Automated Bedform Identification - A Meta-Analysis of Current Methods and the Heterogeneity of their Outputs

Leon Scheiber[1], Judith Yttje Zomer[2], Li Wang[3], Julia Cisneros[4], Ronald Gutierrez Llantoy[5], and Alice Lefebvre[6]

[1]Ludwig-Franzius-Institute
[2]Wageningen University
[3]State Key Laboratory for Estuarine and Coastal Research, East China Normal University, Shanghai, China
[4]Virginia Tech
[5]Pontificia Universidad Católica de Perú
[6]MARUM - University of Bremen

December 27, 2023

Abstract

Ongoing efforts to characterize underwater dunes have led to a considerable number of freely available tools that identify these bedforms in a (semi-)automated way. However, these tools differ with regard to their research focus and appear to produce results that are far from unequivocal. We scrutinize this assumption by comparing the results of five recently published dune identification tools in a comprehensive meta-analysis. Specifically, we analyse dune populations identified in three bathymetries under diverse flow conditions and compare the resulting dune characteristics in a quantitative manner. Besides the impact of underlying definitions, it is shown that the main heterogeneity arises from the consideration of a secondary dune scale, which has a significant influence on statistical distributions. Based on the quantitative results, we discuss the individual strengths and limitations of each algorithm, with the aim of outlining adequate fields of application. Yet, the concerted bedform analysis and subsequent combination of results have another benefit: the creation of a benchmarking data set which is inherently less biased by individual focus and therefore a valuable instrument for future validations. Nevertheless, it is apparent that the available tools are still very specific and that end-users would profit by their merging into a universal and modular toolbox.

Automated Bedform Identification – A Meta-Analysis of Current Methods and the Heterogeneity of their Outputs

Leon Scheiber*[,1], Judith Zomer[2], Li Wang[3], Julia Cisneros[4], Ronald R. Gutierrez[5] & Alice Lefebvre[6]

[1] Ludwig-Franzius-Institute for Hydraulics, Estuarine and Coastal Engineering, Leibniz University Hannover, 30167 Hanover, Germany; scheiber@lufi.uni-hannover.de; OrcID: 0000-0001-7989-7639.

[2] Hydrology and Environmental Hydraulic Group, Department of Environmental Sciences, Wageningen University & Research, Wageningen, Netherlands; OrcID: 0000-0002-4385-8782.

[3] State Key Laboratory for Estuarine and Coastal Research, East China Normal University, Shanghai, China; OrcID: 0000-0001-8665-4061.

[4] Department of Geosciences, Virginia Tech, Blacksburg, Virginia, USA; OrcID: 0000-0001-6451-4180.

[5] Departamento de Ingeniería (GERDIS, GEOSED), Pontificia Universidad Católica del Perú, Lima, Peru; OrcID: 0000-0003-0150-7439.

[6] MARUM-Center for Marine Environmental Sciences, University of Bremen, Bremen, Germany; OrcID: 0000-0002-9234-8279.

Corresponding author: Leon Scheiber (scheiber@lufi.uni-hannover.de)

Key Points:

- Five recent dune identification tools are compared in a meta-analysis assessing three bathymetries from diverse flow environments
- Results differ in sampling size, computation time and dune characteristics, but mainly regarding the allowance of small-scale bedforms
- The combination of results can be used as benchmarking data in future validations because it is less biased by individual research focus

Keywords: geomorphology, underwater dunes, bedform analysis, dune identification, meta-analysis

Abstract

Ongoing efforts to characterize underwater dunes have led to a considerable number of freely available tools that identify these bedforms in a (semi-)automated way. However, these tools differ with regard to their research focus and appear to produce results that are far from unequivocal. We scrutinize this assumption by comparing the results of five recently published dune identification tools in a comprehensive meta-analysis. Specifically, we analyse dune populations identified in three bathymetries under diverse flow conditions and compare the resulting dune characteristics in a quantitative manner. Besides the impact of underlying definitions, it is shown that the main heterogeneity arises from the consideration of a secondary dune scale, which has a significant influence on statistical distributions. Based on the quantitative results, we discuss the individual strengths and limitations of each algorithm, with the aim of outlining adequate fields of application. Yet, the concerted bedform analysis and subsequent combination of results have another benefit: the creation of a benchmarking data set which is inherently less biased by individual focus and therefore a valuable instrument for future validations. Nevertheless, it is apparent that the available tools are still very specific and that end-users would profit by their merging into a universal and modular toolbox.

Plain Language Summary

In this paper, we present a comparison of several recently published bedform analysis tools that were created to measure the size and shape of underwater dunes. We compare these tools and how they performed on three different seabed elevation maps that detail dunes in a river, an estuary, and a laboratory flume. We focus on the resulting measurements of dune morphology, such as dune height and length. We show that the consideration of a secondary dune scale has a significant influence on the statistical distributions that describe the dune populations measured. With this knowledge, we offer a discussion of the strengths and limitations of each tool and examples of their proper application. We also offer the combination of the measurements from each tool as a benchmarking dataset that can be used for future tool creations and method validation. Finally, we recommend each tool be used with specific needs in mind and a universal and modular toolbox should be created that incorporates all available options for dune identification.

1 Introduction

Underwater dunes and ripples are a particular type of planetary landform. These so-called bedforms develop at the interface of a flow field and a movable sediment layer. They can be observed in the most diverse environments on Earth and other planetary bodies: from the deep-sea and continental shelves (Cukur et al., 2022; Franzetti et al., 2013; Reeder et al., 2011) over tidally-constrained basins (Armstrong et al., 2021; Hu et al., 2021) to inland streams and rivers (Le Guern et al., 2021; Wu et al., 2021) and even across the barren landscapes of Mars and Titan (Breed et al., 1979; Lorenz et al., 2006). The formation and dynamic behavior of underwater dunes have wide implications for hydrological and morphological processes. For instance, they can allow conclusions about local flow conditions at present (Lefebvre et al., 2011a; Parsons et al., 2005) and, through paleo-hydraulic analyses, the conditions of ancient environments (Hartley & Owen, 2022; Myrow et al., 2018). Their migration is an indicator of downstream bed load transport, which represents a critical component in the balance of erosion and accretion as sediment is transported from the highlands to the lowlands of our world (Jordan et al., 2019; Nittrouer et al., 2008). Furthermore, bedforms run the risk of interfering with man-made structures, such as offshore pipelines, navigation channels, transportation tunnels or bridge piers (Amsler & Garcia, 1997; Bruschi et al., 2014; Huizinga, 2017; Scheiber et al., 2021a; Xu et al., 2010). Last but not least, the natural flow and grain size variation within dune fields adds value to marine ecosystems and is therefore an important part of habitat mapping (Greene et al., 2020; Meijer et al., 2022). On these grounds, bedforms are of interest to a diverse community of researchers from both natural and applied sciences (Lefebvre & Winter, 2021).

Most of our knowledge of bedforms stems from experimental flows in laboratory flumes, where crests (and troughs) tend to be perpendicular to the main flow direction and the measuring of bedform characteristics is therefore straightforward, but dunes in the field typically have a more complex shape (Best, 2005). For instance, multiple scales of bedforms can co-exist in so-called compound dunes, where smaller bedforms are superimposed on larger primary dunes (Ashley, 1990). Other cases of complex bathymetric processes are dune amalgamation or calving (Bradley & Venditti, 2021). The observation of such processes is possible by growing amounts of three-dimensional bathymetric data, obtained via high-resolution multibeam echo-sounding (MBES) through support of federal or scientific surveys. Early methodologies to study underwater dunes can be differentiated into two main categories: geostatistical assessments of bed elevation profiles (Simons et al., 1965) and spectral analyses that translate rhythmic bedforms into sinusoidal or wavelet components (Nordin & Algert, 1966). Since then, the rapid increase in computational power during the last decades has fueled the publication of numerous methodologies for a (semi-)automated identification and characterization of bedforms. Table 1 shows a selection of recent publications in this respect, of which several build on a combination of the two aforementioned approaches. More recently, attempts have also been made to integrate geomorphometric and object-based methods (Pike, 2000), yet focusing on mathematical surface properties rather than describing individual bedforms. It should be noted that many identification methods are only designed to analyze specific data sets and therefore not provided for further use.

Table 1. Automated bedform analyses - (non-exhaustive) list of recent publications presenting computer-aided routines for the identification and characterization of bedforms from bathymetric data. Those publications with an asterisk are considered in the meta-analysis that is the basis of this study.

Spectral / wavelet-based	**Spatial / geo-statistical**	**Geomorphometric / Object-based**	**Combined**
Winter and Ernstsen (2007)	Wilbers and Brinke (2003)	Ogor (2018)	van Dijk et al. (2008)
Lefebvre et al. (2011b)	van der Mark et al. (2008)	Di Stefano and Mayer (2018)	Cazenave et al. (2013)
Gutierrez et al. (2013)	Ganti et al. (2013) ?	Cassol et al. (2022)	Wang et al. (2020)*
Lisimenka and Kubicki (2017)	Cisneros et al. (2020)*	Hansen et al. (2022)	Lefebvre et al. (2022)*
Gutierrez et al. (2018)	Schippa and Cavalieri (2021)	Lebrec et al. (2022)	Zomer et al. (2022)*
Lee et al. (2021)?	Scheiber et al.	Cassol et al. (2021)	

	(2021a)*	
	Lee et al. (2021)	Núñez-González et al. (2021)
	Lebrec et al. (2022)	

It is generally praiseworthy that so many researchers have contributed to the automation of bedform analyses and continue to provide access to readily applicable algorithms. This allows practitioners from both public authorities and neighboring fields of research to focus on their objectives more specifically. However, the sheer amount of options leaves end users with the agony of choice as to what tool should be used under which conditions, and even more so, because different algorithms have been found to produce significantly different results (Scheiber et al., 2021b). To address this shortcoming in current research and, hopefully, enhance future bedform studies, we have designed a meta-analysis, in which five identification tools are applied to three bathymetric benchmarking data sets. Based on this methodology, our meta-analysis aims to:

i. quantify the range of differences in obtainable results,

ii. discuss inherent biases resulting from different focuses,

iii. recommend fields of application for future end users.

This paper begins with a short description of the specificities in dune identification, a definition of all relevant dune characteristics and a presentation of the assessed bathymetries. Following this methodological section 2, we present statistical analyses that epitomize the heterogeneity of dune tracking outputs with regard to sampling sizes, dune scales and geometries (section 3 Results). Thereafter, we investigate possible causes and jointly derive guidelines for the sound application of individual tools in section 4 Discussion. Finally, key findings are summarized in 5 Conclusions and a short outlook is given addressing implications of this study and further research needs.

2 Materials and Methods

In order to systematically compare the available options for an automated detection of bedforms and their characterization, we teamed up in an international working group. To ensure correct usage, each co-author, who represents one of five recently published identification methods, applied their respective method to three independent bathymetries. These benchmarking data sets comprise dunes that formed in uniform river flow, under tidally-constrained conditions and in flume experiments, thus representing a wide range of environments, flow directions and scales.

2.1 Bedform identification

The methods applied in this study comprise both spectral and statistical approaches. They generally follow the explanations given in the independent research articles, where they have first been published. However, they can be differentiated depending on their specific objectives and the way that crests and troughs are identified (Table 2).

Table 2. Overview of recently published dune identification algorithms considered in this meta-analysis listed in chronological order.

Publication	Main focus	Crest identification
Cisneros et al. (2020)	scale separation; dune shape	local extremes and resolution scale slope changes (local aspect changes)
Wang et al. (2020)	scale separation; compound dunes	max. between zero-crossing (spectral/wavelet analysis; filtering)
Scheiber et al. (2021a)	compound dunes	local extremes (findpeaks.m)
Lefebvre et al. (2022)	dune shape	local extremes (min. curvature; min. elevation)
Zomer et al. (2022)	scale separation	max. between zero-crossing (after LOESS filtering)

In particular, two of the publications focus especially on the shape of dunes and their spatial variability, i.e. ensuring a correct delineation of the flow-traverse crests and troughs that define a dune (Cisneros et al., 2020; Lefebvre et al., 2022). However, the two methods differ in terms of how crests and troughs are identified in a given bathymetric map. For instance, Cisneros et al. (2020) start by calculating a matrix of aspect directions from a 3x3 moving window and then identify changes from stoss- to lee-side. Lefebvre et al. (2022), in contrast, build upon the detection of continuous crest objects, which have a minimum curvature below a certain threshold. A second focus can be seen in the separation of bedform scales as inherent to compound dunes. In the case of Wang et al. (2020), this is accomplished by applying an initial two-dimensional Fourier transform, followed by wavelet analysis and multiple filtering techniques including circular high-pass and robust spline filtering applied to rotated bed elevation profiles (BEPs). The subsequent zero-crossing analysis is comparable to the one employed by Zomer et al. (2022). However, Zomer et al. (2022) use a different method to separate bedform scales. In this approach, the primary bedform morphology is fitted using a "locally estimated scatter plot smoothing" (LOESS) algorithm, combined with a sigmoid function to correctly fit the steep lee slopes of primary dunes. The tool by Scheiber et al. (2021a) focuses on exactly these compound bed features, but relies on an iterative identification of local extremes in order to describe bedforms on all existing scales and as thoroughly as possible. All of the methods are implemented in Mathworks' MatLab and were operated, here, by the respective developer. They can be obtained from the corresponding authors or from different online repositories as listed in the data availability statement.

2.2 Bedform characteristics

Once crests and troughs have been delineated across the bathymetry under investigation, the dimensions of the corresponding dunes can be measured. In this context, it is most common that each dune is defined by one distinct crest and its two neighboring troughs. Although other definitions exist, e.g. by two crests or two slopes, these variants were not regarded as useful for the objectives of this study. Moreover, as bedforms typically show an asymmetric shape, one can further distinguish between the stoss side, i.e. the slope facing the formative flow, and the lee side downstream (see Figure 1). On this basis, several characteristics can be calculated describing the size and shape of the dune, first and foremost its height H and length L (also called spacing or wavelength). Besides these Latin descriptors, the Greek letters η and λ are used by many authors. Even if these characteristics seem intuitive, the current literature features a wide variety of possible definitions (Figure 1).

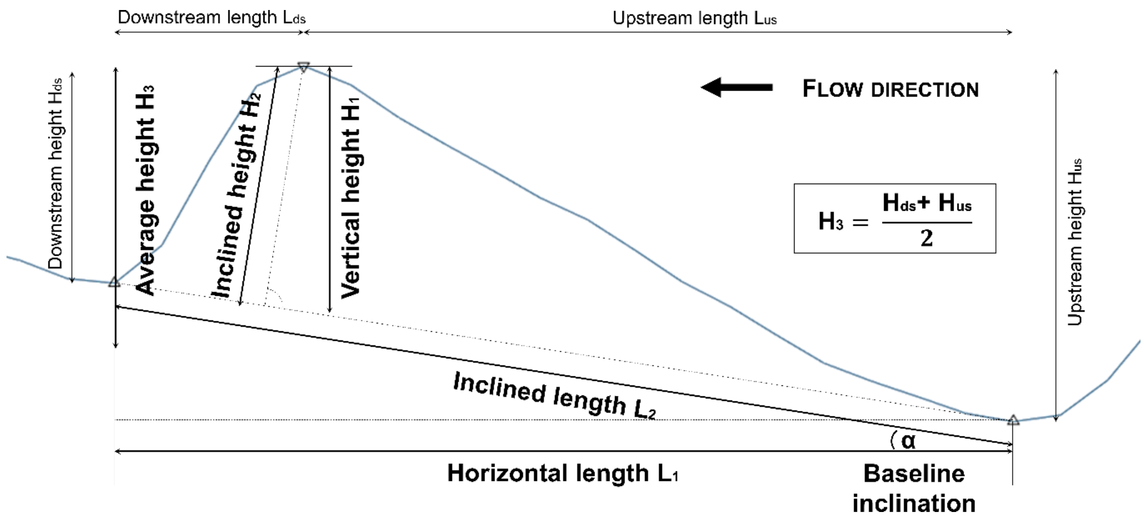

Figure 1. HEIGHT/LENGTH DEFINITIONS - An asymmetric dune as classically defined by one distinct crest and two adjacent troughs. Similar to dune length, the calculation of dune heights can allow for the general inclination of a bedform or not.

The diversity in geometric definitions is also reflected in the identification tools considered in this study. Specifically, Lefebvre et al. (2022) and Zomer et al. (2022) calculate dune height as the average vertical depth between the defining crest and the adjacent troughs (H_3 in Figure 1), whereas Wang et al. (2020) and Scheiber et al. (2021a) opted for an orthogonal distance between the crest and a baseline between the troughs (H_2 in Figure 1). Cisneros et al. (2020), in contrast, use the height difference between crest and downstream trough to define dune height. In addition to these height definitions, **Error! Reference source not found.** incorporates the respective lengths and gives information about the corresponding descriptors. To understand the implications of this methodological difference, we conducted an independent sensitivity study before the actual meta-analysis (Scheiber & Lefebvre, 2023).

Table 3. List of candidate dune identification tools and their geometric definition of dune height and length. Also note that not only two characteristics are inconsistent among the participants, but even the corresponding descriptors.

Publication	Height definition	Length definition	Descriptor
Cisneros et al. (2020)	Downstream vertical depth	Horizontal distance	Latin
Wang et al. (2020)	Inclined distance	Inclined distance	Greek
Scheiber et al. (2021a)	Inclined distance	Inclined distance	Greek
Lefebvre et al.	Average vertical	Horizontal distance	Latin

(2022)	depth		
Zomer et al. (2022)	Average vertical depth	Horizontal distance	Greek

Besides dune height and length, another typically assessed characteristic is the aspect ratio H/L, which can be used to describe the general steepness of a bedform. The longitudinal section in **Figure *1***, however, makes clear why this measure should not be mistaken for an average or even maximum dune side slope. The inclination of stoss (upstream) and lee sides (downstream length) is also dependent on their length ratio; measures for the corresponding relationship are known as dune asymmetry. Although the definitions of dune asymmetry differ as well, they always relate to the relational positions of crests and troughs and are therefore neglected in the following statistical comparison of dune identification results.

2.3 Bathymetries

Most dune identification tools are developed with a particular target region in mind or their algorithms are, at least to some degree, shaped by the character of the calibration data. Because this may cause a considerable bias when comparing the different methods, we compiled a benchmarking data set, which is meant to represent a variety of typical environments for bedform analyses. This includes unidirectional river flow, reversing tidal currents and controlled flume experiments (**Table *4***). In detail, the bed elevation data used in this study stem from a field campaign in the Rio Paraná in Argentina (Parsons et al., 2005), navigational safety surveys along the Weser tidal inlet channel in Germany (Lefebvre et al., 2022), and from the River Dynamics Laboratory at Simon Fraser University in Canada (Bradley & Venditti, 2019), respectively.

Table 4. Overview of benchmarking data sets: The chosen bathymetries represent three typical environments for bedform analyses, including unidirectional river flow from the Rio Paraná, tidally-constrained conditions from the Weser and data from flume experiments at the Simon Fraser University. The table also contains the reference of the first publication and the assessed spatial resolution as well as the corresponding physical constraints.

Waterbody	**Paraná River**	**Weser Estuary**	**Laboratory Flume**
First publication	Parsons et al. (2005)	Lefebvre et al. (2022)	Bradley and Venditti (2019)
Location	Argentina	Germany	Canada
Flow conditions	Unidirectional river flow	Reversing tidal currents	Uniform laboratory flow
Spatial resolution	0.25 m	2 m	0.03 m
Median grain size	220 µm	400 µm	550 µm
Average depth	5-12 m	16-12 m	0.15 m
Average velocity	1.3 m/s	1.0 m/s	0.6 m/s

After analysing the original data sets in an initial performance test, the spatial extents of all three bathymetries were limited to a subset of 450 x 100 m for the field data and 450 x

100 cm for the flume data, respectively. In some of the following comparative statistics, the flume data was assumed to be scaled by a factor of 1:100, i.e. centimetre extents were treated as metres, to improve readability. **Figure *2*** juxtaposes the bathymetric subsets and gives an impression of how unlike the contained bedforms can be in size and shape. While the longitudinal section of the Rio Paraná shows strongly asymmetric dunes of a rounded shape (Figure 2 a), the Weser bathymetry (of the same length) is characterized by much less asymmetric or in some cases even symmetrical and sharp-crested dunes (Figure 2 b). The bedforms in the flume data set are least homogenous in shape, which may be attributed to the rather short-term hydraulic forcing (Figure 2 c and Table 4). Compound dunes consisting of large-scale primary dunes and multiple, superimposed secondary dunes can be observed in the Parana and flume data. Based on visual inspection, the Weser bathymetry does not feature a secondary dune scale – possibly due to the lower data resolution. A summary of physical constraints, in particular median sediment grain size, average depth and flow velocities, is given in Table 4.

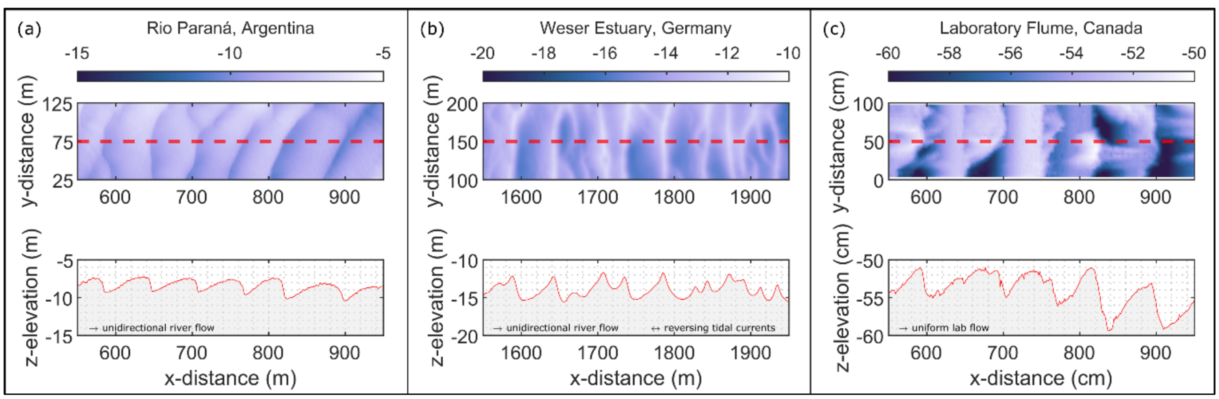

Figure 2. BATHYMETRIC DATA - Surface elevation plot (top) and longitudinal section along the red dashed line in the middle (bottom) of the three benchmarking data sets. These bathymetries comprise (a) round-shaped river dunes from the Rio Paraná. Argentina, (b) steep and tidally-constrained dunes from the Weser Estuary, Germany, and (c) scaled bedforms from the River Dynamics Laboratory at Simon Fraser University, Canada. The vertical exaggeration for all longitudinal sections is x:z = 1:10.

2.4 Comparative statistics

Given that all three bathymetries were assessed with five independent dune identification algorithms, we yield a total of 15 results data sets. These include the location of the crests and troughs that define the identified dunes as well as their heights and lengths. All other characteristics are based on and can be derived from these parameters. Moreover, combinations of height and length are assumed to be specific enough (if saved with sufficient accuracy) to retrace individual bedforms across the data sets. For this reason, the statistical comparison listed as the first objective of this study focuses on the probability distributions of dune heights and lengths, which is accomplished in both one- and two-dimensional statistics (distribution of heights or lengths vs. distribution of height/length pairs). The differences between these distributions are furthermore quantified by applying two statistical measures: the Wasserstein metric (WS) and the Jensen–Shannon divergence (JSD). The WS is a well-established way to measure the (dis-)similarity between different probability distributions. Also known as “earth mover's distance", it combines the distance and volume under two probability curves (by analogy, two piles of earth) into one "effort" function valid if one is transformed into the other. The resulting "minimum effort" is the Wasserstein distance between the two probability functions (Hitchcock, 1941). The JSD, in contrast, draws on the concept of relative entropies to express how well a probabilistic function describes a target function (Briët & Harremoës, 2009). Other than the WS distance, the JS divergence can yield values between 1 and 0 with smaller values indicating a higher resemblance between two functions. For an overview of these and other available measures to describe probabilistic differences, interested readers may be referred to the summary of Liu and Xiao (2022). The combination of WS and JSD is a useful means to quantify differences in a direct comparison of individual results data sets and thus understand which identification tools perform in a similar manner.

3 Results

This meta-analysis elucidates the heterogeneity of dune identification outputs from several perspectives. After a short evaluation of the effects of geometrical definitions, we compare the general performance of the considered tools in terms of the number of identified bedforms and corresponding computation times. In the second step, statistical variations regarding the frequency of heights and lengths are discussed. After that, we directly compare the results of all five algorithms by juxtaposing their height/length distributions in systematic difference plots and by the corresponding WS and JSD values. By harmonizing the number of outputs through resampling, we are finally able to summarize all results into one synthesis data set. The benefits of this summary data and its potential for future studies are discussed afterwards.

3.1 Sensitivity to height and length definitions

Before the actual comparison of dune identification results, we first assessed how sensitive the statistical characteristics of a given dune field are with regard to the different ways to calculate height and length (Scheiber and Lefebvre. 2023). In this sensitivity study, we assessed an independent data set and compared the corresponding results for the most common geometric definitions. In particular, we compared vertical with averaged dune heights and horizontal with inclined dune lengths, respectively. The two histograms in Figure 3 illustrate the divergence of these options in terms of the relative (percentage) difference. The three shades of blue color represent different dune scales as proposed by Ashley (1990). According to this study case, dune lengths are hardly impacted by the differentiation between horizontal and inclined distances between troughs (cf. Figure 3 a). More than 9 out of 10 dunes show a relative length difference below 1.1% and virtually no length results differ by more than 10%. This is also reflected in the mean values which differ only by 3 cm.

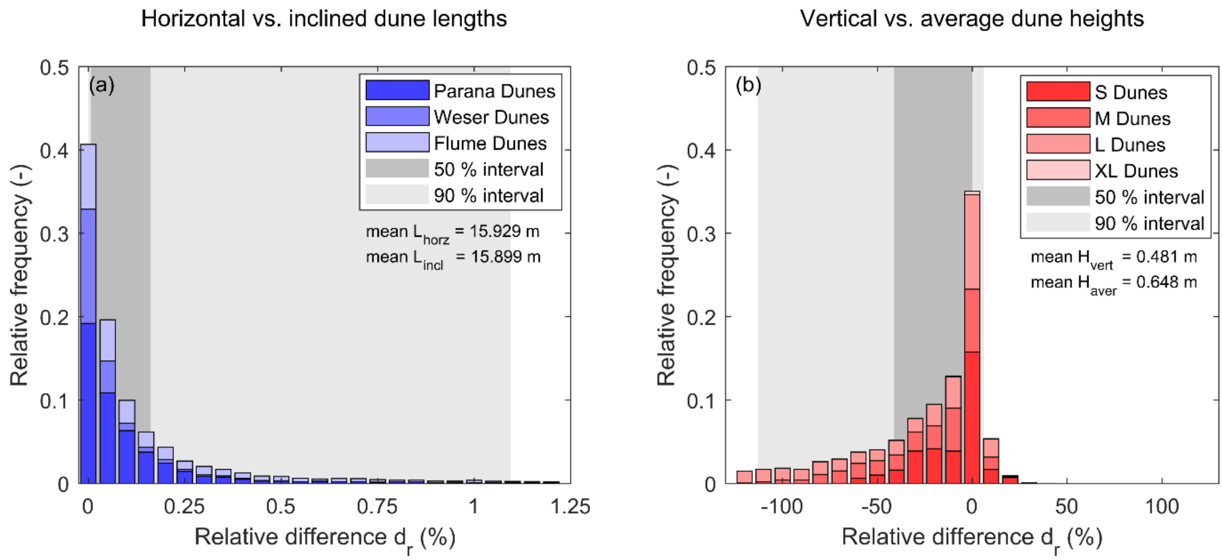

Figure 3. DEFINITIONAL SENSITIVITY – Relative (percentage) difference of vertical versus average dune height (left) and horizontal versus inclined dune length (right), respectively. Three shades of blue color represent dune scales, from small over medium to large dunes according to the nomenclature by Ashley (1990); light and dark grey patches in the background represent 50 and 90% intervals, respectively.

Dune heights, in contrast, show a much larger variation depending on whether they are calculated as the vertical distance between crest and trough baseline or as the average crest-trough distance. For instance, the 90% interval spans from –112.9% to +6.18% and the 50% interval from -41.2% to +6.2%, respectively. The difference in overall mean heights amounts to 16.7 cm, which is in the order of magnitude of a small dune. To avoid any bias resulting from the different geometric definitions, we standardized all dune characteristics. That is, all lengths were (re-)calculated using horizontal lengths (consistent with L_1) and dune heights using average heights (consistent with H_3), respectively, before proceeding with the comparative statistics.

3.2 General performance

The first evident difference in the analytical outputs is the absolute number of identified bedforms (Figure 4). It should be noted that this initial performance test included bathymetries larger than the presented extent. While the smallest amount of bedforms in the Paraná data is found by Lefebvre, it is Wang who found the least bedforms in the Weser and Lefebvre again in the flume data set. In contrast, the highest number of bedforms is found by Scheiber in all cases. Taking the results of all three study sites into account, the number of bedforms identified by this algorithm is about 25-times higher than the amount from Lefebvre with this ratio being even higher for individual data sets. This finding already points to the different focuses inherent to the compared methods, in particular the consideration of different dune scales. Regarding their computational effort, the methods also vary considerably (Figure 4 b). For most data sets, the calculations by Scheiber show the shortest relative computation times, i.e. average time needed to identify and measure a single bedform. On the other hand, Wang required the longest computation times per bedform indicating a more sophisticated workflow. Notwithstanding that each analysis was carried out using different hardware setups and that, therefore, we cannot rule out an influence of computational capacities, the individual methods doubtlessly rely on processing steps of different number and complexity. These methodological differences, which are mainly determined by the scientific focus, are also reflected in the statistical variation of dune characteristics.

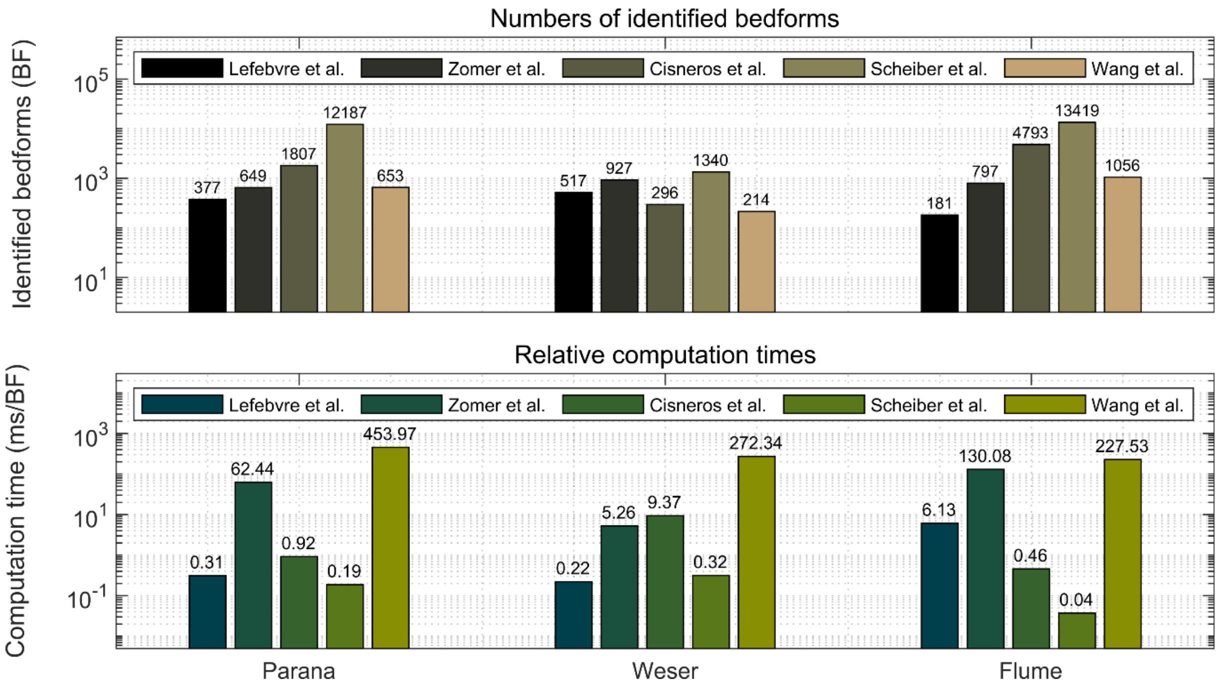

Figure 4. GENERAL PERFORMANCE - a) Absolute numbers of identified bedforms for each of the three bathymetries. b) Relative computation times in milliseconds/bedform indicate the varying computational complexity of the five identification tools under consideration.

3.3 Statistical variation

Although applied to the same bathymetric data sets, the five dune identification tools produced significantly different statistical results. Besides the sampling sizes, the two most essential dune characteristics, dune heights and lengths, varied in their distribution. This can be perceived from Figure 5, which displays the statistical variation of height and length results from all three bathymetries as a combination of box plots and violins. In the box plots, black horizontal lines represent median values and lower and upper box edges refer to 25/75th percentiles, respectively. When contrasting the median heights (H50) and lengths (L50) of the methods, we can observe two groups regarding the results for Rio Paraná (Figure 5; left panel). While the median heights of Lefebvre, Zomer and Wang are in the order of 1.5 m and the corresponding L50 is 58 m, the values from Cisneros and Scheiber are H50≈0.2 m and L50≈7 m, respectively. What is interesting is that results from the first three methods are mainly limited to dune lengths greater than 30 m, but the latter two methods did identify both small and large bedforms. Similarly, 50% height intervals in the case of the flume bathymetry (Figure 5; right panel) range from 2.5 m to 5.5 m for Lefebvre, Zomer and Wang, but from 0.5 m to 1.5 m for Cisneros and Scheiber – a strong indication that different dune scales were considered. Only for the Weser bathymetry, where no compound dunes were visible, box plot ranges are a bit more homogeneous except that Cisneros reports relatively smaller but longer bedforms.

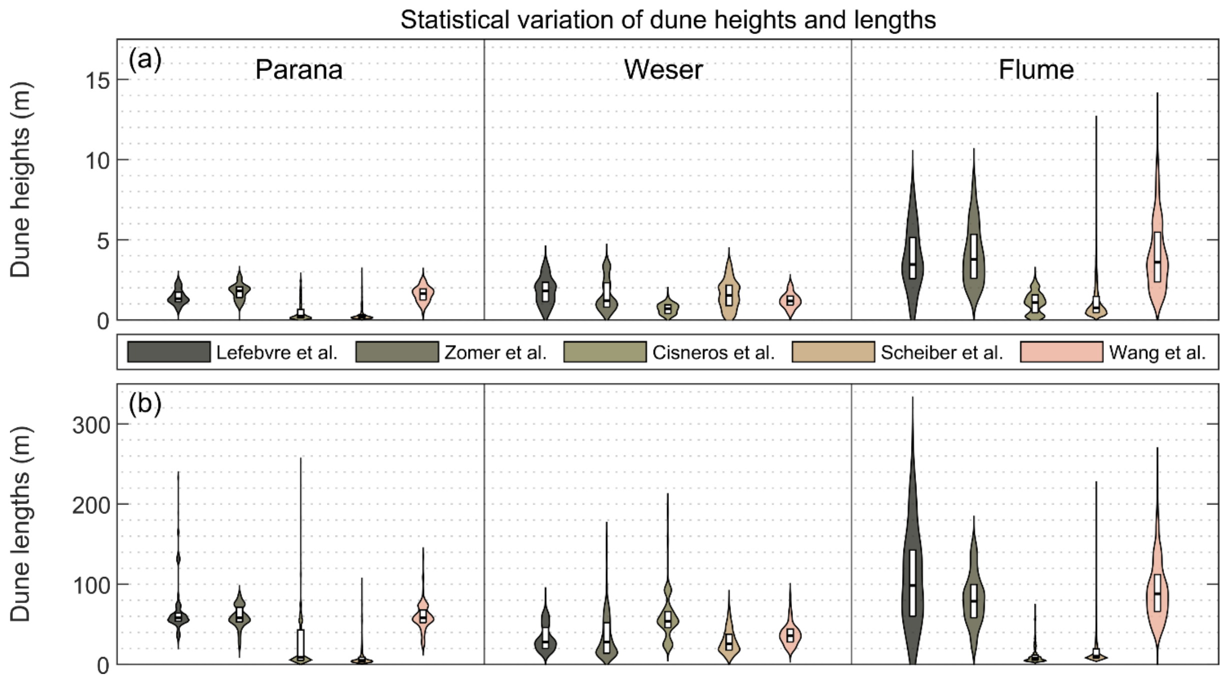

Figure 5. STATISTICAL DISTRIBUTIONS - Variation of dune heights (a) and dune lengths (b) displayed as a combination of box plots and violins. The 25/75th percentiles are shown as box edges and median values as black horizontal levels in between. The shape of the surrounding violins, stretching between minimum and maximum values, represents the probability density of identified dune characteristics. Please note that results for the flume bathymetry were scaled by a factor of 100 for better readability here. The visualization is based on the "daviolinplot" function provided by Karvelis (2023).

The corresponding violin shapes represent the continuous probability density of identified bedforms and thus extend to the most extreme values. Based on these extents and assuming that no algorithm produced a considerable number of artefacts, we can conclude that two dune scales are present in the Paraná (Figure 5; left panel). The algorithms differ in so far as Lefebvre, Zomer and Wang focus on primary dunes, whereas Cisneros and Scheiber additionally considered a significant number of secondary dunes in their assessment, whose abundance results in spinner-shaped violins with a thin spike at high and a distinct bulge at low values, respectively. In this connection, it should be noted that Zomer deliberately excluded secondary bedforms in the Parana (and the flume) dune fields considering the data resolution not sufficient to resolve secondary bedforms. The dune heights reported by Lefebvre, Zomer and Wang show a distinct peak between 1 m and 2 m, whereas the distributions by Cisneros and Scheiber peak below 0.25 m. The lower diagram, in contrast, suggests that most of these (secondary) dunes have a length below 10 m, with a considerable number of primary dunes included in a long upper tail. In the results by Lefebvre, Zomer and Wang, lengths peak between 50 m and 80 m, but all three distributions show a positive, i.e. leftward skew. A similar grouping can be reported for the flume data (Figure 5; right panel) with the exception that Cisneros seems to identify only small and medium dunes here. These are constrained to about 3 m in height and 80 m in length, while the remaining four algorithms report significantly higher values. In particular, peaks in the range of 2-7 m and 50-150 m pertain to larger primary dunes, whereas smaller secondary dunes can be expected below 1 m in height and 20 m in length, respectively. It is interesting to see here that only the results by Scheiber include both primary and secondary bedforms, whereas Cisneros only reports the smaller and Lefebvre, Zomer and Wang focus on the larger bedform scale. It should be noted again that all dimensions in the flume data are actually in centimetre scale and were only converted to allow comparability. Regarding the Weser bathymetry (Figure 5; middle panel), the violins paint again a more homogeneous picture than the other dune fields with generally wider distributions peaking between 0.9-2.1 m and 15-55 m, respectively. Most of these distributions show a left skew as well. However, the results by Zomer are in this case more comparable to the ones by Scheiber, because they suggest relatively smaller geometries than Lefebvre and Wang. Only the length results by Cisneros include three peaks which rather suggests statistical gaps than multiple dune scales. All in all, the depicted relative frequencies corroborate the notion of major heterogeneities in the identification results, yet also provide evidence about their causes. While the methods by Lefebvre and Wang clearly focus on the identification of primary dunes, Cisneros, Scheiber and partially Zomer allow for the co-existence of primary and secondary dunes. These two dune scales appear nearly evenly distributed in the data of Zomer, whereas Cisneros and Scheiber found significantly more secondary dunes. Although this juxtaposition helps to understand the inherent research focus that influenced the development of individual identification tools, the illustration is mainly limited to a visual comparison.

To assess the quantitative differences between individual dune identification results, a direct and comprehensive comparison is needed. This can be achieved by combining height and length results for all three bathymetries into one two-dimensional probability density plot for each identification method. Following the double-logarithmic scatter diagram by Flemming (1988), the resulting probability functions relate to surfaces constructed by merging the height and length distributions. Along the main diagonal of **Figure *6***, the shapes or, more precisely, the peaks of these functions are depicted in the form of contour lines. In the second step, this presentation allows us to directly compare pairs of probability distributions through difference plots below the main diagonal. The recurring red-to-blue colormaps in this visualization express the residual surface when subtracting the probability function to the right from the one at the top as written in the respective panel titles. It therefore highlights any disagreement between the compared probabilities by bright red (more abundant at the top) and bright blue (more abundant at the right distributions) colour. Above the main diagonal, in turn, this agreement/disagreement is quantified by means of the Jensen-Shannon divergence (JSD) and the Wasserstein metric (WS), where smaller values correspond with higher agreement and vice versa. Both metrics are referred to dune height and length individually in order to ensure an independent interpretation of these two dimensions.

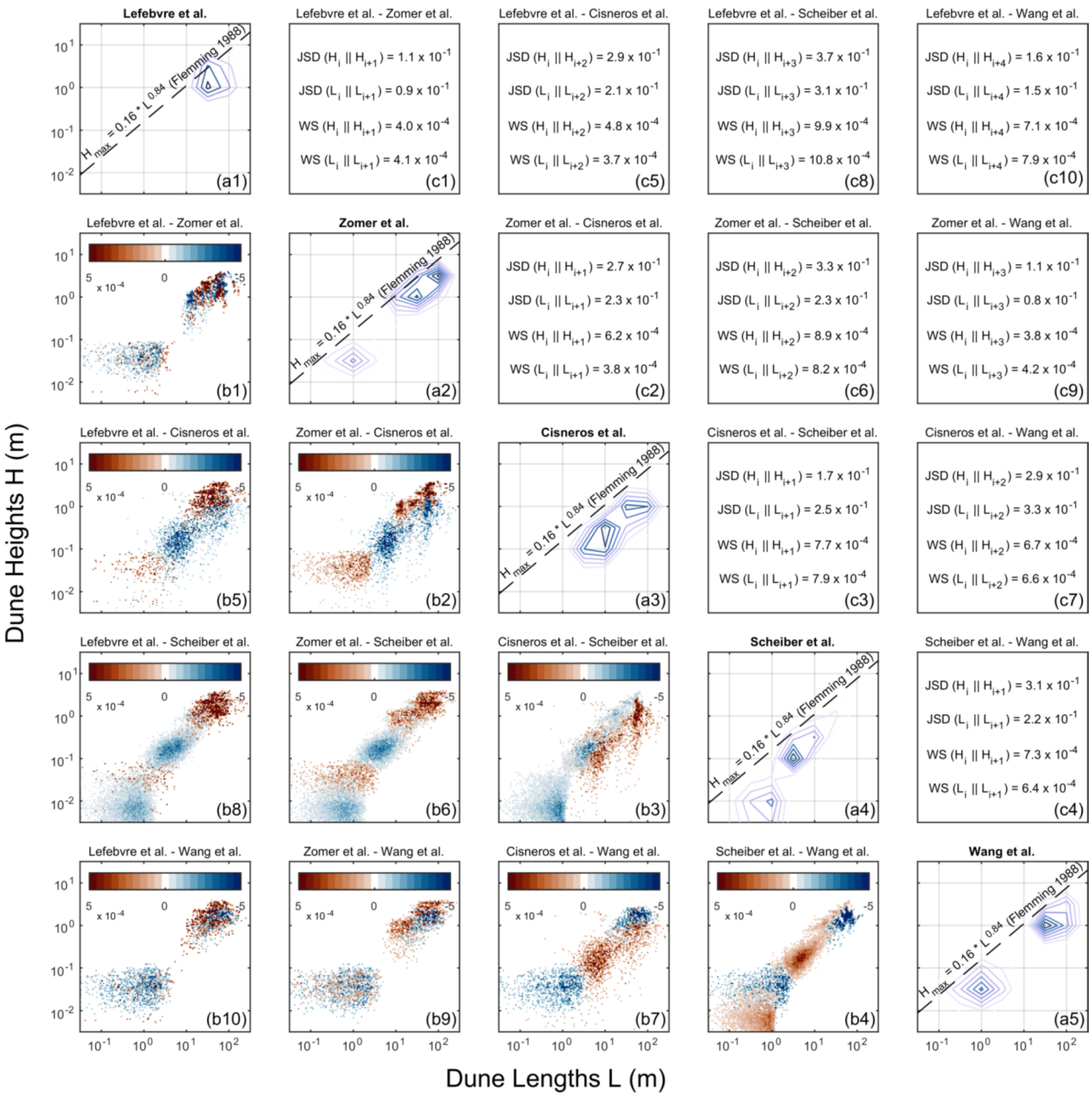

Figure 6. DIRECT COMPARISONS – The main diagonal (a1-a5) gives a double-logarithmic presentation of dune height/ length pairs for each of the five assessed identification methods. The depicted contour lines describe the shape of the corresponding probability density functions. Below that main diagonal, red-to-blue colored plots (b1-b10) illustrate the differences when subtracting the results to the right from the ones on top of the respective subplot. Above the main diagonal, analogous comparisons (above vs. left method) are quantified by two statistical metrics (c1-c10): the Jensen-Shannon divergence (JSD) and the Wasserstein metric (WS). Both metrics refer to dune heights and lengths independently. Unlike in the previous two illustrations, flume results are no longer scaled here.

In general, all previously reported dune populations of the individual bathymetries can be retraced in this combined data set. According to the contour plots (a1-a5) aligned along the main diagonal of **Figure *6***, the smallest scale of dunes (L≈1 m, H≈1 cm) is reported by Scheiber, directly followed by the lower of two peaks in the results by Zomer and by Wang (L≈1 m, H≈3 cm). Medium dunes (3 m<L<10 m, 10 cm<H<30 cm), in turn, are visible in the results by Cisneros and Scheiber. Finally, large to very large dunes (20 m<L<200 m, 50 cm<H<200 cm) cause a distinct peak in all distributions but the one from Scheiber. Moreover, this visualization indicates which size of dunes was most abundant according to the individual identification method. However, it should be noted that this illustration highlights frequency peaks rather than allowing for the complete spectrum of results. When directly comparing these distributions based on the difference plots below the main diagonal (b1-b10), the best congruence can be observed for Zomer versus Wang, directly followed by Lefebvre versus Zomer and by Lefebvre versus Wang. Apparently, these three algorithms produce very similar results, which is also reflected in the corresponding metrics above the main diagonal (c1-c10). In this regard, the JSD gives a good indication of the similarity between the distribution shapes, whereas the WS also accounts for the volume below the probability functions. Both metrics are in the same order of magnitude for the three methods. The highest values apply to Lefebvre versus Scheiber whose results consequently appear least congruent. Overall, it becomes evident that the five identification methods mainly differ depending on their direct (scale-separation) or indirect (detection limit) consideration of small-scale secondary dunes. The analyses focussing on large-scale primary dunes (either determined by the algorithm itself or by user-defined settings), i.e. those by Lefebvre, Zomer and Wang, generally identify bedforms on a comparable co-domain. To some degree, the same holds true for the comparison of Cisneros versus Scheiber who both included a high number of secondary dunes, with the latter reporting about 10-times more bedforms than the former. The presented analyses suggest that it makes a crucial difference, in which way secondary bedforms are considered, because their occurrence may significantly affect statistical outcomes. Given that half of the assessed algorithms gave special attention to the smaller dune scale, while the other focused on primary dunes, a synthesis of identification results can give robust insights into and allow a more profound interpretation of the overall composition of natural bedforms as contained in a combined data set.

3.4 Synthesis of results

The bathymetries of Rio Paraná, Weser and Simon Fraser University flume represent a diverse data set with regard to dune scales, physical forcing and measurement resolution. The independent but concerted assessment of such a comprehensive data set, however, is an unprecedented endeavour. Their combination provides the opportunity for a statistical analysis of bedform composition, which is significantly less biased by the proficiency and expectations of single authors than conventional investigations. To this end, and to some extent as a mere by-product of the presented meta-analysis, **Figure *7*** contains the data from all three flow environments and all five methods as transparent probability contours. Building on the diagram of dune height and length pairs by Flemming (1988), panel (a) synthesizes this data into one combined probability distribution, visualized by a 99% contour with a solid black outline. Due to the diversity of the assessed bathymetries, this contour spans from approximately 5 cm to 100 m in length and 0.2 cm to 5 m in height. This is also visible in the two adjacent panels **Figure *7*** (b) and (c), which have to be seen as the lateral projections of the two-dimensional probability density function from (a).

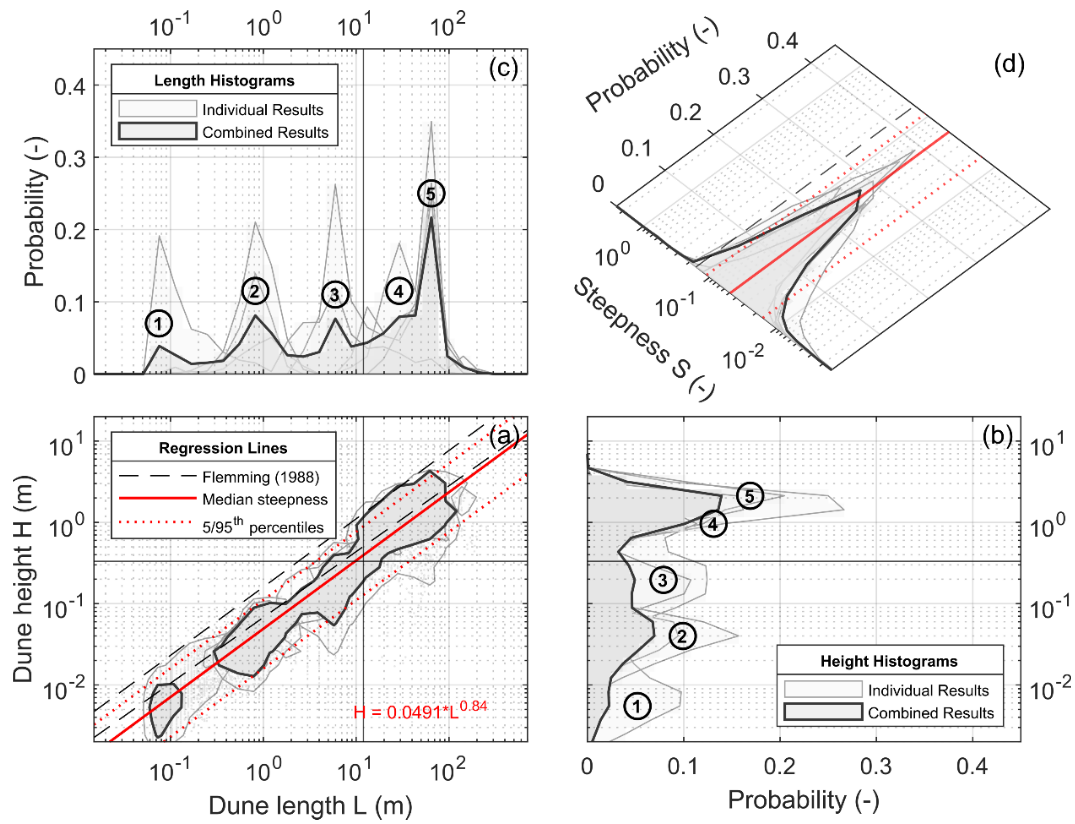

Figure 7. COMBINED RESULTS - All dune height/length pairs identified for all three study sites and all five methods as well as their combined distribution. Subplot (a) follows the double-logarithmic scatter diagram by Flemming (1988) and is complemented by filled contours. The prominent area with a black outline represents the synthesis of all individual results. This synthesis data is further described by a power law included in red. In addition, parallel dotted lines represent the 5/95th percentiles of this steepness relationship. Subplot (b) is the lateral projection of the two-dimensional probability function in (a) and thus visualizes the distribution of dune heights with five peaks Analogously, subplot (c) shows the distribution of dune lengths, which includes five peaks as well. Finally, subplot (d) provides the distribution of dune steepness according to the red power function parallel to the historic H_{max}. It also features the 5/95th percentiles, which frame this unimodal distribution.

Like in the top view, both graphs contain the individual results as transparent shapes and their synthesis with a black outline in the front. Median values are given as fine horizontal and vertical lines, respectively. The inspection of these height and length distributions reveals the peaks of four distinct (and one concealed) bedform populations contained in this combined data set. From small to large, these peaks are (1) secondary and (2) primary dunes in the flume, (3) secondary dunes in the Paraná, (4) non-compound dunes in the Weser amalgamating with (5) primary dunes in the Paraná, respectively. It is interesting to see that the combination of these co-domains spans the full spectrum between the smallest ripples and very large dunes as defined by Ashley (1990). What is more, they basically cover the co-domain of bedforms evaluated by Flemming (1988) and their upper boundary is well confined by the maximum regression line (upper dashed line) suggested in this classical reference ($H_{max} = 0.16 \cdot L^{0.84}$ according to

Flemming (1988)). Acknowledging that this function describes a universal relationship for the natural limitation of dune steepness, we suggest adding a new perspective on bedform geometries by evaluating this steepness in panel (d). This steepness is not the same as the aspect ratio H / L but follows the relationship $S = H / L^{0.84}$ included in panel (a). Similar to the projections of dune height (b) and length (c), this rotated axis shows the projected distribution of dune steepness. This steepness appears to be nearly symmetric and the median S for the combined data set amounts to 0.0491. It is worth noting that the classic aspect ratio H/L is of comparable size in this case amounting to 0.0416 on average. Transferring this value back to the top view of height/length pairs (panel a: solid red line), we can see that in this case, the global mean regression line by Flemming (1988) lies very close but slightly higher, suggesting the presence of dunes which are not fully developed. The negligible skewness, only discernible by the $5/95^{th}$ percentiles in the steepness distribution (panel d: dotted red lines), points in the same direction. Unlike dune heights and lengths, whose statistical description is necessarily distorted as a result of the different dune scales, the distribution in this top right panel appears almost normally distributed, which gives strong evidence about the natural development of dunes that grow until a universal equilibrium steepness is reached.

4 Discussion

The comparison of five exemplary dune identification tools allowed us to estimate the range of their sampling sizes, computation times and bedform characteristics – in other words, the heterogeneity of expectable outputs. After their quantitative description in the previous section, these findings will be interpreted and contextualized in the following. Subsequently, methodological differences, resulting from the distinct focus of each identification tool, as well as individual strengths and limitations will be summarized in order to indicate fields of application. At the end, opportunities for a transfer of the findings from this meta-analysis are discussed and some open research questions are raised.

4.1 Interpretation and deductions

Although best efforts have been made to standardize the inputs and outputs for this meta-analysis, it is obvious that significant differences exist not only in terms of the individual methodology for dune identification but also regarding the assumptions that define our understanding of these bedforms. First of all, the existing identification methods are inconsistent regarding the geometric definitions of even the most basic characteristics, such as dune heights and lengths. The arising differences are neglectable in the case of horizontal vs. inclined lengths due to sufficiently small inclination angles, but definitely considerable when comparing vertical with average heights. In this case, and especially if asymmetric inclined dunes are measured, determined heights can nearly double. Detailed investigations regarding this influence of geometric definitions were conducted by Scheiber and Lefebvre (2023) and the reported sensitivity should be taken into account in any bedform-related analysis. However, given that all of the implemented definitions are based on clear reasoning, we can only emphasize that definitions should always be chosen in due consideration of the specific research question and, even more importantly, that such decisions are documented and discussed openly. This was also considered in the presented meta-analysis, where we standardized all individual identification results before comparing height and length data.

In the initial performance test, the sheer number of individually identified bedforms and even more so the relative computation times differed significantly enough to require a logarithmic scale for visualization. This epitomizes the heterogeneity that was present throughout our analyses. One essential explanation for this finding is the different consideration of compound dunes, i.e. bedforms that consist of multiple scales of dunes. It should be noted that small-scale secondary bedforms populating the stoss-side (upstream) slope of larger primary dunes are naturally more abundant than their host, simply because they require less space. As a consequence, researchers who include these secondary bedforms in their algorithm will detect significantly more dunes. This, in turn, will lead to completely different frequency distributions and corresponding statistics, which puts into question the validity of a direct comparison between these two perspectives and quantitative bedform statistics in general. Given that three out of five algorithms in our meta-analysis include a separation of scales, while two algorithms do not, it was not possible to allow for this issue in a balanced manner.

In the presented results, two groups formed depending on the prevalence of secondary dunes. Specifically, the results from Lefebvre, Zomer and Wang as well as the ones from Scheiber and Cisneros were strongly connected in the case of compound dunes. Accordingly, the results for the non-compound Weser dunes appeared most homogeneous. But even for the other two bathymetries, the relative differences of median values within the two groups did not exceed 25%, which is less than could be expected from preliminary tests. However, some uncertainty arises from the model scale of the flume experiments, because individual methods are trained for specific aspect ratios which are far from natural here. An elegant way of addressing the co-existence of primary and secondary dunes is implemented in the algorithms by Zomer and by Cisneros, who both use separate output variables for the two dune scales. Nevertheless, differences remain with regard to the lower dune identification limit, i.e. the height and length of the smallest identifiable dune. The definition of this threshold is not only dependent on the research focus but also on data resolution and general physics. In signal processing, the smallest identifiable wavelength is determined by (the inverse of) the so-called Nyquist frequency, which is ½ of the sampling rate. In bedform studies, however, this theoretical threshold collides with the practical accuracy of bathymetric data typically obtained via echo-sounding devices from a floating vessel. In order to avoid mistaking artefactual noise for actual ripples, we suggest using a threshold of at least 5-times the available horizontal resolution. For published algorithms, easy-to-change options should be provided which facilitate a manual adjustment of this threshold if different standards are required. It should also be noted that, for the present case, different thresholds were applied, which certainly had an impact on the presented statistics and their comparison.

Moreover, some identification results can be neglected based on our physical understanding of bedforms. This refers to bedforms which are either extraordinarily steep or very flat. Both cases are physically questionable and, in this connection, a systematic limiting of frequency distributions may help alleviate this problem. For instance, the steepness distribution introduced in **Figure *7*** shows that in our study the 99% interval of the combined results data coincides fairly well with the maximum regression line established by Flemming (1988). The parallel steepness curve, in turn, is slightly below the historic mean value regression, which is a repeatedly reported finding in similar comparisons (e.g. Bartholdy et al., 2002; Lisimenka & Kubicki, 2017). Considering that the minor skewness visible in panel (d) does not hold for the Weser data set, we can deduce that dunes in this environment are closest to a maximum or equilibrium steepness suggesting the steadiest flow conditions. Moreover, the 5th percentile line is of great help in identifying unreasonably flat bedforms, which cannot be determined from height or length results alone. We therefore recommend considering dune steepness according to the above definition ($S = H / L^{0.84}$) as a third key characteristic of bedform analyses, which can be useful when filtering preliminary identification results.

4.2 Method distinction/clarifications

In the pursuit of an automated and mostly objective description of bedforms from a given bathymetric map, we can differentiate five general processing steps: i) the determination of a dominant crest orientation, ii) the optional separation of different dune scales, iii) the identification of individual bedforms based on crests and troughs, iv) the calculation of corresponding geometric characteristics, such as heights and lengths, and v) the delineation of dune objects in all three dimensions. Although not all of these steps are required in every study and others can be conducted manually, they are still instrumental building blocks that can be found in the five methods compared in this meta-analysis. In the bid to outline guidelines for the correct utilization of these methods, the following paragraphs summarize their distinct working principles and objectives thus narrowing down potential fields of application.

The identification tool presented by Lefebvre et al. (2021) focuses on the identification in 2D (ii) and 3D (v) and the measurement of bedforms (iv). It was originally developed for bathymetric maps of a fairway channel in the Weser Estuary. No scale separation was needed because of a relatively coarse resolution of 2 m which prevents the recognition of small-scale bedforms. Because of the constrained environment, it is assumed that the main flow direction follows the channel, and the main crest direction is perpendicular to the main flow direction. The crestlines were detected as objects with a low curvature and the trough lines as minimum elevation between crestlines. That way, the crest and trough lines can be analysed (direction, variability, etc.). The method was developed to produce fast and not overly accurate results because a very large dataset had to be analysed. The method is particularly adapted to very large datasets with relatively low resolution. Furthermore, it is likely that the minimum curvature method is most accurate over sharp crests (such as those in estuaries) and might be less accurate for rounded crests (usually developing in unidirectional flows).

The bedform separation and identification tool presented by Zomer et al. (2022) was developed to quantify the properties of large primary and smaller superimposed bedforms contained in bed elevation data from the Dutch Waal river. The separation of dune scales (ii) is performed by decomposition of the data using a LOESS algorithm. Steep lee slopes of primary dunes are preserved by implementing breaks in the LOESS fit and the steep slopes are subsequently approximated with a sigmoid function fit. Primary and secondary dune identification (iii) is done based on zero-crossing. Dune characteristics that can be calculated (iv) include height, length, steepness, and maximum lee-side slope angle. Results are grouped into secondary and primary bedforms and the tool also allows for filtering of the results based on user-defined conditions. The tool is appropriate for data sets with multiple (two or more) well-defined bedform scales. An important advantage of the method is that steep lee-side slopes of primary dunes are well preserved in the filtered signal. This is relevant for studies focusing on lee side slope characteristics or environments where secondary bedforms are present on the lee side of primary dunes. A user can also use either the bedform separation or identification and combine it with other methods, allowing flexible tailoring to both the data set and the purpose of analysis.

The Bedform Analysis Method for Bathymetric Information (BAMBI) was introduced by Cisneros et al. (2020) in order to analyse the slopes of dune lee sides in the world's largest rivers. In the order of processing, BAMBI comprises iii) the identification of dunes based on local extremes, iv) the calculation of geometric characteristics, ii) the separation of dune scales, and to some extent v) the delineation of dunes as 3D objects. In contrast to many other approaches, morphological measurements from BAMBI encompass the mean and maximum lee side angle (steep slope) and the relative height of the steepest lee side slope for each dune. The tool is therefore particularly well-suited for analysing processes depending on the dune shape, such as nearfield hydrodynamics. Moreover, the investigations in this study have shown that BAMBI detects significantly more secondary dunes than formative primary dunes. The resulting left-skewed height distribution can be assumed to reflect the natural inventory of morphological features, which makes the approach comparable to the one by Scheiber et al. (2021a).

The Bedform Identification Algorithm (BIA) presented by Scheiber et al. (2021a) focuses on the identification (iii) and measurement of bedforms (iv). After a manual determination of the dominant dune orientation (i), it circumvents scale separation (ii) by an iterative assessment of longitudinal bed elevation profiles. Based on the length classes suggested by Ashley (1990) this process is repeated five times ensuring that dunes of all prevailing scales are identified before deleting duplicate values. The fact that this method yields bi-modal distributions in both bathymetries with compound dunes corroborates its advantages when information about the full spectrum of bedform scales is needed. However, other than in the methods explicitly aiming for scale separation (e.g. Cisneros et al., 2020; Zomer et al., 2022), results are not grouped in primary and secondary dunes by default but can be assigned to the underlying length classes if desired. The performance test in Figure 4 illustrates that this approach is particularly efficient when it comes to computational costs. It can therefore be of good use for assessments of very large data sets. Nevertheless, the consideration of secondary bedforms has a significant impact on statistical distribution as shown in this study. It therefore requires a careful definition of the lower detection limit, which should agree with the respective research objectives and allow for the inherent measurement inaccuracies.

The automated procedure to calculate the morphological parameters of superimposed rhythmic bedforms presented by Wang et al. (2020) is probably the most sophisticated of the five compared algorithms. Combining Fourier, wavelet and zero-crossing analyses, the strength of this approach lies in its objectivity which is ensured by a complete automation of processing steps, including (i) the determination of dominant wavelengths and dune orientation, (ii) the separation of superimposed and primary dune scales and (iii) the identification of bedforms and (iv) calculation of their characteristics. As shown in a validation case, this methodology succeeds even if superimposed bedforms were oriented almost perpendicular to the primary dunes. Even though this level of processing comes at a high computational cost, as illustrated in the comparison of general performances (section 3.2), the approach requires almost no (subjective) decisions by the user due to its complete automation and therefore bares the smallest risk for human bias or misapplication.

4.3 Transferability and outlook

It is surprising how differently the involved authors approached the task of bedform identification and how the individual methods performed. Fortunately, the resulting statistics of bedform height and length did not diverge disproportionately as long as we differentiate by the (optional) consideration of secondary bedforms. The differences, which can be observed otherwise, are primarily a matter of statistical distortion. This explanation soothes the initial scepticism about heterogeneities in the available identification algorithms, which could finally be re-validated.

Beyond this study, both the utilized bathymetries and the summary of individual identification results provide a unique data set and an added value on their own. We assume that the influence of remaining biases, which may accidentally have been built into the individual identification tools, is mitigated by the combination of their outputs. The final results should hence be closer to the objective truth and are therefore perfectly fit for future benchmarking efforts. To this end, we are happy to provide the complete data set as a supplement to this article.

Future studies may also reconsider the importance of the upper regression line by Flemming (1988), which proved to be a universal maximum ratio between bedform heights and lengths. We argue that this ratio is inherent to all bedform populations and that, in fact, its parallel translation (by a factor of $S = H / L^{0.84}$) can be an indicator of the stage of dune development towards the natural maximum steepness. For the presented benchmarking data, this approach showed better goodness of fit ($R^2 = 0.60$) than the historic mean regression line ($R^2 = 0.58$). Although custom power laws can (necessarily) capture the individual data even better, the definition of steepness as a comparable measure of dune growth can be of use beyond this study. The consideration of this additional geometric characteristic, besides height and length, not only helps to identify numerical artefacts. It can also be used to uncover bedform populations subject to unsteady morpho-dynamic conditions, if the steepness distribution is particularly skewed. This allows insights into the temporal development of bedform geometries derived from a single snapshot.

What is more, all of the discussed methods work on a transect basis, that is they assess two-dimensional bed elevation profiles to calculate height and length. By evaluating multiple parallel profiles across a given bathymetry, it can be ensured that each bedform is sampled at different sections, which increases the robustness of these methods. The obtained distributions are certainly less prone to the impact of outliers, in spite of inevitable “edge effects” occurring at the transect ends (Gogolewski, 2020). However, average/representative characteristics are until now not associated with specific, three-dimensional dunes. This shortcoming is targeted by recent studies of object-based dune identification (Cassol et al., 2022; Lebrec et al., 2022). Nevertheless, the question remains which characteristics are representative of a laterally changing bedform. In our opinion, this can only be addressed by treating bedforms as the three-dimensional entities they are but associating these objects with characteristics of a quantified variation. To this end, the reporting of 50% and 90% frequency intervals and the consideration of dune steepness can be useful elements of future studies.

From the plethora of open-access dune identification tools, only five methods were compared in this study. However, our results have shown that there is a strong need to standardize the most useful approaches and centralize them, at best, in one universal and open-access toolbox. This toolbox should facilitate seamless dune identification, allowing users to choose the most suitable approach for each of the individual processing steps in a modular way. This would enable both experts and non-experts to test and utilize different methods, while ensuring a unified approach to calculating dune characteristics. Ultimately, this standardization would facilitate the creation of comparable and consolidated data sets in order to dwell less on methodological details and rather advance our understanding of the morphological processes that create and shape bedforms in all kinds of environments.

5 Conclusions

This study compared five recently published dune identification algorithms in a comprehensive meta-analysis. It was shown that the absolute number of bedforms detected by the available tools can differ by two orders of magnitude and the required computation times by four orders of magnitude, respectively. But also, the determined bedform characteristics, such as dune heights and lengths, differed significantly. Considering that even the underlying definitions of these characteristics are not identical in all tools (and resulting differences can sum up to the height of a small dune), an initial standardization was imperative. The subsequently determined statistical distributions for three benchmarking data sets from diverse flow environments revealed two groups among the considered approaches. Within these groups, the relative difference in median heights and lengths did not exceed 25%, but between the groups, statistics looked much more heterogeneous. The observed differences in bedform characteristics mainly originate from the unlike consideration of secondary dunes which are superimposed on larger primary dunes. Depending on whether this secondary (and naturally more abundant) dune scale is included in the identification process or not, statistical distributions tend to show a strong left skew. This general difference in dune identification affected all parts of the results and, consequently, was also visible when directly comparing the methods with each other. Based on two statistical metrics, the Jensen-Shannon divergence and the Wasserstein metric, we could show that a high (quantitative) resemblance is given between algorithms with the same perspective on secondary dunes. However, if secondary dunes are taken into account, it is essential to distinguish these from random measuring inaccuracies. In this respect, we recommend using a minimum detectable dune length of 5-times the horizontal measuring resolution. Apart from a mere quantification of differences, the concerted analysis of the three dune fields and subsequent synthesis of results generated a unique benchmarking data set. This by-product of our meta-analysis is inherently less subjected to the bias of individual focus and can therefore be of good use to any future identification algorithm. In addition, the distribution of these combined results aligns very well with the maximum regression line proposed by Flemming (1988), albeit at slightly smaller mean values. We acknowledge $H_{max} = S * L^{0.84}$ as a universal relationship between dune height and length but suggest including the distribution of steepness S as an additional proxy to describe dune growth in future bedform studies. In summary, the presented meta-analysis was able to provide insights into the performance of recent dune identification tools and quantify the heterogeneity of their outputs as well as clarify individual strengths and limitations that determine optimum fields of application. To support end-users in their analyses, we see the need for a universal toolbox which centralizes different approaches in one interface and allows experts and non-experts to detect and characterize bedforms in a unified yet modular way

Acknowledgements

First of all, we are indebted to those persons providing access to the bathymetric data we used in this study, namely Daniel R. Parsons for the Parana, Sven Dunker and the Waterways and Shipping Authority Weser-Jade-Nordsee for the Weser as well as Ryan W. Bradley and Jeremy G. Venditti for the flume data sets. Moreover, we want to express our gratitude towards Stefan

Heidekrüger for his patient guidance with regard to the comparison of frequency distributions and Arnd J. Obert for his technical support considering the Jensen-Shannon divergence. And, finally, we wish to thank Bishal Khatri for his dedicated help as a student assistant at Leibniz University Hannover.

Data Availability

The three bathymetric data sets used in this study are available according to the original publications cited in the text. The five dune identification algorithms that were applied to analyse the bedforms in these bathymetries can be found in the following GitHub repositories:

- Lefebvre (https://github.com/DrAliceLefebvre/Weser_Bedform_Analysis_Codes),
- Zomer (https://github.com/j-zomer/BedformSeparation_Identification),
- Cisneros (https://github.com/juliamorphology/bambi_0.1.0),
- Scheiber (https://github.com/LeonSchei/BIA-BedformIdentificationAlgorithm),
- Wang (on personal request).

The morphometric results from these individual approaches as well as the combined results data set can be requested from the corresponding author.

References

Amsler, M. L., & Garcia, M. H. (1997). Sand-dune geometry of large rivers during floods: Discussion. *Journal of Hydraulic Engineering-ASCE*, *123*(6), 582–585.

Armstrong, C., Howe, J. A., Dale, A., & Allen, C. (2021). Bathymetric observations of an extreme tidal flow: Approaches to the Gulf of Corryvreckan, western Scotland, UK. *Continental Shelf Research*, *217*, 104347. https://doi.org/10.1016/j.csr.2021.104347

Ashley, G. M. (1990). Classification of large-scale subaqueous bedforms; a new look at an old problem. *Journal of Sedimentary Research*, *60*(1), 160–172. https://doi.org/10.2110/jsr.60.160

Bartholdy, J [J.], Bartholomae, A., & Flemming, B. W [B. W.] (2002). Grain-size control of large compound flow-transverse bedforms in a tidal inlet of the Danish Wadden Sea. *Marine Geology*, *188*(3-4), 391–413. https://doi.org/10.1016/s0025-3227(02)00419-x

Best, J [Jim] (2005). The fluid dynamics of river dunes: A review and some future research directions. *Journal of Geophysical Research: Earth Surface*, *110*(F4), n/a-n/a. https://doi.org/10.1029/2004JF000218

Bradley, R. W., & Venditti, J. G. (2019). Transport Scaling of Dune Dimensions in Shallow Flows. *Journal of Geophysical Research: Earth Surface*, *124*(2), 526–547. https://doi.org/10.1029/2018JF004832

Bradley, R. W., & Venditti, J. G. (2021). Mechanisms of Dune Growth and Decay in Rivers. *Geophysical Research Letters*, *48*(20). https://doi.org/10.1029/2021GL094572

Breed, C. S., Grolier, M. J., & McCauley, J. F. (1979). Morphology and distribution of common 'sand' dunes on Mars: Comparison with the Earth. *Journal of Geophysical Research*, *84*(B14), 8183–8204. https://doi.org/10.1029/JB084iB14p08183

Briët, J., & Harremoës, P. (2009). Properties of classical and quantum Jensen-Shannon divergence. *Physical Review a*, *79*(5), 52311. https://doi.org/10.1103/PhysRevA.79.052311

Bruschi, R., Bughi, S., & Drago, M. (2014). *In-place stability and integrity of offshore pipelines crossing or resting on active bedforms or loose or soft soils* (3rd ed., Vol. 13). https://www.researchgate.net/profile/michele-drago/publication/292928692_ln-place_stability_and_integrity_of_offshore_pipelines_crossing_or_resting_on_active_bedforms_or_loose_or_soft_soils/links/56f54e3408ae95e8b6d1d4f4/ln-place-stability-and-integrity-of-offshore-pipelines-crossing-or-resting-on-active-bedforms-or-loose-or-soft-soils.pdf#page=25

Cassol, W. N., Daniel, S., & Guilbert, É. (2021). A Segmentation Approach to Identify Underwater Dunes from Digital Bathymetric Models. *Geosciences*, *11*(9), 361. https://doi.org/10.3390/geosciences11090361

Cassol, W. N., Daniel, S., & Guilbert, É. (2022). An Approach for the Automatic Characterization of Underwater Dunes in Fluviomarine Context. *Geosciences*, *12*(2), 89. https://doi.org/10.3390/geosciences12020089

Cazenave, P. W., Dix, J. K., Lambkin, D. O., & McNeill, L. C. (2013). A method for semi-automated objective quantification of linear bedforms from multi-scale digital elevation models. *Earth Surface Processes and Landforms*, *38*(3), 221–236. https://doi.org/10.1002/esp.3269

Cisneros, J., Best, J [Jim], van Dijk, T., Almeida, R. P. de, Amsler, M., Boldt, J., Freitas, B., Galeazzi, C., Huizinga, R., Ianniruberto, M., Ma, H., Nittrouer, J. A., Oberg, K., Orfeo, O [Oscar], Parsons, D [Dan], Szupiany, R., Wang, P., & Zhang, Y [Yuanfeng] (2020). Dunes in the world's big rivers are characterized by low-angle lee-side slopes and a complex shape. *Nature Geoscience*, *13*(2), 156–162. https://doi.org/10.1038/s41561-019-0511-7

Cukur, D., Kong, G.-S., Buchs, D. M., Lee, G.-S., Kim, S.-P., Um, I.-K., Chun, J.-H., Kim, B.-Y., Seo, J. E., Chae, H. S., & Horozal, S. (2022). Upslope migrating sand waves on sediment-starved shelves: An example from the southeastern continental margin of the Korean Peninsula. *Marine Geology*, *444*, 106728. https://doi.org/10.1016/j.margeo.2021.106728

Di Stefano, M., & Mayer, L. (2018). An Automatic Procedure for the Quantitative Characterization of Submarine Bedforms. *Geosciences*, *8*(1), 28. https://doi.org/10.3390/geosciences8010028

Flemming, B. W [Burghard W.] (1988). On the classification of subaquatic flow-transverse bedforms. *Bochumer Geologische Und Geotechnische Arbeiten*, *29*, 44–47. https://www.researchgate.net/profile/Burghard-Flemming/publication/271909932_On_the_classification_of_subaquatic_flow-transverse_bedforms/links/54d679a80cf2970e4e6dc399/On-the-classification-of-subaquatic-flow-transverse-bedforms.pdf

Franzetti, M., Le Roy, P., Delacourt, C., Garlan, T., Cancouët, R., Sukhovich, A., & Deschamps, A. (2013). Giant dune morphologies and dynamics in a deep continental shelf environment: Example of the banc du four (Western Brittany, France). *Marine Geology*, *346*, 17–30. https://doi.org/10.1016/j.margeo.2013.07.014

Ganti, V., Paola, C., & Foufoula-Georgiou, E. (2013). Kinematic controls on the geometry of the preserved cross sets. *Journal of Geophysical Research: Earth Surface*, *118*(3), 1296–1307. https://doi.org/10.1002/jgrf.20094

Gogolewski, D. (2020). Influence of the edge effect on the wavelet analysis process. *Measurement*, *152*, 107314. https://doi.org/10.1016/j.measurement.2019.107314

Greene, H. G., Baker, M., & Aschoff, J. (2020). A dynamic bedforms habitat for the forage fish Pacific sand lance, San Juan Islands, WA, United States. In *Seafloor Geomorphology as Benthic Habitat* (pp. 267–279). Elsevier. https://doi.org/10.1016/B978-0-12-814960-7.00014-2

Gutierrez, R. R., Abad, J. D., Parsons, D. R [Daniel R.], & Best, J. L [James L.] (2013). Discrimination of bed form scales using robust spline filters and wavelet transforms: Methods and application to synthetic signals and bed forms of the Río Paraná, Argentina. *Journal of Geophysical Research: Earth Surface*, *118*(3), 1400–1418. https://doi.org/10.1002/jgrf.20102

Gutierrez, R. R., Mallma, J. A., Núñez-González, F., Link, O., & Abad, J. D. (2018). Bedforms-ATM, an open source software to analyze the scale-based hierarchies and dimensionality of natural bed forms. *SoftwareX*, *7*, 184–189. https://doi.org/10.1016/j.softx.2018.06.001

Hansen, L. Ø., Ernstsen, V. B., Andersen, M. S., Al-Hamdani, Z., Becker, M., Andresen, K. J., Trinhammer, P. L., Vang, T., Bartholdy, J [Jesper], & Kroon, A. (2022). A multi-scale geomorphometric approach to semi-automated classification of seabed morphology of a dynamic and complex marine meander bend. *Geomorphology*, *413*, 108371. https://doi.org/10.1016/j.geomorph.2022.108371

Hartley, A. J., & Owen, A. (2022). Paleohydraulic analysis of an ancient distributive fluvial system. *Journal of Sedimentary Research*, *92*(5), 445–459. https://doi.org/10.2110/jsr.2021.062

Hitchcock, F. L. (1941). The Distribution of a Product from Several Sources to Numerous Localities. *Journal of Mathematics and Physics*, *20*(1-4), 224–230. https://doi.org/10.1002/sapm1941201224

Hu, H., Yang, Z., Yin, D., Cheng, H., Hackney, C. R., & Parsons, D. R [Daniel R.] (2021). The combined effect of discharge and tides on low-angle dune evolution at the tidal current limit of the Changjiang Estuary. *Geomorphology*, *392*, 107917. https://doi.org/10.1016/j.geomorph.2021.107917

Huizinga, R. J. (2017). *Scientific Investigations Report.* https://doi.org/10.3133/sir20175076

Jordan, C., Tiede, J., Lojek, O., Visscher, J., Apel, H., Nguyen, H. Q., Quang, C. N. X., & Schlurmann, T. (2019). Sand mining in the Mekong Delta revisited - current scales of local sediment deficits. *Scientific Reports*, *9*(1), 17823. https://doi.org/10.1038/s41598-019-53804-z

Karvelis, P. (2023). *daviolinplot* (Version 3.2.1) [Computer software]. GitHub. https://github.com/frank-pk/DataViz/releases/tag/v3.2.1

Le Guern, J., Rodrigues, S., Geay, T., Zanker, S., Hauet, A., Tassi, P., Claude, N., Jugé, P., Duperray, A., & Vervynck, L. (2021). Relevance of acoustic methods to quantify bedload transport and bedform dynamics in a large sandy-gravel-bed river. *Earth Surface Dynamics*, *9*(3), 423–444. https://doi.org/10.5194/esurf-9-423-2021

Lebrec, U., Riera, R., Paumard, V., O'Leary, M. J., & Lang, S. C. (2022). Automatic Mapping and Characterisation of Linear Depositional Bedforms: Theory and Application Using Bathymetry from the North West Shelf of Australia. *Remote Sensing*, *14*(2), 280. https://doi.org/10.3390/rs14020280

Lee, J., Musa, M., & Guala, M. (2021). Scale-Dependent Bedform Migration and Deformation in the Physical and Spectral Domains. *Journal of Geophysical Research: Earth Surface*, *126*(5). https://doi.org/10.1029/2020JF005811

Lefebvre, A., Ernstsen, V. B., & Winter, C. (2011a). Influence of compound bedforms on hydraulic roughness in a tidal environment. *Ocean Dynamics*, *61*(12), 2201–2210. https://doi.org/10.1007/s10236-011-0476-6

Lefebvre, A., Ernstsen Verner B., & Winter, C. (2011b). *Bedform characterization through 2D spectral analysis*. https://www.jstor.org/stable/26482278

Lefebvre, A., Herrling, G., Becker, M., Zorndt, A., Krämer, K., & Winter, C. (2022). Morphology of estuarine bedforms, Weser Estuary, Germany. *Earth Surface Processes and Landforms*, *47*(1), 242–256. https://doi.org/10.1002/esp.5243

Lefebvre, A., & Winter, C. (2021). Special issue: Marine and river bedforms dynamics, MARID VI. *Earth Surface Processes and Landforms*, *46*(9), 1646–1651. https://doi.org/10.1002/esp.5091

Lisimenka, A., & Kubicki, A. (2017). Estimation of dimensions and orientation of multiple riverine dune generations using spectral moments. *Geo-Marine Letters*, *37*(1), 59–74. https://doi.org/10.1007/s00367-016-0475-1

Liu, P., & Xiao, F. (2022). Earth Mover's divergence of belief function. *Computational and Applied Mathematics*, *41*(6). https://doi.org/10.1007/s40314-022-02000-3

Lorenz, R. D., Wall, S., Radebaugh, J., Boubin, G., Reffet, E., Janssen, M., Stofan, E., Lopes, R., Kirk, R., Elachi, C., Lunine, J., Mitchell, K., Paganelli, F., Soderblom, L., Wood, C., Wye, L., Zebker, H., Anderson, Y., Ostro, S., . . . West, R. (2006). The sand seas of Titan: Cassini RADAR observations of longitudinal dunes. *Science (New York, N.Y.)*, *312*(5774), 724–727. https://doi.org/10.1126/science.1123257

Meijer, K. J., Franken, O., van der Heide, T., Holthuijsen, S. J., Visser, W., Govers, L. L., & Olff, H. (2022). Characterizing bedforms in shallow seas as an integrative predictor of seafloor stability and the occurrence of macrozoobenthic species. *Remote Sensing in Ecology and Conservation*, Article rse2.312. Advance online publication. https://doi.org/10.1002/rse2.312

Myrow, P. M., Lamb, M. P., & Ewing, R. C. (2018). Rapid sea level rise in the aftermath of a Neoproterozoic snowball Earth. *Science (New York, N.Y.)*, *360*(6389), 649–651. https://doi.org/10.1126/science.aap8612

Nittrouer, J. A., Allison, M. A., & Campanella, R. (2008). Bedform transport rates for the lowermost Mississippi River. *Journal of Geophysical Research*, *113*(F3). https://doi.org/10.1029/2007JF000795

Nordin, C. F., & Algert, J. H. (1966). Spectral Analysis of Sand Waves. *Journal of the Hydraulics Division*, *92*(5), 95–114. https://doi.org/10.1061/JYCEAJ.0001533

Núñez-González, F., Hesse, D., Ettmer, B., Gutierrez, R. R., & Link, O. (2021). Development and validation of a novel metric for describing the three-dimensionality of bed forms. *Geomorphology*, *390*, 107856. https://doi.org/10.1016/j.geomorph.2021.107856

Ogor, J. (2018). *Design of algorithms for the automatic characterization of marine dune morphology and dynamics*. Dissertation. https://www.theses.fr/2018enta0005

Parsons, D. R [D. R.], Best, J. L [J. L.], Orfeo, O [O.], Hardy, R. J., Kostaschuk, R., & Lane, S. N. (2005). Morphology and flow fields of three-dimensional dunes, Rio Paraná, Argentina: Results from simultaneous multibeam echo sounding and acoustic Doppler current profiling. *Journal of Geophysical Research: Earth Surface*, *110*(F4), n/a-n/a. https://doi.org/10.1029/2004JF000231

Pike, R. J. (2000). Geomorphometry - diversity in quantitative surface analysis. *Progress in Physical Geography*, *24*(1), 1–20. https://doi.org/10.1191/030913300674449511

Reeder, D. B., Ma, B. B., & Yang, Y. J. (2011). Very large subaqueous sand dunes on the upper continental slope in the South China Sea generated by episodic, shoaling deep-water internal solitary waves. *Marine Geology*, *279*(1-4), 12–18. https://doi.org/10.1016/j.margeo.2010.10.009

Scheiber, L., & Lefebvre, A. (2023). The influence of different geometric definitions on dune characteristics. In MARID VII (Chair), *Marine and River Dune Dynamics*, Renne, France.

Scheiber, L., Lojek, O., Götschenberg, A., Visscher, J., & Schlurmann, T. (2021a). Robust methods for the decomposition and interpretation of compound dunes applied to a complex hydromorphological setting. *Earth Surface Processes and Landforms*, *46*(2), 478–489. https://doi.org/10.1002/esp.5040

Scheiber, L., Zomer, J., Wang, L., Cisneros, J., Gutierrez, R., & Lefebvre, A. (2021b). *Spoilt for Choice – When to Use Which Bedform Identification Tool for What Purpose?* https://doi.org/10.1002/essoar.10509747.1

Schippa, L., & Cavalieri, I. (2021). *Sustainable WATER RESOURCES MANAGEMENT: Effective Approaches for River Basins and Urban Catchments*. Semiautomatic algorithm for the initerpretation of the bedforms and statistical analysis (XI, 250, 169). WIT Press.

Simons, D. B., Richardson, E. V., & Nordin, C. F. (1965). *Bedload Equation for Ripples and Dunes*. U.S. Government Printing Office.

van der Mark, C. F., Blom, A., & Hulscher, S. J. M. H. (2008). Quantification of variability in bedform geometry. *Journal of Geophysical Research*, *113*(F3). https://doi.org/10.1029/2007JF000940

van Dijk, T. A. G. P., Lindenbergh, R. C., & Egberts, P. J. P. (2008). Separating bathymetric data representing multiscale rhythmic bed forms: A geostatistical and spectral method compared. *Journal of Geophysical Research*, *113*(F4). https://doi.org/10.1029/2007JF000950

Wang, L., Yu, Q., Zhang, Y [Yongzhan], Flemming, B. W [Burghard W.], Wang, Y., & Gao, S. (2020). An automated procedure to calculate the morphological parameters of superimposed rhythmic bedforms. *Earth Surface Processes and Landforms*, *45*(14), 3496–3509. https://doi.org/10.1002/esp.4983

Wilbers, A., & Brinke, W. ten (2003). The response of subaqueous dunes to floods in sand and gravel bed reaches of the Dutch Rhine. *Sedimentology*, *50*(6), 1013–1034. https://doi.org/10.1046/j.1365-3091.2003.00585.x

Winter, C., & Ernstsen, V. B. (2007). *Spectral analysis of compound dunes*. https://www.researchgate.net/profile/verner-ernstsen/publication/256475581_spectral_analysis_of_compound_dunes/links/02e7e522f73e4c650a000000/spectral-analysis-of-compound-dunes.pdf

Wu, S., Xu, Y. J., Wang, B., & Cheng, H. (2021). Riverbed dune morphology of the Lowermost Mississippi River – Implications of leeside slope, flow resistance and bedload transport in a large alluvial river. *Geomorphology*, *385*, 107733. https://doi.org/10.1016/j.geomorph.2021.107733

Xu, J., Li, G., Dong, P., & Shi, J. (2010). Bedform evolution around a submarine pipeline and its effects on wave-induced forces under regular waves. *Ocean Engineering*, *37*(2-3), 304–313. https://doi.org/10.1016/j.oceaneng.2009.10.002

Zomer, J. Y., Naqshband, S., & Hoitink, A. J. F. (2022). Short communication: A tool for determining multiscale bedform characteristics from bed elevation data. *Earth Surface Dynamics*, *10*(5), 865–874. https://doi.org/10.5194/esurf-10-865-2022