Using a statistical efficiency methodology for predictors' selection in the bedload transport problem: A high gradient experimental channel case

Article *in* AEJ - Alexandria Engineering Journal · December 2021

DOI: 10.1016/j.aej.2021.11.052

CITATIONS
0

READS
49

8 authors, including:

Veronica Carrillo
University of Cuenca
9 PUBLICATIONS **4** CITATIONS
SEE PROFILE

Daniel Emilio Mendoza
University of Cuenca
17 PUBLICATIONS **32** CITATIONS
SEE PROFILE

Felipe Cisneros
University of Cuenca
70 PUBLICATIONS **1,372** CITATIONS
SEE PROFILE

Luis Timbe
University of Cuenca
36 PUBLICATIONS **303** CITATIONS
SEE PROFILE

Alexandria Engineering Journal (2021) xxx, xxx–xxx

HOSTED BY

Alexandria University

Alexandria Engineering Journal

www.elsevier.com/locate/aej
www.sciencedirect.com

Using a statistical efficiency methodology for predictors' selection in the bedload transport problem: A high gradient experimental channel case

Veronica Carrillo [a,*], Daniel Mendoza [a], John Petrie [b], Pedro Matovelle [a], Sebastian Torres [a], Esteban Pacheco [a], Felipe Cisneros [a], Luis Timbe [c]

[a] *Laboratorio de Hidráulica y Dinámica de Fluidos/Departamento de Ingeniería Civil, Facultad de Ingeniería, Universidad de Cuenca, Cuenca, Ecuador*
[b] *Hydraulic Engineer, United Stated Army Corps of Engineers, Los Angeles, CA, USA*
[c] *Departamento de Recursos Hídricos y Ciencias Ambientales, Universidad de Cuenca, Cuenca, Ecuador*

Received 22 June 2021; revised 7 November 2021; accepted 20 November 2021

KEYWORDS

High gradient;
Bedload transport;
Akaike-information-criterion;
Best predictors;
Laboratory experiments

Abstract Bedload transport rates for high-gradient gravel bed rivers has been studied through a physical model that replicated the typical features of these channels. A stepwise regression was performed to identify the best predictors from a set of independent variables. As independent variables channel slope, the ratio of area occupied by large particles to the total plan area, flow discharge, mean flow depth, mean flow velocity, water surface velocity, boundary shear stress, and shear velocity were considered. Different characteristic diameters (d_{16}, d_{50}, d_{84}, and d_{90}) were used to nondimensionalize the variables as well as to test the influence of grain size. A linear and a potential model were obtained for each characteristic diameter. Based on the correlation coefficients (R^2) with the data used to build the models, the d_{50} and d_{84} linear and potential models were selected to perform further analysis. A set of independent data was used to verify the selected models. Better performance was observed for the potential models with 96% of the data falling within ø order of the magnitude bands both for d_{50} and d_{84} . R^2 for the d_{50} and d_{84} potential models were 0.63 and 0.76, respectively. Therefore, the d_{84} potential model can be selected as the present study representative model.

1. Introduction

Mountainous drainage systems are composed mainly of steep rough-bedded channels. It has been demonstrated that this type of channel behaves differently from milder gradient streams [42,44,67,66]. Specific features of high-gradient gravel

* Corresponding author.
E-mail address: veronica.carrillo@ucuenca.edu.ec (V. Carrillo).
Peer review under responsibility of Faculty of Engineering, Alexandria University.

https://doi.org/10.1016/j.aej.2021.11.052
1110-0168 © 2021 THE AUTHORS. Published by Elsevier BV on behalf of Faculty of Engineering, Alexandria University.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Nomenclature

AIC	Akaike-Information-Criterion
d_s	characteristic diameter (m)
d_{16}	characteristic grain diameter for which 16% is finer by weight (m)
d_{50}	characteristic grain diameter for which 50% is finer by weight (m)
d_{84}	characteristic grain diameter for which 84% is finer by weight (m)
d_{90}	characteristic grain diameter for which 90% is finer by weight (m)
Fr	Froude number
g	gravity acceleration (m/s^2)
GSD	grain size distribution
l_m	model length
l_p	prototype length
q	dimensionless volumetric bedload transport rate per unit width
q_s	volumetric bedload transport rate per unit width m^3/s/m or m^2/s
Q	dimensionless flow discharge
Q	flow discharge (m^3/s)
R	sediment submerged specific gravity $R = \frac{\rho_s - \rho}{\rho}$
Re	Reynolds number
Rh	Hydraulic radius (m)
S_o	bed slope (m/m)
V	dimensionless velocity
V	velocity (m/s)
Vm	dimensionless mean flow velocity
v_m	model mean flow velocity (m/s)
v_p	prototype mean flow velocity (m/s)
Vws	dimensionless water surface velocity
$V*$	dimensionless shear velocity
w_s	settling velocity (m/s)
Y	dimensionless flow depth
Y	flow depth (m)
γ	specific weight of water (N/m^3)
λ_l	geometric scale
λ_v	velocity scale
ν	kinematic viscosity (m^2/s)
ρ	density of water (kg/m^3)
ρ_s	sediment density (kg/m^3)
τ	dimensionless boundary shear stress
τ_o	boundary shear stress $\tau_o = \gamma YSo$ (N/m^2)
$\%LP$	ratio between large particles plan area and total area

bed streams include the presence of coarse bed material, low values of relative submergence, wide range of bed material particle sizes, high variation of channel slope as well as high slope values, channel bed forms, and highly variable channel geometry [26,31,41,42,45,66] . The accurate determination of bedload transport rates still remains a major concern in the study of high-gradient rivers [3]. Processes such as the quantification of sediment loads, channel evolution, and river management depend on the ability to predict transport rates accurately [40]. The high variability and complexity of the parameters involved in sediment transport do not allow for obtaining estimates of transport rates through theoretical or semi-empirical models with a precision greater than within an order of magnitude [54,53]. When general bedload transport equations are applied to estimate fields or laboratory transport rates, errors of several orders of magnitude are common. When models developed specifically for steep gravel bed rivers are used, these errors are reduced. However, differences up to three orders of magnitude are still present [27,42,51,54].

Many studies have been performed with the aim to develop a physically-based model that accurately describes bedload transport [5,8,20,37,48] . However, the complex interaction between flow and sediment particles, highly variable flow conditions due to turbulence and channel flow resistance, and the presence of large particles that are immobile under certain flows but become mobile above a flow threshold, make it difficult to include all the parameters and their variability (in time and space) in a simplified model [4,3,15,25,54,53]. Recent studies have developed experimental procedures to incorporate some of this variability [11]. Yager et al. [67] performed flume experiments to test a modified equation that includes the effects of the reduction of shear stress, available for sediment transport, by the presence of large elements in the channel that move only under extreme hydrological events. Yager et al. [66] developed a correction to the Parker (1990) equation to account for the effects of steps that absorb a part of the boundary shear stress resulting in an increase in sediment flow resistance. Frey and Church [25] performed an analysis of gravel sediment transport in terms of a granular perspective to introduce the particular physics of these types of flows to the evaluation of gravel transport. Ghilardi et al. [26] completed flume experiments to demonstrate that the relation between sediment diameter and large size particles distribution (in space) play an important role in bedload transport. Moreover, they stated that the transport process for steep gravel channels is better described in terms of discharge rather than shear stress. They also included the direct influence of bed slope, thus, affirming that this parameter has the highest impact on transport capacity.

As a step forward to the consideration of more general field conditions, the present study developed an experimental analysis that included the simulation of bedload transport with natural-shaped sediment particles with a grain size distribution, and randomly distributed (in space) large size sediment particles based on what has been observed in typical high-gradient rivers of Southern Ecuador. Specifically, the experimental configuration and parameters considered for the present study were based on the features observed in a high-gradient river. The Tabacay River was selected because of its high gradient and sediment transport features such as slope, bed material sizes, and transport capacity. It is a mountainous high-gradient river that originates in the Ecuadorian Andes. Its drainage area is 66.5 km^2. The river has a total length of 8 km. Thirteen cross sections were surveyed to characterize the river throughout its length[11]. Measured longitudinal slopes range from 2% to 10% with an average of 5%. The

mean annual discharge is 1.5 m^3/s [36]. Parameters such as bed particle diameters and flow discharge were scaled from river values (prototype) to flume values (model) and vice versa, to relate the magnitude of the events performed in the flume experiments with field values. The aim of the present study was to obtain a bedload transport rate relation based on measured rates obtained from a laboratory experimental set up that was intended to include most of the variability present in field. Additionally, the obtained model was validated with an independent data set to verify its applicability.

2. Model identification for bedload transport rates

Bedload transport is a complex process that depends not only on many hydraulic and geometric parameters and physical properties, but also on the interaction between them. Hydraulic parameters include parameters such as mean flow velocity, mean water depth, boundary shear stress; geometric parameters include channel slope, characteristic bed material diameter. Water and sediment physical properties include density and fluid viscosity. The interactions between these parameters contribute to the high uncertainties reported in the estimation of bedload transport rates [49,69] . To include the most relevant variables involved in sediment transport mechanics, dimensionless parameters are used to develop empirical or theoretical relations to estimate bedload transport rates [69]. For the present study, the relations used to obtain dimensionless hydraulic parameters as proposed by Parker [46] are presented below.

$$q = \frac{q_s}{R^{1/2} g^{1/2} {d_s}^{3/2}} \tag{1}$$

$$\tau = \frac{\tau_o}{\rho R g d_s} \tag{2}$$

$$Q = \frac{Q}{R^{1/2} g^{1/2} {d_s}^{5/2}} \tag{3}$$

$$Y = \frac{Y}{d_s} \tag{4}$$

$$V = \frac{V}{R^{1/2} g^{1/2} {d_s}^{1/2}} \tag{5}$$

Where, q is dimensionless volumetric bedload transport rate per unit width, q_s is volumetric bedload transport rate per unit width ($m^3/s/m$ or m^2/s), R is sediment submerged specific gravity

$R = \frac{\rho_s - \rho}{\rho}$, ρ is water density (kg/m^3), ρ_s is sediment density (kg/m^3), g is gravity acceleration (m/s^2), d_s is characteristic diameter (m), τ is dimensionless boundary shear stress, $\tau_o = \gamma YSo$ is boundary shear stress (N/m^2), γ is water specific weight (N/m^3), Y is flow depth (m), S_o is bed slope (m/m), Q is dimensionless flow discharge, Q is flow discharge (m^3/s), Y is dimensionless flow depth, V is dimensionless velocity, and V is velocity (m/s).

Commonly, equations for bedload transport rate are expressed as functions of τ . However, stream power, unit discharge, and energy dissipation rates can be also used to describe bedload transport rates [5,16,21,24,55,68] . Recent studies have identified that for gravel bed rivers, there are additional parameters that impact transport rates. Even though slope is indirectly considered through parameters such as boundary shear stress or stream power, the direct inclusion of slope in case of high-gradient rivers has improved the estimates of transport rates [15,61] . Additionally, the presence of large particles in the river bed has been demonstrated to impact the transport capacity by decreasing the tractive force (through boundary shear stress) available to transport sediment [26,67] . Therefore, with the aim to include most of the variables and processes that define bedload transport rates, the slope S_o, ratio between large particles plan area and total area $\%LP$, dimensionless discharge Q, dimensionless flow depth Y, dimensionless mean velocity Vm, dimensionless water surface velocity Vws, dimensionless boundary shear stress τ, and dimensionless shear velocity $V*$ were considered as appropriate independent variables. $\%LP$ has been selected as a parameter to represent the influence of large particles on the riverbed. The set of dimensionless variables has been selected considering the inclusion of hydraulic and geometric parameters, and physical sediment and fluid properties, and to represent the most common interaction between them.

An important parameter for both the sediment transport process and to express variables in the dimensionless form is the bed material characteristic diameter (d_s). Characteristic diameters d_{16} (16% finer), d_{50} (50% finer), d_{84} (84% finer), and d_{90} (90% finer) have been frequently used to represent the grain size distribution (GSD) of riverbed material [14,23,34,37,38,47,63,64,65] . Generally, a characteristic diameter that represents the entire GSD is used to study the fluviomorphology of rivers to include the influence of the GSD and to keep final models and relations as simple, practical, and replicable as possible [29,32]. Depending on the process being analyzed (bed roughness, flow resistance, sediment transport, etc.), different grain diameters have been selected as characteristic [7,23,57] . Therefore, for the purposes herein, the definition of a characteristic diameter was based on the bedload transport conditions, whose selection criteria rely on each diameter's role on the estimation of transport rates [51]. Here; d_{16}, d_{50}, d_{84}, and d_{90} were considered as potential characteristic diameters and used to put the independent variables in dimensionless forms. Thus, characteristic diameter was also considered an independent variable. Thereby, a more appropriate characteristic diameter for bedload sediment transport can be determined through a joint analysis with the other parameters that determine transport rates.

Many studies have evaluated the prediction capacity of bedload transport equations [27,30,35,54] . Among these studies, the only generalized conclusion is that there is no model or equation that performs well consistently and that the degree of model complexity and calibration is not always related with the performance of the model. However, simpler models have shown reasonable accuracy for the description of bedload transport. Additionally, physically-based models may not be able to account for the wide range of parameters and their fluctuations [4,6,13,49,52,60]. The inclusion of additional independent variables in the bedload sediment transport has been mostly applied as correction factors [15,26] . Therefore, it has not been possible to identify whether the relationship of each of these variables with bedload transport is linear [26].

In the present study, linear and potential models (through logarithmic transformation) were used to investigate the set

of variables' configurations that efficiently describe the transport phenomenon. As a simplified alternative to the highly complex theory of sediment transport, a more straightforward multiple linear regression has been used as a surrogate framework herein, which works under a proper stochastic formality that helped to establish a set of parsimonious variables describing the sediment transport rates. [4,22] . A metric for measuring such parsimony into a linear model is the Akaike-Information-Criterion (AIC), which determines the efficiency of the model by rewarding the predictability and penalizing the number of predictors included [1–2]. In an ideal case, optimizing the AIC parameter by searching among all the combination spaces of variables into a linear model will establish the desired set of most parsimonious predictors. Nonetheless, a Stepwise regression technique could support such identification more efficiently from the perspective of computing requirements. Stepwise regression is an automatic procedure for identifying the predictor variables through the addition/-subtraction of explanatory variables until a desired criterion has been reached (the AIC criterion, in this particular case) [28]. Forward, backward, and bidirectional selection procedures are frequently described as the most common algorithms of searching, and for the purposes herein, we used the three of them. The final model would be the one with the lowest AIC performance. The methodology searches through a wide range of variables' combinations, from a model described by a single constant to a model including all the available variables considered. The number of steps for the searching was established as five thousand, as a maximum. The model was reported as significative under a p-value less than 0.05 in the F-test.

Since some of the involved variables are certainly not independent of each other, there is a latent risk of multicollinearity and equifinality. To enhance this issue, an option could be the inclusion of uncorrelated variables a priori, selected by using some technique before the stepwise procedure – for example, a Principal-Component-Analysis (PCA). However, with respect to the authors' criteria, the aforementioned would not be convenient in this particular case because the linear model is considered as a surrogating technique describing a complex nonlinear phenomenon here. Thus, multiple linear frameworks could be interpreted as the first term in Tylor's expansion, hopefully capturing the main variability of the series as a joint contribution of all the predictors. In that sense, although an individual ranking of the importance of the predictors has a meaning on the linear scheme, it would be less meaningful in the phenomenon because of its nonlinear nature. Moreover, some correlation between the predictors could become conceptually substantial in a context beyond the linear scheme.

Because of the above, it will be convenient to include all the information available for the stepwise procedure. However, it is not desirable to have predictors with a high degree of collinearity. Therefore, the variance inflation factor, VIF (measure of the interaction and/or correlation between "independent" variables) should be checked out [33]. We established a VIF threshold as a maximum of ten ($VIF_{max} = 10$), which is often considered reasonable [33]. Of course, another more demanding criterion is also suggested by considering lower VIF values [59]. However, the relaxation of this parameter for the proposed value here allowed to avoid multicollinearity, and at the same time, preserved important information in the selected set of parameters, which was desirable for other purposes such as conceptual or physical interpretations in a broader range of analysis.

Moreover, because we used the model for identifying the set of variables that most probably describe the phenomenon jointly, we suggested checking out the presence of equifinality. Conveniently, this is possible by applying the three stepwise algorithms in this case (i.e., forward, backward, and bidirectional algorithms). A different set of predictors may be chosen by different stepwise strategies. In principle, the lower AIC information would ultimately establish the best model option among these. However, when the AIC parameter is similar and the sets "seriously" differ among them, it constitutes the evidence of equifinality. Thus, attention should be paid in that regard because it means that multiple groups of parameters can describe the phenomenon equally well, which would trivialize the purpose of the linear model outlined herein.

Discerning the cases on which the set of parameters identified with each of the three methodologies (forward, backward, and bidirectional) "significantly" differ among the three stepwise procedures that have a similar AIC implies some degree of subjective evaluation that entails the expert's criteria. However, since the linear scheme has shown consistency in the context of the sediment transport for certain experimental ranges and definitions [49], we hypothesize that a non-trivial selection will imply that the three models do not differ in more than two parameters in their chosen sets. On the other hand, if the latter criteria were not fulfilled, non-information could be excluded, establishing the set as the union of all the sets of identified variables for further discussion. However, it should be noted that the last established criterion is a suggestion in the context of the problem being treated here. In the end, the output variable q will be linearly approximated by some combination of variables. The algorithm to implement the stepwise procedure is freely available in "R" [58].

Finally, the analysis was carried out considering both forms (linear and potential) along with eq. (6) and eq. (7), and for each representative diameter to construct the dimensionless variables for the modeling. Since some evidence argues that the sediment transport phenomenon could respond to a potential nature [60], expression (7) will also constitute a conceptual validation of the simpler linear scheme used to identify the set of parameters. In that context, we compared both framework schemes to support the validity of the proposed methodology.

$$q = \mathrm{C} + \mathrm{a}S_o + b\%LP + c\,Q + d\,Y + e\,Vm + f\,Vws + g\,\tau + h\,V* \tag{6}$$

$$q = \mathrm{C}{\cdot}S_o{}^{a}{\cdot}\%LP^{b} \cdot Q^{c} \cdot Y^{d} \cdot Vm^{e} \cdot Vws^{f} \cdot \tau^{g} \cdot V*^{h} \tag{7}$$

3. Experimental design

The experiments were performed in the Hydraulics and Fluid Dynamics Laboratory of the Civil Engineering Department of the University of Cuenca. A tilting flume, 0.30-m-wide and 0.45-m-high, was used to create a mobile bed channel. The flume had a total length of 12 m; however, a length of 3.0 m was used as the test zone to represent the mobile bed. The test zone was located 5 m downstream from the entrance weir. The end of the test zone was 4 m from the end of the channel. Water was supplied to the flume channel through a recirculation system. At the entrance of the flume, a V-notch

weir allowed the estimation of the discharge. Discharge measurements were estimated from the calibrated head-discharge relationship (maximum error of 5%) and the measured water depth (with a point gauge with a resolution of ± 0.1 mm). A schematic of the experimental setup is presented in Fig. 1. The discharge range was defined based on the system capacity but mainly considering events from low to moderate intensity, corresponding to prototype values of 0.8 m^3/s to 3 m^3/s that represent 50% to 200% of the mean annual discharge of the Tabacay River. Moreover, these values represent events up to bankfull conditions for most of the thirteen cross sections that were analyzed. Bankfull discharge for the thirteen cross sections varies from 5.84 to 0.28 m^3/s and has a mean value of 2.41 m^3/s[11]. These types of events have been considered because available bedload equations generally tend to show poor performance under these conditions [6]. The range of experimental discharges are presented in Table 1 as well as their corresponding prototype values.

The material used was prepared based on the material of the thirteen cross sections characterized in the Tabacay River. For each cross section, bathymetry and GSD for the mobile bed material were obtained (without considering the large size less-mobile particles). The differences reported between the GSD obtained for each of the thirteen cross sections were not of a considerable magnitude. Median characteristic diameter d_{50} ranged from 10.8 to 38.6 mm. Therefore, the final GSD to be used for the experiments was determined by a least square regression based on all the thirteen GSDs measured in field. For the experiments, the final GSD was scaled. To represent the bedload transport in a scaled model, the dimensionless volumetric bedload transport rate q for the model had to be equal to the corresponding values in prototype [50]. This was achieved for a Froude scaled undistorted model by determining the sediment size corresponding to the scaled sediment settling velocity (through Froude scaling) rather than the geometric scaled sediment size [50]. The following relation was used to determine the settling velocity [19,62] .

$$w_s = \frac{v}{d_s}\left[\sqrt{\left(10.36^2 + 1.049 d_*^3\right)} - 10.36\right] \tag{8}$$

$$d_* = \left(\frac{gR}{v^2}\right)^{1/3} d_s \tag{9}$$

Where, w_s is the setting velocity (m/s) and v is the kinematic viscosity (m^2/s).

The geometric scale $\lambda_l = l_m/l_p$ (where $l_m$and l_p are model length and prototype length, respectively) was defined based on the characteristic channel width measured at the thirteen cross sections to be modeled in a 0.3 m wide channel. Thus, $\lambda_l = 0.12$. It has been shown that for high-gradient rivers, the top width is not highly sensitive to changes in discharge [11,18]. Therefore, the experiments in a rectangular channel do not represent a simplification that can compromise the accuracy of the model's measurements as a representative of the field parameters. Based on Froude scaling, the ratio between the model and prototype velocities ($\lambda_v = v_m/v_p$) to scale settling velocity $\lambda_v = 0.35$. The original and scaled GSD for the mobile bed material are presented in Fig. 2.

The model GSD was built with sediment extracted from the field sites on the Tabacay River. This material has a specific gravity of 2265 kg/m^3. To prepare the mobile bed for the experiments, first a fixed layer of the material with the model GSD was placed at the base of the channel to represent the roughness found in the natural river. The mobile bed for each experiment was located on top of this fixed layer. The height of this mobile bed was 12 times the d_{90} of the model GSD presented in Fig. 2. In addition to the mobile bed material, particles were also considered for the experiments. In order to

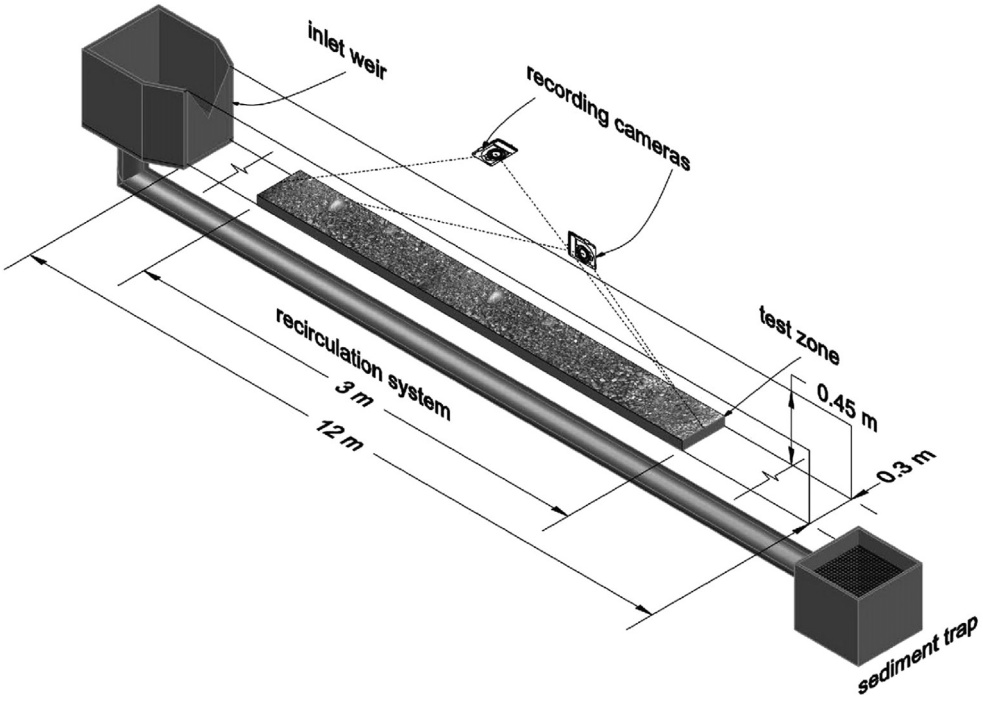

Fig. 1 Schematic of the experimental setup.

Table 1 Experimental parameters variation.

	Discharge $Q(l/s)$	Channel Slope $So(\%)$	Mean flow depth $Y(cm)$	Froude Number Fr [1]	Reynolds Number Re [2]	$\%LP(\%)$	Relative Submergence Y/d_{84}
Model	0.44 – 13.50	4	0.70 – 3.80	0.24 – 3.06	$4.9x10^3$ – $1.4x10^5$	5, 15, 25, 35	1.4 – 3.75
Prototype	87.83 – 2705.80		5.83 – 31.67		$1.1x10^5$ – $2.7x10^6$		0.89 – 3.87
Model	0.71 – 12.05	5	1.15 – 4.15	0.50 – 2.88	$8.9x10^3$ – $91.0x10^5$	5, 15, 25, 35	1.25 – 4.40
Prototype	141.79 – 2415.31		9.58 – 34.58		$2.1x10^5$ – $2.4x10^6$		1.28 – 4.52
Model	0.72 – 15.34	6	1.43 – 4.00	0.57 – 2.48	$9.0x10^3$ – $1.2.10^5$	5, 15, 25, 35	1.55 – 4.56
Prototype	143.53 – 3075.95		11.94 – 33.33		$2.2x10^5$ – $3.0x10^6$		1.60 – 4.69
Model	0.74 – 15.39	7	0.95 – 3.75	0.49 – 3.01	$8.6x10^3$ – $1.3x10^5$	5, 15, 25, 35	1.17 – 4.27
Prototype	149.00 – 3084.81		7.92 – 31.25		$2.1x10^5$ – $3.1x10^6$		1.21 – 4.39
Model	0.51 – 14.13	8	0.90 – 3.50	0.54 – 2.74	$4.8x10^3$ – $1.2x10^5$	5, 15, 25, 35	1.06 – 4.60
Prototype	101.67 – 2832.23		7.50 – 29.17		$1.2x10^5$ – $2.9x10^6$		1.09 – 4.76
Model	0.54 – 14.67	9	0.90 – 4.50	0.46 – 2.55	$5.1x10^3$ – $1.2x10$	5, 15, 25, 35	0.97 – 5.17
Prototype	108.03 – 2491.06		7.50 – 37.50		$1.2x10^5$ – $2.9x10^6$		1.00 – 5.32
Model	0.94 – 14.39	10	1.00 – 3.40	0.39 – 2.75	$8.5x10^3$ – $1.2x10^5$	5, 15, 25, 35	1.07 – 3.93
Prototype	187.53 – 2885.69		8.33 – 28.33		$2.0x10^5$ – $2.9x10^6$		1.10 – 4.05

[1] $Fr = \frac{V}{\sqrt{gY}}$.
[2] $Re = \frac{\sqrt{gY}Rh}{\nu}$.
Where, V is mean flow velocity and Rh is the hydraulic radius.

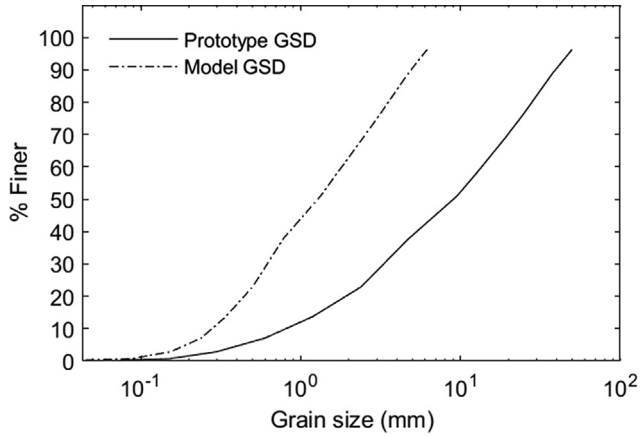

Fig. 2 Grain size distribution for prototype (field measured) and model (scaled for laboratory experimentation).

Fig. 3 Schematic for the determination of the ratio between the projected area occupied by large size particles and total plan area. The hatched areas represent large, immobile particles. (ESC 1:20).

consider a more general condition, random spatial distribution and protrusion of these particles was considered. The criteria used to determine the amount of the large particles was the ratio between the area occupied by these particles and the total plan area defined as %LP. The ratios considered for the experiment developed in the present study varied from 0.05 to 0.35 with an increase of 0.10 representing 5%, 15%, 25% and 35% of the plan area occupied by the large almost immobile particles. This range was determined from an analysis of the ratios observed in each of the thirteen cross sections. An example of this is shown in Fig. 3. For the section shown in Fig. 3, the percentage of plan area occupied by the large particles (hatched area) is 16%. The size of the large particles varied from 8 to 12 times d_{90} (90% finer from the model GSD in Fig. 2). These sizes were determined based on the large particles present in the thirteen cross sections. An average range of the variation identified from the thirteen cross sections was used for the laboratory experiments. Based on the %LP, a new GSD including the large particles was obtained for each case.

For each experimental run, the sediment was not fed during the experiment. A mobile sediment layer was located in the test zone and allowed to be transported until all the material left the test zone. This condition was considered because a set of calibration experiments (performed in a preceding work[12] demonstrated no dependence on the sediment feed rate for the final transport rate. Each experiment was considered to begin once the flow was established i.e., no variation in steady parameters, such as discharge, above sediment flow depth was registered. At this point, the time began to be registered. Some sediment particles were transported before the flow was established. However, these particles were not considered in determining the transport rates. Each run ended when the entire mobile sediment layer was transported out of the test zone.

To determine the experimental transport rates, the transported material was collected in a sediment trap at the end of the flume. The material was then dried and weighed. Trans-

port rate was calculated as weight of the material transported divided by the experiment time. Discharge, channel slope, and percentage of large particles varied for the different experimental runs. With combinations of all the independent variables, a total of 140 experiments were performed. In Table 1 presents the range of these parameters. For each experiment, the discharge, flow depth, surface velocity, and water temperature were measured. Each combination of parameters was run three times to ensure the validity of the flow hydraulics and transport rates measured. Surface velocity measurements were performed with a camera located on top of the channel and tracers (plastic spheres), as observed in Fig. 1. A lateral camera was also used to verify the flow depths throughout the experiment. The flow depth measurements were made with a point gauge (resolution of ± 0.1 mm) and verified with the values observed in the videos. As observed in Table 1, relative submergence in the model and prototype are similar. The slight differences that were observed were caused by the fact that sediment diameter is not scaled directly with the geometric scale λ_l .

4. Model verification

To analyze the prediction capacity of the selected models, the set of data developed by [61] was used. This data set is comprised of 77 sediment transport rates measurements and their corresponding hydraulic parameters. Table 2 presents a summary of these hydraulic parameters. Vws and $\%LP$ are not reported. However, these parameters are not commonly reported for bedload transport datasets. A relation between mean flow velocity and water surface velocity were obtained based on the data of the present study. This relation was used to estimate Vws based on Vm . For $\%LP$, a value of 0.5 was used. A value of fifty percent in area occupied by large particles could be considered relatively high. This can be seen as a desirable condition to be tested. Since the presence of large and almost immobile particles has been identified as one of the reasons for inaccuracies in bedload transport rates predictions [67], 2012), d_{84} is not reported either. However, it was estimated based on d_{90} and d_{50} .

5. Results and discussion

Table 3 presents a summary of all the experimental runs (140 total). It presents the results corresponding to the minimum and maximum discharge and the minimum and maximum % LP for all the slopes analyzed. The table does not show intermediate values of discharge and %LP. The experiment number (# **experiment**) presented in Table 3 is referenced with the total number of experiments. Intermediate numbers such as 2, 3, 5 – 16, etc. correspond to the discharge and %LP intermediate values. Therefore, they don't appear in Table 3. The best predictors for bedload transport rate were determined through stepwise regression. A linear and a potential model were determined for each characteristic diameter (d_{16}, d_{50}, d_{84}, and d_{90}) with the most representative parameters. The resulting parameters for each model are presented in Table 4.

$\boldsymbol{d_{14}}$

$$q = -2.9723 + 11.4304S_o + 0.0002\,Q + 0.0521\,Vm + 0.3073 \times Vws + 2.9343\,V*$$

$$q = 25.2041{S_o}^{1.7725}\,{Vm}^{1.6447}\,{V*}^{1.1863}$$

$\boldsymbol{d_{50}}$

$$q = -1.1417 + 3.4858S_o - 0.2188\%LP + 0.0002\,Q + 0.0717 \times Vm + 0.128\,Vws + 1.9086\,V*$$

$$q = 22.8516{S_o}^{1.7401}\,{Vm}^{1.6447}\,{V*}^{1.2441}$$

$\boldsymbol{d_{84}}$

$$q = -0.6420 + 0.4472S_o - 0.0816\%LP + 0.0632\,Vm + 0.0832 \times Vws + 1.7111\,V*$$

$$q = 0.6706{S_o}^{2.1940}\,Q^{0.8201}\,{Vm}^{0.7001}$$

$\boldsymbol{d_{90}}$

$$q = -0.5587 + 0.4985S_o - 0.0814\%LP + 0.0623\,Vm + 0.0641 \times Vws + 1.5789\,V*$$

$$q = 0.9721{S_o}^{2.2791}\,Q^{0.8048}\,{Vm}^{0.7171}$$

Slope (S_o) and dimensionless mean flow velocity (Vm) appear in all equations indicating that regardless of the form of the equations and the characteristic diameter, these two parameters have a strong influence on the final transport rate. In high-gradient rivers, the influence of slope has been largely proven [15,31,39,43,49,52,60,61,67,69] . Vm includes information of discharge, cross section characteristics, and roughness parameters that also define sediment transport. For %LP, no influence was registered for the potential models. For the linear models, as the %LP increases, the sediment transport rate

Table 2 Smart [61] experimental results.

	$d_{90}(mm)$	$d_{50}(mm)$	$Q(l/s)$	$So(\%)$	$Y(cm)$	$q_s(m^2/s)$
Near uniform GSD	5.20	4.20	4.84 – 25.86	3.00 – 20.00	3.10 – 6.80	$0.46\text{x}10^{-3}$ – $19.19\text{x}10^{-3}$
	12.10	10.50	9.32 – 49.20	3.40 – 20.00	1.90 – 7.90	$0.13\text{x}10^{-3}$ – $27.82\text{x}10^{-3}$
Non– uniform GSD	4.60	2.00	4.80 – 25.42	5.00 – 20.00	2.80 – 5.90	$1.90\text{x}10^{-3}$ – $30\text{-}30\text{x}10^{-3}$
	11.00	4.30	4.80 – 28.12	3.00 – 20.00	2.90 – 8.40	$0.34\text{x}10^{-3}$ – $33.00\text{x}10^{-3}$

Table 3 Summary of the experimental results.

# experiment	$So(\%)$	$\%LP(\%)$	$Q(l/s)$	$Vm(m/s)$	$Vs(m/s)$	$Y(cm)$	$q_s(m^2/s)$
1	4	0.05	0.44	0.13	0.38	1.55	$6.92\ x10^{-7}$
4	4	0.35	0.47	0.10	0.42	1.85	$6.65\ x10^{-6}$
17	4	0.05	13.39	1.31	1.02	3.40	$1.14\ x10^{-3}$
20	4	0.35	13.35	1.27	0.93	3.50	$1.01\ x10^{-3}$
21	5	0.05	0.72	0.22	0.22	1.75	$1.92\ x10^{-5}$
24	5	0.35	0.74	0.23	0.33	1.45	$6.74\ x10^{-6}$
37	5	0.05	11.99	1.11	0.91	3.60	$1.38\ x10^{-3}$
40	5	0.35	12.04	0.97	0.94	4.15	$1.77\ x10^{-3}$
41	6	0.05	0.72	0.26	0.30	1.85	$5.79\ x10^{-5}$
44	6	0.35	0.76	0.23	0.25	1.55	$2.22\ x10^{-5}$
57	6	0.05	15.33	1.29	1.05	3.95	$2.18\ x10^{-3}$
60	6	0.35	15.24	1.27	1.01	4.00	$1.84\ x10^{-3}$
61	7	0.05	0.77	0.25	0.36	1.25	$1.29\ x10^{-4}$
64	7	0.35	0.74	0.33	0.32	1.88	$6.24\ x10^{-5}$
77	7	0.05	15.37	1.58	1.23	3.25	$2.21\ x10^{-3}$
80	7	0.35	15.37	1.44	1.15	3.55	$2.33\ x10^{-3}$
81	8	0.05	0.51	0.17	0.43	1.00	$1.35\ x10^{-4}$
84	8	0.35	0.51	0.23	0.40	1.50	$7.73\ x10^{-5}$
97	8	0.05	14.13	1.50	1.54	3.15	$2.52\ x10^{-3}$
100	8	0.35	14.08	1.51	1.40	3.10	$1.94\ x10^{-3}$
101	9	0.05	0.67	0.17	0.65	1.30	$2.34\ x10^{-4}$
104	9	0.35	0.54	0.16	0.22	1.10	$2.16\ x10^{-4}$
117	9	0.05	14.67	1.09	1.28	4.50	$3.37\ x10^{-3}$
120	9	0.35	14.61	1.30	1.13	3.75	$2.97\ x10^{-3}$
121	10	0.05	0.99	0.18	0.59	1.80	$5.58\ x10^{-4}$
124	10	0.35	0.94	0.17	0.54	1.85	$3.74\ x10^{-4}$
137	10	0.05	14.38	1.41	2.12	3.40	$4.40\ x10^{-3}$
140	10	0.35	14.36	1.50	2.08	3.20	$3.33\ x10^{-3}$

Table 4 Linear and potential model parameters for each characteristic diameter considered.

Characteristic diameter	Type of model	C	a	b	c	d	e	f	g	h	Coefficient of determination R^2
$\boldsymbol{d_{14}}$	linear[1]	−2.9723	11.4304	0.0000	0.0002	0.0000	0.0521	0.3073	0.0000	2.9343	0.938
	potential[2]	25.2041	1.7725	0.0000	0.0000	0.0000	1.6447	0.0000	0.0000	1.1863	0.883
$\boldsymbol{d_{50}}$	linear[1]	−1.1417	3.4858	−0.2188	0.0002	0.0000	0.0717	0.1280	0.0000	1.9086	0.941
	potential[2]	22.8516	1.7401	0.0000	0.0000	0.0000	1.6447	0.0000	0.0000	1.2441	0.874
$\boldsymbol{d_{84}}$	linear[1]	−0.6420	0.4472	−0.0816	0.0000	0.0000	0.0632	0.0832	0.0000	1.7111	0.927
	potential[2]	0.6706	2.1940	0.0000	0.8201	0.0000	0.7001	0.0000	0.0000	0.0000	0.867
$\boldsymbol{d_{90}}$	linear[1]	−0.5587	0.4985	−0.0814	0.0000	0.0000	0.0623	0.0641	0.0000	1.5789	0.927
	potential[2]	0.9721	2.2791	0.0000	0.8048	0.0000	0.7171	0.0000	0.0000	0.0000	0.869

[1] The values presented correspond to the parameters of eq. (6).
[2] The values presented correspond to the parameters of eq. (7).

decreases due to the negative coefficient. This is expected since large particles absorb a portion of the total shear stress, leaving less shear stress available for transporting smaller sediment.

The influence of Q is almost negligible for smaller characteristic diameters and linear models. As the characteristic diameter increases for the potential models, the impact of discharge became significant. This can be understood considering that in order to put the discharge in a dimensionless form, the characteristic diameter is used to the fifth power; and in high-gradient rivers, large characteristic diameters become more significant due to the low values of relative submergence [9]. Further, Vws was present only in the linear models. As reported, the impact of Vws on bedload transport decreases as the characteristic diameter increases, indicating that as the roughness increases (greater characteristic diameter), the water surface velocity decreases its impact on sediment transport rates. In the case of $V*$ for the potential models, it was present only for the d_{14} and d_{50} models, and its exponents are similar. On the other hand, $V*$ for linear models had a higher impact on smaller characteristic diameters. This impact decreases as the characteristic diameter increases. Further, this is in accordance

with the fact that for a larger bed material size, less shear stress is available for sediment transport. Therefore, less transport rate is obtained for larger characteristic diameter.

Dimensionless flow depth (also relative submergence) Y and dimensionless boundary shear stress τ didn't appear in any model. The influence of flow depth in bedload transport rate may be included more efficiently in other variables such as mean flow velocity, water surface velocity, and/or shear velocity. Even though τ is not included in any model, $V*$ was present in almost all the models with the exception of two potential models (d_{84} and d_{90}). Boundary shear stress and shear velocity are equivalent parameters i.e., both include the same information (flow parameters) but have different forms. Additionally, the relation between the two variables is not linear. Therefore, their contribution to the variation of the independent variable (sediment transport rate) could be different and one could be more relevant than the other. Shear velocity can be seen as the transformed form of boundary shear stress. Furthermore, since both variables include the same information on the flow process and are highly correlated, neither model includes both variables. This is the result of defining a limit value of the VIF ($VIF = 10$). According to the stepwise regression methodology, shear velocity is a better predictor for bedload transport rate.

Different parameters were selected as the best predictors to represent bedload transport rates based on the results from the stepwise regression for the two types of models (linear and potential) and the characteristic diameter. Differences were observed between potential and linear models. However, all the dimensionless variables in the potential models for each characteristic diameter are also present in the linear models. The linear methodology can, thus, be seen as a preselection methodology to perform a more conceptual (potential) analysis with less variables. Based on characteristic diameters, changes were mostly registered for the parameters exponents/-coefficients magnitudes. However, the differences observed between d_{84} and d_{90} models could be negligible in both potential and linear models. Potential models could be seen as more compact and concise models. As it is observed in Table 4, all potential models included only three representative parameters. The linear models included five or six parameters. This could be due to the nonlinear nature of sediment transport. For linear models to represent a nonlinear process, they may require more variables. As a measure of each model's prediction capacity, their corresponding determination coefficients (R^2) were obtained (see Table 3). Linear models for all characteristic diameters report higher determination coefficients. For the linear models, the highest R^2 was obtained for the d_{50} model.

Data from measurements and predicted values (for each model developed) are plotted in Fig. 4 for each one of the models determined. The figure presents the equality line as well as one order and ø order of magnitude bands. In general, less scatter is seen for linear models, compared to the potential models. However, linear models show larger scatter for the lower transport rates. Previously developed models are believed to be less accurate at lower transport rates [10,17,56]. Generally, they tend to overpredict measured rates. Here, the observed scatter did not present a tendency to either overestimate or underestimate the measured values. Moreover, most of the data—with few exceptions—fell within the ø order of magnitude bands. In the case of potential models, the scatter was high but more uniformly distributed throughout the range of transport rate values. No differences were reported between high and low transport rates. For potential models, the data mostly fell within the ø order of magnitude bands. Based on the determination coefficients and the comparison between measured and calculated bedload transport rates presented in Fig. 4, the linear model developed with d_{50} could be seen as the model that better describes the measured data. However, the linear model developed with d_{84} present less scatter for the lower transport rates. Moreover, d_{84} is commonly used as the characteristic diameter for gravel bed rivers [14,51,65] . Additionally, considering that the potential models for d_{50} and d_{84} didn't present a tendency to have a higher scatter for lower dimensionless transport rates and that this can give the potential models a wider range of application, the d_{50} and d_{84} linear and potential models were selected to perform additional verification.

Considering that bedload transport is a complex and highly variable process, the accuracy within one order of magnitude has been defined as an acceptable reference to verify a transport rate prediction [3,42,51] . Fig. 5 presents the comparison between measured and calculated transport rates for d_{50} and d_{84} linear and potential models with the data from Smart [61] (Table 2). The line of equality and the lines defining one and ø orders of magnitude are also presented to verify the level of accuracy of each model. As observed in Fig. 5, for the d_{50} linear model, 3% of the total measurements (2 measurements) fell out of the one order of magnitude bands. Therefore, 97% of the data were within the one order of magnitude range. Considering the ø order of magnitude range, 27% of the measurements fell out of the range. Therefore, 73% of the data fell within ø order of magnitude. For the d_{50} potential model, 96% of the data were within the one order of magnitude range. A considerable difference was reported for the ø order of magnitude range. For the (d_{50}) potential model, 96% of the total measurements fell within the ø order of magnitude range. This represents an advantage of the potential model over the linear model.

For the d_{84} linear model, 5% of the total measurements (4 measurements) fell out of the one order of magnitude bands. Therefore, 95% of the data fell within the one order of magnitude range. Considering the ø order of magnitude range, 32% of the measurements fell out of the range. Thus, 68% of the data fell within the ø order of magnitude range. For the d_{84} potential model, 96% of the data were within the one order of magnitude range. As observed for the d_{50} model, a considerably higher prediction capacity was observed for the ø order of magnitude range for d_{84} potential model. 96% of the total measurements fell within the ø order of magnitude range. The R^2 obtained for Smart [61] data and the predicted values with d_{50} linear and potential models were 0.58 and 0.63, respectively. For the d_{84} linear and potential models, R^2 were 0.28 and 0.76, respectively. Considering the inherent fluctuation of bedload sediment transport, being able to predict a set of data within a ø order of magnitude is an encouraging result.

For both linear models, the percentage of the underestimation of measured transport rates was higher than the percentage of overestimation. The percentage of underestimation for the d_{50} model, was 69%. For the d_{84} model the percentage of underestimation was 73%. Additionally, the data appeared

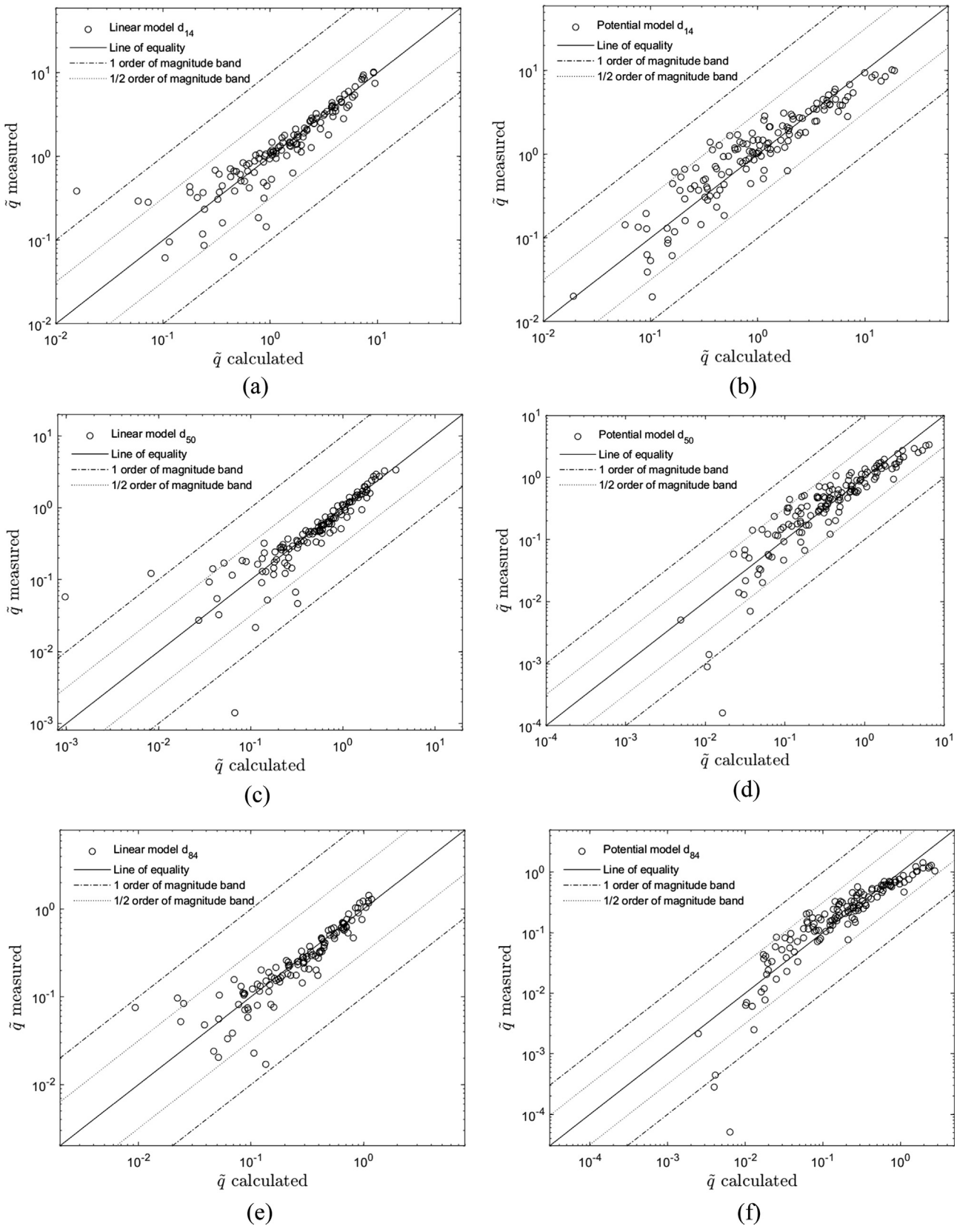

Fig. 4 Comparison between dimensionless transport rates measured and calculated with d_{14} as characteristic diameter (a) linear model and (b) potential model, with d_{50} (c) linear model and (d) potential model, with d_{84} (e) linear model and (f) potential model, and with d_{90} (g) linear model and (h) potential model.

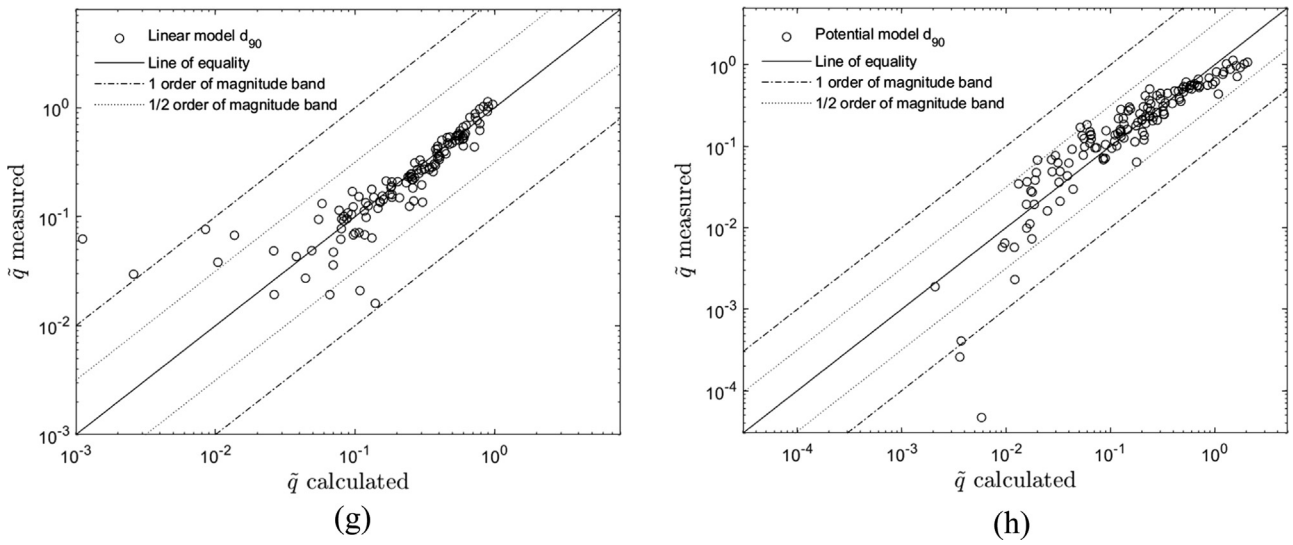

Fig. 4 (*continued*)

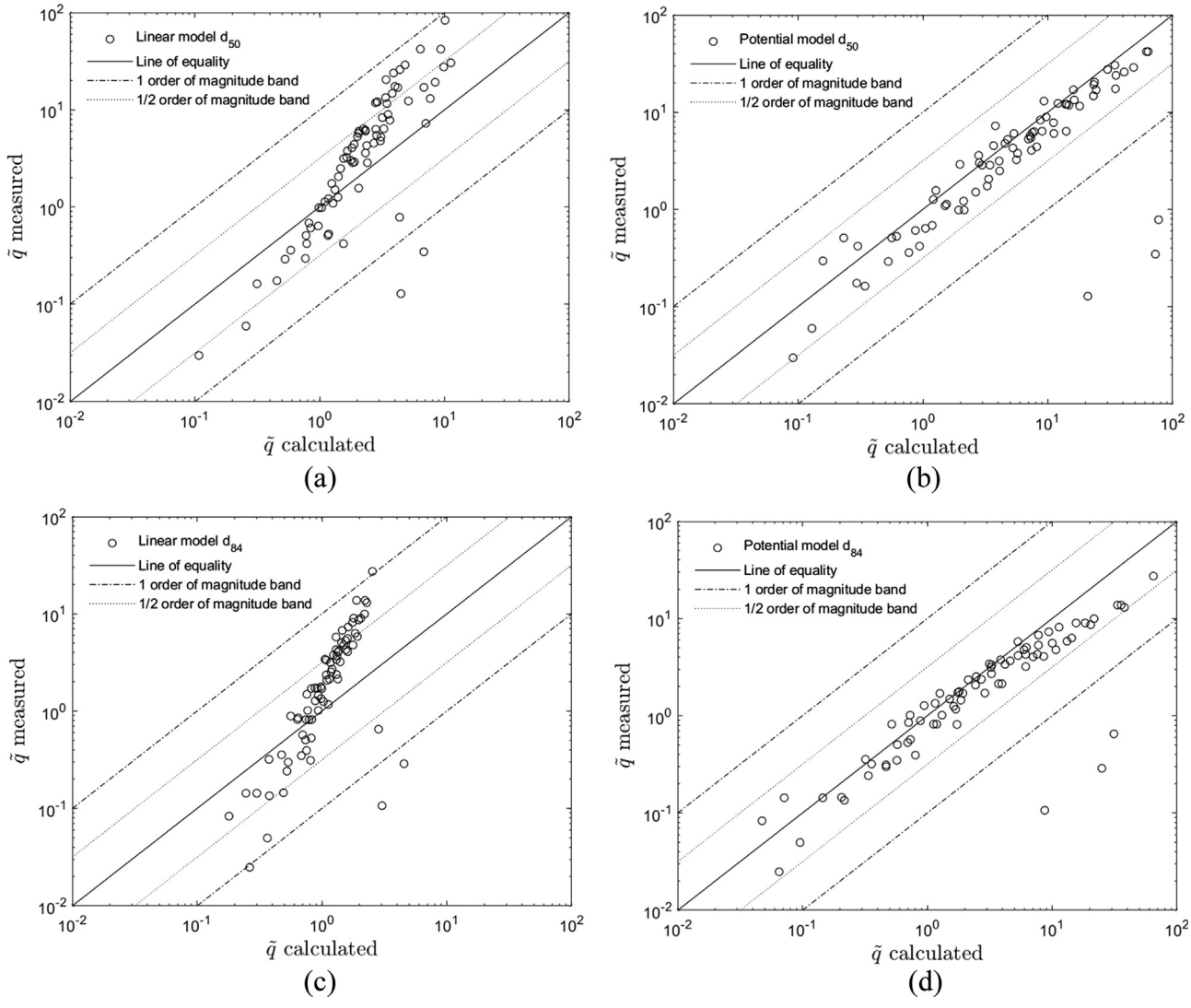

Fig. 5 Comparison between dimensionless transport rates [61] experimental data and their corresponding calculated values with (a) d_{50} linear model, (b) d_{50} potential model, and (c) d_{84} linear model, and (d) d_{84} potential model.

to cluster around a line with a slope greater than 1 (line of equality). Even though, for potential models, most of the data were below line of equality (79% of the data for d_{50} and 81% for d_{84}) representing higher calculated transport rates, the behavior (around the line of equality) was better for the potential models, compared to the linear models.

Bedload sediment transport has been proven to present a nonlinear relation with the independent variables that produces it[53]. As observed for the measured data, a linear relation performed better than the exponential data in terms of the determination coefficient R^2 . However, a difference on the goodness of fit is observed between the lower and higher transport rates. The weakest performance was registered for lower transport rates that could be associated with stronger nonlinear effects. Therefore, a range of applicability can be defined for the linear models (d_{50} and d_{84}). Considering the ø order of magnitude, linear models can be appropriate to determinate dimensionless transport rates from 10^{-1} to 4 for the d_{50} linear model and from 10^{-2} to 2 for the d_{84} linear model. When independent data were used to verify the models obtained, the performance of the potential model increased considerably with respect to linear models, making the nonlinearity effects more apparent. Considering the ø order of magnitude, it can be said that the potential model performed appropriately for the entire range of the data analyzed ($10^{-2} \leq q \leq 10^{2}$ for d_{50} and d_{84}).

Based on the comparisons developed between model performance for the data used to build the models as well as independent data and the fact that bedload sediment transport is a nonlinear process, both the linear and potential models performed considerably well under determined conditions. For the data used to develop the models, linear models performed better for narrower intervals. Based on the fact that a higher level of the models' complexity doesn't always represent a higher level of the models' performance, linear models have been proposed as simplified relations to describe experimental data [49]. Therefore, for the intervals defined by the data considered in the present study ($10^{-1} \leq q \leq 4$ for d_{50} linear model and from $10^{-2} \leq q \leq 2$ for d_{84} linear model), the d_{50} and d_{84} linear models can be used to estimate bedload transport rates. For the potential models (d_{50} and d_{84}), slightly lower values of the R^2 were registered (for the data used to develop the models). However, the data scatter is uniform along the range of data variation. Furthermore, the data mostly fell within the ø order of magnitude bands. For the independent set of data, linear models reported considerably lower R^2 than the potential models. Even though the data fell within one order of magnitude bands, the tendency observed for linear models doesn't adjust the line of equality. Moreover, the potential models showed a better behavior around the line of equality and cause almost all data (96%) to fall within ø order of magnitude bands.

As observed, differences were obtained in the performance of the linear and potential models. Linear models can be used within the range in which their performance has been verified as satisfactory. Potential models have been verified with an appropriate level of accuracy for the entire range of variation of the two sets of data considered in the present study. Therefore, if a single model should be selected, based on the overall behavior in terms of accuracy and ranges of application, the d_{84} potential model could be assumed as the most adequate model.

6. Conclusions

An experimental study was performed to establish a relation to estimate bedload sediment transport rates based on the features of a high-gradient gravel bed river. To ensure that most of the parameters that define bedload transport rates were included in the model, the following were considered as independent variables: channel slope (S_o), area occupied by large particles/total plan area ($\%LP$), dimensionless form of flow discharge (Q), mean flow depth (Y), mean flow velocity (Vm), water surface velocity (Vws), boundary shear stress (τ), and shear velocity ($V*$). To express the variables in their dimensionless form, particle diameters d_{14}, d_{50}, d_{84}, and d_{90} were considered as the characteristic diameters. A Stepwise regression analysis was performed to determine the best predictors from the set of independent variables. Through the AIC parameter, simple but accurate models were obtained based on the determination of relevant parameters and their most efficient combination. The use of the AIC favors the optimization of computational calculations. This can be an efficient methodology to apply in data sets with an extensive number of parameters. While the set of independent variables used in the present work was not too large, this study can nevertheless be used to verify that a physical process can be characterized with a model built based on data through the inclusion of the relevant variables.

Based on the best predictors, linear and potential models with d_{14}, d_{50}, d_{84}, and d_{90} were obtained from the experimental data. The linear models had higher determination coefficients that reflected less scatter around the equality line when measured and the calculated rates were compared ($R^2 = 0.94$ for d_{50} and $R^2 = 0.93$ for d_{84}). However, higher scatter was observed for the lower values of dimensionless transport rates. Regardless of the scatter, most of the data were within ø order of magnitude bands. This is encouraging considering that one order of magnitude have been defined as the limit for an acceptable prediction of bedload transport rates due to their complexity and inherent fluctuations [3,51]). An applicability range could be defined for the linear models. When $10^{-1} \leq q \leq 4$ (for d_{50}) and $10^{-2} \leq q \leq 2$ (for d_{84}), the linear models could be used to estimate dimensionless bedload transport rates. The potential models presented more scatter ($R^2 = 0.87$ for d_{50} and d_{84}, however, it was uniformly distributed throughout the range of dimensionless transport rates values, i.e., no different scattering was observed between low and high dimensionless transport rates.

The linear and potential models for the d_{50}, and d_{84} characteristic diameters were selected to verify these models with an independent set of data from Smart [61]. Encouraging results were obtained from this verification; 73% and 68% of the data fell within ø order of magnitude bands for the d_{50} ; and d_{84} linear models, respectively. For the d_{50} linear model, 97% of the data fell within one order of magnitude. In the case of the d_{84} linear model, 95% of the data was within one order of magnitude. However, the tendency observed for the linear models did not agree well with the equality line. A slope greater than unity was sharply defined by the data. For the potential models, the results were even more encouraging, as 96% of the data were within ø and one order of magnitude bands for both d_{50} and d_{84}, with the data clustering around the equality line. The R^2 for the potential models were higher than those for the linear models. Therefore, for the range of the data from Smart

[61], the potential models reported better performance. Thus, the range for acceptable performance for the potential models could be defined for the entire dimensionless transport rate variation, i.e., $10^{-4} \leq q \leq 10^{1}$.

Even though bedload transport has been identified to be nonlinear, discrete parts of the function can be approximated by lines because the interaction between variables may include the nonlinearity of the process. This can lead to the conclusion that linear methodology can be used as a first approximation for the model build process, which may be updated as nonlinear based just on the parameters determined as significant in the linear analysis. Thus, the use of resources (time and computational expenses) can be optimized.

Considering that potential models have a wider range of applications and that their R^2 were not considerably smaller than the linear models through the model building process, while also factoring in that sediment transport has been characterized as nonlinear, the d_{84} potential model (with R^2 higher than the d_{50} for the validation data) can be considered as the one with the higher overall performance. However, for the ranges $10^{-1} \leq q \leq 4$ (for d_{50}) and $10^{-2} \leq q \leq 2$ (for d_{84}), the linear models presented here can be used to estimate transport rates with a slightly higher accuracy.

The methodology developed in the present study has allowed the determination of a bedload transport rate model. From the several possible variables that can impact the final rates, the most relevant were selected in their most efficient combination to keep the final model sufficiently simple but accurate. Linear methodology can be seen as a preselection step to optimize the parameters used for a more conceptual potential model determination. Moreover, parsimonious models were obtained and proved to perform considerably well for an independent set of data, thereby verifying that the procedure applied was able to obtain a model that could capture—at least on some level—the nature of the phenomenon analyzed. Since this methodology has given encouraging results with experimental data, it can be applied to analyze field data as well.

7. Data availability

The data presented in this study are available on request from the corresponding author, V.C.

CRediT authorship contribution statement

Veronica Carrillo: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Daniel Mendoza:** Formal analysis, Writing – review & editing. **John Petrie:** Formal analysis, Writing – review & editing. **Pedro Matovelle:** Methodology, Investigation. **Sebastian Torres:** Methodology, Investigation. **Daniel Mendoza:** Formal analysis, Writing – review & editing. **Esteban Pacheco:** Formal analysis, Writing – review & editing. **Felipe Cisneros:** Formal analysis, Writing – review & editing. **Luis Timbe:** Formal analysis, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Akaike, H. (1973). "Information theory and an extension of the maximum likelihood principle." *2nd International Symposium on Information Theory*, Akadémiai Kiadó, Budapest, Hungary, 267–281.
[2] H. Akaike, A new look at the statistical model identification, IEEE Trans. Autom. Control 19 (6) (1974) 716–723.
[3] C. Ancey, Bedload transport: a walk between randomness and determinism. Part 2. Challenges and prospects, J. Hydraul. Res. 58 (1) (2020) 18–33.
[4] C. Ancey, T. Böhm, M. Jodeau, P. Frey, Statistical description of sediment transport experiments, Phys. Rev. E 74 (1) (2006) 011302.
[5] Bagnold, R. A. (1980). "An empirical correlation of bedload transport rates in flumes and natural rivers." *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 372(1751), 453–473.
[6] J.J. Barry, J.M. Buffington, J.G. King, A general power equation for predicting bed load transport rates in gravel bed rivers, Water Resour. Res. 40 (10) (2004), https://doi.org/10.1029/2004WR003190.
[7] J.C. Bathurst, Flow resistance estimation in mountain rivers, J. Hydraul. Eng. 111 (4) (1985) 625–643.
[8] J.C. Bathurst, Critical conditions for bed material movement in steep, boulder-bed streams, IAHS-AISH Publication (1987) 309–318.
[9] J.C. Bathurst, Environmental river flow hydraulics, Applied fluvial geomorphology for river engineering and management, United Kingdom, 1997, pp. 69–103.
[10] J.C. Bathurst, Effect of coarse surface layer on bed-load transport, J. Hydraul. Eng. 133 (11) (2007) 1192–1205.
[11] V.M. Carrillo, J.E. Petrie, F.E. Cisneros, L.M. Timbe, Application of hydraulic geometry to high gradient rivers in Southern Ecuador, Water Resour. Res. 57 (3) (2021), https://doi.org/10.1029/2019WR026756.
[12] V. Carrillo, J. Petrie, L. Timbe, E. Pacheco, W. Astudillo, C. Padilla, F. Cisneros, Validation of an experimental procedure to determine bedload transport rates in steep channels with coarse sediment, Water 13 (5) (2021) 672.
[13] N.-S. Cheng, Exponential formula for bedload transport, J. Hydraul. Eng. 128 (10) (2002) 942–946.
[14] N.-S. Cheng, Representative grain size and equivalent roughness height of a sediment bed, J. Hydraul. Eng. 142 (1) (2016) 06015016, https://doi.org/10.1061/(ASCE)HY.1943-7900.0001069.
[15] N.-S. Cheng, X. Chen, Slope correction for calculation of bedload sediment transport rates in steep channels, J. Hydraul. Eng. 140 (6) (2014) 04014018, https://doi.org/10.1061/(ASCE)HY.1943-7900.0000843.
[16] N.-S. Cheng, Y. Lu, M. Wei, Simple formulation of bedload sediment transport rate based on novel definition of pickup probability, J. Hydraul. Eng. 146 (9) (2020) 06020012, https://doi.org/10.1061/(ASCE)HY.1943-7900.0001797.
[17] V. D'Agostino, M.A. Lenzi, Bedload transport in the instrumented catchment of the Rio Cordon, CATENA 36 (3) (1999) 191–204.
[18] G.C.L. David, E. Wohl, S.E. Yochum, B.P. Bledsoe, Controls on at-a-station hydraulic geometry in steep headwater streams, Colorado, USA, Earth Surf. Proc. Land. 35 (15) (2010) 1820–1837.
[19] S. Dey, S. Zeeshan Ali, E. Padhi, Terminal fall velocity: the legacy of Stokes from the perspective of fluvial hydraulics, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 475 (2228) (2019) 20190277.
[20] H.A. Einstein, Formulas for the transportation of bed load, Trans. ASCE 107 (1942) 561–573.

[21] D. Fang, Book review of: Sediment transport theory and practice by T. C. Yang, Int. J. Sedim. Res. 13 (1998) 74–83.
[22] S.L. Fathel, D.J. Furbish, M.W. Schmeeckle, Experimental evidence of statistical ensemble behavior in bed load sediment transport, J. Geophys. Res. Earth Surf. 120 (11) (2015) 2298–2317.
[23] Ferguson. (2007). "Flow resistance equations for gravel- and boulder-bed streams." *Water Resources Research*, 43(5), n/a–n/a.
[24] R. Fernandez Luque, R. Van Beek, Erosion and transport of bed-load sediment, J. Hydraul. Res. 14 (2) (1976) 127–144.
[25] Frey, P., and Church, M. (2012). "Gravel transport in granular perspective." *Gravel-Bed Rivers*, M. Church, P. M. Biron, and A. G. Roy, eds., John Wiley & Sons, Ltd, Chichester, UK, 37–55.
[26] T. Ghilardi, M. Franca, A. Schleiss, Sediment transport in steep channels with large roughness elements, in: A. Schleiss, G. de Cesare, M. Franca, M. Pfister (Eds.), River Flow 2014, CRC Press, 2014, pp. 899–907.
[27] Basil Gomez, Michael Church, An assessment of bed load sediment transport formulae for gravel bed rivers, Water Resour. Res. 25 (6) (1989) 1161–1186.
[28] R.R. Hocking, A biometrics invited paper. The analysis and selection of variables in linear regression, Biometrics 32 (1) (1976) 1.
[29] Honderich, T. (Ed.). (2005). *The Oxford Companion to Philosophy*. Oxford University Press, Oxford ; New York.
[30] C. Juez, J. Murillo, P. García-Navarro, Numerical assessment of bed-load discharge formulations for transient flow in 1D and 2D situations, J. Hydroinf. 15 (4) (2013) 1234–1257.
[31] Carmelo Juez, Sandra Soares-Frazao, Javier Murillo, Pilar García-Navarro, Experimental and numerical simulation of bed load transport over steep slopes, J. Hydraul. Res. 55 (4) (2017) 455–469.
[32] Donald W. Knight, River hydraulics – a view from midstream, J. Hydraul. Res. 51 (1) (2013) 2–18.
[33] M.H. Kutner, C.J. Nachtsheim, J. Neter, "Simultaneous inferences and other topics in regression analysis". *Applied Linear Regression Models*, 4th ed., McGraw-Hill Irwin, New York, NY, 2004, pp. 168–170.
[34] J.T. Limerinos, Determination of the manning coefficient from measured bed roughness in natural channels, Water Supply Paper, Report, 1970.
[35] Yvonne Martin, Evaluation of bed load transport formulae using field evidence from the Vedder River, British Columbia, Geomorphology 53 (1-2) (2003) 75–95.
[36] Medina Cueva, G. L. (2008). "Gestión y Manejo de los Recursos Hídricos en la Cuenca del río Tabacay Cantón Azogues - Ecuador." Universidad de Cuenca, Cuenca – Ecuador.
[37] E. Meyer-Peter, R. Müller, Formulas for bed-load transport, Proc. 2nd Meeting IAHR (1948) 39–64.
[38] Robert G. Millar, Grain and form resistance in gravel-bed rivers Résistances de grain et de forme dans les rivières à graviers, J. Hydraul. Res. 37 (3) (1999) 303–312.
[39] E.R. Mueller, Morphologically based model of bed load transport capacity in a headwater stream, J. Geophys. Res. 110 (F2) (2005) F02016.
[40] Erich R. Mueller, John Pitlick, Jonathan M. Nelson, Variation in the reference shields stress for bed load transport in gravel-bed streams and rivers: variation in the reference shields stress, Water Resour. Res. 41 (4) (2005), https://doi.org/10.1029/2004WR003692.
[41] Jasna Muskatirovic, Analysis of bedload transport characteristics of Idaho streams and rivers, Earth Surf. Proc. Land. 33 (11) (2008) 1757–1768.
[42] Manuel Nitsche, Dieter Rickenmann, Jens M. Turowski, Alexandre Badoux, James W. Kirchner, Evaluation of bedload transport predictions using flow resistance equations to account for macro-roughness in steep mountain streams:evaluation of bedload transport predictions, Water Resour. Res. 47 (8) (2011), https://doi.org/10.1029/2011WR010645.
[43] M.C. Palucis, M.P. Lamb, What controls channel form in steep mountain streams? Controls on steep channel form, Geophys. Res. Lett. 44 (14) (2017) 7245–7255.
[44] Marisa C. Palucis, Thomas P. Ulizio, Brian Fuller, Michael P. Lamb, Flow resistance, sediment transport, and bedform development in a steep gravel-bedded river flume, Geomorphology 320 (2018) 111–126.
[45] Papanicolaou, T., and Kramer, C. (2006). "The role of relative submergence on cluster microtopography and bedload predictions in mountain streams." *River, Coastal and Estuarine Morphodynamics: Proceedings of the 4th IAHR Symposium on River, Coastal and Estuarine Morphodynamics (RCEM 2005, Urbana, Illinois, USA, 4-7 October 2005)*, G. Parker and M. Garcia, eds., Taylor & Francis.
[46] Gary Parker, Hydraulic geometry of active gravel rivers, J. Hydraulics Division 105 (9) (1979) 1185–1201.
[47] G. Parker, Transport of gravel and sediment mixtures, in: M. García Practice (Ed.), Sedimentation Engineering: Processes, Measurements, Modeling, American Society of Civil Engineers, Reston, VA, 2008.
[48] Gary Parker, Peter C. Klingeman, David G. McLean, Bedload and size distribution in paved gravel-bed streams, J. Hydraulics Division 108 (4) (1982) 544–571.
[49] M. Poorhosein, H. Afzalimehr, J. Sui, V.P. Singh, S. Azareh, Empirical bed load transport equations, Int. J. Hydraulic Engineering 3 (3) (2014) 93–101.
[50] Pugh, C. A. (2008). "Sediment transport scaling for physical models." *Sedimentation Engineering: Processes, Measurements, Modeling, and Practice*, M. Garcia, ed., American Society of Civil Engineers, Reston, VA.
[51] A. Recking, A comparison between flume and field bed load transport data and consequences for surface-based bed load transport prediction: flume and field bed load transport, Water Resour. Res. 46 (3) (2010), https://doi.org/10.1029/2009WR008007.
[52] A. Recking, Simple method for calculating reach-averaged bed-load transport, J. Hydraul. Eng. 139 (1) (2013) 70–75.
[53] A. Recking, An analysis of nonlinearity effects on bed load transport prediction, J. Geophys. Res. Earth Surf. 118 (3) (2013) 1264–1281.
[54] Alain Recking, Frédéric Liébault, Christophe Peteuil, Thomas Jolimet, Testing bedload transport equations with consideration of time scales: bedload modelling, Earth Surf. Proc. Land. 37 (7) (2012) 774–789.
[55] Rickenmann, D. (1990). "Bedload transport capacity of slurry flows at steep slopes." ETH Zurich.
[56] Dieter Rickenmann, Comparison of bed load transport in torrents and gravel bed streams, Water Resour. Res. 37 (12) (2001) 3295–3305.
[57] Dieter Rickenmann, Alain Recking, Evaluation of flow resistance in gravel-bed rivers through a large field data set: evaluation of flow resistance equations, Water Resour. Res. 47 (7) (2011), https://doi.org/10.1029/2010WR009793.
[58] Ripley, B., Venables, B., Hornik, K., Gebhardt, A., and Firth, D. (2013). *Package 'MASS' CRAN Repos. https://cran.r-project.org/web/packages/MASS/MASS.pdf*.
[59] S. Sheather, A Modern Approach to Regression with R, Springer Texts in Statistics, Springer, New York, New York, NY, 2009.
[60] S.K. Sinnakaudan, A. Ab Ghani, M.S. Ahmad, N.A. Zakaria, Multiple linear regression model for total bed material load prediction, J. Hydraul. Eng. 132 (5) (2006) 521–528.
[61] Graeme M. Smart, Sediment transport formula for steep channels, J. Hydraul. Eng. 110 (3) (1984) 267–276.

[62] R. Soulsby, Dynamics of Marine Sands, Thomas Telford, London, 1997.
[63] L.C. van Rijn, Unified view of sediment transport by currents and waves. I: initiation of motion, bed roughness, and bed-load transport, J. Hydraul. Eng. 133 (6) (2007) 649–667.
[64] Peter J. Whiting, William E. Dietrich, Boundary shear stress and roughness over mobile alluvial beds, J. Hydraul. Eng. 116 (12) (1990) 1495–1511.
[65] Ellen E. Wohl, Andrew Wilcox, Channel geometry of mountain streams in New Zealand, J. Hydrol. 300 (1-4) (2005) 252–266.
[66] E.M. Yager, W.E. Dietrich, J.W. Kirchner, B.W. McArdell, Prediction of sediment transport in step-pool channels, Water Resour. Res. 48 (1) (2012), https://doi.org/10.1029/2011WR010829.
[67] Yager, E. M., Kirchner, J. W., and Dietrich, W. E. (2007). "Calculating bed load transport in steep boulder bed channels: sediment transport in steep channels." *Water Resour. Res.*, 43 (7), n/a–n/a.
[68] C.T. Yang, C. Huang, Applicability of sediment transport formulas, Int. J. Sedim. Res. 16 (3) (2001) 335–353.
[69] Shu-Qing Yang, Sediment transport capacity in rivers, J. Hydraul. Res. 43 (2) (2005) 131–138.