Predicting Dominance of Sand Transport by Waves, Tides, and Their Interactions on Sandy Continental Shelves

E. V. King1,2, D. C. Conley1, G. Masselink1, and N. Leonardi3

1Coastal Processes Research Group, School of Biological and Marine Sciences, University of Plymouth, Plymouth, UK, 2Scottish Environment Protection Agency, Glasgow, UK, 3Department of Geography and Planning, School of Environmental Sciences, University of Liverpool, Liverpool, UK

Abstract  Waves and tidal currents resuspend and transport shelf sediments, influencing sediment distributions and bedform morphology with implications for various disciplines including benthic habitats, marine operations, and marine spatial planning. Shelf-scale assessments of wave-tide-dominance of sand transport tend not to fully include wave-tide interactions, which nonlinearly enhance bed shear stress and apparent roughness, change the current profile, modulate wave forcing, and can dominate net sand transport. Assessment of the contribution of wave-tide interactions to net sand transport requires computationally/labor intensive coupled numerical modeling, making comparison between regions or climate conditions challenging. Using the Northwest European Shelf, we show the dominant forcing mode and potential magnitude of net sand transport is predictable from readily available, uncoupled wave, tide, and morphological data in a computationally efficient manner using a k-Nearest Neighbor algorithm. Shelf areas exhibit different dominant forcing modes for similar wave exceedance conditions, related to differences in depth, grain size, tide range, and wave exposure. Wave-tide interactions dominate across most areas in energetic combined conditions. Meso-macro tidal areas exhibit tide-dominance while shallow, fine-grained, microtidal regions show wave-dominance over a statistically representative year, with wave-tide interactions dominating extensively >30 m depth. Sediment transport mode strongly affects seabed morphology. Sand wave geometry varies significantly between predicted dominance classes with increased wave length and asymmetry, and decreased height, for increasing wave-dominance. This approach efficiently indicates where simple noninteractive wave and tide processes may be sufficient for modeling sediment transport, and enables efficient interregional comparisons and sensitivity testing to changing climate conditions with applications globally.

Plain Language Summary  The transport of sand across the marine environment is important to understand, as it influences the fate of sediments, pollutants, and can affect seabed habitats. In marine settings, sand transport results primarily from the forces exerted by the tide and waves, and these forces interact in a nonlinear way. Numerical models can be used to calculate sand transport rates, however to understand what processes are driving sand transport under different conditions and across large areas requires complex modeling which takes time and resources. Here, we show we can predict the magnitude and dominant forcing using a machine learning algorithm trained with readily available data for the Northwest European Shelf. Different forces drive net sand transport depending on water depth, sand grain size, tide range, and wave exposure. Areas with the largest tides are dominated by tidal forces over a year, while shallow areas with fine sand which are exposed to energetic waves are dominated by wave forces. We show that sand waves on the seabed increase in length, become more asymmetrical, and decrease in height when waves dominate sand transport during storms.

1. Introduction

Residual (net) sediment transport patterns influence the transport and fate of continental shelf sediments, influencing sediment distributions and morphological evolution (Harris & Collins, 1991; King et al., 2019; Leonardi & Plater, 2017; Pingree & Griffiths, 1979; Pingree & Le Cann, 1989; Stride, 1963; van der Molen, 2002; Xu et al., 2016; Zhang et al., 2016). Waves and tidal currents result in resuspension and transport of shelf sediments (Carter & Heath, 1975; Pattiaratchi & Collins, 1988; Thompson et al., 2019), influencing
sand wave morphology (Damen et al., 2018a, 2018b; Wang et al., 2019) with implications for marine spatial planning of pipelines and cables for windfarms and offshore renewable energy (Cheng et al., 2020; Németh et al., 2003; Roetert et al., 2017), mobilization and dispersal of sediments and contaminants (e.g., due to dredge disposal and bottom trawling; Cieśliakiewicz et al., 2018; Mengual et al., 2019; Uncles et al., 2020), and the fate of shoreface nourishments (Luijendijk et al., 2017). Stress waves and sand transport driven by tides and waves influence benthic communities through disturbance, while also acting as a vector for recolonization (Aldridge et al., 2015; Bricenho et al., 2015; Dernie et al., 2003; Hall, 1994; Harris, 2014; Levin, 1995; Reiss et al., 2010). The relative impact of wave and tidal forcing influences sand wave morphology and migration rates (Campmans et al., 2018a, 2018b; Damen et al., 2018a, 2018b; Van Dijk & Kleinbans, 2005), causing potential disturbance and affecting the distribution of benthic communities (Damveld et al., 2018, 2020; Harris, 2014). Predictive habitat suitability modeling requires an understanding of physical disturbance regimes and knowledge of the dominant drivers of sand transport at the shelf scale is important (Harris, 2014).

Assessments of the relative impact of waves and tidal currents on the bed across sandy continental shelves have been conducted. Bricenho et al. (2015) mapped the relative impact of tides and storm events at the bed across the NW European Shelf over a 10-year period. South West exposed coasts and shallow water areas were found to be most at risk from large waves and thus most likely to show wave-dominated transport, and modeling suggests the maximum benthic force is wave dominated (Bricenho et al., 2015). The detailed distribution of physical disturbance shows a complex relationship between depth, tidal stress, wave fetch, and grain size, with large uncertainty (Aldridge et al., 2015). Porter-Smith et al. (2004) classify the Australian continental shelf based on sediment threshold of motion exceedance from tidal currents and swell waves with classes ranging through waves-only, wave-dominated, mixed, tide-dominated, and tide-only. Van der Molen (2002) considers the relative impact of waves, winds, and tides on sand transport in the Southern North Sea. However, at present, shelf-scale analyses of dominant forcing modes for sand transport do not consider wave-tide interactions. Wave-tide interactions (WTI) nonlinearly enhance bed shear stress and apparent roughness due to interaction between wave and tidal bottom boundary layers, influence the vertical current profile, and modulate wave forcing through tidal elevation changes (Fredsoe, 1984; Grant & Madsen, 1979, 1986; Hopkins et al., 2015; Kemp & Simmons, 1982, 1983; Klopman, 1994; Nielsen, 1992; Olabarrieta et al., 2010; Tambroni et al., 2015; Umeyama, 2005).

Boundary layer processes dominated by WTI are fundamentally different from those dominated by either waves or tides, and WTI can dominate net sand transport across large areas of the shelf over a tidal cycle (King et al., 2019). Analyses excluding WTI may underestimate net sand transport under combined wave and tide conditions where WTI can dominate. A classification scheme was proposed by King et al. (2019) for net sand transport per tidal cycle to account for contributions of waves, tides, and WTI (accounting for radiation stresses, Stoko’s drift, enhanced bottom-friction and bed shear stress, refraction, current-induced Doppler shift, tidal modulation of wave heights and wave blocking); however, this currently requires computationally expensive coupled numerical modeling to assess. A computationally efficient method to assess the dominant sand transport mode and magnitude will enable efficient inter-regional comparison of the role of waves, tides, and WTI on sand transport at scale and under varied or changing climate forcing. This enables efficient assessment of where simple noninteractive wave and tide processes may be sufficient to model sand transport, particularly relevant where application of a model or parameterization is predicated on dominance of waves (e.g., parameterizations of headland bypassing; King et al., Under Review; McCarroll et al., 2021), or tides (e.g., models of sand wave morphological evolution in tide-dominated environments; Besio et al., 2008). It also enables efficient assessment of the role of combined wave and tidal processes on seafloor morphology, such as by comparing dominant processes with observed sand wave geometries (e.g., Damen et al., 2018a, 2018b). It is therefore beneficial to develop a means to quickly assess the dominant sand transport mode on sandy continental shelves without the need for computationally expensive numerical modeling. This study will consider a homogeneous sand bed, and its implications in relation to mixed sediments, including the influence of graded sediment transport, hiding-exposure effects, and bio-stabilization (McCarron et al., 2019; Thompson et al., 2019; Van Oyen & Blondeaux, 2009), are explored in Section 4.2.
This study aims to apply a data-driven method to predict the dominant sand transport drivers and sand transport magnitude on sandy continental shelves using the classification scheme of King et al., (2019). This will allow assessment of the importance of WTI to sand transport on sandy continental shelves with a computationally efficient method versus fully coupled hydrodynamic modeling. Machine learning algorithms are being increasingly used for data-driven applications in the geosciences to take advantage of increased availability of large, complex, and multivariate data sets over variable spatial resolutions and coverages, enabling efficient and repeatable analyses to be conducted (Lary et al., 2016; Kanevski et al., 2009). The k-Nearest Neighbor (kNN) algorithm has been employed for prediction of seafloor properties in the geosciences including seafloor total organic carbon (Lee et al., 2019), isochore thickness (Lee et al., 2020), and sediment accumulation rates (Restrepo et al., 2020). Other applications of machine learning algorithms in the geosciences include predictions of seafloor sediment porosity (Martin et al., 2015) and seafloor fluid expulsion anomalies (Phrampus et al., 2020). The kNN algorithm is one of the simplest machine learning algorithms, and can be used in geospatial classification prediction (Kanevski et al., 2009).

To achieve the aim of this study, we will pursue the following objectives: (a) determine a list of readily available environmental and morphological variables with predictive capacity for the dominant sand transport mode and order of magnitude; (b) use results of sand transport rates obtained through a validated numerical model to train a kNN classifier for dominant sand transport class and order of magnitude; (c) collate environmental and morphological predictors across a sandy continental shelf with highly varied environmental conditions; and (d) use the trained kNN classifier to assess the dominant transport mode and sand transport magnitude across the shelf.

2. Methods

2.1. Study Region

The Northwest European continental shelf (Figure 1) was selected for this study due to a combination of ready availability of environmental and morphological variables covering the entire continental shelf area, herein referred to as “shelf scale” data (Graham et al., 2018; O’Dea et al., 2012; Tonani et al., 2019; Tonani & Sautler, 2020; Wilson et al., 2018), a highly varied tidal regime ranging from macrotidal to microtidal (Pingree & Griffiths, 1979), a varied wave climate ranging from regions exposed to a potential 7000 km fetch dominated by long-period swell waves (e.g., Celtic Shelf; Collins, 1987; Draper, 1967; Scott et al., 2016) to regions sheltered from the Atlantic swell and dominated by wind-waves (e.g., Netherlands Shelf; van der Molen, 2002). This continental shelf has a predominantly sand bed with median sand fraction grain size ranging from fine to coarse sand (Figures 1b and 1c; Wilson et al., 2018). Broadly separated into sand, mud and gravel, sand comprises 82.1% of the shelf area, mud 6.4%, and gravel 11.4%, based on Folk (1954) using the data of Wilson et al. (2018). More detailed sediment classification maps of the region can be found in Mitchell et al. (2019). The shelf area has a wealth of literature examining environmental drivers of benthic disturbance (Aldridge et al., 2015; Bricheno et al., 2015; Thompson et al., 2019), sand transport (Harris & Coleman, 1998; King et al., 2019; Leonardi & Plater, 2017; Pingree & Griffiths, 1979; Uncles, 2010; van der Molen, 2002) and bedform morphodynamics (Cheng et al., 2020; Damen et al., 2018a, 2018b; Ward et al., 2015). These factors make this an ideal region to examine the performance of a method for predicting the dominant driver of sand transport at the shelf scale.

2.2. Classification Scheme

Previous modeling work by King et al. (2019) simulated net sand transport per tidal cycle across a macro-mesotidal section of the Celtic Shelf (Figure 1a) using Delft3D (Booij et al., 1999; Lesser et al., 2004) in a depth-averaged mode using the sand transport formulation of Van Rijn (2007a, 2007b). Delft3D in a depth-averaged mode has previously been used successfully to simulate sand transport processes including the impact of WTI on the inner shelf (Hansen et al., 2013; Hopkins et al., 2015; King et al., 2019; Luijendijk et al., 2017; McCarroll et al., 2018; Ridderinkhof et al., 2016). Simulations were performed for spring and neap tides and median and extreme (1% exceedance) waves from two modal directions with waves and tides simulated together and in isolation, to allow calculation of individual wave, tide, and WTI components of net sand transport.
Figure 1. Maps of (a) depth, (b) median sand fraction grain size, and (c) sand as a percentage of available sediment for the Northwest European Continental Shelf. Depths are taken from the FOAM-AMM7 model (O’Dea et al., 2012), while sediment characteristics are taken from Wilson et al. (2018). Selected shelf areas for later comparison are indicated and named in (a). The colorbar of (a) is clipped at 300 m to better illustrate the shelf. Below 300 m, contours are shown (light gray lines). The extent of the model domain of King et al. (2019), used to train the kNN model, is indicated as “Delft3D Domain”.
King et al. (2019) derived a classification scheme to categorize sand transport between wave-dominated, tide-dominated, and nonlinear-dominated, where nonlinear refers to the added nonlinear effects on sand transport due to WTI (Figure 2). Classes are determined by two ratios:

\[ R_1 = \frac{Q_T}{Q_W + Q_N} \]  

\[ R_2 = \frac{Q_W}{Q_N} \]  

\[ Q_N = Q_{ALL} - Q_W - Q_T \]

where \( R_1 \) represents the ratio of the tide-only net sand transport magnitude \( Q_T \) to the combined wave-only net transport magnitude \( Q_W \) and the component attributed to nonlinear WTI \( Q_N \). This determines the relative influence of waves (including nonlinear interactions) versus tides, determined by subtracting the tidal component from the coupled wave + tide net transport magnitude \( Q_{ALL} \). Ratio \( R_2 \) represents the relative contribution of nonlinear interactions \( Q_N \) versus waves alone \( Q_W \). This allows the contribution of tides, waves, and the added nonlinear effects due to wave-tide interactions to be quantified, visualized, and compared. This classification scheme considers net sand transport per tidal cycle, and the class can change under different combinations of wave and tidal forcing. Classification changes under different conditions qualitatively matched modeled shifts in sand transport direction (King et al., 2019; Pattiaratchi & Collins, 1988), supporting the predicted shift in the dominant mode of net sand transport.

This classification scheme results in three dominant modes of net sand transport (wave-dominated, \( W \), tide-dominated, \( T \), and nonlinear-dominated, \( N \), where the respective forcing is responsible for at least 75% of the net sand transport magnitude. When the dominant class is responsible for >50% of net sand transport, but <75%, a subdominant class is defined (noted using lowercase letters between brackets behind the dominant class). At present, this scheme requires results from coupled and uncoupled numerical simulations of net sand transport to calculate. Using the simulations of King et al. (2019), it is possible to extract the sand transport dominance class, the net sand transport magnitude, and the corresponding environmental variables for use in a predictive model. The following section will examine kNN as a classification prediction method, based on defined predictor variables, which we will apply to this classification scheme (Section 2.3).

### 2.3. K-Nearest Neighbor (kNN)

The kNN algorithm works on the principle that areas with similar conditions are likely to share the same class. The kNN algorithm requires a predictand (the variable or class we want to predict) and a set of defined
predictors (variables we have measured or estimated). The algorithm is trained on the predictor data associated with known values of the predictand. The algorithm is then used to predict unseen data where the predictand is unknown by calculating the distance to the “k” nearest neighbors in parameter space to the new data, where “k” is the number of nearest points the algorithm uses for its calculation. The implementation used in this study is included in the MATLAB Statistics and Machine Learning Toolbox (MathWorks, 2020). The predicted class is the class with the minimum estimated cost, determined as a function of the probability that the new data come from a particular class and the expected cost of misclassification for each observation. Numerous search methods exist for determining the nearest neighbors for use in the algorithm. In this study, a Kd-tree is used to perform the nearest neighbor search, saving computation time as only a subset of the distances to points need to be calculated. Data were standardized and distances calculated using a city-block distance metric (sum of absolute distances) using all available predictors with $k = 7$, as this provided optimal accuracy while minimizing the value of $k$ to avoid smoothing the data (see Table S1).

Model performance was determined using five-fold cross-validation of the training data set. This entails splitting the data set into five equal parts, and iteratively training the model on four of five parts, while validating using the fifth part by calculating the percentage of observations which was classified correctly, changing the validation to fifth each time. The final model accuracy is an average of the five cross-validation scores. This method mitigates the likelihood of overfitting (Kanevski et al., 2009; Lee et al., 2019).

The choice of predictors is motivated by data availability, physical relevance to the prediction of the net sand transport forcing mode and magnitude, spatio-temporal resolution, as well as predictive value of each potential predictor. To assess the value of individual predictors, each predictor was tested in isolation to predict the class and order of magnitude of the net sand transport. The accuracy of each predictor was then compared with the predictive accuracy of an array of random numbers, to test whether predictors had greater predictive value than random noise. This was achieved by training a kNN classifier on each predictor in isolation, and comparing the accuracy of this classifier with a classifier trained on uniformly distributed random noise. The selection of predictors, including their predictive accuracy, is described below (Section 2.4).

### 2.4. Environmental Predictors

Environmental predictors across the NW European Shelf used in this study are shown in Table 1a, including their sources and resolution (spatial, temporal) where applicable. With these data sources defined, the model scenarios conducted to generate training data are included in Table 1b, including the range of the parameters used. Modeled scenarios were conducted as described in King et al. (2019), calculating net sand transport for wave-only, tide-only, and wave + tide forcing over springs and neaps at 1 km resolution for an approx. $350 \times 240$ km region of the Celtic shelf with variable wave exposure and meso-megatidal regime. A full model description and validation is also presented therein. Additional scenarios were conducted in addition to those described in King et al. (2019) to include more intermediate wave conditions and a range of grain sizes. Mixed size fractions (e.g., sand-gravel mixtures) were not considered, and this is discussed in Section 4.2. Dominant transport classes were calculated as in Figure 2, and order of magnitude of net sand transport was determined from the coupled wave + tide simulations. Predictors for training were determined from the uncoupled simulations to ensure WTI were not included in the predictor variables, replicating the uncoupled nature of the shelf-scale models.

An example of the relationship between tidal range $TR$, maximum tidal current speed $U_{\text{max}}$, relative wave height $H_r/h$ and the sand transport dominance classes of King et al. (2019) is shown in Figure 3. The modeled $TR$ and $U_{\text{max}}$ are shown as a function of $H_r/h$ with class indicated by color (Figures 3a and 3b). Tide-dominated areas exhibit low relative wave heights and stronger tidal currents, while wave-dominated areas are the inverse. Nonlinear-dominated areas occupy the mixed energy section of the parameter space. A three-predictor kNN classifier is shown in Figure 3c, indicating the classification boundaries for relative to the three predictors: new data falling within this parameter space will be classified accordingly. This is a simplified classifier for 3D visualization, whereas the final classifier has eight dimensions (see Table 2).

Each of the eight predictors in Table 1 was tested in isolation and compared with classifications predicted by an array of random numbers to determine its predictive value. For a predictor to be accepted, it is required to have an accuracy greater than that of the random array, as in Lee et al. (2019). The predictive accuracy of...
Table 1
(a) Environmental Predictors Across the NW European Shelf; (b) Environmental Predictors and Scenarios Used in Delft3D Simulations to Generate Training Data

| Predictor Name | Symbol | Units | Source | Spatial resolution | Temporal Resolution | Interpolation | Processing |
|----------------|--------|-------|--------|--------------------|--------------------|---------------|------------|
| Significant wave height | $H_s$ | m | Tonani and Saulter (2020) | 1.5 km | 1 h | Interpolated to 7 km grid | Mean $H_s$ per tidal cycle |
| Peak period | $T_p$ | s | As $H_s$ | 1.5 km | 1 h | Interpolated to 7 km grid | Mean $T_p$ per tidal cycle |
| Wave Power | $P$ | W | As $H_s$ | 1.5 km | 1 h | Interpolated to 7 km grid | Mean $P$ per tidal cycle |
| Depth$^a$ | $h$ | m | O’Dea et al. (2012) | 7 km | - | Converted to MSL2000 | - |
| Relative wave height | $H_s/h$ | - | As $H_s$ and $h$ | 7 km | 1 h | - | Mean $H_s/h$ per tidal cycle |
| Tide range | $TR$ | m | Graham et al. (2018); O’Dea et al. (2012) | 7 km | 1 h | - | Determined per tidal cycle |
| Max tidal current | $U_{\text{max}}$ | ms$^{-1}$ | As $TR$ | 7 km | 1 h | - | Determined per tidal cycle |
| Angle between waves and currents | $\theta$ | Deg | As $H_s$ and $TR$ | 7 km | 1 h | - | Mean wave direction and max tidal current direction |
| Median grain size | $D_{50}$ | $\mu$m | Wilson et al. (2018) | 0.125° | - | Interpolated to 7 km grid | - |

| Scenario | $H_s\cap T_p$ joint exceedance probability | $H_s$ min, median, max (m) | $T_p$ min, median, max (s) | Tide condition | $TR$ min, median, max (m) | $U_{\text{max}}$ min, median, max (ms$^{-1}$) | $D_{50}$ (μm) | N° data |
|----------|------------------------------------------|-----------------------------|----------------------------|---------------|-----------------------------|---------------------------------|----------------|--------|
| 1        | 1%                                      | 0.2, 7.1, 8.5               | 5.9, 17.6, 19.0            | Springs       | 1.8, 3.0, 7.8               | 0.03, 0.7, 3.6                  | 125             | 44,861 |
| 2$^b$    | 1%                                      | 0.2, 7.1, 8.5               | 5.9, 17.6, 19.0            | Springs       | 1.8, 3.0, 7.8               | 0.03, 0.7, 3.6                  | 330             | 44,683 |
| 3        | 1%                                      | 0.2, 7.1, 8.5               | 5.9, 17.6, 19.0            | Springs       | 1.8, 3.0, 7.8               | 0.03, 0.7, 3.6                  | 750             | 43,582 |
| 4        | 1%                                      | 0.2, 7.1, 8.5               | 5.9, 17.6, 19.0            | Neaps         | 0.6, 1.2, 4.2               | 0.02, 0.3, 1.5                  | 125             | 44,566 |
| 5$^b$    | 1%                                      | 0.2, 7.1, 8.5               | 5.9, 17.6, 19.0            | Neaps         | 0.6, 1.2, 4.2               | 0.02, 0.3, 1.5                  | 330             | 43,652 |
| 6        | 1%                                      | 0.2, 7.1, 8.5               | 5.9, 17.6, 19.0            | Neaps         | 0.6, 1.2, 4.2               | 0.02, 0.3, 1.5                  | 750             | 39,972 |
| 7        | 10%                                     | 0.1, 4.1, 4.8               | 5.0, 14.9, 15.4            | Springs       | 1.8, 3.0, 7.8               | 0.03, 0.7, 3.6                  | 125             | 44,577 |
| 8        | 10%                                     | 0.1, 4.1, 4.8               | 5.0, 14.9, 15.4            | Springs       | 1.8, 3.0, 7.8               | 0.03, 0.7, 3.6                  | 330             | 44,272 |
| 9        | 10%                                     | 0.1, 4.1, 4.8               | 5.0, 14.9, 15.4            | Springs       | 1.8, 3.0, 7.8               | 0.03, 0.7, 3.6                  | 750             | 41,709 |
| 10       | 10%                                     | 0.1, 4.1, 4.8               | 5.0, 14.9, 15.4            | Neaps         | 0.6, 1.2, 4.2               | 0.02, 0.3, 1.5                  | 125             | 41,885 |
| 11       | 10%                                     | 0.1, 4.1, 4.8               | 5.0, 14.9, 15.4            | Neaps         | 0.6, 1.2, 4.2               | 0.02, 0.3, 1.5                  | 330             | 39,175 |
| 12       | 10%                                     | 0.1, 4.1, 4.8               | 5.0, 14.9, 15.4            | Neaps         | 0.6, 1.2, 4.2               | 0.02, 0.3, 1.5                  | 750             | 16,637 |
| 13       | 50%                                     | 0.1, 1.9, 2.1               | 3.4,10.5,10.8              | Springs       | 1.8, 3.0, 7.8               | 0.03, 0.7, 3.6                  | 125             | 43,380 |
| 14$^b$   | 50%                                     | 0.1, 1.9, 2.1               | 3.4,10.5,10.8              | Springs       | 1.8, 3.0, 7.8               | 0.03, 0.7, 3.6                  | 330             | 41,224 |
| 15       | 50%                                     | 0.1, 1.9, 2.1               | 3.4,10.5,10.8              | Springs       | 1.8, 3.0, 7.8               | 0.03, 0.7, 3.6                  | 750             | 30,842 |
| 16       | 50%                                     | 0.1, 1.9, 2.1               | 3.4,10.5,10.8              | Neaps         | 0.6, 1.2, 4.2               | 0.02, 0.3, 1.5                  | 125             | 13,415 |
| 17$^b$   | 50%                                     | 0.1, 1.9, 2.1               | 3.4,10.5,10.8              | Neaps         | 0.6, 1.2, 4.2               | 0.02, 0.3, 1.5                  | 330             | 10,274 |
| 18       | 50%                                     | 0.1, 1.9, 2.1               | 3.4,10.5,10.8              | Neaps         | 0.6, 1.2, 4.2               | 0.02, 0.3, 1.5                  | 750             | 5,265  |

**Summary:**

- $^a$Depth used as a predictor combined in $H_s/h$
- $^b$Scenarios described in King et al. (2019)
Figure 3.
Table 2
Predictive Accuracy of kNN Models Trained on the Environmental Predictors Compared With Calculated Dominance Classes and Order of Magnitudes From Model Data

| Test number | Variable(s)          | Symbol | Accuracy: Dominant class (King et al., 2019) % | Accuracy: Order of magnitude (OOM) % | Difference relative to random array for class % | Difference relative to random array for OOM % |
|-------------|----------------------|--------|-----------------------------------------------|-------------------------------------|-----------------------------------------------|---------------------------------------------|
| 1           | Random array         | $R_{nd}$ | 30.2                                         | 27.2                                | -                                             | -                                           |
| 2           | Significant wave height | $H_s$  | 58.3                                         | 42.3                                | +28.1                                         | +15.1                                       |
| 3           | Peak period          | $T_p$  | 49.2                                         | 27.4                                | +19.0                                         | +0.2                                        |
| 4           | Wave Power           | $P$    | 58.5                                         | 42.3                                | +28.3                                         | +15.1                                       |
| 5           | Relative wave height | $H_s/h$ | 58.2                                         | 42.1                                | +28.0                                         | +14.9                                       |
| 6           | Tide range           | $TR$   | 49.1                                         | 28.7                                | +18.9                                         | +1.5                                        |
| 7           | Max tidal current    | $U_{max}$ | 49.0                                         | 28.7                                | +18.8                                         | +1.5                                        |
| 8           | Angle between waves and currents | $\theta$ | 43.0                                         | 35.1                                | +12.8                                         | +7.9                                        |
| 9           | Median grain size    | $D_{50}$ | 9.4                                          | 24.5                                | −20.8                                         | −2.7                                        |
| 10          | All from tests 2 to 8 | -      | 60.8                                         | 44.5                                | +30.6                                         | +17.3                                       |
| 11          | As 10 + Rnd          | -      | 73.1                                         | 53.6                                | +42.9 (+12.3)$^b$                             | +26.4 (+9.1)$^b$                            |
| 12$^a$      | As 10 + $D_{50}$     | -      | 81.9                                         | 90.8                                | +51.7 (+21.1)$^b$                             | +63.6 (+46.3)$^b$                           |

Note. Accuracy is determined from five-fold cross-validation of the training data set, and is calculated for a random number array (test 1), individual predictors (tests 2–9), and the combined predictors to further test $D_{50}$ (tests 10–12). The accuracy of the final kNN prediction with all predictors is shown (test 12).

$^a$Difference relative to test number 10.

$^b$Test 12 represents the accuracy of the final kNN model used.

Each predictor is shown in Table 2 for the dominance class and order of magnitude. The only variable with a lower predictive value than random noise in isolation was median grain size $D_{50}$ (test 9). To further test $D_{50}$, accuracy of the k-NN prediction was tested alongside the other predictors with and without $D_{50}$ (tests 10 and 12) and also with and without the random array (tests 10 and 11). It was found that in conjunction with the other predictors, $D_{50}$ provided a greater improvement in accuracy (class: 21.1%; magnitude: 46.3%) than the random array (class: 12.3%; magnitude: 9.1%), and was vital for an accurate prediction of the dominant class and order of magnitude (Table 2), therefore $D_{50}$ was included as a predictor. Final predictive accuracy was 81.9% for class and 90.8% for magnitude, and most misclassified data were only out by one class. Further sensitivity analyses of classification accuracy to the volume of training data available, the available variables (e.g., exclusion of wave data), and sensitivity to training using only data from shallow/deep waters is presented in the Supporting Information (Tables S2–S4; Text S1).

Tidal predictors (tide range $TR$, maximum current speed $U_{max}$) are shown across the NW European shelf in Figures 4a–4d for springs and neaps. A distribution of $TR$ over a statistically representative year is shown in Figures 4e and 4f at two locations marked with triangles in subplots a–b. The distribution of $TR$ was calculated across each node of the NW European shelf area over 1 year. Areas below the shelf break were excluded from analysis as they were below the maximum depth in the training data. Similarly, wave predictors are shown in Figures 5a–5d. These predictors are shown for 1% and 50% joint exceedance of $H_s$ and $T_p$ as determined from a fitted joint probability gumbel copula distribution (Genest & Favre, 2007) at each node across the domain over 1 year, using generalized extreme value and gamma marginal distributions for $H_s$ and $T_p$ respectively. Wave direction was taken as the mean wave direction over the year. Wave heights are in agreement with wave conditions for similar exceedances modeled by Bricheno et al. (2015). Depth was taken from the AMM7 model for calculation of $H_s/h$, while grain size was determined from the synthetic
Figure 4. Maps of representative tide conditions across the NW European Shelf. Histograms show distributions of tide range (TR) normalised by the maximum tide range over 1 year for two locations indicated by white triangles in subplots (a) and (b) for their respective columns. Fitted probability distribution functions are shown (red curves).
Figure 5. Maps of representative wave statistics. (a and b) Significant wave height $H_s$ for 1% exceedance and 50% exceedance probability. (c and d) Peak period $T_p$ for 1% exceedance and 50% exceedance probability. (e and f) Joint probability distribution function (blue) and cumulative distribution function (red dashed) contours for two locations over 1 year. Selected $H_s$ and $T_p$ for the 1% and 50% exceedance probability are indicated by a + and x respectively, taken at the point on the CDF contour with the maximum probability density interpolated from the PDF. The locations used for subplots (e and f) are indicated by white triangles in subplots (a and b) respectively.
map created by Wilson et al. (2018; Figure 1). All variables were resampled where necessary to the AMM7 model grid at 7 km resolution. The fitted distributions of the tide range (e.g., Figures 4e and 4f) and joint $H_t$ and $T_p$ (e.g., Figures 5e and 5f) enable the generation of tide and wave forcing data for a statistically representative year, assuming wave and tide condition are independent, keeping water depth and grain size constant, and using the mean wave direction and maximum tidal current direction as an indicator of the direction difference between waves and the tidal major axis.

3. Results

In this section, we present the results of the kNN classification across the NW European shelf for different environmental conditions, and examine the influence of different conditions on the shelf areas presented in Figure 1a. We go on to present the determination of the dominant sand transport class and order of magnitude over a statistically representative year.

3.1. Environmental Forcing Controls on Sand Transport Across the Shelf

Results from the kNN prediction for different environmental forcing conditions are presented as maps in Figure 6. The dominant class, indicating the dominant driver of sand transport, and the potential order of magnitude of net sand transport are presented for spring (Figure 6a, 6b, 6e and 6f) and neap (Figure 6c, 6d, 6g and 6h) tides under median (50% exceedance; Figures 6a–6d) and extreme (1% exceedance; Figures 6e–6h) wave forcing as characterized for each node on the shelf area (see Figures 4 and 5). Regions greater than 140 meters in depth are excluded to avoid extrapolation, as these exceed the largest depth in the training model and are deep enough that wave impacts are likely to be minimal.

Coastal areas around the UK are generally tide-dominated at spring tides and median wave forcing, with the second largest predicted order of magnitude of potential net sand transport (Figures 6a and 6b), exceeded only by the extreme waves at spring tide conditions (Figures 6e and 6f). This includes large areas of the meso-macrotidal Celtic shelf, UK East Coast and the Irish Sea. Under median waves, the influence of nonlinear wave-tide interactions was most prevalent in microtidal shelf areas below 30 m depth, with net transport up to several orders of magnitude lower than shallower or meso-macrotidal regions. Only microtidal, shallow, wave-exposed areas such as Dogger Bank and the DE-DK Shelf show wave-dominance in median wave conditions. The lowest magnitudes across the shelf are found for median waves at neaps, where only the shallow, exposed areas of the NL and DE-DK Shelves show elevated net sand transport driven by waves (Figures 6c and 6d). Sand transport is effectively switched off in areas other than the NL and DE-DK shelves under these low energy conditions.

In the highest energy conditions with extreme waves at springs, macro-meso tidal areas show dominance of WTI, while waves dominate sand transport in the Eastern North Sea where tidal currents are weaker (Figures 6e and 6f). Sand transport is dominated by waves across this shelf area during extreme waves at neaps, with the greatest magnitudes in finer grained, shallow, and wave exposed areas of the NL and DE-DK Shelves in the Eastern North Sea (Figures 6g and 6h). This wave-dominance is despite these areas having lower wave energy at this exceedance than more swell exposed regions (e.g., Celtic Shelf), indicating the importance of grain-size and water depth as controls. The next section explores the influence of environmental forcing conditions in more detail for the different shelf areas.

3.2. Environmental Forcing Controls on Sand Transport for Shelf Sub-Areas

A sensitivity analysis for different shelf areas was conducted by changing the environmental forcing conditions, including tidal condition, wave exceedance and grain size, and calculating the average class across each shelf area. To determine an average class, the kNN-predicted classification for each node within the designated region (Figure 1a) was converted to a representative pair of ratios $R_1$ and $R_2$ (Equations 1 and 2; Figure 2). Values of $R_1$ and $R_2$ were taken as the center value of each classification bin, while end values (e.g., for $R_1 > 3$ in tide-dominated conditions) were assumed to be dominant by a factor 6 in their respective direction (e.g., $R_1 = 6$ for tide-dominated transport). The mean $R_1$ and $R_2$ of all nodes within each region were calculated, weighted by the predicted net transport magnitude. These results are presented in Figure 7.
Figure 6. Results from the KNN predictions for different conditions presented as maps, including the dominant transport mode classification (left column) and order of magnitude (right column). Colors on the right column are on a logarithmic scale. Extreme (1% exceedance; Ex) and median (50% exceedance; Med) wave forcing is shown at springs (Spr) and neaps (Neap).
Figure 7. Sensitivity analysis for different shelf areas under changing environmental forcing conditions (UK – United Kingdom, NL – Netherlands, DE-DK – Denmark-German, NO – Norwegian). “Ex” denotes 1% exceedance “Extreme” wave forcing, “Med” denotes 50% exceedance “Median” wave forcing. $D_{NQ}$ denotes the $N$th quantile of the sediment $D_{50}$ diameter as distributed through the specified region. (b) The influence of grain size on the predicted classification for the UK East Coast region. Red and blue symbols indicate the class for the 2.5th and 97.5th centiles $D_{50}$ in the region, respectively. Other shelf areas show the class for the median $D_{50}$ in these regions. The 2.5th, 50th and 97.5th centile $D_{50}$ values are indicated on linear scales next to the classification triangle for each region.
The same wave and tidal forcing conditions are presented as shown in Figure 6. Symbols are placed within the classification triangle according to the regional mean R1 and R2 for that forcing condition. The influence of grain size variation is shown in Figure 7b. This is an indication of the variability in the response throughout the region arising from the spatial variability of grain size (D50). The dominant class was calculated for the median, 2.5th and 97.5th centiles of D50 through each region. Sand transport was more tidally dominated for finer grain sizes, due to easier resuspension. For clarity, the results for the other regions are shown for the median D50 through that region, with an indication of the variability in grain size shown on a scale.

Environmental forcing conditions are the primary control on the dominant net sand transport mode, with grain size moderating this. Different shelf areas exhibit different responses to changing forcing. Shelf areas adjacent to the UK (Irish Sea, Celtic Sea & UK-East Coast) are tidally dominated for median wave forcing at spring tides, while the NL Shelf and DE-DK Shelf show a significant tidal influence. Dogger Bank and the microtidal area of the NO Shelf have relatively low tidal sand transport magnitudes (Qnet ≲ 0.01 m^3 m^{-1} cycle^{-1}; Figure 6b) and are classified as nonlinear-dominated. Under median waves at neaps, tidal sand transport is low across the shelf and nonlinear interactions drive the sand transport that does occur. For extreme waves at springs, sand transport in most areas is dominated by WTI with the exception of the DE-DK Shelf which is wave dominated with a subdominant impact from WTI. Under extreme waves at neaps, most areas shift to wave-dominated sand transport. The macrotidal Celtic Sea and the relatively sheltered Irish Sea retain a subdominant contribution from WTI in these conditions. The next step is to determine which forces drive net sand transport over a statistically representative year, and the order of magnitude of that sand transport, taking the full annual distribution of waves and tides into account.

### 3.3. Dominance and Magnitude of Net Sand Transport Over a Year

Using the fitted tide range distribution (e.g., Figures 4e and 4f), and the fitted copula joint probability distribution for significant wave height and peak period (e.g., Figures 5e and 5f) for each node across the shelf, it was possible to generate tide and wave forcing data for a statistically representative year of semi-diurnal tidal cycles. By assuming independence between wave condition and tide condition, keeping water depth and grain size constant, and using the mean wave direction and tidal maximum current direction, it was possible to tabulate a representative set of predictors over a statistically representative year. These were then used to determine a classification and order of magnitude for each tidal cycle. The sum of the order of magnitude over the statistically representative year gives a sense of the magnitude of potential net sand transport across the shelf over one year, while the classification for each node was determined as the class for which the maximum net sand transport occurred over the year. Results are shown in Figure 8.

Net sand transport ranges from approx. 10 m^3 m^{-1} y^{-1} in deeper, microtidal areas of the NO Shelf, to up to 10,000 m^3 m^{-1} y^{-1} in more wave exposed areas of the DE-DK Shelf and the macrotidal areas of the south west English Channel. Much of the shelf surrounding the UK is tidally dominated, while deeper areas of the shelf, including much of the Celtic Sea and NO Shelf, are dominated by the nonlinear effects of WTI. Shallow, fine grained areas of Dogger Bank and the DE-DK shelf are dominated by wave-driven sand-transport, reflecting the lower tidal velocities across these regions. The NL Shelf is also dominated by the nonlinear effects of WTI, reflecting stronger tidal currents and coarser grain size than Dogger Bank and the DE-DK Shelf (Figures 1b and 4). This does not consider wind-driven net sand transport, nor the influence of sand-mud or sand-gravel mixtures. Areas with very low fractions of sand (Figure 1c) are included in these figures, and therefore these results should be considered for the potential net sand transport magnitude assuming continual availability of sand at the bed. These points are discussed in detail in Section 4.2. In addition, a comparison to the observed sand wave morphology is made in the Discussion (Section 4.1).

### 4. Discussion

The magnitude of net sand transport and relative dominance of waves, tides, and their nonlinear interactions was predicted for the Northwest European Continental Shelf using a kNN approach trained on extensive numerical modeling data on the Celtic Shelf area using a coupled hydrodynamic, wave and sand transport model (King et al., 2019). This shelf area has a highly varied tidal climate ranging from micro-
mega-tidal, varying degrees of wave exposure and a highly energetic wave climate (Harris & Coleman, 1998). These factors result in a varied parameter space with which to test the application of this kNN classification approach while generating insights into the dynamics of sand transport across this shelf.
The dominance of waves on the DE-DK Shelf and Dogger Bank, and dominance of tides along the UK East Coast predicted here is in agreement with modeling of wave, wind, and tidal sand transport in the North Sea (van der Molen, 2002), lending confidence to the predictions of the kNN model. This paper builds upon the previous work by considering the influence of WTI, indicating that the nonlinear wave-tide interaction dominates along the Dutch Shelf and deeper areas of the Celtic Sea and the Norwegian Shelf. This paper also presents a computationally efficient method for estimating the dominant processes influencing net sand transport, and its magnitude, for different environmental forcing conditions using readily available data. In the next section we examine a potential application of this method to look at the influence of environmental forcing parameters on sand wave morphology. We then discuss other applications, limitations, and future work that arises from this.

4.1. Comparison With Sand Wave Morphology

Modeling of sand wave dynamics is important for offshore renewable energy industrial activities and studies have been conducted to understand their dynamics in the Dutch North Sea and elsewhere (Cheng et al., 2020; Damen et al., 2018a, 2018b; Roert et al., 2017; Van Oyen et al., 2011; van Santen et al., 2011; Wang et al., 2019). Tidal sand waves are also habitat to benthic species whose spatial distribution is dependent upon sand wave morphology, with feedback effects on sand wave evolution (Damveld et al., 2018, 2020).

Surface waves affect sand wave growth, wave length, and migration, reducing sand wave height and increase wave length (Campmans et al., 2018a, 2018b). Damen et al. (2018a) examined sand waves on the Netherlands Shelf, finding weaker than expected correlation of sand wave height with $H_s$ possibly due to the interdependent and opposite acting correlations between $H_s$, water depth, and sand wave height (Campmans et al., 2018a, 2018b; Houthuys et al., 1994; Van Dijk & Kleinhans, 2005). They find that it is more reliable to consider the impact of the waves at the bed, for example, using the Shields parameter.

Tidal currents are known to positively correlate with spatial frequency (Damen et al., 2018a, 2018b; Van Santen et al., 2011). Damen et al. (2018a) find weak correlation between tidal currents and sand wave height. It is important to consider the level of suspended sediment transport as a control on sand wave length and height (Borsje et al., 2014; Damen et al., 2018a, 2018b). This could be a future application of this kNN method, to predict the balance between suspended and bedload sand transport under variable forcing conditions.

Here, we use the same trained kNN classifier as presented earlier to predict the dominant transport mode across the same region considered by Damen et al. (2018a). Where possible, predictor data used were taken from the data set of Damen et al. (2018b). These included 1% exceedance $H_s$, M2 current amplitude (instead of the maximum tidal current) and grain size $D_{50}$. Tide range, current mean direction, and wave mean direction were interpolated from the shelf-scale predictors used earlier, and $T_p$ was interpolated from the 1% exceedance $T_p$ (Figure 5c). The predicted transport class was determined at 1 km resolution at the same locations as the data presented in Damen et al. (2018a) and this is presented in Figure 9a. Under these conditions we predict the dominance of nonlinear WTI in the southeast of the sand wavefield, moving to wave-dominance in the northwest.

The height, wave length, and asymmetry of the sand waves were binned for each classification and compared between classes (Figures 9b–9d). This resulted in comparison of 9,161 data points each representing sand wave characteristics over a 1-km$^2$ area. Results suggest that sand wave height is lowest in the wave-dominated regions, and larger in regions dominated by nonlinear WTI. Similarly, wave length and asymmetry appear to increase with an increase in wave-dominance. The statistical dissimilarity of the sand wave populations in each class was tested using the Kolmogorov-Smirnov (KS) test. Distributions of sand wave characteristics were found to be unique between classes at the 95% confidence level. A second one-sided KS test was performed to test the hypotheses that sand wave height decreases moving from nonlinear interaction-dominated to wave-dominated sand transport, and that wave length and asymmetry increase. These hypotheses were found to be true at the 95% confidence level, and P-values are included in Figures 9e–9g. These results are in agreement with previous research into wave and tidal influences on sand wave height, wave length, and asymmetry (Campmans et al., 2018a, 2018b; Damen et al., 2018a, 2018b), lending confidence to the results of the kNN prediction and indicating that WTI may play a significant role influencing
Figure 9.
sand wave morphology, and this classification scheme has a predictive power for sand wave morphology on sandy continental shelves. This prediction is based on the most energetic wave and tidal conditions. The annual classification determined in Figure 8 indicates that this region is dominated by nonlinear WTI on an annual scale, suggesting that the more energetic conditions play a significant role in controlling sand wave morphology, with increased wave-dominance under storm conditions limiting sand wave heights. While sandwave asymmetry may be indicative of the net sand transport direction, sandwave growth is dependent on local residual currents driving sediment to the sandwave crest (Besio et al., 2006; Campmans et al., 2018a, 2018b; Huischer, 1996; van Santen et al., 2011), and gross sand transport may be a stronger indicator for this as these local residuals do not depend on the regional net sand transport rate. It is feasible to extend and apply this method to the prediction of gross sand transport.

4.2. Assumptions, Limitations and Future Work

In this study we show that the magnitude of net sand transport and the relative contribution from waves, tides, and nonlinear WTI is amenable to estimation using readily available wave and tidal data utilizing a kNN classification prediction approach. The kNN method itself does not account for the physical relationships between predictors and the resultant classification, relying instead on the associations between predictors and classifications in the parameter space. This implies that the trained classifier will only be representative of the physical processes represented in the training data. The trained classifier cannot therefore be used to extrapolate outside the range and physics represented in the data used to train it, however it can be applied in other regions. Classifier accuracy is expected to be poor when used to extrapolate beyond the limitations of its training data, therefore care should be taken when applying this method to identify potential biases and limits (such as depth limitations) of the training data in order not to extrapolate. This may represent a limitation in data-poor regions. A sensitivity analysis to available variables for training the classifier, the size of the training data set, and classifier accuracy when trained in exclusively shallow/deep waters is presented in the Supporting Information (Tables S2–S4; Text S1). Here, we discuss the processes represented in the model used to generate the training data in this study, and the implications of those not represented.

Data used to train this kNN predictor were generated by a well validated numerical model of coupled hydrodynamics, waves and sand transport (King et al., 2019). The range of each predictor in the training data is shown in Table 1. Sand transport rates are determined using the formulation of van Rijn (2007a, 2007b), therefore, the predictor is representative of the physics included therein. Importantly, baroclinic and wind-driven currents are not included in the training model. This paper considers processes at the shelf scale, and due to the resolution of the forcing variables it should be considered to represent an estimate of the dominant sand transport processes on the continental shelf, and does not consider processes landward of the shoreface (approx. 15 m) (e.g., Hamon-Kerivel et al., 2020; Héquette et al., 2008).

Important wind speed events can interact constructively or destructively with tidal currents to influence sand transport rates, depending on the relative angle of wind-driven currents to the tidal current direction (Héquette et al., 2008). Wind-driven currents are weak on the Celtic Shelf (Pingree & Le Cann, 1989), and wind-driven residual currents across the NW European Shelf are likely to be most significant at neaps when tidal currents are weakest (Pingree & Griffiths, 1980), with the strongest wind-driven residuals present in the Southern North Sea. van der Molen (2002) discusses wind-driven sand transport relative to tides and wind waves in the Southern North Sea, finding wind-driven flows contributing significantly to net sand transport where tidal currents are small, alongside wave-driven currents. The areas defined by van der Molen (2002) as storm dominant (winds + waves) qualitatively agree with the wave-dominated areas of the NL

Figure 9. Application of classification prediction to sand wave physical characteristics averaged per square kilometer as per Damen et al. (2018a). (a) Sand transport dominant class across the Netherlands (NL) Shelf determined with a mix of environmental data from Damen et al. (2018a) and other predictors as described earlier, interpolated to each square kilometer (1 pixel = 1 km²). (b–d) Box plots showing sand wave height, wave length, and spatial frequency respectively for each dominant class. Plots indicate the median, 25th and 75th percentiles and whiskers indicate 1.5 times the IQR beyond the 75th or 25th percentile. (e–g) P-values from a 2-sample, 1-sided Kolmogorov-Smirnov test, testing if the data are significantly lower in magnitude in more wave-dominated conditions (height, spatial frequency – e and g), or greater in magnitude in the more wave-dominated condition (wave length – f) at the 95% confidence level.
Shelf under energetic wave and tidal forcing presented in Figure 9. Their tide dominated area corresponds to the nonlinear wave-tide interaction dominated part of the shelf, and it is noted that wave-tide interaction is not fully represented in their modeling. While wind-driven circulations are beyond the scope of this study, this kNN method could be extended using a coupled training model to isolate the relative influence of wind-driven circulations on net sand transport and incorporate these into the classification.

Baroclinic circulations are not considered in this study either. Van Leeuwen et al. (2015) classify the North Sea by stratification regime. The regions of greatest net sand transport predicted here correspond qualitatively with areas either permanently mixed or intermittently stratified conditions, with seasonally stratified conditions affecting the deeper, microtidal areas of the North Sea which are predicted to have a lower magnitude of net sand transport. In winter, the NW European shelf area considered in this study is well mixed while areas such as the UK East Coast, the NL Shelf, the DE-DK Shelf, and English Channel tend to remain well mixed or show weak stratification through spring, summer, and autumn (Holt et al., 2010), and therefore baroclinic effects are not expected to influence significantly the prediction of this model in these regions.

An additional limitation is that this study only considers a pure sand bed, whereas sand-mud and sand-gravel mixtures affect sand resuspension (McCarron et al., 2019; Thompson et al., 2019). Graded sediment transport resulting from heterogeneous, bimodal sand distributions may also affect the wave length of sand waves (Van Oyen & Blondeaux, 2009). In sand-gravel mixtures, the hiding-exposure effect increases the critical shear stress required to mobilize the sand fraction, its effect becoming more significant for mixtures of >10% gravel (McCarron et al., 2019). Much of the North Sea sediment is comprised of >90% sand (Figure 1c), and this effect is most likely to impact predictions on shelf areas with a higher coarse grain size fraction such as the Celtic Sea. While we also do not consider biological effects on sediment resuspension, Thompson et al. (2019) show physical sediment characteristics to be more significant than biological factors in controlling bed stability. The purpose of this kNN classification method is to be applicable with readily available hydrodynamic and morphological data, therefore consideration of nonuniform grain size distributions, the effect of mixed sand-mud or sand-gravel substrates, and biological effects would necessarily add complexity to the predictive model and therefore limit its use by introducing a data requirement which may not be readily available to coastal practitioners. The method could be extended to include the effects of mixed grain size fractions in future.

This paper considers net sand transport rates, which depend upon additional variables to those eventually used in the trained classifier in this paper, such as tidal asymmetry and mean currents (Leonardi & Plater, 2017; Pingree & Griffiths, 1979; Stride, 1963; Uncles, 1982, 2010; Zimmerman, 1978). It was decided not to include these additional variables in order to keep the data requirements of the classification to a minimum and ensure the classifier relied on more readily available data. The modeling of King et al. (2019) includes these effects in the numerical modeling, and while this classifier does not use tidal asymmetry and mean currents as a predictor, it performs well when validated against the calculated net sand transport classes and magnitudes of King et al. (2019), with an accuracy of 81.9% and 90.8%, respectively.

The benefit of this method is to enable a rapid assessment of the dominant processes affecting net sand transport, and its magnitude, without the need for a computationally expensive numerical model. We show that the classification scheme of King et al. (2019) has a predictive value for sand wave morphology on the NL Shelf, as a further application of this method. While this paper considers shelf-scale processes, this classification scheme can be applied to other sand transport processes in the nearshore, such as headland bypassing (King et al., Under Review). Careful consideration must be taken to examine the validity of the assumptions made in the model used for training, and how well these represent net sand transport drivers in the area under consideration. This includes consideration of natural processes including wind-driven/ baroclinic circulations (Héquette et al., 2008), anthropogenic processes such as trawling and dredging (Cieślikiewicz et al., 2018; Mengual et al., 2019; Uncles et al., 2020), and potential future changes in forcing conditions due to climate change (Meucci et al., 2020; Young & Ribal, 2019). The computational efficiency of this method relative to running a coupled wave-tide numerical model enables quick assessment to be made of the influence of changing environmental conditions such as upward trends in storminess across central, western, and northern Europe (Castelle et al., 2018; Donal et al., 2011) on the magnitude and dominant forces driving the net transport of sand on sandy continental shelves, with potential applications globally.
5. Conclusions

In this paper we apply a data-driven method to predict the dominant sand transport drivers and magnitude across a sandy continental shelf. We use k-Nearest Neighbor classification prediction trained with data from coupled hydrodynamic, wave and sediment transport modeling on a subdomain of the shelf to predict the sand transport magnitude and mode across the entire shelf, using readily available wave, tide, and morphological data. Key findings of this paper include:

1. The relative dominance of waves, tides, and nonlinear effects due to wave-tide interactions in the net transport of sand over a tidal cycle, as well as net sand transport magnitude, can be predicted with readily available data. These are: significant wave height, peak period, mean wave direction, wave power, tide range, maximum tidal current speed and direction, water depth, and median grain size.
2. Wave and tidal conditions are primary controls on net sand transport mode and magnitude, while grain size is a secondary control.
3. Different shelf areas exhibit different dominant drivers of net sand transport for similar exceedance conditions, relating to differences in water depth, grain size, tide range, and wave exposure between regions.
4. Tides are the dominant or sub-dominant forcing for net sand transport across most shelf areas for median waves at springs. For extreme waves at springs, most areas show dominance of the nonlinear effects of wave-tide interactions. At neaps, with median waves, sand transport is very low across the shelf, driven by wave-tide interaction where it does occur. Extreme waves at neaps result in wave-dominated sand transport in most areas of the shelf, while wave-tide interactions influence sand transport in deeper or macrotidal regions.
5. Sand transport magnitude and dominance was predicted for a statistically representative year based on distributions of tide range and $H_s$-$T_p$ joint-probability calculated across the shelf. Potential net sand transport shows tidal dominance in meso-macrotidal waters around the UK, wave-dominance on Dogger Bank and the German/Denmark Shelf, and dominance of wave-tide interactions on the Netherlands shelf and in deeper areas of the North Sea and Celtic Sea.
6. The kNN prediction was applied at higher resolution to the Netherlands shelf area, and classes for energetic (conditions 1% exceedance waves at spring tide) were compared with sand wave morphology across the region with data obtained from Damen et al. (2018b). Sand wave height is shown to significantly (95% confidence) reduce with greater wave-dominance, while sand wave length and asymmetry significantly increase. Sand wave morphologic parameters were significantly different between predicted classes at the 95% confidence level.
7. This paper presents a computationally efficient method to determine an initial estimate of the dominant driving forces and magnitude of net sand transport on sandy continental shelves, enabling efficient large-scale comparison between different regions and testing of the influence of changing environmental forcing on net sand transport with applications globally.

Data Availability Statement

Sand wave data used in this study are available at https://doi.org/10.4121/uuid:0d7e016d-2182-46ea-bc19-cdfda5c20308 and the authors thank Damen et al. (2018b) for making this valuable data set available. The other data on which this paper is based are publicly available from the corresponding author and are made available online via the University of Plymouth PEARL open access research repository at http://hdl.handle.net/10026.1/17601.

References

Aldridge, J. N., Parker, E. R., Bricheno, L. M., Green, S. L., & van der Molen, J. (2015). Assessment of the physical disturbance of the northern European Continental shelf seabed by waves and currents. Continental Shelf Research, 108, 121–140. https://doi.org/10.1016/j.csr.2015.03.004
Besio, G., Blondeaux, P., Brocchini, M., Hulscher, S. J. M. H., Idier, D., Knapen, M. A. F., et al. (2008). The morphodynamics of tidal sand waves: A model overview. Coastal Engineering, 55(7–8), 657–670. https://doi.org/10.1016/j.coastaleng.2007.11.004
Besio, G., Blondeaux, P., & Vittori, G. (2006). On the formation of sand waves and sand banks. Journal of Fluid Mechanics, 557, 1–27. https://doi.org/10.1017/S00221120060090256
Booij, N., Holthuijsen, L. H., & Ris, R. C. (1999). A third-generation wave model for coastal regions 1. Model description and validation. Journal of Geophysical Research, 104(C4), 7649–7666. https://doi.org/10.1029/98JC02622

Acknowledgments

The authors acknowledge the UK Hydrographic Office for VORF corrections. The authors acknowledge the MET Office (Andy Saulter) for the hydrodynamic, bathymetric, and wave forcing data and NOAA for atmospheric pressure and wind data, and EMODnet bathymetry Consortium for the EMODnet Digital Bathymetry (DTM, 2016). This research was supported by the NERC-funded BLUECoast Project (NE/N015525/1).
Scott, T., Masselink, G., O’Hare, T., Sautler, A., Poate, T., Russell, P., et al. (2016). The extreme 2013/2014 winter storms: Beach recovery along the southwest coast of England. *Marine Geology*, 382, 224–241. https://doi.org/10.1016/j.margeo.2015.05.010

Stride, A. H. (1963). Current-swept sea floors near the southern half of Great Britain. *Quarterly Journal of the Geological Society*, 119, 175–197. https://doi.org/10.1144/gsjgs.119.1.0175

Tambroni, N., Blondeaux, P., & Giovanna, V. (2015). A simple model of wave-current interaction. *Journal of Fluid Mechanics*, 775, 328–348. https://doi.org/10.1017/jfm.2015.306

Thompson, C. E. L., Williams, M. E., Amoudry, L., Hull, T., Reynolds, S., Panton, A., & Fones, G. R. (2019). Benthic controls of resuspension in UK shelf seas: Implications for resuspension frequency. *Continental Shelf Research*, 185, 3–15. https://doi.org/10.1016/j.csr.2017.12.005

Tonani, M., & Sautler, A. (2020). For NWS ocean waves reanalysis product NWSHELF_REANALYSIS_WAV_004_015. Product user manual. Issue 1.0. Copernicus Marine Environment Monitoring Service.

Tonani, M., Sykes, P., King, R. R., McConnell, N., Pequignet, A.-C., O’Dea, E., et al. (2019). The impact of a new high-resolution ocean model on the Met Office North-West European Shelf forecasting system. *Ocean Science*, 15, 1133–1158. https://doi.org/10.5194/os-15-1133-2019

Umeayama, M., (2005). Reynolds stresses and velocity distributions in a wave–current coexisting environment. *Journal of Waterway, Port, Coastal and Ocean Engineering*, 131, 203–212.

Uncles, R. J. (1982). Computed and observed residual currents in the Bristol Channel. *Oceanologica Acta*, 5(1), 11–20. http://archimer.ifremer.fr/doc/00120/23148/

Uncles, R. J. (2010). Physical properties and processes in the Bristol Channel and Severn Estuary. *Marine Pollution Bulletin*, 61(1-3), 5–20. https://doi.org/10.1016/j.marpolbul.2009.12.010

Uncles, R. J., Clark, J. R., Bedington, M., & Torres, R. (2020). Chapter 31 - On sediment dispersal in the Whitsand Bay marine conservation zone: Neighbour to a closed dredge-spoil disposal site. In J. Humphreys, & R. W. E. Clark (Eds.), *Marine protected areas* (pp. 599–629). Elsevier. https://doi.org/10.1016/B978-0-08-102698-4.00031-9

van der Molen, J. (2002). The influence of tides, wind and waves on the net sand transport in the North Sea. *Continental Shelf Research*, 22(18-19), 2739–2762. https://doi.org/10.1016/S0278-4343(02)00312-3

van Dijk, T. A. G. P., & Kleinhans, M. G. (2005). Processes controlling the dynamics of compound sand waves in the North Sea, Netherlands. *Journal of Geophysical Research*, 110, F04S10. https://doi.org/10.1029/2004JF000173

van Leeuwen, S., Tett, P., Mills, D., & van der Molen, J. (2015). Stratified and nonstratified areas in the North Sea: Long-term variability and biological and policy implications. *Journal of Geophysical Research: Oceans*, 120, 4670–4686. https://doi.org/10.1002/2014JC010485

van Rijn, L. C., (2007b). Unified view of sediment transport by currents and waves. II: Initiation of motion, bed roughness, and bed-load transport. *Journal of Hydraulic Engineering*, 133(6), 649–667. https://doi.org/10.1061/(ASCE)0733-9429(2007)133:6(649)

Van Oyen, T., & Blondeaux, P. (2009). Tidal sand wave formation: Influence of graded suspended sediment transport. *Journal of Geophysical Research*, 114, C07004. https://doi.org/10.1029/2008JC005136

Van Oyen, T., de Swart, H., & Blondeaux, P. (2011). Formation of rhythmic sorted bed forms on the continental shelf: An idealised model. *Journal of Fluid Mechanics*, 684, 475–508. https://doi.org/10.1017/jfm.2011.312

van Santen, R. B., de Swart, H. E., & Van Dijk, T. A. G. P. (2011). Sensitivity of tidal sand wave characteristics to environmental parameters: A combined data analysis and modelling approach. *Continental Shelf Research*, 31(9), 966–978. https://doi.org/10.1016/j.csr.2011.03.003

Wang, Z., Liang, B., Wu, G., & Borsje, B. W. (2019). Modeling the formation and migration of sand waves: The role of tidal forcing, sediment size and bed slope effects. *Continental Shelf Research*, 190, 103866. https://doi.org/10.1016/j.csr.2019.103866

Ward, S. L., Neill, J. P., Van Langemeh, K. J. J., & Scourse, J. D. (2015). Classifying seabed sediment type using simulated tidal-induced bed shear stress. *Marine Geology*, 367, 94–104. https://doi.org/10.1016/j.margeo.2015.05.010

Wilson, R. J., Spiers, D. C., Sabatino, A., & Heath, M. R. (2018). A synthetic map of the north-west European Shelf sedimentary environment for applications in marine science. *Earth System Science Data*, 10(1), 109–130. https://doi.org/10.5194/essd-10-109-2018

Xu, K., Mickey, R. C., Chen, Q., Harris, C. K., Hetland, R. D., Hu, K., & Wang, J. (2016). Shelf sediment transport during hurricanes Katrina and Rita. *Computers & Geosciences*, 90(B), 24–39. https://doi.org/10.1016/j.cageo.2015.10.009

Young, I. R., & Ribal, A. (2019). Multiplatform evaluation of global trends in wind speed and wave height. *Science*, 364(6440), 548–552. https://doi.org/10.1126/science.aav9527

Zhang, W., Cui, Y., Santos, A. I., & Hanebuth, T. J. J. (2016). Storm-driven bottom sediment transport on a high-energy narrow shelf (NW Iberia) and development of mud depocenters. *Journal of Geophysical Research: Oceans*, 121, 5751–5772. https://doi.org/10.1002/2015JC0101526

Zimmerman, J. T. F. (1978). Topographic generation of residual circulation by oscillatory (tidal) currents. *Geophysical & Astrophysical Fluid Dynamics*, 11, 35–47. https://doi.org/10.1080/03091927808242650