

Performance evaluation of wave input reduction techniques for modeling inter-annual sandbar dynamics

de Queiroz, Bruna; Scheel, Freek; Caires, Sofia; Walstra, Dirk Jan; Olij, Derrick; Yoo, Jeseon; Reniers, Ad; de Boer, Wiebe

DOI

[10.3390/jmse7050148](https://doi.org/10.3390/jmse7050148)

Publication date

2019

Document Version

Final published version

Published in

Journal of Marine Science and Engineering

Citation (APA)

de Queiroz, B., Scheel, F., Caires, S., Walstra, D. J., Olij, D., Yoo, J., Reniers, A., & de Boer, W. (2019). Performance evaluation of wave input reduction techniques for modeling inter-annual sandbar dynamics. *Journal of Marine Science and Engineering*, 7(5), Article 148. <https://doi.org/10.3390/jmse7050148>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

Article

Performance Evaluation of Wave Input Reduction Techniques for Modeling Inter-Annual Sandbar Dynamics

Bruna de Queiroz ^{1,*}, Freek Scheel ¹, Sofia Caires ¹, Dirk-Jan Walstra ^{1,2}, Derrick Olij ³, Jeseon Yoo ⁴, Ad Reniers ² and Wiebe de Boer ^{1,2}

¹ Deltares, P.O. Box 177, 2600 MH Delft, The Netherlands; Freek.Scheel@deltares.nl (F.S.);

Sofia.Caires@deltares.nl (S.C.); DirkJan.Walstra@deltares.nl (D.-J.W.); Wiebe.deBoer@deltares.nl (W.d.B.)

² Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands; A.J.H.M.Reniers@tudelft.nl

³ PwC Strategy&, 1066 JR Amsterdam, The Netherlands; derrick.olij@gmail.com

⁴ Korean Institute of Ocean Science and Technology (KIOST), Yeongdo Haeyangro 385, Busan 49111, Korea; jyoo@kiost.ac.kr

* Correspondence: brunaquei@gmail.com

Received: 8 April 2019; Accepted: 8 May 2019; Published: 15 May 2019



Abstract: In process-based numerical models, reducing the amount of input parameters, known as input reduction (IR), is often required to reduce the computational effort of these models and to enable long-term, ensemble predictions. Currently, a comprehensive performance assessment of IR-methods is lacking, which hampers guidance on selecting suitable methods and settings in practice. In this study, we investigated the performance of 10 IR-methods and 36 subvariants for wave climate reduction to model the inter-annual evolution of nearshore bars. The performance of reduced wave climates is evaluated by means of a brute force simulation based on the full climate. Additionally, we tested how the performance is affected by the number of wave conditions, sequencing, and duration of the reduced wave climate. We found that the Sediment Transport Bins method is the most promising method. Furthermore, we found that the resolution in directional space is more important for the performance than the resolution in wave height. The results show that a reduced wave climate with fewer conditions applied on a smaller timescale performs better in terms of morphology than a climate with more conditions applied on a longer timescale. The findings of this study can be applied as initial guidelines for selecting input reduction methods at other locations, in other models, or for other domains.

Keywords: input reduction; morphodynamics; sandbars; process-based modeling; wave climate; sequencing; Markov Chain; Monte Carlo

1. Introduction

Understanding and predicting the evolution of coastal morphology is important in coastal engineering, because of its implications for coastal safety, the environment, and the economy. For instance, coastal morphodynamics influence the occurrence of rip currents affecting swimmer safety, the protection of the hinterland from coastal flooding and erosion, and the establishment and development of coastal ecosystems. Often, process-based morphodynamic models are used to predict coastal evolution. These models account for a wide range of coastal processes, such as waves, currents, sediment transport, and morphology, which results in a high level of complexity and, consequently, extensive computational effort.

As complexity and computational effort increase, process or input reduction is necessary to obtain feasible computational times for engineering applications. In a morphodynamic sense, input reduction (IR) can be defined as the selection of a reduced set of representative forcing conditions that lead to accurate approximations of the long-term morphological evolution [1]. A robust input reduction method should preserve some natural variability of the environment to be able to represent the full set of conditions accurately [2,3]. In coastal environments, waves are typically the dominating forcing conditions (i.e., for wave-dominated coasts). Accurate modeling of the nearshore morphodynamics requires selecting representative wave conditions that capture the variation in both wave height and direction, including mild and extreme events.

Mainly two categories of wave input reduction methods exist: binning and clustering methods. Binning methods divide the wave conditions into bins, sometimes using a specific weight target, such as longshore sediment transport. Clustering methods cluster the wave conditions according to their statistical similarity. Binning methods have been previously investigated by [2,3]. However, to the best of our knowledge, clustering methods applied to inter-annual morphological predictions have not yet been addressed.

In addition to the selection of the representative wave conditions, the number, the duration, and sequencing of the wave conditions also affect the IR performance [4]. Sequencing of wave conditions refers to the order in which the representative wave conditions occur in the model. The chronology of the storm conditions is likely to affect the morphological response of the sandbar due to non-linear effects. The sequencing of the representative wave conditions can be performed by random, systematic or Markov Chain sequencing methods. Random sequencing draws wave conditions randomly from the reduced wave climate [2,3]. Systematic sequencing orders the wave conditions according to wave height (e.g., descending or ascending) and incident wave angle with respect to the shore-normal (e.g., either from positive to negative or vice versa, see [2]). Markov Chain sequencing utilizes the wave chronology of the full dataset to order the representative wave conditions. To the best of our knowledge, Markov Chain sequencing methods applied to inter-annual morphological predictions have not yet been addressed in the literature.

Although input reduction is a common practice in morphodynamic modeling [2,3,5–9], a comprehensive study on the performance of different IR-methods is lacking. In this study, we investigate the performance of 10 input reduction techniques with 36 subvariants (i.e., different initializations and input variables), including both binning and clustering methods. To this end, we use a cross-shore sandbar behavior model forced with measured wave time-series to simulate the morphological evolution of a beach profile for the cases of Noordwijk in the Netherlands (Figure 1a) and Anmok in South Korea (Figure 1b). For the most promising method, we systematically assess the performance with respect to the number of wave conditions, the sequencing method, and the duration of reduced wave climate (following [2]). As the cross-shore model is computationally inexpensive, we are able to test a wide range of input reduction methods and settings to derive guidelines to select a suitable input reduction setup for future studies.

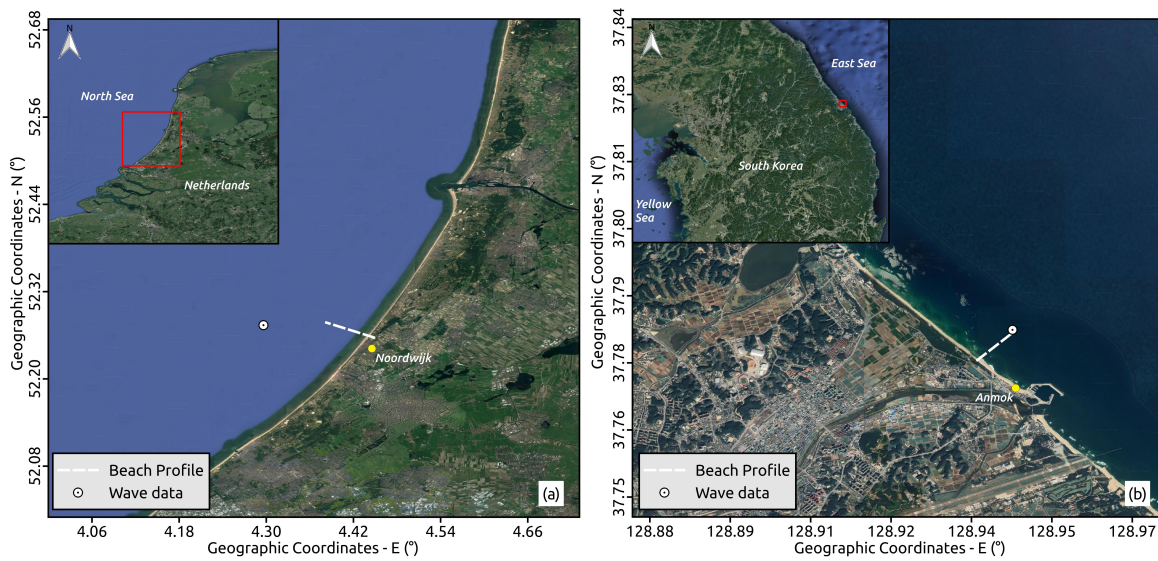


Figure 1. Location of the simulated beach profile and the wave observations used as forcing of the models of the study area, Noordwijk, the Netherlands (a) and the validation area, Anmok, South Korea (b). Map data: Google.

2. Overall Approach

2.1. Research Steps for Testing IR-Methods and Settings

We divided the performance assessment into three steps (see Figure 2): (1) selecting the most promising IR-method; (2) testing the influence of different settings on the performance of the most promising method (i.e., the number of conditions, the sequencing method, and the duration of the reduced wave climate); and (3) verifying the optimal settings for a second case study. For the first two steps we used the case study of Noordwijk in the Netherlands (Figure 1a) and for the third step, the case study of Anmok in South Korea (Figure 1b).

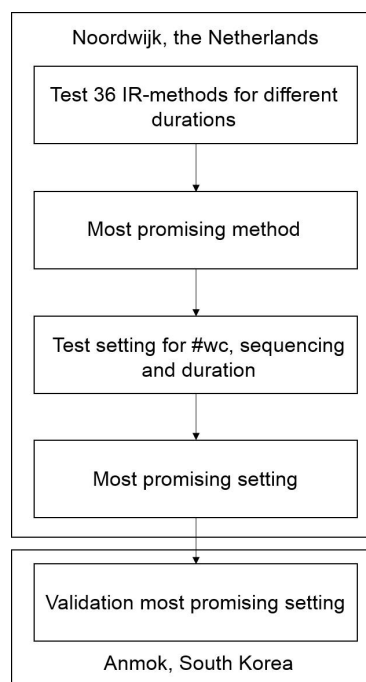


Figure 2. Flowchart of the research outline of this study, #wc is number of wave conditions.

First, we assessed the performance of 36 IR-method variations (see Table 1) for different simulation times for the case study of Noordwijk. For this assessment, we kept the number of wave conditions constant at 12 and used only random sequencing. The descriptions of the IR-methods included in this study are provided in Section 3. Second, we tested the influence of the number of wave conditions, the sequencing method, and the duration of the wave climate for the most promising IR-method found in step 1 (i.e., the Sediment Transport Bins Method). As the variables are interconnected, we tested different combinations of these variables. The tested settings are described in Section 4. Finally, we verified the most promising settings from the Noordwijk case by applying it to Anmok beach. The tested IR-methods and settings are summarized in Table 1.

Table 1. Overview of simulated input reduction (IR)-methods and variations. The meaning of the variable symbols is as follows: S_y = longshore sediment transport; T_p = wave peak period; θ = wave angle; H_{rms} = root-mean-square wave height.

Type	Method	Variation	Input Variables	Cluster Initiation	
Binning	Conditions with the Largest Transport Contribution (CLTCM)		CLTCM	S_y, T_p, θ	-
	Fixed Bins (FBM)	FBM1	H_{rms}, θ	-	
		FBM2	$H_{rms}^{2.5}, \theta$	-	
		FBM3	S_y, θ	-	
	Energy Flux (EFM)	EFM	H_{rms}, θ	-	
	Sediment Transport Bins Method (STBM)	STBM	S_y, θ	-	
	The Representative Wave Approach (RWA)	RWA	H_{rms}, θ	-	
	Maximum Dissimilarity Algorithm (MDA)	MDA	H_{rms}, T_p, θ	-	
	Grouping with Equal Sediment Influence (GESIM)	GESIM	S_y, T_p, θ	MDA	
	Clustering	Crisp k-means (CKM)	CKM1	H_{rms}, T_p, θ	K-means++
CKM2			$H_{rms}^{2.5}, T_p, \theta$	K-means++	
CKM3			S_y, T_p, θ	K-means++	
CKM4			H_{rms}, T_p, θ	MDA	
CKM5			$H_{rms}^{2.5}, T_p, \theta$	MDA	
CKM6			S_y, T_p, θ	MDA	
CKM7			H_{rms}, T_p, θ	Fixed Bins	
CKM8			$H_{rms}^{2.5}, T_p, \theta$	Fixed Bins	
CKM9			S_y, T_p, θ	Fixed Bins	
Fuzzy k-means (FKM)		FKM1	H_{rms}, T_p, θ	K-means++	
		FKM2	$H_{rms}^{2.5}, T_p, \theta$	K-means++	
		FKM3	S_y, T_p, θ	K-means++	
		FKM4	H_{rms}, T_p, θ	MDA	
		FKM5	$H_{rms}^{2.5}, T_p, \theta$	MDA	
		FKM6	S_y, T_p, θ	MDA	
		FKM7	H_{rms}, T_p, θ	Fixed Bins	
		FKM8	$H_{rms}^{2.5}, T_p, \theta$	Fixed Bins	
		FKM9	S_y, T_p, θ	Fixed Bins	
K-harmonic means (KHM)		KHM1	H_{rms}, T_p, θ	K-means++	
		KHM2	$H_{rms}^{2.5}, T_p, \theta$	K-means++	
		KHM3	S_y, T_p, θ	K-means++	
		KHM4	H_{rms}, T_p, θ	MDA	
		KHM5	$H_{rms}^{2.5}, T_p, \theta$	MDA	
		KHM6	S_y, T_p, θ	MDA	
		KHM7	H_{rms}, T_p, θ	Fixed Bins	
		KHM8	$H_{rms}^{2.5}, T_p, \theta$	Fixed Bins	
		KHM9	S_y, T_p, θ	Fixed Bins	

2.2. Performance Evaluation

We assessed the performance of the 36 IR-method variants using a cross-shore bar sandbar behavior model (i.e., UNIBEST-TC), which was calibrated by [10] for Noordwijk beach in the Netherlands. We used this 3.3 years-long brute force model simulation (i.e., forced with the full measured wave time series) as the benchmark to assess the performance of the IR-methods. For the quantitative performance, we used the cumulative skill score described in [11] for the cross-shore profile at the end of the 3.3 years simulation.

The skill score (R) of the reduced wave climate model predictions (z_{red}) in relation to the brute force predictions (z_{full}) is determined for all bar profiles in time (t) over the cross-shore distance (x) on the bar profile ($x_1 - x_{end}$).

$$R(t) = 1 - \frac{\epsilon_{red}^2(t)}{\epsilon_{full}^2(t)} \tag{1}$$

$$\epsilon_{red}^2(t) = \sum_{x=x_1}^{x=x_{end}} \sum_{t=0}^{t=t_{end}} (z_{red}(x, t) - z_{full}(x, t))^2 \tag{2}$$

$$\epsilon_{full}^2(t) = \sum_{x=x_1}^{x=x_{end}} \sum_{t=0}^{t=t_{end}} (z_{full}(x, t) - z_{full}(x, t = 0))^2 \tag{3}$$

$R = 1$ means a perfect match (i.e., no deviations) between the morphological prediction with the reduced wave climate and the brute force model. $R \leq 1$ indicates discrepancies between the morphological prediction with the reduced wave conditions and with the brute force model. Because of the tendency of the skill score R to reward predictions that underestimate the overall magnitude of bed changes [12], a qualitative assessment was also performed by visual comparison of the final profile of the reduced and full wave climate.

For the validation of the most promising settings, we applied the results obtained from Noordwijk case in a calibrated 0.8 years-long brute-force model of Anmok beach. The model was forced with wave time series measured 850 m offshore (i.e., at ca. 20 m water depth) with a temporal resolution of 1 h. The calibration was performed by means of the skill score (Equations (1)–(3))11 and visual comparison with the beach profile obtained from a survey campaign. The model settings for the UNIBEST-TC model are listed in Table 2.

Table 2. Calibration parameters of the brute force model of Anmok beach.

Parameter	Value
Grid Resolution (dx)	10 m–100 m
Time-step (dt)	0.04167 days
Median grain diameter (D_{50})	400 μm
Breaker-delay (λ)	1
Angle of repose 1 (ϕ_1)	1.5
Cross-shore location of ϕ_1 (XF_1)	400 m
Angle of repose 2 (ϕ_2)	0.1
Cross-shore location of ϕ_2 (XF_2)	150 m
Current-related roughness (k_c)	0.005593
Wave-related roughness (k_w)	0.00045

3. Tested Input Reduction Methods

We selected five binning and five clustering methods for the performance assessment (see Table 1). In most of these methods, the wave conditions are clustered based on their spectral parameters, such as the root-mean-square wave height (H_{rms}), peak period (T_p), and wave direction (θ). However, we also clustered the conditions with respect to their contribution to the sediment transport by substituting H_{rms} by the associated longshore sediment transport, S_y (when known) or H_{rms}^p , where p is the power that represents the non-linear relation between wave height and sediment transport. Typically, p varies from $p = 2$ to $p = 3$. In this study, $p = 2.5$ was applied. All variations in terms of input variables are shown in Table 1.

In each bin or cluster, the representative wave condition is defined as the centroid of that bin or cluster. For the spectral wave parameters, the centroids are defined as the average of the wave

conditions within a cluster or bin. For H_{rms}^p , the definition of the centroids within a bin is defined by a non-linear weighting formula for the wave height Equation (4).

$$H_{rms,j}^{rep} = \left(\frac{\sum_{i:x_i \in C_j} (f_i H_{rms,i}^p)}{\sum_{i:x_i \in C_j} f_i} \right)^{1/p} \tag{4}$$

where f_i is the frequency of occurrence of the root-mean-square wave height $H_{rms,i}$ of a wave condition x_i belonging to the cluster or bin j with observations C_j . For S_y the centroids are defined by the average of the wave conditions within the bin or cluster. To obtain the values of H_{rms} the nearest wave condition of the centroid is used. The sub-sections below discuss the principles of the selected IR-methods.

3.1. Binning Methods

3.1.1. Conditions with the Largest Transport Contribution Method

The Conditions with the Largest Transport Contribution method (CLTCM) [13] selects only the wave conditions with the highest contributions to the longshore sediment transport. Initially, the wave conditions are binned into a larger number of wave height and wave direction bins than the desired number of wave conditions for the reduced wave climate. The sediment transport contribution of each bin is determined, and the k bins with the highest sediment transport contribution are selected as the representative wave conditions. This method uses the sediment transport rates as the input variable (see Table 1) and, hence, requires the transport rates corresponding to the wave conditions to be known before the input reduction is executed.

3.1.2. Fixed Bins Method

The Fixed Bins method (FBM) ([3]) divides the wave conditions in pre-defined wave height and wave direction bins. The algorithm first divides the wave conditions in directional bins with uniform resolution. Next, each directional bin is divided into wave height bins according to its range of wave height. This results in wave height bins that can vary among the directional classes (see Figure 3a).

3.1.3. Energy Flux Method

The Energy Flux method (EFM) ([3]) divides the wave conditions in pre-defined wave direction and wave height bins with equal amount of energy flux (E_f).

$$E_f = \left(\frac{\rho g H_s^2}{8} \right) c_g \tag{5}$$

where ρ is the water density (assumed to be 1025 kg/m³), g is the gravity acceleration ($g = 9.81 \text{ m/s}^2$), H = deep water wave height and c_g = wave group celerity in deep water. The EFM generates a higher bin resolution for conditions with more wave energy and a lower resolution for conditions with less wave energy (see Figure 3b). The wave height of the representative wave conditions is defined as the inverse function of the average energy flux of each bin while wave period and wave direction are defined as the average of the wave conditions in a bin.

3.1.4. Sediment Transport Bins Method

Similar to the EFM, Sediment Transport Bins method (STBM) divides the wave data in pre-defined wave direction and wave height bins with equal weight, but the weight is determined by the longshore sediment transport obtained from the brute force simulation. In contrast to the EFM, the definition of the directional bins starts from the shore-normal angle such that wave conditions that cause opposite sediment transport rates do not average out within a bin (see Figure 3c).

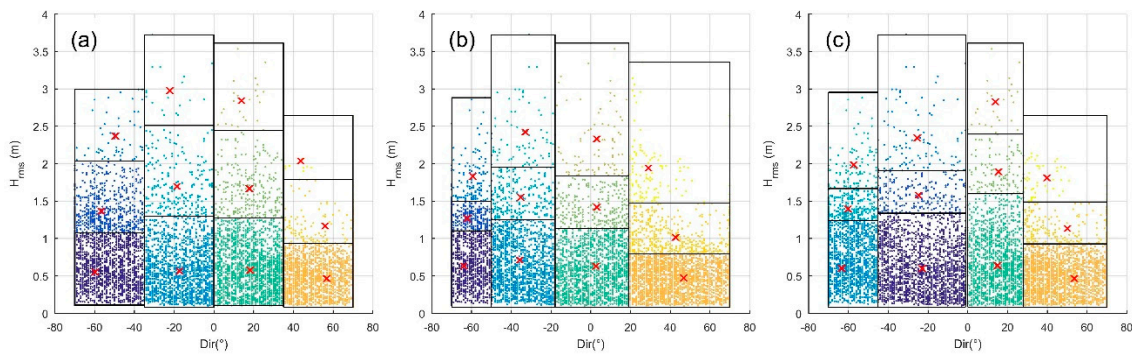


Figure 3. Examples of a selection of Fixed Bins (a), Energy Flux (b), and Sediment Transport Bins (c) methods with 12 representative wave conditions. The red crosses are the representative wave conditions, and the small dots are the wave data. The colors and black lines indicate the bins.

3.1.5. Representative Wave Approach

The Representative Wave approach (RWA) is adapted from [6] and divides the wave data into bins over time. In this paper, we divided the wave data into seasons. For each section, the representative wave condition is the average of the wave conditions in that bin. This is the only method that preserves the chronology of the original wave dataset.

3.2. Clustering Methods

Table 1 provides an overview of the clustering methods and variations that were selected for this study. The clustering methods use the normalized 3-dimensional (e.g., H_{rms} , T_p , θ) Euclidean distance as a measure of similarity between the wave conditions. The closer the distance between the wave conditions, the more similar to each other they are. Similar to the binning methods, $H_{rms}^{2.5}$ or S_y can be used as an alternative input variable for H_{rms} . We tested these variations for the crisp k-means, fuzzy k-means, and k-harmonic means methods (see Table 1). The similarity of the wave conditions depends on the cluster initiation. Therefore, we tested different cluster initiations for the clustering methods: fixed bins, maximum dissimilarity algorithm, and K-means++ algorithm [14].

3.2.1. Maximum Dissimilarity Algorithm

The Maximum Dissimilarity algorithm (MDA) (see [15]) creates a subset of k centroids that represents the full diversity of the wave data by maximizing the dissimilarity between the vectors in the subset. To measure the dissimilarity between vectors, we used the MaxMin Algorithm [16] with the efficient algorithm of [17]. The first centroid is the wave condition with the maximum distance to all other wave data. After the first centroid is excluded from the dataset, the second centroid is the wave condition with the maximum distance to the first centroid. The subsequent centroids are the wave conditions with the maximum distance among the minimum distance of the remaining wave conditions to the previous centroids.

3.2.2. Grouping with Equal Sediment Influence Method

The Grouping with Equal Sediment Influence method (GESIM) aggregates wave conditions in clusters with approximately the same sediment transport contribution [13]. Therefore, it uses only S_y , T_p , and θ as input variables. It has the same principle as the STBM, but it aggregates the wave conditions into clusters instead of dividing them into bins. GESIM starts by selecting k initial wave conditions as individual clusters using the MDA. Subsequently, in every iteration, each cluster incorporates the closest observation to the cluster until a total sediment transport threshold is reached. The threshold is defined by the total sediment transport divided by the number of representative cases (k). When wave conditions cannot join a cluster anymore, the remaining wave conditions join the cluster to which they have the smallest distance. In the end, this results in k clusters that represent

approximately the same amount of sediment transport. The centroid of each cluster is defined as the average of the wave conditions in the cluster.

3.2.3. Crisp K-Means Method

The Crisp K-Means method (CKM), also known as K-means, is one of the most widely used clustering methods [18,19]. It starts with k initial centroids that are defined randomly with weights based on the distance of the wave conditions through the K-means++ algorithm [14]. Then, every wave condition is assigned to the cluster it is closest to. The CKM has a hard membership function which means that wave conditions can only be a member of one cluster. Next, the centroids are updated by averaging the wave conditions that constitute the clusters. This procedure is repeated iteratively until the difference between the current and previous centroids is smaller than a user-defined accuracy criterion (see Figure 4a). More details can be found in [20].

3.2.4. Fuzzy K-Means Method

The Fuzzy K-Means method (FKM) (see [21]) is similar to the CKM, but with a soft membership function. Therefore, wave conditions can be assigned to more than one cluster. This means that all wave conditions have some influence on the definition of the centroids determined by the fuzzy membership function. Initially, the centroids are defined as in the CKM. Then, the fuzzy membership function ($M_{i,j}$) of each wave condition is calculated for every cluster.

$$M_{i,j} = \frac{(1/\|x_i, v_j\|^2)^{1/(o-1)}}{\sum_{j=1}^k (1/\|x_i, v_j\|^2)^{1/(o-1)}} \tag{6}$$

where $\|x_i, v_j\|$ is the Euclidean distance between wave conditions (x_i) and centroids (v_j), i being the wave observation index of the full dataset, j the cluster index, and k the number of clusters (i.e., number of representative wave conditions). The fuzzy parameter o , where $o > 1$, is case specific and requires calibration.

Based on sensitivity analyses, we used $o = 1.5$. The new centroids are defined as the weighted average of the wave conditions using the fuzzy membership as weight. In this way, wave conditions closer to the previous centroid have a higher influence on the definition of the next centroid. This iterative process is repeated iteratively until the algorithm converges towards a stable solution (see Figure 4b).

3.2.5. K-Harmonic Means

The K-Harmonic means (KHM) (see [22,23]) has the same procedure as the FKM, but the weight used for the definition of the centroids is defined by a dynamic weighting function (K_i).

$$K_i = \frac{(\|x_i, v_j\|)^{-o-2}}{\sum_{j=1}^k (\|x_i, v_j\|)^{-o-2}} \times \frac{\sum_{j=1}^k \|x_i - v_j\|^{-o-2}}{(\sum_{j=1}^k \|x_i - v_j\|^{-o})^2} \tag{7}$$

In this case $o \geq 2$. Higher dimensions of the dataset require a larger value for o [23]. The parameter o is case specific and calibration is required to define it. Based on sensitivity analysis we used $o = 4.2$. The dynamic weight leads to a larger influence of outliers on selecting the centroids rather than wave conditions that are closer to the centroids (Figure 4c).

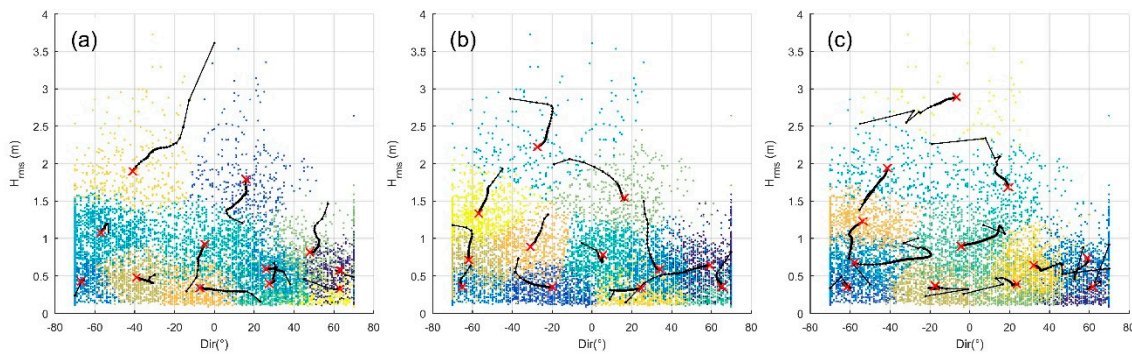


Figure 4. Examples of a selection of Crisp k-means (a), Fuzzy k-means (b), and K-harmonic (c) methods with 12 representative wave conditions. The red crosses are the representative wave conditions, and the small dots are the wave data. The colors indicate the clusters. The black lines represent the path followed by the centroids during the iterative process.

4. Tested Settings

4.1. Number of Representative Wave Conditions

The performance of an IR-method depends on the number of representative wave conditions (k) included in the reduced wave climate. Therefore, we tested the influence of the number of wave conditions on the performance of the most promising method. For binning methods, the number of representative wave conditions is defined by the combination of the number of directional bins ($ndir$) and the number of wave height bins ($nhrms$). The resolution of $ndir$ and $nhrms$ affects the performance of the input reduction method given by the distinct effects that wave height and wave direction have on sediment transport and, thus, morphology. In step 1 (i.e., testing different IR-methods), we used, $k = 12$, $ndir = 4$, and $nhrms = 3$. For the sensitivity testing, we varied k from 8 to 32 for different combinations of $ndir$ and $nhrms$ (see Table 3).

Table 3. Simulated number of wave conditions.

Label	Number of Representative Wave Conditions (k)	$ndir$	$nhrms$
$k = 8$	8	4	2
$k = 10$	10	2	5
$k = 16$	16	4	4
$k = 24A$	24	6	4
$k = 24B$	24	4	6
$k = 24C$	24	8	3
$k = 32A$	32	8	4
$k = 32B$	32	4	8

4.2. Sequencing Methods

The sequencing of the wave conditions can have a major impact on the performance of the morphological predictions due to the non-linear response of morphology to wave [2]. Ideally, the sequencing of representative wave conditions should resemble the natural variability of the full wave climate. Therefore, we tested the influence of different sequencing methods on the performance of the most promising IR-method. The sequencing methods simulated are listed in Table 4. Figure 5 illustrates the sequencing methods applied to the STBM with $k = 12$ and $T_{wc} = 301$ days. For the random sequencing and Monte Carlo methods, we used five replicates to limit the effect of the random initial choice on the performance of the method. Note that the Markov Chain sequencing has no repetitions since it does not contain randomness and that the reduced climate of the Monte Carlo Markov Chain with repetition sequencing is not repeated four times as in the other methods.

Table 4. Simulated sequencing methods.

Name	Sequencing Method
S1	Random (five replicates)
S2	Markov Chain
S3	Monte Carlo Markov Chain (five replicates)
S4	Monte Carlo Markov Chain with repetition (five replicates)

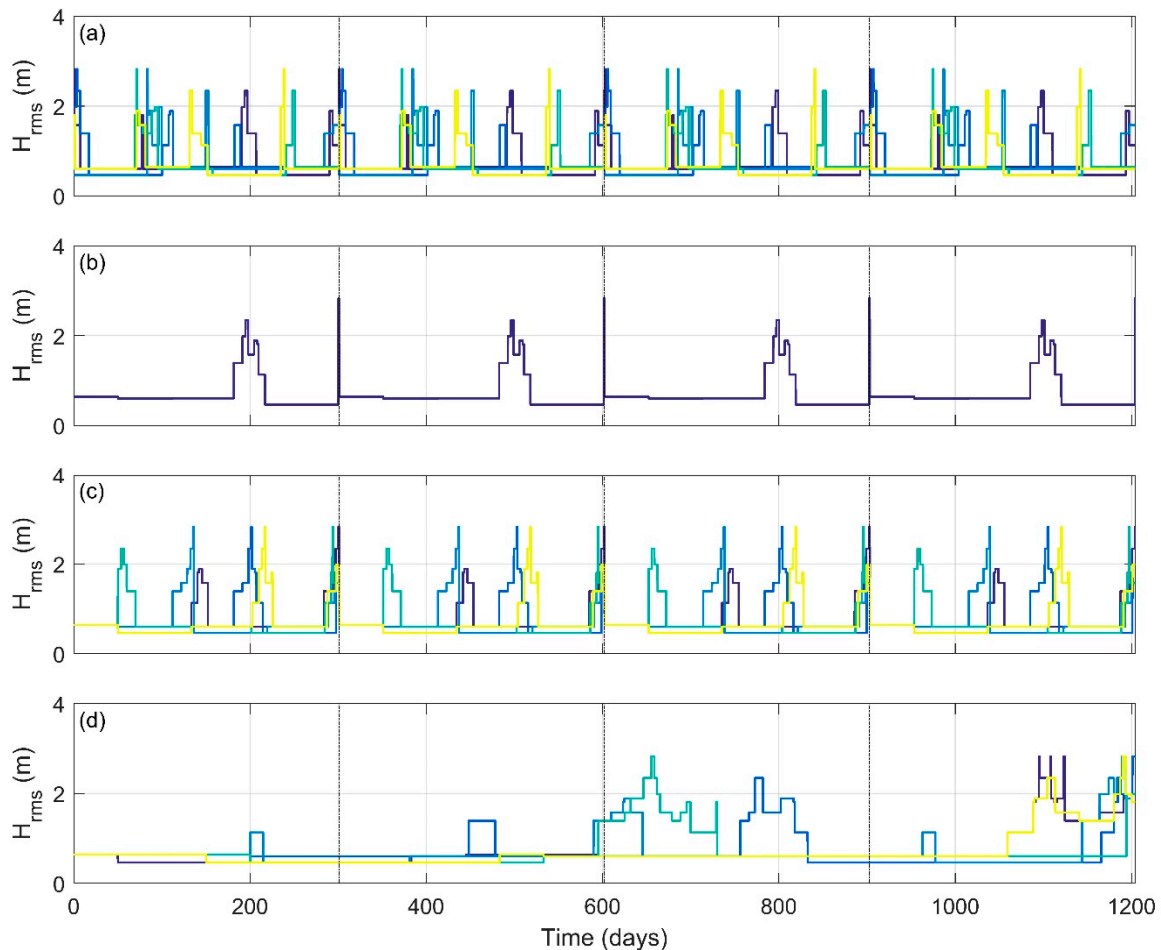


Figure 5. Wave height time-series of the reduced wave climate by the Sediment Transport Bins method (STBM) with a duration of 301 days sequenced by random (a), Markov Chain (b), Monte Carlo Markov Chain (c), and Monte Carlo Markov Chain with repetition (d) methods with $k = 12$. The colored lines represent the five repetitions. The dashed line marks the duration of each reduced wave climate.

4.2.1. Random Sequencing

Random sequencing orders the representative wave conditions randomly. First, the representative wave conditions are assigned integers ranging from 1 to k . Next, a random permutation of the integers is performed. The integers are then sorted in ascending order and their respective representative wave conditions are sequenced accordingly. The random sequence is performed with five repetitions for each method, except for the RWA that has its sequence determined by the chronology of the dataset.

4.2.2. Markov Chain Sequencing

The Markov Chain sequencing (MC) orders the representative wave conditions in the way they most likely would occur in the full dataset. The procedure is described as follows:

1. Number the representative wave conditions stored in the database V from 1 to k ;

- Determine for every wave condition from the full dataset (X) which of the representative wave conditions in V is most similar to it. In this step, a new vector F is created with size $N \times 1$ (N = number of observations of the full dataset), in which the number of the wave conditions that is most similar to each observation is stored (see Equation (8)).

$$F_i = \sum_{j=1}^k jI[\|x_i, v_j\| = \min(x_i, V)], j = 1, \dots, k, i = 1, \dots, N \tag{8}$$

where I is a true–false indicator that is 1 when the equation between brackets is true and 0 when it is false, x_i is the wave observation in the full dataset, and v_j the representative wave condition.

- Determine the Markov transitions for the wave conditions in F . The Markov transitions are stored in a Markov transition matrix M of size $k \times k$, where k is the number of representative wave conditions (see Equation (9)).

$$M(m, n) = P(F_{t+1} = n | F_t = m) = \sum_{t=1}^{N-1} \frac{I[F_t = m]I[F_{t+1} = n]}{N - 1}, m = n = 1, \dots, k \tag{9}$$

where P is the transition probability of a representative wave condition n in F given a representative wave condition m with transition index t .

- Define two time series matrices: A_s and A_{NS} . A_s starts empty and will contain the numbers that are assigned to the wave conditions in step 1 in the sequence determined by the algorithm. A_{NS} contains the numbers assigned to the wave conditions in step 1 at the start of the algorithm. When a wave condition is selected by the algorithm, its number will be deleted from matrix A_{NS} and added to the matrix A_s .
- Define the first wave condition ($A_{s,1}$) as the most similar one to the initial wave condition in the observation dataset (see Equation (10)).

$$A_{s,1} = \sum_{j=1}^k jI[\|x_1, v_j\| = \min(\|x_1, V\|)], j = 1, \dots, k \tag{10}$$

where $\|x_1, v_j\|$ and $\|x_1, V\|$ are the Euclidean distances between the initial observation of the full dataset (x_1) and a representative wave condition (v_j) and all representative wave conditions in the reduced dataset (V), respectively. The number assigned to the initial wave condition ($A_{s,1}$) is deleted from the matrix A_{NS} , which reduces its size to $(k - 1)$.

- The next wave condition to be selected for the reduced time series ($A_{s,t}$) is the one with the highest Markov transition probability (M), conditional on the previous selected wave condition ($A_{s,t-1}$) and available in the matrix A_{NS} (see Equation (11)).

$$A_{s,t} = \sum_{q=1}^{k-t+1} A_{NS,q}I[M(A_{s,t-1}, A_{NS,q}) = \max(M(A_{s,t-1}, A_{NS}))] \tag{11}$$

where t is a transition index and q is the transition probability index of M .

- Reorder the wave conditions in the database V according to their assigned numbers in matrix A_s .

4.2.3. Monte Carlo Markov Chain Sequencing

The Monte Carlo Markov Chain sequencing (MCMC) orders the representative wave conditions randomly corresponding to the Markov transition probabilities. Since the MCMC sequencing contains randomness, it was performed with five repetitions. The procedure of this sequencing method is described as follows:

1. Follow steps 1 to 2 from MC (cf. above).
2. Determine the cumulative Markov transitions ($\sum_{n=1}^k P(F_{t+1} = n|F_t = m)$) for the wave conditions in F . The cumulative Markov transitions are stored in a Markov transition matrix M of size $k \times k$:

$$M(m, n) = \sum_{n=1}^k P(F_{t+1} = n|F_t = m) = \sum_{n=1}^k \sum_{t=1}^{N-1} \frac{I[F_t = m]I[F_{t+1} = n]}{N-1}, m = n = 1, \dots, k \quad (12)$$

3. Define two time series matrices: A_s and A_{NS} as in step 4 of Section 4.2.2;
4. Define the first wave condition ($A_{s,1}$) as the most similar one to the initial wave condition in the observation dataset as in Equation (10). The Markov transition probability of the initial wave condition ($A_{s,1}$) is reduced from the cumulative Markov transition matrix M and the remaining cumulative probabilities are normalized. Moreover, the number assigned to the initial wave condition ($A_{s,1}$) will now be deleted from the matrix A_{NS} , hence, its size reduces to $(k - 1) \times 1$.
5. Draw a random number between 0 and 1 (R_t). The next wave condition to be selected ($A_{s,t}$) is the first occurrence with the Markov transition probability containing the random number previously defined:

$$A_{s,t} = \sum_{q=1}^{k-t+1} A_{NS,q} I \left[M(A_{s,t-1}, A_{NS,q}) \geq R_t = \min(M(A_{s,t-1}, A_{NS}) \geq R_t) \right], R_t = rand[0, 1] \quad (13)$$

6. Subtract the Markov transition probability of the selected wave condition $A_{s,t}$ from the cumulative Markov transition matrix M and normalize the remaining probabilities:

$$M(m, n) = \frac{M(m, n \geq A_{s,t}) - M(m, A_{s,t})}{\max(M(m, n))}, m = n = 1, \dots, k \quad (14)$$

7. Exclude the selected wave condition $A_{s,t}$ from the matrix A_{NS} .
8. Reorder the wave conditions in the database V according to their assigned numbers in matrix A_s .

4.2.4. Monte Carlo Markov Chain with Repetition Sequencing

The Monte Carlo Markov Chain with repetition sequencing (MCMCR) has the same principle as the MCMC sequencing. However, instead of excluding the selected wave condition $A_{s,t}$ immediately, it allows the wave case to be repeated NR times, where NR is the number of repetitions of the reduced wave climate (see Section 4.3). Hence, in this sequencing method, the reduced wave climate is not entirely repeated, but the wave conditions are allowed to persist NR times. The MCMCR sequencing was performed with 5 repetitions.

4.3. Wave Climate Duration

The wave climate duration indicates the timescale for which the reduced wave climate is applied. The durations that we tested are listed in Table 5. These values are defined according to [2]. The duration of the reduced wave climate (T_{wc}) is determined by the sum of the durations of its conditions (see Equation (15)).

$$T_{wc} = \sum_{j=1}^k T_{wc,j} = \sum_{j=1}^k f_{rep,j} \frac{T_R}{N_R} \quad (15)$$

where T_R = input reduction period (for Noordwijk, $T_R = 3.3$ years), and N_R = number of repetitions of the reduced wave climate. Ideally, the reduced wave climate should resemble the natural variability of the full wave climate. The duration of each representative wave condition should be long enough for the morphology to adjust to the hydrodynamic conditions, though not too

long to prevent unrealistic irreversible morphological disturbances. The duration of the wave climate is associated with the required computational time through the number of transitions, $NoT = NoC * N_R - 1$, where $NoC =$ number of reduced wave conditions. The higher the number of transitions, the higher the computational demand as models need to spin up between wave conditions ([2]).

Table 5. Simulated wave climate durations.

Duration of Wave Climate (days)	Number of Repetitions
$T_{wc} = 1205$	$N_R = 1$
$T_{wc} = 602$	$N_R = 2$
$T_{wc} = 401$	$N_R = 3$
$T_{wc} = 301$	$N_R = 4$
$T_{wc} = 241$	$N_R = 5$
$T_{wc} = 134$	$N_R = 6$

5. Results

5.1. Performance Evaluation of Input Reduction Methods

The performance of the IR-variants in terms of average cumulative skill score is presented in Figure 6. The CLTCM and MDA methods have no performance score as the selected representative wave conditions resulted in such unrealistic morphological changes that the model simulations crashed. Therefore, these two methods are not considered suitable for input reduction. For the remaining 34 variants, we separately discuss the results for the binning (Section 5.1.1) and clustering (Section 5.1.2) methods and reflect on the influence of the duration of the reduced wave climate.

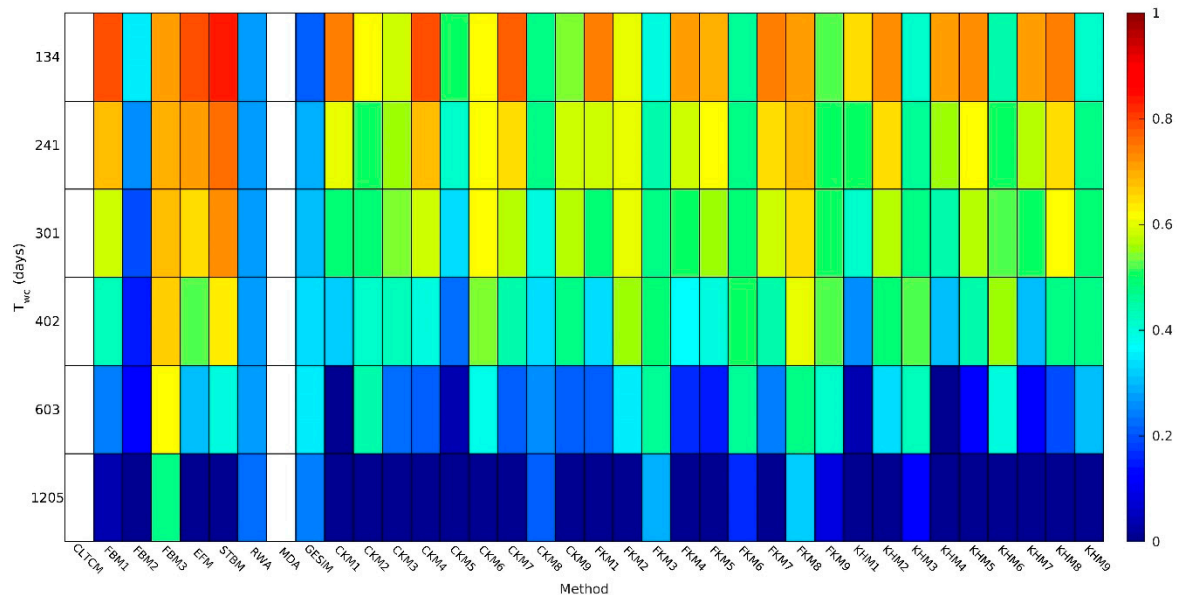


Figure 6. Average cumulative skill score of the five random sequences of the simulated methods.

5.1.1. Binning Methods

Overall, the binning methods perform better than the clustering methods (see Figure 6). In terms of skill score, the STBM and EFM are the most promising methods. The STBM performs best both in terms of skill score (see Figure 6) and the modeled morphological evolution (see Figure 7). The better performance of the STBM is probably related to the bin definition: The STBM weighs both wave height and wave direction, while the EFM only weighs wave height. Moreover, the STBM does not allow opposite wave contributions to the sediment transport to average out (its directional bins definition

starts from the shore-normal wave angle) while the EFM does. Therefore, the STBM is selected as the most promising method to test the influence of the number of wave conditions, sequencing, and wave climate duration on the performance.

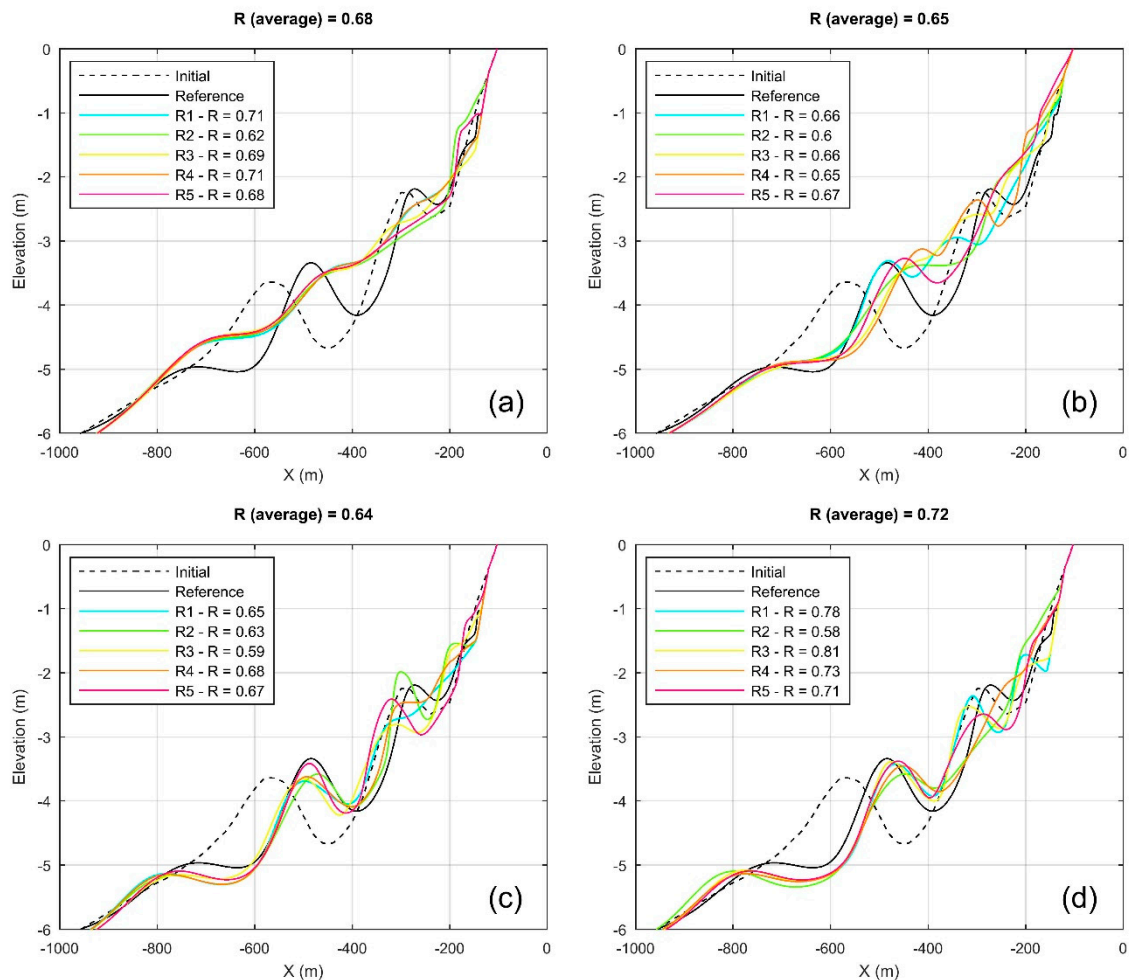


Figure 7. Initial and final profiles with the reduced wave climate of methods Fixed Bins method 3 (FBM3) (a), Fuzzy K-Means method 8 (FKM8) (b), Energy Flux method (EFM) (c), and Sediment Transport Bins method (STBM) (d) with $k = 12$ and $T_{wc} = 301$ days. The black lines are the initial (dashed) and final (solid) profiles from the brute force model. The colored lines are the final profiles of each random sequence.

The performance of the input reduction increases for shorter wave climate durations. The skill is poor when the reduced wave climate is applied for the total duration of the brute force model ($T_{wc} = 1205$ days). This is attributed to unrealistically long durations for extreme wave conditions which result in unrealistic and irreversible morphological changes. We selected $T_{wc} = 301$ days (i.e., approximately a yearly timescale) as the optimal balance between performance (i.e., reflecting the wave climate variability) and computational effort (i.e., number of transitions).

Among the Fixed Bins methods, FBM1 has relatively good results, but only when the reduced wave climate is repeated often ($T_{wc} = 301$ and smaller). FBM2 presents very poor performance as it does not select low wave height cases due to the weighting function. FBM3 shows consistent good performance for all durations of wave climate. However, its morphological response is very poor (see Figure 7a). This is a result of the skill score rewarding predictions that underestimate the overall magnitude of bed changes [12]. The RWA performs poorly; the average conditions of the seasons tend to be similar, resulting in a poor selection of representative wave conditions.

5.1.2. Clustering Methods

The clustering methods perform generally worse than the most promising binning methods (i.e., EFM and STBM). The clustering methods rely primarily on the frequency of occurrence of the observations, which leads to an over-representation of frequently occurring low wave conditions in the selection of representative waves. Since the morphology is highly dependent on energetic conditions, this reduces the performance of the clustering methods. Of all clustering methods, the Crisp k-means methods (CKM) tend to perform better than the others. Among the CKMs, the cluster initiation MDA performs slightly better than K-means++ or Fixed Bins. The results are best with S_y as the input variable instead of H_{rms} or $H_{rms}^{2.5}$.

For the Fuzzy k-means and K-harmonic means methods, the cluster initiation of Fixed Bins shows the best performance. These methods performed better with $H_{rms}^{2.5}$ as the input variable, because there is a balance between the weighting function and the high dependency on the dense cloud of observations intrinsic of these methods. The weighting function leads to a selection of centroids with higher wave height, while the dependency on the observation's frequency of occurrence leads to a selection of centroids with low wave height. When using S_y as the input variable, these methods have lower skill due to the absence of an inverse function for the sediment transport used in this study (i.e., the nearest wave condition to the centroid is selected), which results in a poor selection of representative wave conditions.

FKM8 shows similar skill scores to the EFM, but the morphological response of FKM8 is evidently worse than EFM (see Figure 7b,c, respectively). This is a result of the skill score rewarding predictions that underestimate the overall magnitude of bed changes [12]. GESIM also performs poorly because it does not aggregate the observations into clusters well: Once a cluster reaches the limit of sediment transport, it 'closes' and the observations closer to this cluster will be aggregated into another one which might be relatively far from the observations. This could lead to a poor selection of representative wave conditions since they are defined as the average of the observations within a cluster.

Additionally, most of the methods that do not present the pattern of improvement of skill score with decreasing duration of the wave climate are associated with input variables $H_{rms}^{2.5}$ and S_y (CKM5, CKM8, FKM3, FKM6, FKM9, KHM3, KHM6). When the selection of the representative wave conditions is initially poor, decreasing T_{wc} does not improve the performance of the method. Whereas when the selection is initially reasonably good, decreasing T_{wc} can improve the performance of the method.

5.2. Performance Evaluation of Input Reduction Settings

The effects of the input reductions settings (i.e., number of wave conditions, sequencing method, and wave climate duration) are only evaluated for the STBM as the most promising IR-method. The skill scores for simulations with a different number of conditions and sequencing methods are shown in Figure 8. Note that, except for the Markov Chain Sequencing (S2), that does not contain randomness, we used the mean skill score of five random replicates for the other sequencing methods. The Monte Carlo Markov Chain Sequencing with repetition (S4) does not apply for $T_{wc} = 1205$ days, as a repetition of wave conditions is not possible when applying it on the full timescale of the reduction period.

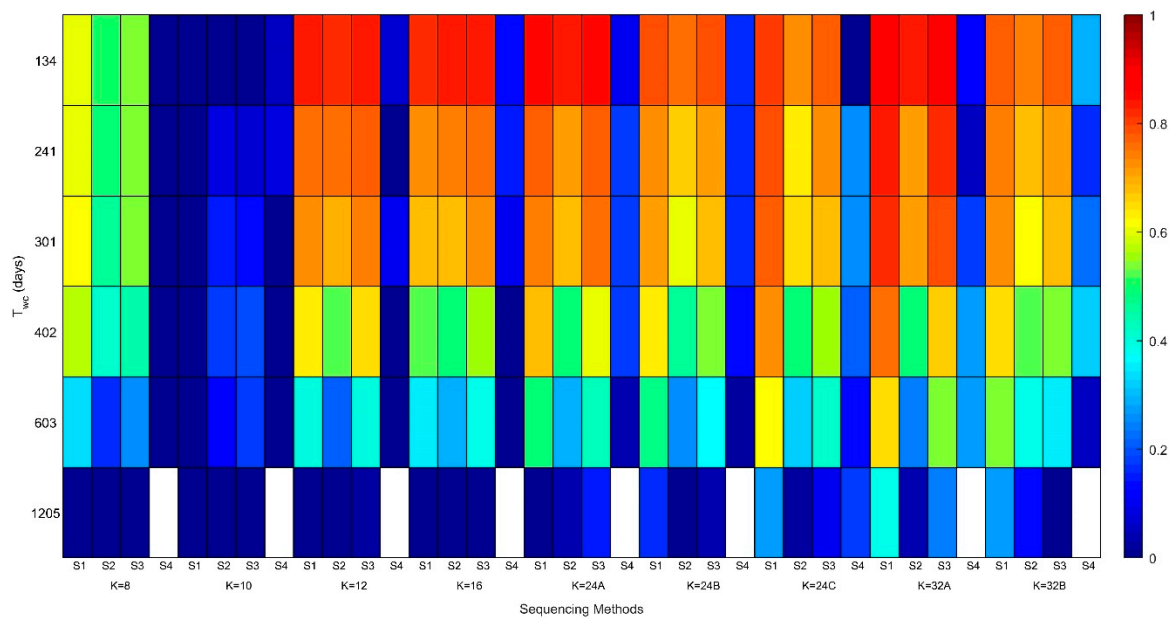


Figure 8. Average cumulative skill score of the five replicates of the different sequencing and number of cases of the Sediment Transport Bins method.

5.2.1. Number of Wave Conditions in Reduced Climate

Overall, the input reduction for $k = 32A$ has the best performance in terms of skill score, whereas $k = 10$ has the worst performance. The performance appears to be related to the resolution in directional bins: $k = 32A$ has the highest number of directional bins (i.e., eight), whereas $k = 10$ has only two directional bins. Increasing the number of cases does not necessarily imply a substantial improvement in skill score, except when the number of directional bins is increased considerably, such as for $k = 32A$ and $k = 24C$. The influence of the wave height on the longshore transport and morphology, even though non-linear, is always proportional. However, the influence of the wave direction on the longshore sediment transport is sinusoidal (i.e., fluctuating around the angle of maximum transport) and, hence, not proportionally increasing or decreasing with wave angle. Therefore, the resolution in directional space appears to be more important than the wave height. For the Noordwijk case in particular, the importance of the directional bins is also related to the strong influence of the cross-shore distribution of the longshore sediment transport on the inter-annual bar morphology [10].

5.2.2. Sequencing of Wave Conditions

Although the random and MCMC sequencing methods generally have similar skill scores, visual inspection of the temporal evolution of the cross-shore profile indicated that random sequencing yields the best results. The MC and MCMCR methods perform poorly, because they tend to aggregate calm conditions at the beginning of the simulation and energetic conditions at the end. This aggregation occurs because the highest probabilities of the Markov Chain transitions remain on the same state. Therefore, the methods that introduce randomness in the sequencing tend to better resemble the natural variations on the wave climate for the case study.

5.2.3. Duration of Reduced Wave Climate

The influence of the duration of the reduced wave climate was further investigated by comparing the STBM with random sequencing for $k = 12$ and $k = 32A$. When applying the reduced wave climate on a smaller timescale (e.g., 134 days), increasing the number of representative cases does not result in much improvement of morphological evolution (see Figure 9). Although the morphological evolution is more in line with the brute force simulation for $k = 32A$ than for $k = 12$ with the same T_{wc} , the number of transitions is much higher for a higher number of wave conditions (see Table 6). Hence, the

slight performance improvement has relatively large computational costs. Therefore, $k = 12$ seems to be more appropriate than $k = 32A$. Furthermore, the morphological evolution of the nearshore bars is better represented by a reduced climate with less representative wave conditions applied for a shorter duration ($k = 12$ with $T_{wc} = 134$ days) than with more representative wave conditions applied for a longer duration ($k = 32A$ with $T_{wc} = 301$ days). Therefore, the wave climate duration is found to be more important for the performance than the number of wave conditions in the reduced climate.

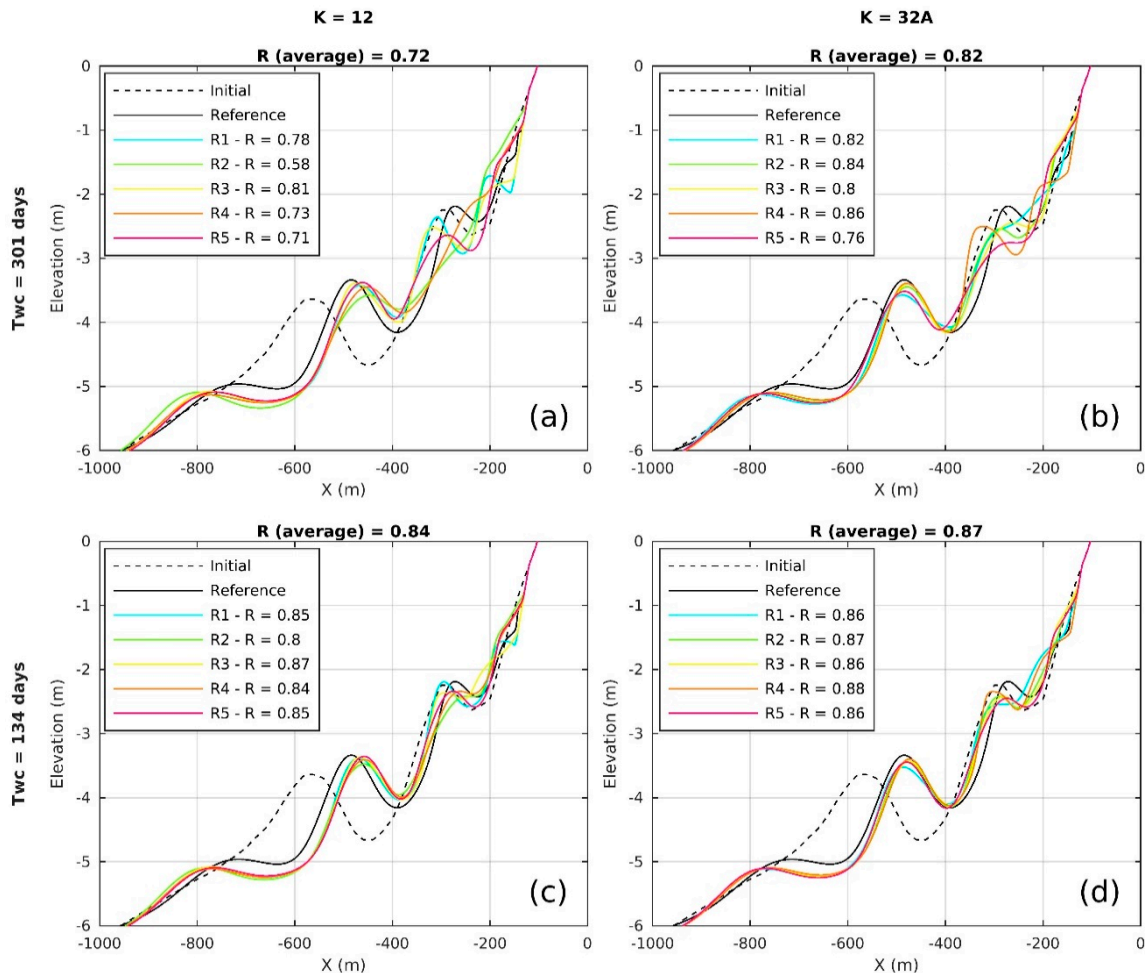


Figure 9. Initial and final profiles with the reduced wave climate of the STBM, random sequencing, $k = 12$ and $T_{wc} = 301$ days (a), $k = 32A$ and $T_{wc} = 301$ days (b), $k = 12$ and $T_{wc} = 134$ days (c), and $k = 32A$ and $T_{wc} = 134$ days (d). The black lines are the initial (dashed) and final (solid) profiles from the brute force model. The colored lines are the final profiles of each random sequence.

Table 6. Number of cases, duration of wave climate, and number of transitions.

Number of Cases (k)	Duration of Wave Climate (T_{wc})	Number of Transitions (NoT)
$k = 12$	$T_{wc} = 301$ ($NR = 4$)	$NoT = 47$
$k = 32 A$	$T_{wc} = 301$ ($NR = 4$)	$NoT = 127$
$k = 12$	$T_{wc} = 134$ ($NR = 9$)	$NoT = 107$
$k = 32 A$	$T_{wc} = 134$ ($NR = 9$)	$NoT = 287$

5.3. Validation with Anmok Beach

The most promising input reduction setup for Noordwijk (i.e., STBM, $k = 12$, random sequencing, and $T_{wc} = 134$ days), was applied to Anmok beach in South Korea (Figure 1). The result show that the final profiles are well represented in amplitude with slight errors in phase (see Figure 10). The

replicate R1 presents small scale undulations in the profile due to the coarse sediment characteristic of the profile. Despite these instabilities, the overall performance of the input reduction in Anmok is not impaired. The skill scores for Anmok are smaller than the ones for Noordwijk. The average skill score is 0.64, with maximum and minimum skill score of 0.75 and 0.57, respectively. This could be due to the smaller timescale analyzed in Anmok ($T_R = 0.8$ years) since in the beginning of the simulation the morphological variation is small causing larger errors and very low skill scores (see Equation (1)), thus, influencing more the cumulative skill score of shorter periods [12]. Nevertheless, the results of the validation were considered satisfactory.

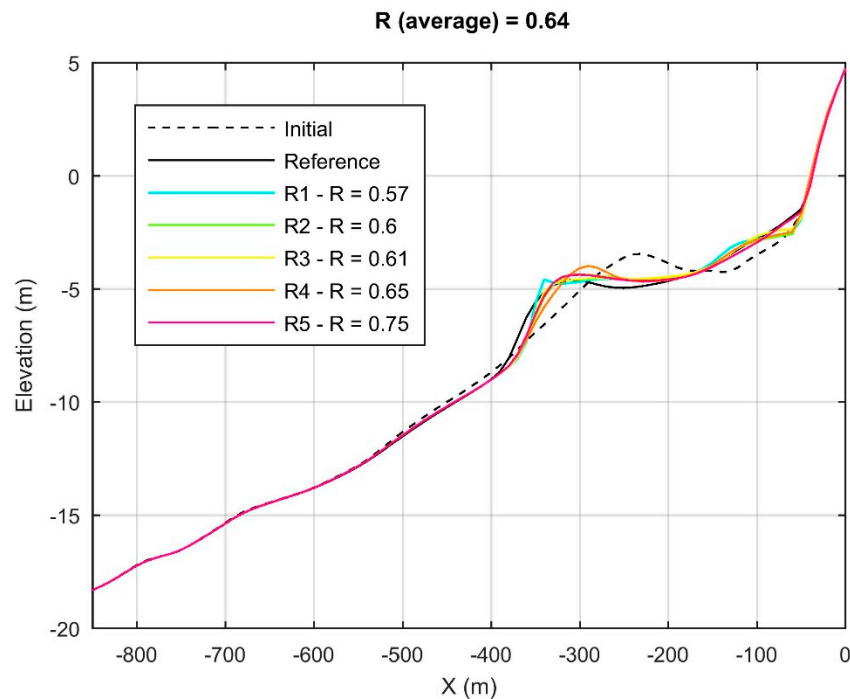


Figure 10. Initial and final profiles with the reduced wave climate of the SBM, random sequencing, $k = 12$ and $T_{wc} = 98$ days. The black lines are the initial (dashed) and final (solid) profiles from the brute force model. The colored lines are the final profiles of each random sequence.

6. Discussion

A good selection of representative wave conditions for morphology should balance mild and energetic conditions as well as direction variability while prioritizing directions that contribute the most to the sediment transport. In our assessment, we found that binning methods perform better than clustering methods. Among the binning methods, the ones that split the wave conditions into bins with equal weight performed better than the ones that split the wave conditions arbitrarily into bins, as long as the reduced wave climate is not very detailed (e.g., $k < 16$, consistent with [2]). On the other hand, [3] found that the EFM performed better than the CERC (Coastal Engineering Research Center) method proposed by them, which is analogous to the STBM used in this study but with the longshore sediment transport calculated by the CERC formula [24]. The difference in the findings of [3] and the present study could be related to the incident wave angle that is not considered in the CERC formula. Therefore, positive and negative transport contributions can cancel themselves out. Yet, the EFM is the second-best method in this study. Note that in the STBM, sediment transport rates obtained from the brute force simulations were used as input, which commonly is not available. In this case, we recommend the use of sediment transport formulas considering different coast angles or other proxies, such as the energy flux. The clustering methods did not perform well because of their high dependency on the most recurrent observations. For Noordwijk, these were the mild wave conditions,

resulting in a lack of energetic conditions. However, for very energetic coasts where the occurrence of mild conditions is not dominant over energetic conditions, the clustering methods may perform better.

The sequencing of wave conditions influences the morphological response of the simulations considerably. The random sequencing showed the best morphological response for the cases studies in this paper since randomly ordered reduced wave time series retained a higher variability than the other methods that use statistical information through Markov Chain probabilities. This is in agreement with the results of [2], who found that randomly ordered synthetic time series performed better than systematic sequencing of wave conditions, such as ascending or descending wave heights combined with wave angles towards positive and negative directions. Despite its good performance, the random sequencing has the drawback that it is completely random without any user control. Furthermore, random sequencing has its limitations since it highly depends on the initial condition (i.e., the initial profile). For instance, a winter profile evolves differently than a summer profile with the same sequence of wave conditions. Since in Noordwijk and Anmok the bar dynamics do not seem to be related to specific storm events, the chronology is limitedly relevant, and random sequencing can be applied. However, [2] found that in Hasaki, where chronology is important, random ordering of synthetic time series did not represent the inter-annual bar evolution very well. For such cases, other sequencing methods may perform better.

Regarding the number of representative wave conditions in the reduced wave climate, we found that $k = 12$ is a good quantity of representative wave conditions given that the duration of the wave climate is of the order of 100 days. This aligns with [3] who indicated $k = 12$ as an optimal quantity. Additionally, $k = 12$ is in agreement with the commonly applied wave climates in morphodynamic simulations which typically make use of about 10 waves conditions [5,25]. The wave climate duration turned out to be a very important aspect of wave climate reduction for morphological applications. This agrees with [2]. The same wave condition applied on different timescales will likely give rise to distinct morphology. Generally, decreasing the reduced wave climate duration improves the performance. The present analysis has as the lower limit. A further decrease on T_{wc} would not necessarily improve the skill of the reduced models even further. There is a lower limit of the wave climate duration associated with the response of morphology to the hydrodynamic forcing. If the duration of the wave climate is too short, there is not enough time for the morphology to adjust to the hydrodynamic conditions. This is not observed in the results of this study because the simulated durations of the wave climate were well above the lower limit, which is around 10 to 20 days according to [2]. Additionally, a further decrease on T_{wc} implies loss of applicability since the number of transitions and, thus, computational time increase when the duration of the reduce wave climate decreases.

In this study, we used the Unibest-TC profile model, due to its reduced computational time that allowed to run the considerable amount of simulations required by our methodology. In reality, brute force computations are feasible with this model. Therefore, input reduction techniques are strictly not necessary. Moreover, in Noordwijk, the alongshore variability is small, so a 1D domain is acceptable. In Anmok beach alongshore variability is present and affects the local morphodynamics. However, the changes in alongshore positions of the crescentic bars are very slow compared to the cross-shore evolution allowing for a 1D approach. For larger timescales, this is not expected to be valid. The findings of this study can still be used as initial guidelines when performing input reduction with different models and domains even for areas where alongshore variability is important.

7. Conclusions

In this paper, the performance of 36 variants of wave input reduction (IR) methods in modeling the interannual sandbar evolution was evaluated. The selection of the proper settings for wave-IR is a balance between the resemblance of the natural variability of the full dataset and computational effort. This study provided insights into the methods and settings that are most promising to reduce computational effort at limited performance loss. The results showed that the Sediment Transport Bins method has the best performance of all 36 methods. Generally, binning methods perform better

than clustering methods. Binning methods using weighted bins based on sediment transport proxies (i.e., longshore sediment transport or energy flux) perform better than those based on wave statistics only. The performance improves for an increased number of representative wave conditions at the expense of less reduction in computational time. Furthermore, a higher resolution in wave direction bins performs better than a higher resolution in bins for the wave height or sediment transport proxy. In terms of sequencing, random sequencing yielded the highest performance for the case studies. However, this is probably related to the limited importance of wave chronology in the case studies analyzed in this paper and, hence, may be different in other case studies. Finally, the performance is sensitive to the duration of the wave climate and, hence, its number of repetitions. The performance is better for a reduced wave climate with fewer representative wave conditions and a higher number of repetitions on a short timescale than for a more detailed wave climate applied on a longer timescale. The insights of this study may help coastal practitioners in performing input reduction more efficiently and effectively in practice.

Author Contributions: The majority of the modeling, analysis, and writing were performed by the first author B.d.Q. All co-authors contributed to the contents of the paper through analysis, discussion and/or writing. W.d.B., D.-J.W., and J.Y. initiated the study. W.d.B. and D.-J.W. contributed significantly to the writing of the paper. J.Y. provide funding and data. F.S., S.C., and D.O. made major contributions to the modeling and performance analysis. A.R. provided input to the setup of the research and reviewed the draft manuscript.

Funding: This research was funded by the research project titled Development of Coastal Erosion Control Technology (or CoMIDAS) funded by the Korean Ministry of Oceans and Fisheries and the Deltares strategic research program Future-proof Coastal Infrastructure and Offshore Renewable Energy.

Acknowledgments: The authors would like to thank Rijkswaterstaat and KIOST (PE99742) for providing the necessary survey and wave data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. De Vriend, H.J.; Capobianco, M.; Chesher, T.; de Swart, H.E.; Latteux, B.; Stive, M.J.F. Approaches to long-term modelling of coastal morphology: A review. *Coast. Eng.* **1993**, *21*, 225–269. [[CrossRef](#)]
2. Walstra, D.J.R.; Hoekstra, R.; Tonnon, P.K.; Ruessink, B.G. Input reduction for long-term morphodynamic simulations in wave-dominated coastal settings. *Coast. Eng.* **2013**, *77*, 57–70. [[CrossRef](#)]
3. Benedet, L.; Dobrochinski, J.P.F.; Walstra, D.J.R.; Klein, A.H.F.; Ranasinghe, R. A morphological modeling study to compare different methods of wave climate schematization and evaluate strategies to reduce erosion losses from a beach nourishment project. *Coast. Eng.* **2016**, *112*, 69–86. [[CrossRef](#)]
4. Southgate, H.N. The effects of wave chronology on medium and long term coastal morphology. *Coast. Eng.* **1995**, *26*, 251–270. [[CrossRef](#)]
5. Van Duin, M.J.P.; Wiersma, N.R.; Walstra, D.J.R.; van Rijn, L.C.; Stive, M.J.F. Nourishing the shoreface: Observations and hindcasting of the Egmond case, The Netherlands. *Coast. Eng.* **2004**, *51*, 813–837. [[CrossRef](#)]
6. Brown, J.M.; Davies, A.G. Methods for medium-term prediction of the net sediment transport by waves and currents in complex coastal regions. *Cont. Shelf Res.* **2009**, *29*, 1502–1514. [[CrossRef](#)]
7. Lesser, G.R. An Approach to Medium-Term Coastal Morphological Modelling. Ph.D. Thesis, UNESCO-IHE & TUDelft, Delft, The Netherlands, 4 June 2009.
8. Roelvink, J.A.; Reniers, A.J.H.M. *A Guide to Modelling Coastal Morphology*; World Scientific Publishing Co. Pte. Ltd.: 5 Toh Tuck Link, Singapore, 2012; pp. 200–215.
9. Antolinez, J.A.A.; Mendez, F.J.; Camus, P.; Vitousek, S.; Gonzales, E.M.; Ruggiero, P.; Barnard, P. A multiscale climate emulator for long-term morphodynamics (MUSCLE-morpho). *J. Geophys. Res. Ocean.* **2016**, *775*–791. [[CrossRef](#)]
10. Walstra, D.J.R.; Reniers, A.J.H.M.; Ranasinghe, R.; Roelvink, J.A.; Ruessink, B.G. On bar growth and decay during interannual net offshore migration. *Coast. Eng.* **2012**, *60*, 190–200. [[CrossRef](#)]
11. Ruessink, B.G.; Kuriyama, Y.; Reniers, A.J.H.M.; Roelvink, J.A.; Walstra, D.J.R. Modeling cross-shore sandbar behavior on the timescale of weeks. *J. Geophys. Res. Earth Surf.* **2007**, *112*, 1–15. [[CrossRef](#)]

12. Bosboom, J.; Reniers, A.J.H.M.; Luijendijk, A.P. On the perception of morphodynamic model skill. *Coast. Eng.* **2014**, *94*, 112–125. [[CrossRef](#)]
13. Olij, D.J.C. Wave Climate Reduction for Medium Term Process Based Morphodynamic Simulations with Application to the Durban Coast. Master's Thesis, Delft University of Technology, Delft, The Netherlands, 2015; 131p.
14. Arthur, D.; Vassilvitskii, S.K. Means++: The Advantages of Careful Seeding. *Proc. Annu. ACM-SIAM Symp. Discret. Algor.* **2007**, *8*, 1025–1027.
15. Kennard, R.W.; Stone, L.A. Computer Aided Design of Experiments. *Technometrics* **1969**, *11*, 137–148. [[CrossRef](#)]
16. Willett, P. Molecular Diversity Techniques for Chemical Databases. Available online: <http://www.informationr.net/ir/2-3/paper19.html> (accessed on 11 November 2018).
17. Polinsky, A.; Feinstein, R.D.; Shi, S.; Kuki, A. Molecular Diversity and Combinatorial Chemistry. In *Libraries and Drug Discovery*; American Chemical Society: Washington, DC, USA, 1996; pp. 219–232.
18. Camus, P.; Mendez, F.J.; Medina, R.; Cofiño, A.S. Analysis of clustering and selection algorithms for the study of multivariate wave climate. *Coast. Eng.* **2011**, *58*, 453–462. [[CrossRef](#)]
19. Macqueen, J. Some methods for classification and analysis of multivariate observations. *Proc. Fifth Berkeley Symp. Math. Stat. Probab.* **1967**, *1*, 281–297.
20. Hastie, T.; Tibshirani, R.; Friedman, J. The Elements of Statistical Learning. *Math. Intell.* **2001**, *27*, 83–85.
21. Chang, C.; Lai, J.; Jeng, M. A Fuzzy K-means Clustering Algorithm Using Cluster Center Displacement. *J. Inf. Sci. Eng.* **2011**, *1009*, 995–1009.
22. Zhang, B.; Hsu, M.; Dayal, U. *Clustering Algorithm K-Harmonic Means—A Data Clustering Algorithm*; Technical Report HPL-1999-124; Hewlett-Packard Labs: Bristol, UK, 1999.
23. Zhang, B. *Generalized K-Harmonic Means—Boosting in Unsupervised Learning*; HP Labs Technical Report HPL-2000-137; HP Laboratories: Palo Alto, CA, USA, 2000.
24. USACE. *Shore Protection Manual*; USACE: Washington, DC, USA, 1984.
25. Grunnet, N.M.; Ruessink, B.G. Morphodynamic response of nearshore bars to a shoreface nourishment. *Coast. Eng.* **2005**, *52*, 119–137. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).