirev.2013.07.008.

Nalpanis, P., Hunt, J.C.R., Barret, C.F., 1993. Saltating particles over flat beds. J. Fluid Mech. 251, 661–685. doi:10.1017/S0022112093003568.

Nicking, W.G., McKenna Neuman, N.C., 1997. Wind tunnel evaluation of a wedge-shaped aeoilian sediment trap. Geomorphology 18, 333–345. doi:doi:10.1016/S0169-555X(96)00040-2.

Phillips, M., 1980. A force balance model for particle entrainment into a fluid stram. J. Phys. D Appl. Phys. 13, 221–233. doi:10.1088/0022-3727/13/2/019.

Pye, K., Tsoar, H., 2009. Aeolian sand and sand dunes. Springer. doi:10.1007/978-3-540-85910-9.

Rizvi, A., 1989. Planning responses to aeolian hazards in arid regions. Journal of King Saud University, Architecture & Planning 1, 59–74.

Roney, J.A., White, B.R., 2004. Definition and measurement of dust aeolian thresholds. J. Geophys. Res. 109. doi:10.1029/2003JF000061.

Saffman, P., 1965. Lift on a small sphere in a slow shear flow. J. Fluid Mech. 22, 385. doi:10.1017/S0022112065000824.

Shao, Y., 2008. Physics and modelling of wind erosion. 2nd ed., Springer. doi:10.1007/978-1-4020-8895-7.

Shao, Y., Lu, H., 2000. A simple expression for wind erosion threshold friction velocity. J. Geophys. Res. 105, 22437–43. doi:10.1029/2000JD900304.

Sheather, S.J., 2004. Density estimation. Statistical Science 4, 588–597. doi:10.1214/088342304000000297.

Sherman, D., Li, B., 2012. Predicting aeolian sand transport rates: A reevaluation of models. Aeolian Res. 3, 371–378. doi:10.1016/j.aeolia.2011.06.002.

Webb, N., Strong, C., 2011. Soil erodibility dynamics and its representation for wind erosion and dust emission models. Aeolian Res. 3, 165–179. doi:10.1016/j.aeolia.2011.03.002.

Zhang, C., Zou, X., Cheng, H., Yang, S., Pan, X., Liu, Y., Dong, G., 2007. Engineering measures to control windblown sand in shiquanhe town,

tibet. J. Wind Eng. Ind. Aerod. 95, 53–70. doi:10.1016/j.jweia.2006.05.006.
Zhang, K., Chanpura, R.A., Mondal, S., Wu, C., Sharma, M.M., Ayoub, J.A., Parlar, M., 2014. Particle size distribution measurement techniques and their relevance or irrelevance to sand control design, in: Proceedings of the 2014 SPE International Symposium and Exhibition on Formation Damage Control, Lafayette, Louisiana, USA, February 26-28, 2014, Society of Petroleum Engineers.
Zimon, A.D., 1982. Adhesion of Dust and Powder. Springer. doi:10.1007/978-1-4615-8576-3.
Zingg, A.W., 1953. Wind tunnel studies of the movement of sedimentary material, in: Proceedings of the Fifth Hydraulic Conference. Studies in Engineering, Iowa City, University of Iowa. pp. 111–35.