Nonlinear Processes in Geophysics

Multivariate analysis of nonlinearity in sandbar behavior

L. Pape and B. G. Ruessink

Department of Physical Geography, Faculty of Geosciences, IMAU, Utrecht University, P.O. Box 80.115, 3508 TC Utrecht, The Netherlands

Abstract. Alongshore sandbars are often present in the nearshore zones of storm-dominated micro- to mesotidal coasts. Sandbar migration is the result of a large number of small-scale physical processes that are generated by the incoming waves and the interaction between the wave-generated processes and the morphology. The presence of nonlinearity in a sandbar system is an important factor determining its predictability. However, not all nonlinearities in the underlying system are equally expressed in the time-series of sandbar observations. Detecting the presence of nonlinearity in sandbar data is complicated by the dependence of sandbar migration on the external wave forcings. Here, a method for detecting nonlinearity in multivariate time-series data is introduced that can reveal the nonlinear nature of the dependencies between system state and forcing variables. First, this method is applied to four synthetic datasets to demonstrate its ability to qualify nonlinearity for all possible combinations of linear and nonlinear relations between two variables. Next, the method is applied to three sandbar datasets consisting of daily-observed cross-shore sandbar positions and hydrodynamic forcings, spanning between 5 and 9 years. Our analysis reveals the presence of nonlinearity in the time-series of sandbar and wave data, and the relative importance of nonlinearity for each variable. The relation between the results of each sandbar case and patterns in bar behavior are discussed, together with the effects of noise. The small effect of nonlinearity implies that long-term prediction of sandbar positions based on wave forcings might not require sophisticated nonlinear models.

1 Introduction

Sandbars are alongshore ridges of sand in up to 10 m water depth along micro- to mesotidal, storm-dominated coasts and strongly affect the nearshore flow field and sediment transport processes. They continuously change their position in response to temporal variability in the offshore wave forcing. During storms intense wave breaking drives strong offshore-directed currents that force the sandbar offshore (e.g. Gallagher et al., 1998). Wave nonlinearity (preponderance of high crests and steep front faces) drives onshore migration during intermediate and low-energy conditions (Hoefel and Elgar, 2003). Until recently, sandbar evolution was often modeled using models based on small-scale physics forced by waves and currents (Roelvink and Brøker, 1993; van Rijn et al., 2003; Ruessink et al., 2007), which we will refer to as process-based models. The increasing amount of remote-sensed sandbar data, for example in the Argus program (Holman and Stanley, 2007), now allows for the use of data-driven models as an additional means for studying sandbar behavior. Data-driven models use limited or no physical knowledge. Instead they extract the relations that describe the characteristics of data from the data itself. Although data-driven models are based on the internal structure found in observations, it is important to understand the nature of the underlying processes and how these processes are reflected in the behavior of the observed variables. Nonlinearity is one of the features of a sandbar system that is clearly present in the underlying physical processes such as wave-induced currents and sediment transport, but it is unclear how this nonlinearity is expressed in the observations made from this system. To increase the understanding of the temporal evolution of sandbars a method is needed that can identify nonlinearity in the time-series of observations from a sandbar system.

The predictability of a variable that is measured from a system depends at least partially on the nonlinearity of the processes governing the system's behavior. When linear models are used to describe a system, irregular behavior can only be attributed to unknown random inputs. However, nonlinearity in the system can be a source of irregularity in itself. Depending on the dynamics of the system and the scale on which it is studied, nonlinearity can lead to behavior that is very difficult to predict, or it can be the source of very regular behavior when the system is studied at the level of a possible attractor it evolves into. Although it might seem

Correspondence to: L. Pape
(l.pape@geo.uu.nl)

obvious that a system in which the underlying physical processes are inherently nonlinear requires a data-driven model that is also nonlinear, it might not be true that every variable that is measured from the system reflects this nonlinearity (Schreiber and Schmitz, 2000). For example, the nonlinearities at small-scale physical processes (described by sediment flux equations) can together cause a system to move towards a simple attractor state that reveals itself only at larger scales (sandbar behavior). There are also reasons that nonlinearities acting in a system at the same scale as the measured variables might not reveal themselves in observations, such as nonlinearities that act in opposite directions and cancel out each other, or nonlinearities that are small compared to measurement or model noise. Together these effects can have considerable consequences for the predictability of the evolution of a system. Although not yet successfully demonstrated, Southgate and Möller (2000) conjecture that process-based models might not be able to make accurate predictions of time-sequences of morphology beyond certain prediction horizons, while it may be possible to make accurate predictions of the evolution of a system based on stable statistical properties of attractor states. It is therefore not only important to know that there is nonlinearity in the physical processes that govern the behavior of a system, but also to recognize whether there is nonlinearity present in the data that are measured from that system.

Detecting nonlinearity can be achieved by comparing the prediction accuracy of different data-driven models that can or cannot cope with nonlinearity. This approach is used in the Deterministic versus Stochastic (DVS) method (Casdagli and Weigend, 1993). In this method the entire continuum between highly local piece-wise linear models and a global linear model is analyzed. The course of the prediction error between the two types of models yields qualitative insights in the dynamics of the underlying system. The DVS method is based on projections of observations in state space. Just like other state space-based methods, the DVS method can only be used for univariate data, or multivariate data with the same physical units. For data with the same physical interpretation a distance measure that is the same in all regions of the state space can easily be defined. However, if the state space of a system contains variables with different physical meanings or no physical interpretation at all, the relative contribution of each variable to the distance measure remains unidentified. The problem of using multivariate data in which different variables might be mutually dependent, can be solved in different ways (Sugihara and May, 1990). One method is to study only a single variable, and consider all other variables as noise, and the relation between the studied variable and other variables as noise processes. A second method is to study a single variable in the parts of the data in which the other variables are constant, or are within a certain range. The third option is to study the effects of other variables on the studied variable in terms of individual events. In this fashion the different modes of behavior of a system or the

time it takes for a system to return to a certain mode can be studied in relation to the changes or values in other variables. What all these methods have in common is that they cannot deal with systems that are driven by continuously changing forcing, such as the nearshore system, which is mainly driven by incoming waves with continuously changing properties. A method that does take into account the response of a system to continuously changing variables is input-output modeling (Casdagli, 1992; Rubin, 1995; Jaffe and Rubin, 1996), in which one (input) variable is used to model another (output) variable. Input-output modeling does however not enable the detection of nonlinearity for the contribution of each variable in the relations between combined variables. Other methods exist that try to find the optimal contribution of each variable as a weight in a weighted distance function (Abraham, 1997; Cao et al., 1998; Garcia and Almeida, 2005). However, such methods are only practical solutions aimed at improving model performance and do not provide additional insights into the relations between observed variables.

In the present work a method is proposed to investigate the contribution of different variables in multivariate embeddings to nonlinearity and determinism. As an extension to the DVS method, not only the continuum between piecewise linear and global linear models is investigated, but also the continuum between the relative contribution of each variable to the distance measure in state space. This method, the Multivariate DVS (MDVS) method, is then applied to three sandbar datasets to determine the amount of nonlinearity in the relations between different observables. Although the method we describe here stems from sandbar research, we anticipate our method to be applicable to time-series of other systems as well.

2   Models

A model of the evolution of a system over time involves a method that describes how the system evolves from one state to the next. A single observation almost never holds complete information of the true underlying system state. Nevertheless, it is possible to reconstruct the system state by using sequences of observations instead of a single observation. In such a representation, sequences of observations are projected in a state space that preservers the properties of the transitions between the states of the system that do not change under smooth coordinate changes (Takens, 1981; Sauer et al., 1991). Although the entire state of a system can be fixed given sequences of sufficient length, it is not always desirable to model each relation between every aspect of the system state. Processes that are important for the behavior under study are often modeled as deterministic rules, while the remaining processes are represented by their corresponding distributions. Together these distributions are referred to as noise.

Methods for detecting nonlinearity in time-series data often rely on embedding the measured variables in state space. The evolution of such a system is given by a rule that specifies what state follows from the current state. In case the observations are made from a system that is to some extent driven by external forcings, the part of the state space corresponding to the external forcing factors cannot be predicted given the current state. The external forcings are not considered as part of the system and behave independent of the dynamics of the system. The sandbar system studied in the present work is an example of a system that is forced by external factors. Sandbar migration is the result of gradients in the transport of sediment caused by the motion of water due to incoming waves or wave-generated processes. Time-series of sandbar positions can be extracted from Argus images or from measured profiles, while wave forcings are measured outside the system using offshore located buoys. Because the wave data are expressed in terms of offshore wave characteristics it is difficult to implement the feedback from the morphology on the waves in a data-driven model. In the present work the wave forcing is therefore not included in the system, but acts on the system from the outside. However, although we cannot measure it, the feedback of the morphology on the waves might be an important factor in sandbar behavior. The importance of the interaction between the morphology and the incoming waves might even be different for different parts of the nearshore. In multiple-barred sandbar systems the outer sandbar responds more directly to incoming waves, while the relation between the offshore wave characteristics and the behavior of inner sandbars is less obvious because of wave breaking on the outer bar. Some mention patterns in sandbar behavior that cannot be directly related to forcings (e.g. Southgate and Möller, 2000; Plant et al., 2006), but identifying such behavior from observations remains difficult (Elgar, 2001). It is therefore unclear if the evolution of a multiple-barred sandbar system can best be described in terms of trajectories of sandbar positions in state space, or as primarily forced by external factors. Time-series data driven by external forcing factors are usually modeled using Box-Jenkins type models (Box and Jenkins, 1970), or nonlinear versions of such models (for an application to sandbar modeling, see Pape et al., 2007). However, methods for detecting nonlinearity are based on the evolution of a system in state space. To detect nonlinearity in time-series of sandbar observations and forcings a method is needed that combines both approaches.

A key assumption in modeling is that a system behaves the same under similar conditions. To use this assumption the terms in the previous sentence need to be further specified. In the models used in the present work the term "condition" comprises the system state and the external forcing. The system state can be represented as a vector in a state space, and the concept of similarity by some distance measure between two vectors. A very simple way to model behavior is to search for a number of neighboring vectors in state space, and predict the average value following these neighbors. This might work for neighbors that are very close in state space, but for increasing numbers of neighbors, the predicted values converge to the mean of the data. It is possible to do better than that by establishing a notion of behavior that involves the underlying *structure* of the relation between two system states. For example, instead of predicting system states, the change in system state could be used, or even better, a parametric model established by linear regression. Linear methods assume that a shift in state space means a proportionally large change in the predicted value. However, for nonlinear relations this is not the case. If the behavior of a system varies in different regions of the state space it is not possible to construct a single accurate global linear model. The DVS method (Casdagli and Weigend, 1993) uses this fact to detect nonlinearity by testing the performance of linear regression models based on increasing numbers of neighboring states. Increasing the number of neighbors on which a model is based will first cause an increase in model performance, because including more samples in the process of model building will reduce noise. Increasing the number of neighbors even further causes samples in more remote parts of the state space to be included in the process of model building as well. A linear model will only benefit from this increase if the behavior of the system is the same in all parts of the state space. For nonlinear systems this is not the case, and the performance of a linear model starts to decrease again when the number of neighbors is increased beyond a certain amount because noise reduction becomes less important than modeling nonlinearity. Therefore, the change in the course of the prediction error for models based on increasing numbers of neighbors can be used as an indication of nonlinearity.

Based on this conceptual simplicity, the DVS method can easily be extended for detecting nonlinearity in multivariate data. In this method, which we call the Multivariate DVS (MDVS) method, the entire multivariate time-series (including the external forcings) are projected as vectors in state space. Models are now based on multivariate neighbors in state space. Instead of predicting all variables, only the variables that represent the system state are predicted. Predicting the external forcings from the internal system dynamics would make no sense, so the external forcings are included in state space, but they are not predicted. When neighbors have to be found in a state space that contains multiple variables, it is not immediately clear what the contribution of each variable to the distance function should be. Although the optimal weight setting for each variable might not be known, the importance of each variable to the nonlinearity can be investigated by testing several weight values. Just as in the DVS method, a linear regression model is established based on the selected neighbors and the values following them. This process is repeated for increasing numbers of neighboring samples and several values of the weights. If the behavior of the system depends linearly on a certain variable, then the behavior of the system is constant in the dimensions related to

this variable. In that case it does not matter whether neighbors are used that are nearby in the dimensions of that variable, so changing the value of the corresponding weight will not cause an increase in performance. On the other hand, if the behavior of the system depends nonlinearly on a certain variable, it should be possible to improve the performance of the model by using neighbors that are close in terms of this variable. Testing the performance of models for different weight values in the distance function gives an indication of the amount of nonlinearity in the system associated with each variable. Because the different variables might have different physical meanings, the exact values of the weights might have no physical interpretation either. Therefore, it is not the exact value of each weight, but the *ratio* between the weights that is important in expressing the contribution of each variable to nonlinearity. To allow for an equal spread of weight ratios, all variables are scaled to unity variance for use with the nearest neighbor function.

Investigating the performance of different models based on different parts of the state space requires the time-series of observations to be represented in terms of the system state and the forcings, such that similar conditions can be distinguished based on a distance function. At the same time it should be possible to use this representation as the basis of a linear regression model that can deal with external forcings. In order to create such a representation, the state of the system and the external forcings are projected into a combined time-delay vector $s[t]$ as:

$$s[t] = (1, \boldsymbol{x}[t], \quad \ldots, \boldsymbol{x}[t-m], \\ \boldsymbol{f}[t+1], \ldots, \boldsymbol{f}[(t-o)+1]), \tag{1}$$

where $\boldsymbol{x}$ are the system state variables, $\boldsymbol{f}$ the forcing variables, t is the time, m is the embedding dimension of the system state variables, and o the embedding dimension of the forcing variables. The first element of the vector is always a 1 to allow for the inclusion of a constant in models that are based on this vector. For the matter of simplicity each element of the system state variables and forcing variables is given the same embedding dimension. The time between subsequent elements in $\boldsymbol{x}$ and $\boldsymbol{f}$, or *lag time*, is set to 1 time unit for the same reason, but Eq. (1) can of course easily be adjusted to implement different embedding dimensions for different variables or different lag times.

A linear model that can deal with external forcings is the auto-regressive model with external forcings (ARX). In this model, the next state of the system in terms of the model parameters and s is

$$x[t+1] = s[t] \cdot \Psi, \tag{2}$$

where Ψ is the weight matrix of the linear model, with a column for each variable in s. Note that in this equation the next state of the system $x[t+1]$, depends partially on the forcings $\boldsymbol{f}[t+1]$, which corresponds to the time-indexing used for the sandbar dataset. The weights Ψ of the linear model can be

established by performing a linear regression on a number of nearest neighbors and their subsequent system states. To select those neighbors, weighted distances between a vector and all other time-delay vectors are computed by element-wise multiplication of a weight vector $\boldsymbol{w}$ with the difference between $s[t]$ and $s[v]$:

$$\boldsymbol{d}(t, v) = \boldsymbol{w} \times (s[t] - s[v]), \tag{3}$$

where the elements in $\boldsymbol{w}$ that correspond to an element in $\boldsymbol{x}$ or $\boldsymbol{f}$ have the same value for each time-index (no time-decay is used because it will obfuscate the performance differences between different numbers of neighbors). When $\boldsymbol{d}(t, v)$ is computed for each $v{\neq}t$, the k nearest neighbors with the smallest Euclidian distances

$$e = \sqrt{\boldsymbol{d}(t, v) \cdot \boldsymbol{d}^{\mathrm{T}}(t, v)}, \tag{4}$$

can be selected and used for model building.

To get an indication how the performance of the linear models changes with increasing numbers of neighbors and different weight ratios, a number of experiments were performed on synthetic time-series with different relations between the forcings and subsequent system states. To create the surrogate time-series, forcings were drawn from a standard normal distribution. Next, the forcings were used to create four different AR(1) time-series: (a) the next system state depends on a linear combination of the forcing and the previous system state; (b) the next system state depends nonlinearly on the forcings but linearly on the previous system state; (c) the next system state depends nonlinearly on the previous state but linearly on the forcings; (d) the next system state depends nonlinearly on both the forcings and the previous system state. The time-series were embedded in state space according to Eq. (1) with $m{=}o{=}1$. Consequently, models were built based on different numbers of neighbors k, and 1-step-ahead out-of-sample predictions were obtained for those models. This process was repeated for several different ratios of the weights in the nearest neighbor function. The weight vectors $\boldsymbol{w}$, containing the system state weight w_x, and the forcing weight w_f, were taken from the set:

$$\frac{w_f}{w_x} \in \left\{ \frac{0}{q}, \frac{1}{q}, \frac{\cdots}{q}, \frac{q-1}{q}, \frac{q}{q}, \frac{q}{q-1}, \frac{q}{q-2}, \frac{q}{\cdots}, \frac{q}{0} \right\}, \tag{5}$$

with $q{=}4$, resulting in a total of 9 weight ratios.

For each of the four datasets, the absolute error of the 1-step-ahead predictions is plotted against the number of neighbors k, and the weight ratios in Fig. 1. As becomes clear from the four graphs, the error first decreases with increasing numbers of neighbors. If all variables act linearly on the system the error will converge to its minimum regardless of the weight ratio (Fig. 1a). When one or both variables act on the system in a nonlinear fashion, the error gets larger when the number of neighbors is increased further (Fig. 1b–d). For the three nonlinear systems, the location of the minimum error with respect to the ratio of the weights can be used to
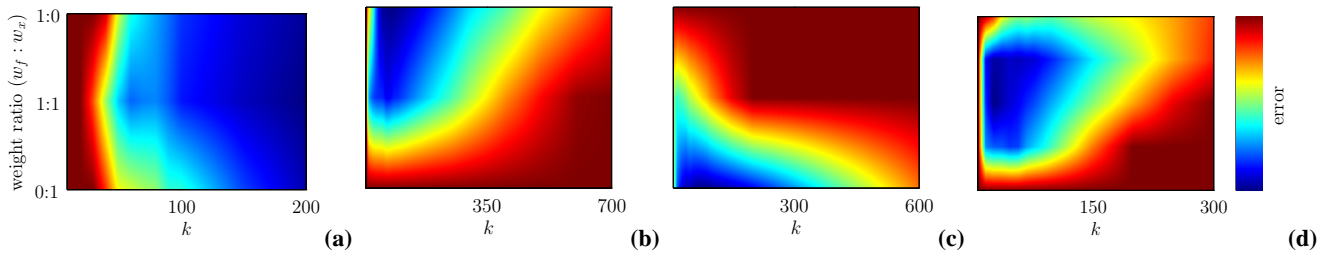
Fig. 1. Course of the error for different numbers of neighbors k and weight ratios $w_f : w_\text{x}$, for: **(a)** linear forcing and linear system state dependencies; **(b)** nonlinear forcing and linear system state dependencies; **(c)** linear forcing and nonlinear system state dependencies; **(d)** nonlinear forcing and nonlinear system state dependencies. Errors increase from blue to red.

infer which variable acts nonlinearly on the system. The best performance for a single nonlinear variable is achieved when only the weight corresponding to that variable has a nonzero value (Fig. 1b, c). Consequently, if both variables act nonlinearly on the system the error will reach its minimum when both weights have nonzero values (Fig. 1d).

3 Experiments

3.1 Sandbar data

In the present work three sandbar datasets from different sources were used. One set consists of in-situ profile measurements, and the other two were obtained using Argus imaging stations (Holman and Stanley, 2007). An Argus station consists of one or more cameras pointed obliquely along the beach, providing an uninterrupted view of the nearshore zone. Each daylight hour, all cameras acquire a time-exposure image (e.g. Fig. 2a) created by averaging over 1200 consecutive images collected at 2 Hz. This averages the individual breaking waves to reveal one or more smooth white bands of breaking waves. The oblique images of the individual cameras can be merged and rectified (Holland et al., 1997) to yield a single planview image (e.g. Fig. 2b). The continuous high-intensity bands that are manifested in the planview images serve as a reasonable estimate for the submerged sandbars (Lippmann and Holman, 1989), and can be extracted from the images by the alongshore tracking of intensity maxima (van Enckevort and Ruessink, 2001) (e.g. Fig. 2b). Since sandbar migration over a single hour is insignificant relative to the accuracy of the rectification and extraction processes, barline extraction is usually performed on a single image at the lowest tide of each day, when the breaking patterns are most pronounced.

The first dataset was gathered by the Argus station at Surfers Paradise, northern Gold Coast, Queensland, Australia (Turner et al., 2004), in the period between 15 July 1999 and 10 April 2007. The second dataset was acquired from the Argus station situated near Egmond, the Netherlands. This dataset starts at 31 May 1999 and continues to 26 April 2007, but data from the years 2000 and 2005 are almost entirely missing due to technical problems.

Barlines were extracted over a stretch of coast of 3 km for the Gold Coast images and 4 km for the Egmond images. Occasionally, sandbar positions could not be computed due to poor image quality (fog or rain droplets on one of the camera lenses), conditions when waves were too low to break, or the malfunctioning of the video acquisition system. To create a continuous time-series dataset, gaps smaller than 2 days were filled by linear interpolation between adjacent breaking-based observations. The majority of the gaps in the data were caused by images that revealed no clear breaker pattern because they were collected during low-energy conditions. Because bar migration is very small under these conditions, we expect that the linear interpolation will have little effect on the results of the nonlinearity detection method. After the extraction process the barlines were averaged in the alongshore direction to yield a single cross-shore position per sandbar each day. From this dataset the periods and sandbars with the best barline accuracy and the most natural sandbar behavior were selected (e.g. the seamost positioned sandbar at Egmond was discarded because waves were seldom high enough to reveal its position; the first part of the inner sandbar at the Gold Coast was left out of the analysis because a beach nourishment was carried out during that period; the decay phases of the outer sandbar at the Gold Coast were not taken into account because the process of extracting barlines from the images was less accurate for these periods). The resulting Gold Coast dataset containing 2143 days of outer sandbar positions and 1914 days of inner sandbar positions is shown in Fig. 3a. The Egmond dataset containing 1865 days of both inner and outer sandbar measurements is depicted in Fig. 3b.

As shown by van Enckevort and Ruessink (2001), the variation in water level is the main cause of the temporal variability in the difference between the video-observed and the actual sandbar position. Sandbar positions that are extracted from images taken at different tidal levels exhibit artificial migration (i.e. migration due to varying tidal levels rather than due to actual sandbar migration). This artificial migration rate depends linearly on the water level difference between two observations (van Enckevort and Ruessink, 2001), and can therefore easily be removed. Each

video-observed sandbar position was projected on a fixed water level (≈ 0.5 m below mean sea level) using in-situ measurements of tidal water levels at Egmond, and the astronomical tide at the Gold Coast, because data on the actual tide (which can differ due to local conditions) was not available for that station.

The third dataset consists of measured water depths obtained from the Hazaki Oceanographical Research Station (HORS), located on the Hazaki coast of Japan facing the Pacific Ocean. At HORS, profile measurements were performed along a 427 m long pier, daily from 1 January 1987 to 1 August 1999, except for weekends and holidays. The profile measurements with a resolution of 5 m, can be used to infer the cross-shore sandbar position. Most of the time there was only a single sandbar, so the position of the sandbar crest was defined as the local maximum elevation of at least 0.3 m at the seamost side of each profile. Although a second, inner sandbar could sometimes be distinguished, the total amount of inner sandbar positions was too small to be considered in the analysis. The parts of the data in which variations in profile height were below the threshold of the local maximum for more than three consecutive days were left out of the analysis. Gaps in the cross-shore sandbar positions time-series smaller than four days were filled using a LOESS interpolation filter (Cleveland and Devlin, 1988). The resulting time-series of 3861 sandbar positions is depicted in Fig. 3c. Note that the gray lines in Figs. 3a–c indicate bad-quality data that were left out of the analysis. Extensive reviews of the HORS dataset are given in Kuriyama (2002) and in Kuriyama and Yanagishima (2006).

Data on the predominant external inputs for driving sandbar variability, the offshore waves, were acquired from several sources. The offshore waves can be represented by their root-mean-squared wave height H_{rms}, peak wave period T_{peak} and wave direction relative to the shore normal ϕ. For the Gold Coast dataset, the variables H_{rms} and T_{peak} were obtained each thirty minutes from the Gold Coast waverider buoy, located about 2 km offshore of the study area, in 16 m water depth. Directional information ϕ was collected hourly by the Brisbane waverider buoy located some 10 km offshore in 70 m water depth, about 100 km north of the study area. The wave climate at Egmond was measured hourly by the IJmuiden06 waverider buoy, located some 40 km offshore in 21 m water depth. At HORS, data on the day-averaged wave heights and periods were available from an ultrasonic wave gage located 5 km offshore in 24 m water depth. Because no directional information was available, the predominant direction of incoming waves at HORS (30°) was used when necessary.

The three variables expressing the properties of the incoming waves were combined into the wave height at breaking (Plant et al., 1999)

$$H_{\text{b}} = \left(\frac{\gamma}{g}\right)^{\frac{1}{5}} \left(H_{\text{rms}}^2 \cdot c_{\text{g}} \cdot \cos(\phi)\right)^{\frac{2}{5}}, \tag{6}$$

where the offshore group velocity c_{g} was computed using linear wave theory involving T_{peak} and the water depth at the location of the offshore wave measurement, the gravitational acceleration g is $9.81 \, \text{m s}^{-2}$, and the wave height to water depth ratio at breaking γ was set to 0.4, a typical field value (Thornton and Guza, 1983). If multiple observations of H_{rms}, T_{peak} and ϕ were available during a day, H_{b} was computed from each 3-tuple and then averaged to yield a single value between two subsequent sandbar observations. The time-averaged H_{b} value was given the time index of the second observation, which explains the formulation in Eq. (1). Missing H_{b} values were filled by linear interpolation. The time-series of the hydrodynamic data are depicted in Fig. 3d–f.

As noted in Sect. 2, the incoming waves have to be considered as an external forcing factor because they are measured offshore. Moreover, the actual wave forcing at different locations in the cross-shore direction might be only weakly related to the offshore-measured forcing. For the outer sandbar this is not very problematic, because the properties of the waves do not change much between the offshore location of the buoy and the location of the outer sandbar. Waves that break on the inner sandbar however, have undergone a complex transformation before their arrival. If the waves are high enough, they break on the outer sandbar first. The waves then reform in the trough between the outer and the inner sandbar, and break again on the inner sandbar. Although no surfzone wave measurements are available, it is still possible to obtain information on changing wave properties in the nearshore by using a wave-transformation model. Such models compute the various changing properties of the waves over a cross-shore profile, based on offshore-measured wave properties. Whereas computing surfzone wave properties for the HORS profile dataset would not make much sense because there is only a single sandbar, very few profile data for the locations in the Argus datasets were available. For the Gold Coast a single profile measurement was performed in June, 2002 and at Egmond five profiles measured annually in 1999–2003 were available.

While it is not possible to compute accurate wave properties at different cross-shore locations for the Argus datasets, the available profiles can still be used to get at least some indication of wave properties in the nearshore zone. The Battjes-Janssen (e.g. Reniers and Battjes, 1997) wave-transformation model with default parameters (Battjes and Janssen, 1978; Battjes and Stive, 1985) was evaluated for each hourly-averaged offshore wave measurement and the profile at the closest profile-measurement date. The root-mean-squared wave height H_{rms} at ≈ 50 m offshore of the outer sandbar, and within the trough between the outer and inner sandbar were then averaged over the time between two subsequent sandbar observations, just as in the computation of H_{b}. The outcome of the Battjes-Janssen model for the two Argus datasets is given in Fig. 3g, h. For completeness, the offshore H_{rms} at HORS is also given in Fig. 3i. The result of

the Battjes-Janssen model, based on a only a few measured profiles, is by no means an accurate representation of the actual wave heights. However, given the available models and information this is the only way to infer at least some properties of the waves at different positions in the profile. This information can then be used to investigate the importance of the external forcing and where it is measured.

As can be inferred from Fig. 3 the most pronounced feature of sandbar behavior is the rapid offshore movement under high wave energy conditions (storms), and the relatively slow onshore return under relatively quiescent wave conditions. There is however a large difference between the sites in the sensitivity of sandbar migration to the height of the wave forcing. The sandbars at the Gold Coast, showing a maximum migration of $75\,\mathrm{m\,d}^{-1}$, are the most responsive to high energy conditions. Whereas the wave height at breaking at HORS can be much larger during the depressions in winter and the typhoons in autumn, the maximum observed sandbar migration at HORS is $50\,\mathrm{m\,d}^{-1}$. At Egmond the sandbars are even less-responsive to high energy conditions, showing a maximum of $30\,\mathrm{m\,d}^{-1}$. Patterns of alternating directions of sandbar migration are linked to the external forcing and might reflect the seasonal weather patterns as well. However, the resulting net migration of each sandbar crest position is always offshore. When a sandbar continues to move offshore, the water depth above the sandbar increases (due to the beach slope) until finally the depth becomes too large for incoming waves to break. At that point the driving force that maintains the sandbar ceases, and the sandbar disintegrates. In the first part of the HORS dataset in Fig. 3c this happens almost every year near the end of the typhoon season, after which a new sandbar is formed close to the shore. This 1-year period cyclic behavior changes after 1993, probably due to changing wave conditions (Kuriyama and Yanagishima, 2006). At the Gold Coast the outer sandbar also decays after one or more major storms (e.g. early 2006 in Fig. 3a). After such an event the existing inner sandbar moves seaward and becomes the new outer sandbar, while a new inner sandbar is formed at the shore. In the period shown in Fig. 3a the total lifespan of a sandbar, from its formation at the shoreline to its decay seaward of the breaker zone, is ≈4 years, giving rise to a periodicity with an interval of ≈2 years. Although at Egmond no sandbar decay event takes place during the studied period, it is known from earlier studies of the Dutch coast that the same pattern of formation and disintegration of sandbars is also present at Egmond, be it at a much slower pace of ≈15 years (Wijnberg and Terwindt, 1995).

As mentioned before, at Egmond there is a third sandbar at the seamost edge of the breaker zone which could not accurately be analyzed from the images. Although here we call the second sandbar counted from the shore the outer sandbar, it is occasionally sheltered from very high waves by the actual outer sandbar, as becomes clear from a comparison between Fig. 3e and h. This means that the second sandbar at Egmond may show behavior that is more intermediate between inner and outer sandbars. Together, all these different sandbars at locations with various boundary conditions are a good representation of the entire spectrum of sandbar behavior and comprise a broad basis for generalizing the results of the nonlinearity analysis.

3.2 Setup

To investigate the nonlinearity related to cross-shore sandbar positions and wave forcing, the MDVS method was applied to the Gold Coast, Egmond and HORS datasets. The system state variable was the cross-shore sandbar position, and H_b and H_{rms} computed at different locations were used as forcing variables. Inner and outer sandbars were studied in separate experiments. The effect of using offshore and locally computed forcings was also studied by using H_{rms} computed at the outer sandbar as forcing for the outer sandbar, and H_{rms} computed in the trough between the outer and inner sandbars as forcing for the inner sandbar. For comparison, H_b was also used as forcing variable in additional experiments.

First, the variables were embedded in state space. Although several methods exist to find optimal values for the lag time and embedding dimension of univariate data (e.g. first zero of autocorrelation, mutual information), finding these values for interdependent multivariate data is not straightforward. The dataset consists of one low-tide image or one profile measurement per day, so a lag time shorter than 1 day could not easily be achieved. A lag time that is too small relative to the timescale of the studied characteristics becomes only a problem when the embedding dimension is very small too (Kantz and Schreiber, 1997). Since several embedding dimensions were tested here, no detailed investigation into the optimal lag time needs to be performed, so the lag time was set to 1 day. Embedding dimensions $m=o=1,\ldots,10$ were tested in separate experiments. The capacity of linear models to simulate nonlinear dynamics increases for larger embedding dimensions. As a result, no noticeable performance difference between linear and nonlinear models was achieved for embedding dimensions larger than 6, so the presentation of the results is limited to $m=o=1,\ldots,6$.

For each sandbar, errors were computed as the mean out-of-sample absolute prediction error over all but the first m samples. Absolute errors were used because they are less sensitive to outliers than methods based on squared differences between forecasts and observations. The prediction error for a model was computed using n-step-ahead prediction. This process starts by running the model with observed system state variables at the first timestep, while model predictions of the system state are used during the next n timesteps. After that, each n timesteps observed system states instead of previously predicted states are used as input. The forcing values are not predicted, so the observed forcing values are provided to the model each timestep. It is expected that the performance difference between linear and nonlinear
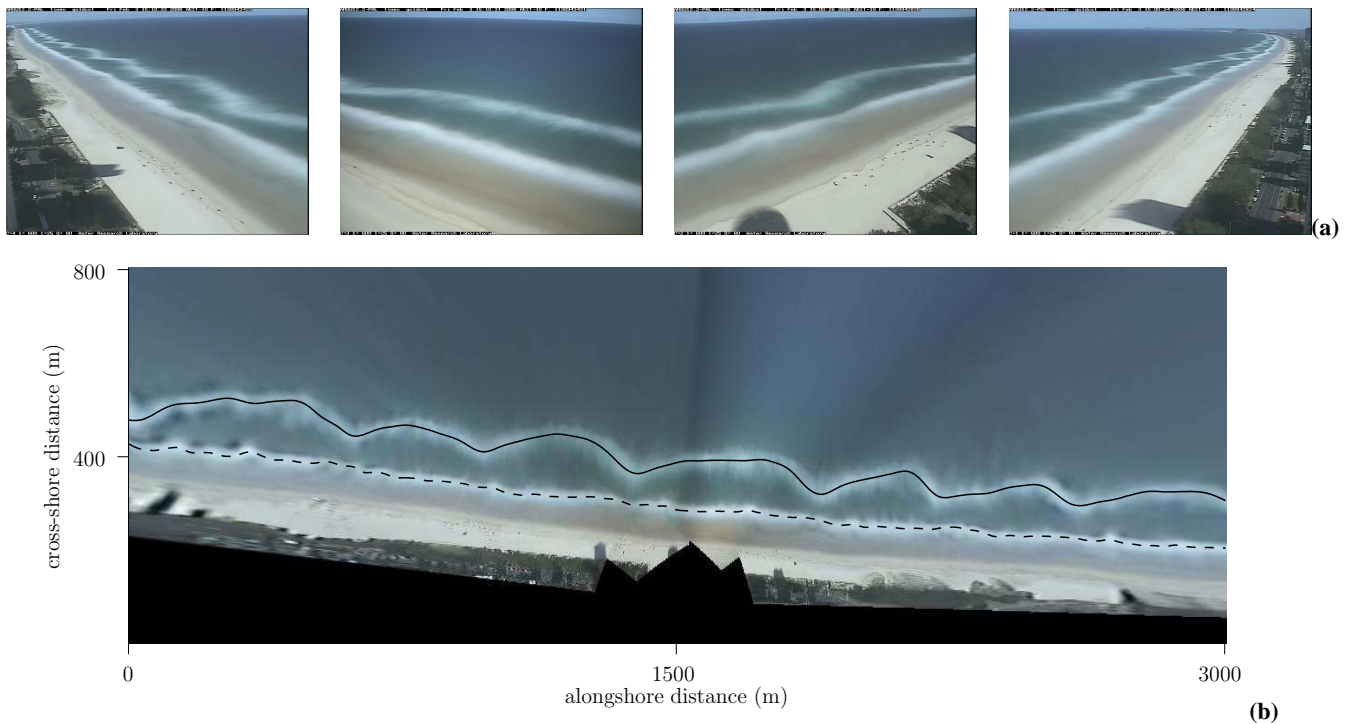
Fig. 2. Argus camera images, merged plan view and tracked barlines.
(a) Time-exposure Argus images of all cameras; the high-intensity bands in each image are due to persistent wave breaking on the sandbars.
(b) Tracked outer (solid) and inner (dashed) barlines in plan view.

models becomes more pronounced with increasing prediction horizons (e.g. Sugihara and May, 1990). Therefore not only the one-lag (day) ahead prediction ($n=1$) was used, as is usual in the DVS method, but also larger prediction horizons up to $n=14$ were tested. An example of observations and actual linear and nonlinear model outputs for even longer prediction horizons is given in Pape et al. (2007). Since we are primarily interested in differences between models based on different numbers of neighbors, the error is represented as the percentage of the error of an ARX model that is based on all samples. The local models and the full ARX model to which they are compared are based on the same lag time, embedding dimension and prediction horizon, ensuring a fair comparison between the local and global models.

Results were obtained for different weight values of the system state weight w_x, and the forcing weight w_f. Weight ratios were taken from the set defined in Eq. (5) with $q=8$, resulting in a total of 17 weight ratios. Figure 4 shows the results of the MDVS method for all sandbars in the three datasets with different forcings. In each image the horizontal axis represents the number of neighbors that was used to build a model for each sample, and the vertical axis the weight ratio between the forcing and system state variable that was used in the nearest neighbor distance function. In this fashion, the course of the error and the location of the minimum error can easily be observed in the pictures. For most embedding dimensions the location of the optimum and

the course of the error were very similar (although the actual values could differ). Therefore, the results were averaged over $m=o=1, \ldots, 6$. As a result, any difference in the location of the optimum among different embedding dimensions becomes visible as a scattered pattern (e.g. Fig. 4f). For shorter prediction horizons the results of different embedding dimensions were reasonably close, but for prediction horizons larger than ≈ 7 days, often no clear optimum or other meaningful patterns were visible. To give an indication how the results of the MDVS method change with increasing prediction horizons, the results for both $n=1$ and $n=7$ are shown in Fig. 4.

3.3 Interpretation of MDVS results

As explained in Sect. 2, increasing the number of samples on which a model is based improves its prediction accuracy due to noise reduction. All images in Fig. 4 exhibit a decrease in error when the number of neighbors is increased to several tens or hundreds. However, when the number of neighbors is increased even further, the errors start to increase again. Piece-wise linear models based on a smaller neighborhood yield more accurate predictions than models based on a large number of neighbors, which is evidence for nonlinearity. As can be inferred from the scales of the colorbars in Fig. 4, the MDVS results of the different sandbars vary in the improvement of nonlinear models with respect to a full linear model. The largest improvement ($\approx 15\%$) is obtained

for the outer sandbar at the Gold Coast (Fig. 4c). The differences are smaller for the other sandbars, ranging from 6% for the inner sandbar at the Gold Coast (Fig. 4d) to only 1.2% for the sandbar at HORS (Fig. 4i). While nonlinear models outperform linear models for all cases considered, the effects of nonlinearity are generally small.

Not only the value of the optimum error percentage varies among different sandbars, also the location of the optimum relative to the weight ratios differs in the images in Fig. 4. In most images the error decreases when the weight of the forcing variable w_f increases (from bottom to top), but starts to decrease again when this weight increases beyond a certain value. It can be deduced from a comparison with Fig. 1d that this is caused by nonlinearity that is related to both the forcing and system state variable. For some results however (e.g. Fig. 4a, d and c), the error does not start to increase again when the $w_f : w_x$ ratio becomes large. In these cases the optimum performance is achieved for $w_x=0$, which means that similarities in the system state are totally discarded in the process of neighbor selection. This situation is the same as the outcome of Fig. 1b, in which only the forcing factor acts nonlinearly on the system. Unlike Fig. 1b, the course of the error in the images in Fig. 4 sometimes also reaches an optimum for a certain k when $w_f=0$ (e.g. Fig. 4d), but that does not necessarily mean that the relation between subsequent system states is nonlinear. The reason this is not taken as an indication for nonlinearity in the system state variable, is that the dependency between the system state and the forcing works in both ways. Although this cannot be true in a strict sense, this can be explained by a further analysis of the lag time in relation to the nature of sandbar behavior. During the onset of a storm, a sandbar moves offshore rapidly and stays there during the rest of the storm. When after the storm the energy of the waves becomes smaller, the sandbar starts to move onshore. Since most storms last for several days, seaward located sandbar positions often coincide with high-energy conditions. When neighbors are chosen for a seaward located sandbar based on sandbar location only, the nearest neighbor algorithm will unintendedly also select high-energy conditions. In other words, the duration of the important events is large compared to the resolution of the data. However, it is still possible to infer something about the importance of each variable to nonlinearity because the results for several embedding dimensions are given. Although the error might show an optimum value over the transect $w_f=0$, the location of the optimum with respect to the weight ratio $w_f : w_x$ indicates the nonlinearity related to that variable. Only when both variables act nonlinearly on the system the error will reach its minimum when both weights have nonzero values. For some of the results (e.g. Fig. 4g, i) this is clearly the case, while for other results the optimum is reached when $w_x=0$ (e.g. Fig. 4a, c).

Figure 4 also shows the results for different forcing variables: H_b; H_{rms} computed at the outer sandbar, which was used as forcing for the outer sandbar; and H_{rms}

computed in the trough between the outer and inner sandbar, used as forcing for the inner sandbar. When H_{rms} was used for the outer sandbar instead of H_b, the location of the optimum with respect to the number of neighbors or the weight ratios did not change much (see Fig. 4a, c, e, g). This was as expected, because the wave height at breaking H_b and H_{rms} computed at the outer sandbar are very well correlated (see Fig. 3). The most noticeable difference between using H_b and H_{rms} as forcings for the inner sandbars (Fig. 4b, d, f, h) is the shift of the minimum error with respect to the weight ratio on the vertical axes. When H_b is used as forcing, the optimum is reached when both w_f and w_x have nonzero values, indicating nonlinearity associated with both variables. Models that use H_{rms} as forcing factor reach their optimum performance when $w_x=0$, so selecting neighbors based on similarity in locally computed forcings only, gave optimal results. In other words, also considering similar system states in the process of model building did not result in a reduction of the prediction error. This indicates that given the locally computed forcing, a linear relation between subsequent sandbar positions suffices.

When nonlinearity is present, the performance difference between linear and nonlinear models will become more pronounced with increasing prediction horizon. For most MDVS results this was indeed the case up to at least $n=10$. Surprisingly, this trend was not found in the results of the outer sandbar at Egmond (Fig. 4e and g), up to at least $n=14$. This can partially be attributed to the fact that the minima of different embedding dimensions become spread out for increasing n, and do not overlap. In other results with scattered minimum error values (e.g. Fig. 4f and i) the performance difference also started to diminish for $n>10$. Using iterative prediction based on previous model outcomes allows the nonlinear models to wander further away from the observed system state. For linear models this happens more slowly. Given the small differences between linear and nonlinear models, and the scattered locations and values of the minimum errors among the different embedding dimensions, this is another possible cause for the decreasing difference between both types of models for large n.

3.4 Discussion of the MDVS results

The MDVS method is meant to detect nonlinearity in the relations between sandbar crest positions and wave forcings. It is however conceivable that some of the detected nonlinearity is actually caused by processes we refer to as noise. During the data accumulation procedure several steps are carried out that might introduce additional nonlinearity in the data. For example, the difference between the position of the highest intensity, which is used as indication for the sandbar crest position, and the in-situ crest position is known to depend nonlinearly on the offshore wave height (van Enckevort and Ruessink, 2001). A change in the wave height between two subsequent image-based observations
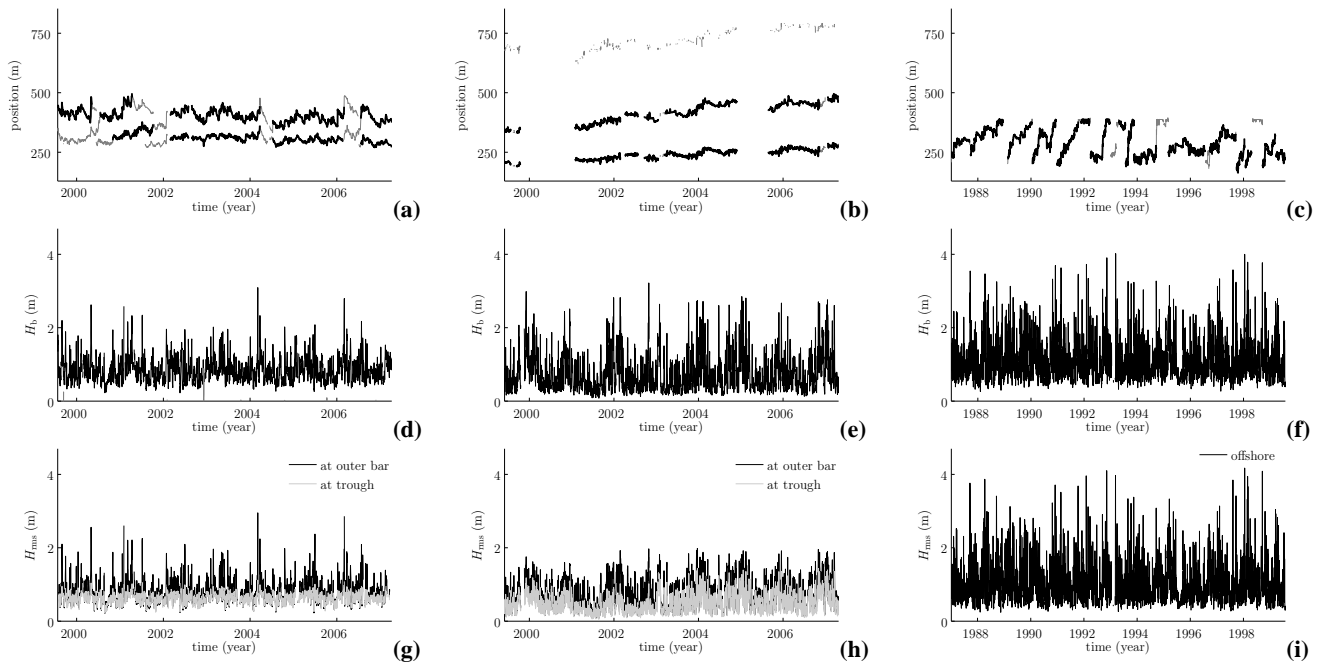
Fig. 3. Overview of the variables in the datasets: **(a)** sandbar positions at the Gold Coast; **(b)** sandbar positions at Egmond; **(c)** sandbar positions at HORS; **(d)** hydrodynamic data at the Gold Coast; **(e)** hydrodynamic data at Egmond; **(f)** hydrodynamic data at HORS; **(g)** H_{rms} over profile at the Gold Coast; **(h)** H_{rms} over profile at Egmond; **(i)** offshore H_{rms} at HORS.

causes a change in both the location of the highest breaking intensity relative to the sandbar crest, and possibly also in the actual sandbar position. The linear part of the migration induced by the change of the highest breaker intensity can be represented by linear models for embedding dimensions larger than one. Still, it might be that the nonlinear part of this relation affects the results of the MDVS outcome.

Another step in the data accumulation procedure involving Argus images, is the averaging process that is applied to find the mean cross-shore sandbar position. As can be seen in Fig. 2, the cross-shore position can vary in the alongshore direction. If the migration of a sandbar depends nonlinearly on its cross-shore position, the different alongshore parts of the sandbar might migrate different distances or even in different directions under the same forcing conditions. Some of this potential nonlinearity is averaged out by taking the mean cross-shore sandbar position. On the other hand, two sandbars with the same mean cross-shore position but different alongshore variability might react different to the same forcing conditions. Part of the nonlinearity found in the observed mean cross-shore position might thus be due to discarding alongshore variability in the models. For example, the outer bar at the Gold Coast often contains a lot of alongshore variability (e.g. Fig. 2), while the difference between linear and nonlinear models also is the largest. However, site-specific issues during data-collection might have introduced additional noise, thus obscuring the differences between sites. At the Gold Coast, breaker patterns are well-pronounced, light and atmospheric conditions are excellent,

and the cameras are steady and almost never malfunction. The images in the Egmond dataset are much more difficult to interpret due to less perfect conditions, especially the vibration of the 50 m high tower in the wind. Sandbar crest positions in the HORS dataset are defined as the most seaward located optimum in the profile, which might not always be well-pronounced. For the outer sandbar at the Gold Coast the effects of nonlinearity are most pronounced, but it might well be that at other sites the effects of nonlinearity are reduced by larger amounts of noise introduced during data gathering.

Together, all these different processes and varying conditions might introduce additional nonlinearity in the observations, but might also cause the system to become more noise-dominated. Since we are dealing with multiple dependencies in multivariate data, existing methods (e.g. Sugihara and May, 1990; Sugihara, 1994) cannot easily be used or adjusted to determine or reduce the effects of noise. However, if the detected nonlinearity was partially caused by nonlinearity in the processes that we refer to as noise, the importance of nonlinearity in the relations between actual sandbar positions and wave forcing is even smaller. At HORS the sandbar location are inferred from actual profiles, which probably introduces less nonlinearities in the data than the extraction procedure for the Argus images. The finding that the importance of nonlinearity at HORS is the smallest supports this. On the other hand, the sandbar positions at HORS, which are only measured at a single alongshore location, might be more noise-dominated than the sandbar positions derived from averaging over 3 km alongshore or more in the Argus images.
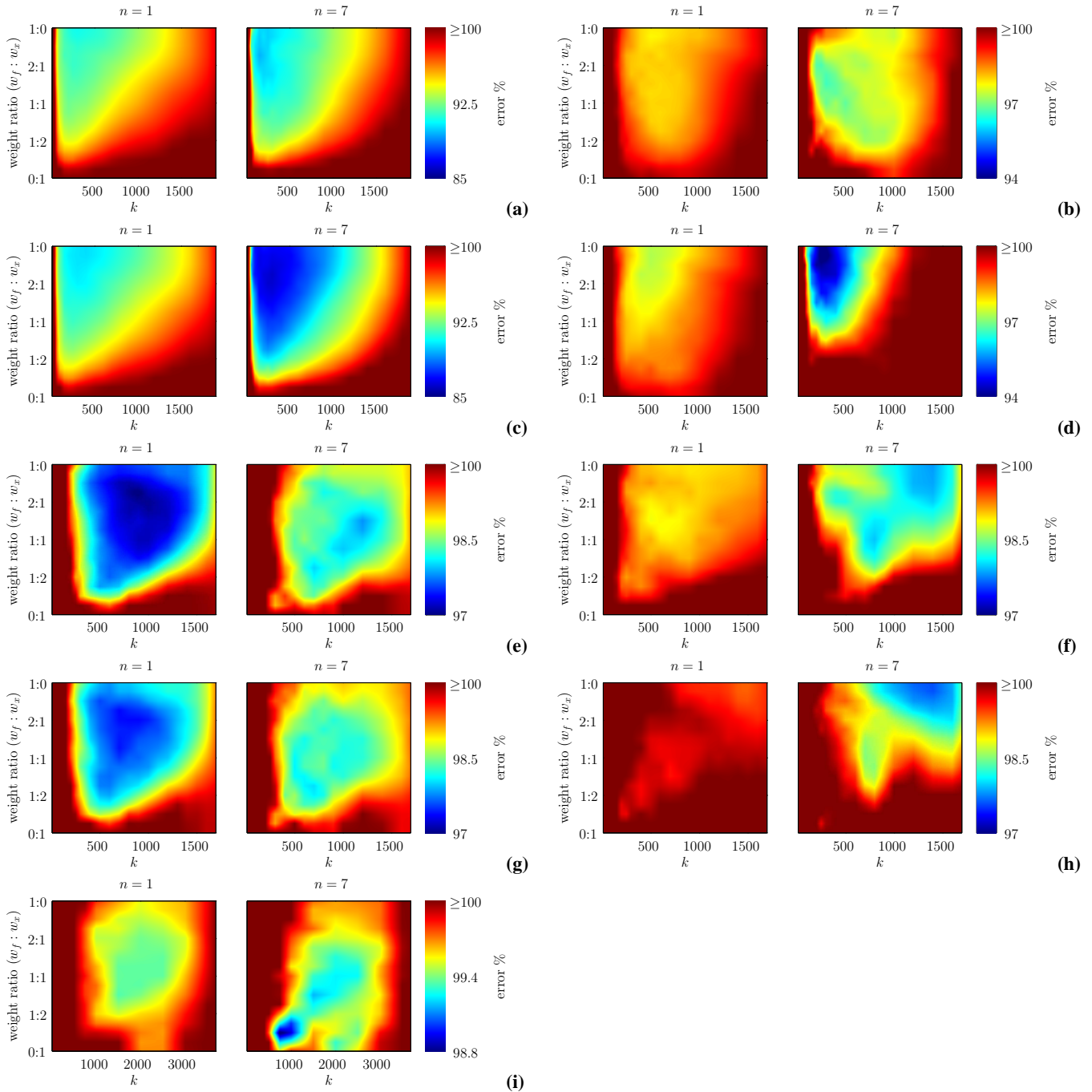
Fig. 4. MDVS results for all sandbars with different forcings (H_b and H_{rms}) and prediction horizons n. Error values are computed as the mean out-of-sample absolute error, averaged over $m=o=1, \ldots, 6$. The images contain the error percentages relative to a full linear model, at each number of neighbors k, and weight ratio $w_f : w_x$. **(a)** Gold Coast, outer sandbar, forcing: H_b. **(b)** Gold Coast, inner sandbar, forcing: H_b. **(c)** Gold Coast, outer sandbar, forcing: H_{rms} at outer sandbar. **(d)** Gold Coast, inner sandbar, forcing: H_{rms} at trough. **(e)** Egmond, outer sandbar, forcing: H_b. **(f)** Egmond, inner sandbar, forcing: H_b. **(g)** Egmond, outer sandbar, forcing: H_{rms} at outer sandbar. **(h)** Egmond, inner sandbar, forcing: H_{rms} at trough. **(i)** HORS, forcing: H_b.

An additional reason for the variations in the importance of nonlinearity among different sandbars might be found in the available amount of similar states. In the MDVS results of all sandbars, the optimum number of neighbors is reached within the available amount of samples. However,

the location of this optimum is not the same for each sandbar. The different periods of cyclic behavior cause differences in the availability of similar states for each sample. For example, at Egmond the locations of the sandbars do not change much within the dataset (Fig. 3b), implying that

many similar states in terms of sandbar position can be found by the nearest neighbor algorithm. Only when models are based on several hundreds of samples, the neighborhood contains enough different samples for the effect of nonlinearity in the system state to become more important than noise reduction (e.g. Fig. 4e, g). On the other hand, a comparison with the HORS dataset in Fig. 3c and its MDVS results in Fig. 4i reveals that differences in optimum k and the relation to sandbar behavior cannot immediately be inferred from the MDVS results. Apart from the differences that arise from data-gathering methods, the available amount of data relative to the length of the cyclic patterns in the sandbar's behavior determines the location of the optimal number of neighbors. While in all datasets used here the available amount of samples is only two or three times as large as the optimal number of neighbors, the true effect of availability of similar states can only be investigated in detail when the amount of cycles in sandbar behavior is close to the optimal k.

As discussed in Sect. 3.3, the MDVS results can be interpreted in terms of the importance of nonlinearity for both the system state and the forcing variable. Sugihara (1994) discusses a case with a single variable and noise, in which the variability induced by the noise process causes the system to behave in a different mode (e.g. move toward or away from a stable state) for different noise levels. Similarly, it might be possible that a sandbar system changes between different modes of behavior, which means that nonlinearities associated with the forcing and the dynamics of the sandbar system are inherently inseparable. If the sandbar system were to change between different modes of behavior, the total behavior of the system would still be nonlinear. Whereas the MDVS results only show nonlinearity for the entire system, that is over all values of the variables, investigating the importance of nonlinearity for different values of each variable might reveal these different modes. To examine the possibility of different modes of behavior, the performance difference between linear and nonlinear models was evaluated at individual values of the cross-shore sandbar position and wave forcing variables. This analysis included all sandbar cases and prediction horizons up to $n=14$, but no relation was found whatsoever. We have to admit that the significance of any relation would have been doubtful given the amount of data and the noise in the available data.

3.5 Implications for modeling

Our findings that linear models can predict observed cross-shore sandbar behavior almost as accurate as nonlinear models implies that simple data-driven models might suffice to model observed cross-shore bar behavior. This is consistent with Pape et al. (2007), who also found small differences between the performance of several linear and nonlinear data-driven models, even over prediction horizons up to 2 year. Also Plant et al. (1999) found that a relatively simple model (calibrated on the observations) could explain up to 80% of

the observed bar position variability. There are two possible causes for the small difference between linear and nonlinear models: the system is actually linear, or the data gathered from the system is dominated by noise.

If cross-shore sandbar behavior is actually linear, there are important implications for the way in which morphological behavior can best be predicted. Process-based models that are used to describe the evolution of the nearshore zone are based on scales that are both temporally and spatially very small compared to a sandbar and its behavior. Whereas process-based models can compute the evolution of the morphology for the entire nearshore zone, the data-driven models used in the present work are based on a small number of larger-scale variables that represent the significant changes in underlying morphology (Plant et al., 2001). The interpretation that sandbar behavior is linear might be founded in the fact that small-scale physical processes can together evolve towards a simple attractor state. While process-based models based on sediment flux computations in fine grids and small timesteps might suffer from numerical instability, or the build-up of errors in the process formulations, statistical models might perform better on larger prediction horizons because they are based on stable statistical properties of attractor states. The long-term evolution of the nearshore zone might therefore be predicted more accurately using relatively simple models based on just a few large-scale variables. For a more elaborate discussion, see Southgate and Möller (2000) and Werner (1999, 2003).

On the other hand, if the small difference between linear and nonlinear models is caused by the presence of noise, the underlying system might still be highly nonlinear. As long as the noise is not reduced, there is no way to rule out the possibility that the underlying system is actually highly nonlinear, meaning that a nonlinear model might be much better in predicting the actual system state. More complex nonlinear models such as the process-based models discussed before might thus be able to model the actual morphology and sandbar position more accurately. There is however no way to validate this claim. Since the datasets used in the present work are among the highest quality and resolution datasets in existence, the accuracy of any type of model would have to be validated against this kind of data. For prediction purposes, the question is not how accurate a model can predict actual sandbar positions, but how accurate it predicts the *observed data*. In case of the Argus data this would not necessarily mean the prediction of the sandbar crest position, but could also be the position of the maximum roller dissipation, the analog of image breaker intensity (Aarninkhof et al., 2005).

4 Conclusions

We developed a multivariate extension to the DVS method of Casdagli and Weigend (1993), termed MDVS, and applied it

to three sandbar datasets, containing a total of five different sandbars and their corresponding external wave forcings. The results of the MDVS method revealed nonlinearity in the relations between the offshore wave forcing and sandbar migration. When wave height estimates at the location of the sandbar were used rather than offshore wave heights, the importance of nonlinearity related to the forcings increased relative to the importance of nonlinearity related to subsequent system states. The differences between nonlinear models and a linear model were however small (the Gold Coast$<$15%, Egmond$<$6% and HORS$<$2%) even when sandbar positions were predicted over several days ahead. It might even be that part of the detected nonlinearity is not present in actual sandbar behavior, but is induced by the data gathering procedure. On the other hand, the presence of noise and the relatively small amount of available data with respect to the period of the cyclic patterns in the data might obfuscate the importance of nonlinearity in the underlying system. Data similar to the sandbar datasets used in the present work are often used for calibrating and validating process-based models. Our finding that linear models are almost as accurate as nonlinear models on the sandbar datasets implies that the prediction of large-scale features in the nearshore zone might not benefit much from the use of such complex nonlinear models.

Edited by: H. A. Dijkstra
Reviewed by: K. Bryan and another anonymous referee

References

Aarninkhof, S., Ruessink, B. G., and Roelvink, J. A.: Nearshore subtidal bathymetry from time-exposure video images, J. Geophys. Res., 110, C06011, doi:10.1029/2004JC002791, 2005.

Abraham, F. D.: Nonlinear coherence in multivariate research: invariants and the reconstruction of attractors, Nonlinear Dynamics, Psychology and Life Sciences, 1, 7–33, 1997.

Battjes, J. A. and Janssen, J. P. F. M.: Energy loss and set-up due to breaking of random waves, in: Proceedings of the 16th Conference on Coastal Engineering, 569–587, ASCE, 1978.

Battjes, J. A. and Stive, M. J. F.: Calibration and verification of a dissipation model for random breaking waves, J. Geophys. Res., 90, 9159–9167, 1985.

Box, G. E. P. and Jenkins, G. M.: Time series analysis: Forecasting and control, Holden-Day, San Francisco, 1970.

Cao, L., Mees, A., and Judd, K.: Dynamics from multivariate time series, Physica D, 121, 75–88, 1998.

Casdagli, M.: A dynamical systems approach to modeling input-output systems, 265–281, Addison-Wesley, 1992.

Casdagli, M. C. and Weigend, A. S.: Exploring the continuum between deterministic and stochastic modeling, in: Time Series Prediction: Forecasting the Future and Understanding the Past, edited by Weigend, A. S. and Gershenfeld, N. A., 347–366, Addison-Wesley, Massachusetts, 1993.

Cleveland, W. S. and Devlin, S. J.: Locally weighted regression: an approach to regression analysis by local fitting, J. Am. Stat. Assoc., 83, 596–610, 1988.

Elgar, S.: Coastal profile at Duck, North Carolina: A cautionary note, J. Geophys. Res., 106, 4625–4627, 2001.

Gallagher, E. L., Elgar, S., and Guza, R. T.: Observations of sand bar evolution on a natural beach, J. Geophys. Res., 103, 3203–3215, 1998.

Garcia, S. P. and Almeida, J. S.: Multivariate phase space reconstruction by nearest neighbor embedding with different time delays, Phys. Rev. E, 72, 027205, doi: 10.1103/PhysRevE.72.027205, 2005.

Hoefel, F. and Elgar, S.: Wave-induced sediment transport and sandbar migration, Science, 299, 1885–1887, 2003.

Holland, K. T., Holman, R. A., Lippmann, T. C., Stanley, J., and Plant, N. G.: Practical use of video imagery in nearshore oceanographic field studies, J. Oceanic Engineering, 22, 81–92, 1997.

Holman, R. A. and Stanley, J.: The history and technical capabilities of Argus, Coast. Eng., 54, 477–491, 2007.

Jaffe, B. E. and Rubin, D. M.: Using nonlinear forecasting to learn the magnitude and phasing of time-varying sediment suspension in the surf zone, J. Geophys. Res., 101, 14 283–14 296, 1996.

Kantz, H. and Schreiber, T.: Nonlinear time series analysis, Cambridge University Press, Cambridge, first Edn., 1997.

Kuriyama, Y.: Medium-term bar behavior and associated sediment transport at Hasaki, Japan, J. Geophys. Res., 107, 3132, doi:10.1029/2001JC000899, 2002.

Kuriyama, Y. and Yanagishima, S.: Medium-term variations of bar properties and their linkages with environmental factors at HORS, Tech. Rep. 4, Kashima port and airport research institute, Kashima, 2006.

Lippmann, T. C. and Holman, R. A.: Quantification of sand bar morphology: a video technique based on wave dissipation, J. Geophys. Res., 94, 995–1011, 1989.

Pape, L., Ruessink, B. G., Wiering, M. A., and Turner, I. L.: Recurrent neural network modeling of nearshore sandbar behavior, Neural Networks, 20, 509–518, 2007.

Plant, N. G., Holman, R. A., Freilich, M. H., and Birkemeier, W. A.: A simple model for interannual sandbar behavior, J. Geophys. Res., 104, 15 755–15 776, 1999.

Plant, N. G., Freilich, M. H., and Holman, R. A.: Role of morphologic feedback in surf zone sandbar response, J. Geophys. Res., 106, 973–989, 2001.

Plant, N. G., Holland, T., and Holman, R. A.: A dynamical attractor governs beach response to storms, Geophys. Res. Lett., 33, L17607, doi:10.1029/2006GL027105, 2006.

Reniers, A. J. H. M. and Battjes, J. A.: A laboratory study of longshore currents over barred and non-barred beaches, Coast. Eng., 30, 1–21, 1997.

Roelvink, J. A. and Brøker, I.: Cross-shore profile models, Coast.

Eng., 21, 163–191, 1993.

Rubin, D. M.: Forecasting techniques, underlying physics and applications, chap. 5, SEPM Short Course No. 36, Society for Sedimentary Geology, 1995.

Ruessink, B. G., Kuriyama, Y., Reniers, A. J. H. M., Roelvink, J. A., and Walstra, D. J. R.: Modeling cross-shore sandbar behavior on the timescale of weeks, J. Geophys. Res., 112, F03010, doi:10.1029/2006JF000730, 2007.

Sauer, T., Yorke, J. A., and Casdagli, M.: Embedology, J. Stat. Phys., 65, 579–616, 1991.

Schreiber, T. and Schmitz, A.: Surrogate time series, Physica D, 142, 346–382, 2000.

Southgate, H. N. and Möller, I.: Fractal properties of beach profile evolution at Duck, North Carolina, J. Geophys. Res., 105, 11 489–11 507, 2000.

Sugihara, G.: Nonlinear Forecasting for the Classification of Natural Time Series, Philosophical Transactions: Physical Sciences and Engineering, 348, 477–495, 1994.

Sugihara, G. and May, R. M.: Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series, Nature, 344, 734–742, 1990.

Takens, F.: Detecting strange attractors in turbulence, in: Dynamical systems and turbulence, Lecture notes in mathematics, 898, 366–381, Springer, Berlin, 1981.

Thornton, E. B. and Guza, R. T.: Transformation of wave height distribution, J. Geophys. Res., 88, 5925–5938, 1983.

Turner, I. L., Aarninkhof, S. G. J., Dronkers, T. D. T., and McGrath, J.: CZM applications of Argus coastal imaging at the Gold Coast, Australia, J. Coastal Res., 20, 739–752, 2004.

van Enckevort, I. M. J. and Ruessink, B. G.: Effects of hydrodynamics and bathymetry on video estimates of nearshore sandbar position, J. Geophys. Res., 106, 16 969–16 979, 2001.

van Rijn, L. C., Walstra, D. J. R., Grasmeijer, B., Sutherland, J., Pan, S., and Sierra, J. P.: The predictability of cross-shore bed evolution of sandy beaches at the time scale of storms and seasons using process-based profile models, Coast. Eng., 47, 295-327, 2003.

Werner, B. T.: Complexity in natural landform patterns, Science, 284, 102–104, 1999.

Werner, B. T.: Modeling landforms as self-organized, hierarchical dynamical systems, Geophysical Monograph, 135, 133–150, AGU Press, 2003.

Wijnberg, K. M. and Terwindt, J. H. J.: Extracting decadal morphological behaviour from high-resolution, long-term bathymetric surveys along the Holland coast using eigenfunction analysis, Mar. Geol., 126, 301–330, 1995.