

Review Article

Guidance on Setup, Calibration, and Validation of Hydrodynamic, Wave, and Sediment Models for Shelf Seas and Estuaries

Jon J. Williams¹ and Luciana S. Esteves²

¹*Mott MacDonald House, 8-10 Sydenham Road, Croydon CR0 2EE, UK*

²*Bournemouth University, Talbot Campus, Poole, Dorset BH12 5BB, UK*

Correspondence should be addressed to Jon J. Williams; jon.williams@mottmac.com

Received 2 August 2017; Accepted 4 October 2017; Published 19 December 2017

Academic Editor: Arnaud Perrot

Copyright © 2017 Jon J. Williams and Luciana S. Esteves. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The paper is motivated by a present lack of clear model performance guidelines for shelf sea and estuarine modellers seeking to demonstrate to clients and end users that a model is fit for purpose. It addresses the common problems associated with data availability, errors, and uncertainty and examines the model build process, including calibration and validation. It also looks at common assumptions, data input requirements, and statistical analyses that can be applied to assess the performance of models of estuaries and shelf seas. Specifically, it takes account of inherent modelling uncertainties and defines metrics of performance based on practical experience. It is intended as a reference point both for numerical modellers and for specialists tasked with interpreting the accuracy and validity of results from hydrodynamic, wave, and sediment models.

1. Introduction

Although a need to standardise model build, calibration, and validation processes around one agreed approach is widely acknowledged, only limited guidance is available (e.g., [1, 2]) and often ambiguous and sometimes conflicting advice is offered in the grey literature (e.g., [3, 4]). A wide variety of different modelling practices are employed by consultants and academics, and frequently insufficient attention is given to the potential errors associated with the measured (and modelled) data used for model calibration and validation. This can result in poor model performance and unreliable model predictions. Without an agreed methodology and a performance standard for model calibration and validation, there is a risk that the quality of different approaches will vary, efforts will be wasted following inefficient or inappropriate calibration methods, and inconsistencies in methodologies will make model intercomparisons problematic.

This paper provides an evidence-based review and presents examples of calibration data sources and of model calibration and validation practices for estuarine and shelf sea models. It is intended to provide guidance to the assessment

and use of model calibration data and to offer procedural clarity and simplification to the model calibration and validation process. In doing so, it acknowledges that some degree of compromise between the complexity of the natural system and the model representation must be reached. For this reason, the paper does not address complex modelling issues around wave-driven currents, littoral drift, and shoreline evolution where specialist models (e.g., the nonhydrostatic version of XBeach and CFD) must be employed.

Since the accuracy of the model calibration depends critically on the calibration data used, attention is given to some of the most common issues associated with data quality. The paper also provides (a) the end users of model data more specialist guidance on modelling approaches, (b) the calibration procedures most frequently applied, and (c) the uncertainty in the model predictions. The paper draws on practical experience of modelling and expands on the earlier and limited guidance on the model calibration and validation that focus on Eulerian point-based criteria defining model performance (e.g., [2, 5, 6]). It also takes account of results and recommendations from modelling case studies where calibration issues have been the focus of the work (e.g., [7–10]).

Specifically, the paper describes (1) general factors that must be considered at the outset of all numerical modelling activities, (2) the quantitative assessment of model performance, (3) data sources and modelling guidelines for hydrodynamic, wave, and noncohesive and cohesive sediment models, and (4) morphological models. Special attention is given to one of the greatest challenges to the modelling community concerned with measuring and modelling sediment transport and associated erosion and accretion. While the focus of the work is based on practical applications of modelling shelf sea and estuarine processes, many of the issues discussed are relevant to a wide range of geophysical models.

2. What Is Model Calibration and Validation?

It is important from the outset to define the terms commonly used by numerical modellers: (a) *calibration* is a process which requires the adjustment of certain model parameters to achieve the best performance of the model for specific locations and applications; (b) *verification* ascertains if the model implements correctly the assumptions made; and (c) *validation* seeks to establish the agreement between the predictions and the observations (e.g., [11]). Validation is achieved by running the model using data covering an alternative period and/or a different location without making any additional adjustment to the model parameters (e.g., [12]). Of course, the accuracy of the model outputs cannot be proved to be greater than the accuracy of the original calibration data used, and validation does not imply verification, nor does verification imply validation. However, in practice, when measured data are available for the system being modelled, validation is often blended with verification [11]. If a comparison of measurements and model results suggests that the predictions from the model are close to the measurements, then the implemented model is assumed to be both a verified implementation of the assumptions and a valid representation of the system being modelled.

Irrespective of the model accuracy, the model calibration must express (a) express the level of agreement achieved; (b) express how realistic is the representation of the processes, and (c) define the criteria by which it has been judged as being fit for purpose. The quantitative assessment of data error, accuracy, and uncertainty in models then defines metrics against which model performance can be judged.

As an illustration of the typical calibration and validation processes applied in most coastal and estuarine models, a schematic diagram of steps followed for a hydrodynamic model is shown in Figure 1. In the initial model run, model parameters are set to the recommended values provided by the modelling software guidance (i.e., “factory settings”). Critical parameters affecting model performance (e.g., bed roughness) are then adjusted to achieve the best possible agreement between model predictions and measurements. Care must be taken to ensure the values set for these adjusted parameters are physically meaningful and appropriate [1]. Achieving a good model calibration for the wrong reasons is as bad as a poorly calibrated model.

A useful first step in the calibration and validation process is the determination of the most sensitive parameters in the

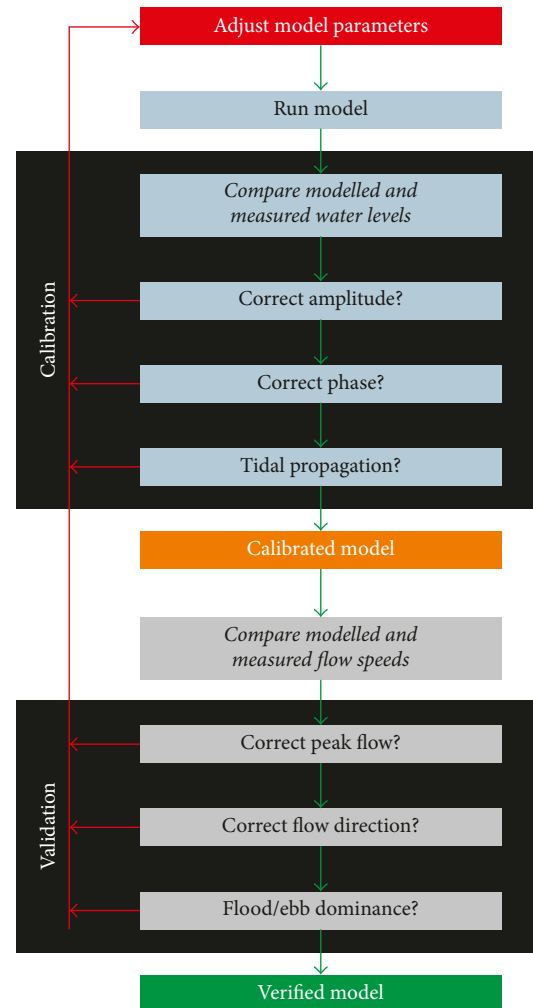


FIGURE 1: Schematic diagram of typical model calibration and validation steps required for a hydrodynamic model.

model. While expert judgment can be helpful, less-experienced model users should undertake sensitivity analyses. Here, the aim is to determine the rate of change in model output with respect to changes in model inputs (parameters). To undertake sensitivity analyses, it is necessary to identify key model parameters and to define the parameter precision required for the calibration (e.g., [13]). Sensitivity analysis approaches can be (a) local, where parameter values are changed one at a time, or (b) global, where all parameters are adjusted simultaneously. Both approaches have drawbacks. For example, the sensitivity of one parameter often depends on the value of other related parameters so that the correct values of other fixed parameters cannot be determined. In global sensitivity analyses, many simulations are required. Despite these drawbacks, both approaches provide insight into the sensitivity of the model parameters and are necessary steps in the model calibration process. However, “manual” calibration of models, where parameters are adjusted in a stepwise fashion, can be very time-consuming and inefficient.

The second step in the calibration process is undertaken to reduce the uncertainty in the model predictions. Normally, this uses carefully selected values for model input parameters

and compares model predictions with observed data for the same conditions. In common with the process described in step 1, this is often done iteratively without any fixed rule and is guided by the experience of the user and knowledge of the processes being modelled. The third step in the calibration process involves validation of the model output of interest (e.g., water level, flow speed, and direction). Validation involves running a model using parameters that were determined during the calibration process and comparing the predictions to observed data not used in the calibration.

The use of automated techniques for model calibration is now widespread (e.g., [14]). Typically, autocalibration procedures rely on Monte Carlo or other sampling schemes to estimate the best choice of values for multiple input parameters. For example, the autocalibration procedure described by van Liew et al. [15] is based on the shuffled complex evolution algorithm of Duan et al. [16], which allows for the calibration of model parameters based on a single objective function. While autocalibration can provide a powerful, labour-saving tool that can be used to substantially reduce the frustration and subjectivity that frequently characterises manual calibrations, care must be exercised when using these approaches to ensure the theoretical boundaries for each specific input parameter are not violated.

Frequently, the evaluation of postcalibration numerical model results is subjective and based on specialist interpretation of graphical output only. Examples of this approach include water-level curves or discharge time series, current vector distributions, and spreading patterns of heat and spills. Indeed, for many practitioners, a good visual fit between model predictions and observations is often sufficient to demonstrate good model performance without the need to quantify this further. Objective measures of model performance are also not new and are used, for example, in the Deltares (semi-) automated model calibration tools and in adjoint modelling (e.g., [17]). However, the increasing complexity of model functionality, and the use of model output by technical end users requiring information on model accuracy to reduce risks, has led to an increasing need for better guidance on how to quantify and evaluate the performance of models. A description of the calibration process applied to a biological-physical model [18] provides a useful example of typical procedures followed in the calibration process.

3. General Considerations

Irrespective of the model being used, there are several generic elements that require consideration prior to and during the model build phase. These elements will each impact on model performance and include bathymetry, bed roughness, model grid setup, the incorporation of specific structures and features, data accuracy and uncertainty, and model boundary conditions.

3.1. Bathymetry. One of the most common problems associated with the calibration of a hydrodynamic model concerns

errors in the underlying bathymetric data. The use of accurate bathymetry is pivotal in all shelf sea and estuarine modelling studies, and effort is required to ensure that the best possible bathymetric information is used. As standard practice, the analysis of bathymetric data should ensure (e.g., through a data review of the study area) that the most recent bathymetry survey data are used. Key features and contours should be checked against historical maps and charts. LiDAR data across water surface must be discarded. Suitable grid dimensions should be determined that reflect the spatial distribution of the bathymetric data, and where data are already gridded, poor interpolations/reductions/extrapolations onto model grids must be identified by reference to the original data sources.

A summary of bathymetric and topographic data requirements for models is shown in Table 1 [3]. Here, a distinction is made between application types, with the most exact being associated with scheme designs (e.g., flood defences) with less accuracy required for appraisal and/or strategy studies. These distinctions are used in other tables and are useful as they define the accuracy of key data required to build a model for different applications. The correct use of the most appropriate data for a given application can save time and effort. While Table 1 reflects the bathymetric and topographic data requirements for modelling estuaries, including specifications for average distances between survey positions, the minimum acceptable channel cross-section spacing, survey age, and the age and resolution for LiDAR data, they provide equally useful guidelines for shelf sea models.

Careful checks on the horizontal and vertical survey datum should always be undertaken prior to any model runs, and models should always aim to use a common reference datum. Typically for vertical positions, national reference points (e.g., Ordnance Datum Newlyn (ODN) in the UK), chart data (related to the lowest astronomical tide or to mean lower low water), or mean sea level (MSL) is used widely. However, while national reference points are useful in local-scale models, MSL has wider utility in larger regional models at all geographical locations. In the UK waters, the Vertical Offshore Reference Frame (VORF (<http://www.ucl.ac.uk/vorf>)) provides spatial maps of values that can be used to convert between vertical data. Similarly, VERTCON [19] and more recently VDATUM (<http://vdatum.noaa.gov/>) in the USA allow vertically transformation of geospatial data among a variety of tidal, orthometric, and ellipsoidal vertical data. In addition, satellite altimetry data can also be used to inform the offset between one or more tidal layers and the relevant satellite or geoid-based datum.

To illustrate a simple datum error, Figure 2(a) shows nearshore bathymetry from a coastal location in southern Portugal with a clear vertical datum problem. This issue is resolved in Figure 2(b) using a simple datum correction. This is a simple case for illustrative purposes only, and often datum errors are more complex and harder to correct.

Other errors can arise in hydrodynamic models due to (a) changes in charting properties (e.g., older Admiralty charts from the UK projected to OSGB which has now changed to WGS84), (b) data types, which have inherent

TABLE 1: Bathymetry and topographic data requirements for models (A = design; B = appraisal; C = strategy; and U = unsatisfactory); from Defra/EA [3].

Type	Average distance between surveyed depths (m)	Minimum channel cross section spacing (km)	Time since survey (years)	LiDAR age (years)	LiDAR resolution (m)
A	≤ 200	0.1	≤ 5	≤ 2	≤ 1.0
B	200–500	0.5	5 to 10	2 to 5	1 to 2
C	500–1000	1.0	10 to 20	5 to 10	2 to 5
U	> 1000	> 1.0	> 20	> 10	> 5

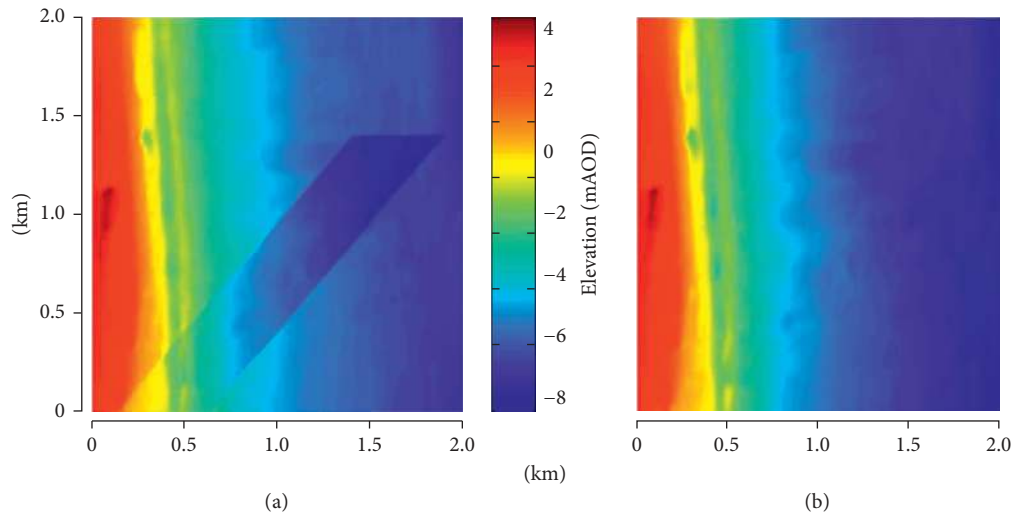


FIGURE 2: (a) Bathymetry with a clear vertical datum problem and (b) the same bathymetry after datum correction.

weaknesses (e.g., poorly interpolated bathymetry which may lead to an underestimation of depth and thus tidal volume), and (c) postprocessing using GIS or other “smoothing” software. To minimise the errors introduced in the model bathymetry, it is recommended initially to visually inspect the raw bathymetry using suitable software (e.g., Matlab or Fledermaus). Abrupt changes in elevations and spikes in data should be treated with caution. It is also helpful to examine gradients, and where possible, to compare interpolated data with known soundings.

Careful consideration should be given to data interpolation since interpolation routines can vary significantly between programs, and the options available (e.g., linear, nearest neighbour, inverse distance weighted, and spline methods) can also result in significantly different answers. Furthermore, some interpolation methods are better suited to sparse data sets (e.g., inverse distance weighted) and others to well-populated data distributions (e.g., nearest neighbour). The selection of the interpolation methods should always recognise this. It is important also to consider the scale of features on the bed that requires resolving in a model. For example, large bedforms, such as sand banks, redirect flows and must be resolved in the model. Smaller bedforms such as sand waves provide a resistance to the flow that can be parameterised through the bed friction term. Eliminating the need to resolve these features individually can reduce the model run time. In other applications, sand waves may need to be resolved on

an individual basis to assess, for example, migration rates and pipeline or cable routes.

Time and care spent ensuring that the underlying bathymetry has been correctly interpolated (datum, projection) and that it is free of spurious values and correctly represents the features of interest will contribute to improving model performance. Without good underlying bathymetric data, the task of trying to calibrate a hydrodynamic model will be extremely difficult, especially in shallow coastal and estuarine areas. For example, Cea and French [20] have investigated how errors in bathymetry can impact on the performance of estuarine shallow water models. They demonstrate that correcting errors in the measured depth can be significantly more efficient than a “classic” calibration approach based only on adjustment of the hydrodynamic roughness of the bed. Their proposed bathymetry calibration framework may offer improved performance from the current generation of numerical models. Further guidance on the use of bathymetry in models is given by Plant et al. [21, 22] and Mourre et al. [23].

3.2. Bed Roughness. The hydrodynamic roughness of the bed (hereafter termed “bed roughness”) is a primary calibration variable for all coastal and estuarine models. It is also essential for modelling other processes accurately such as sediment transport and wave attenuation. Irrespective of the method chosen for defining bed roughness, values are

TABLE 2: Empirically derived values of the drag coefficient (at 1 m above the bed) for different bottom types; from Soulsby [50].

Bed type	C_{100}
Mud	0.0022
Mud/sand	0.0030
Silt/sand	0.0016
Sand (unrippled)	0.0026
Sand (rippled)	0.0061
Sand/shell	0.0024
Sand/gravel	0.0024
Mud/sand/gravel	0.0024
Gravel	0.0047

typically manipulated iteratively by the user within the ranges reported in the literature. Any bed roughness can be generally assigned to a so-called equivalent sand roughness, k_s [24]. The equivalent sand roughness depends on the arrangement (pattern), distance (density), and shape of the roughness elements such as sand grains and ripples. However, in most models, bed roughness is typically parameterised by (a) a drag coefficient defined at a specified height above the bed, (b) the Manning number, n [25], or (c) the Chézy number, C [26].

To illustrate the range of drag coefficient values appropriate for different estuarine and coastal environments, Table 2 shows empirically derived values of the drag coefficient C_{100} measured at 1 m above the bed for different bottom types. In the absence of data to define the bed roughness accurately, these “typical” values are often employed in model applications. However, in many cases, this is an oversimplification, and care must be taken to obtain as much information as possible about bed characteristics so that appropriate bed roughness values can be assigned.

In “industry-standard” models such as MIKE21 and Delft3D, “roughness maps” can be used to define the spatial distribution of bed roughness values across the model domain. A good account of this approach is given by Lefebvre and Lyons [27]. Figure 3 presents a typical example of a bed roughness map showing the spatial distribution of (a) the measured median grain size, D_{50} , obtained from seabed samples and (b) the derived drag coefficient, C_d , which accounts for D_{50} and bedforms detected in a multibeam survey.

Bed roughness has been mapped by an ADCP (e.g., [28]), and high-resolution bathymetry and granulometry samples have been used by Huybrechts et al. [29] to derive bed roughness maps used in a TELEMAC model (cf. [30, 31]). Recently, the use of high-resolution multibeam sonar has provided bathymetric data at a resolution of less than 1 m and revealed the details of sea bed features (e.g., DORIS (<http://www.dorsetwildlifetrust.org.uk/doris.html>)) as well as provided information on sediment properties. The use of these data in modelling studies is currently experimental and requires high computing power to resolve the details. However, it offers the possibility of better defining bed roughness and thus may contribute significantly to reducing the effort needed for model calibration.

3.3. Model Grid Setup. The selection and setup of the model grid is a very important initial stage in the model build process. While some models still employ regular grid structures, most models now employ some form of flexible mesh usually comprising triangular elements. This approach allows high resolution of areas of interest and lower resolution over areas where bathymetry and/or processes are largely spatially invariant. In virtually all coastal and estuarine applications, the use of flexible mesh models provides the best model grid solution (e.g., [32]).

Taking a generic estuarine model as an example, some key points about model grids emerge: (a) the model grid should be designed to ensure the grid resolution can define the main morphological features (including structures) that could have influence on hydrodynamics; (b) narrow channels and banks should have at least three grid cells (preferably 5) to determine the base or crest widths; (c) as far as possible areas of increased grid resolution should follow the course of the main channels, particularly in a curvilinear grid; and (d) when considering the location of upstream and downstream boundaries in the model, boundaries should not be fixed too close to areas of interest. However, this may be constrained by the actual aims of the modelling as well as available boundary data. Further examples of the issues arising when defining a model grid are given by Hsu et al. [33], Kernkamp et al. [34], Liu and Ren [35], and Maynard and Johnson [36].

Table 3 shows an example of the typical model grid resolution required for estuarine models intended for studies of water levels and flow velocities [3]. Table 3 indicates also the minimum number of model grid points required to correctly represent features such as channels and sand banks in the model. Similar grid requirements apply equally to coastal models. It is noted that a fine grid resolution (<2 m) is required to correctly represent the deliberate breaching of flood defences when designing managed realignment schemes.

During the process of grid generation, irrespective of the modelling software used, interpolation of the bathymetry will take place. Again using the example of a generic estuarine model to illustrate some key points, the following checks should be made on completion of bathymetric interpolation processes: (a) the gridded bathymetry must show the same characteristics as the original bathymetry; (b) the gridding process must not displace and/or narrow/constrict channels; (c) different interpolation methods should be assessed; (d) channels must not be widened or narrowed, particularly when these make up a considerable proportion of the estuary cross section; and (e) depths adjacent to the boundaries should be inspected to ensure correct interpolation has occurred.

A further key point to note concerns the spatial resolution in the computational grid of a given model. Typically, a model prediction is only applicable at the spatial resolution defined by the computational grid. In contrast, the measurements, typically used to calibrate and verify model predictions, are obtained at a single location and represent the local environment only. Thus, when comparing a model prediction with a measured value at a point, consideration must be given to the tolerance of this spatial resolution. Most

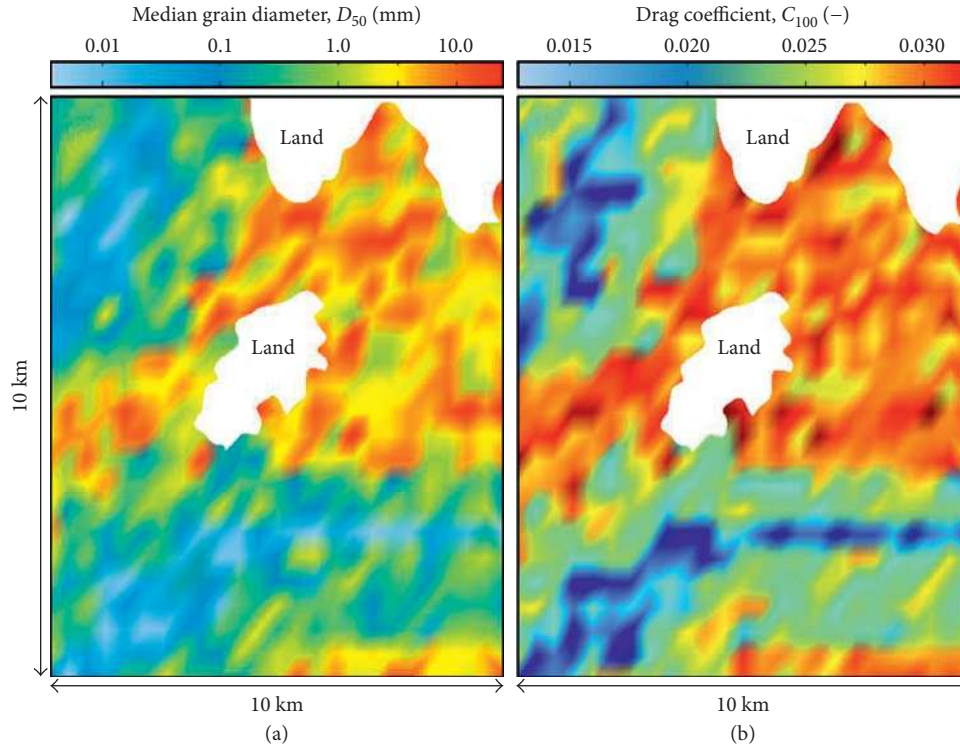


FIGURE 3: (a) Measured median grain size, D_{50} , obtained from seabed samples and (b) the derived drag coefficient, C_d , which accounts for D_{50} and bedforms detected in a multibeam survey.

TABLE 3: Recommended model grid resolution for flood studies in estuaries (A = design; B = appraisal; C = strategy; and U = unsatisfactory); from Defra/EA [3].

Type	Resolution for water level (m)	Resolution for flow velocity (m)	Grid point needed to resolve key features [†]
A	≤ 5	≤ 2	> 8
B	5 to 10	2 to 4	5 to 7
C	10 to 25	4 to 6	3 to 4
U	> 25	> 6	< 3

[†]Channels, sand banks, etc.

often, the grid cell most closely collocated with the declared deployment location is used in model/measurement inter-comparisons. If the grid size is large, the single point measurement may not be representative of the wider area in the model grid element. Conversely, if the instrument is deployed on a mooring line that has the capacity to move (such as a wave buoy), then, through time, the instrument may pass through multiple grid points if the model has a fine-scale resolution.

In some circumstances, data extraction from a grid cell adjacent to the measurement location may better represent the actual conditions at the measurement point if that point lies close to the boundary of a model element. Depending on the model grid resolution, it is advisable therefore to extract model outputs from all grid cells adjacent to the measurement location and to make comparisons with the observations.

Taking the modelling of fluvial bedforms as an example, El Kheishy et al. [37] discuss the selection of an appropriate model grid. In common with many modelling applications,

a compromise must be reached between the resolution needed to define the bathymetry accurately and the consequent execution time required for a certain grid resolution. Their study showed that the apparent bed resistance (shear stress) and bedform steepness decreased with increasing grid spacing. Increasing the grid spacing also created artificial bedform fields giving rise to grid-dependent resistance. The model grid therefore has a significant influence on the model predictions. It is therefore important to be aware of these issues when interpreting model results and to check them whenever possible against all available data sources.

3.4. Model Boundary Conditions. Experienced modelling practitioners will ensure that the intended boundary type is being used at each open boundary and that the cell notation and order of data are correct. Indeed, most industry-standard models (e.g., Delft3D) give a visual representation of the boundaries for checking purposes. It is recommended practice to align boundaries with the dominant flow direction,

tidal characteristics (avoiding amphidromes), waves, or geographic features.

The input data to estuarine and shelf sea model boundaries typically fall into 2 types: (a) a water level and (b) a flux (discharge). A water-level boundary is normally obtained from existing models or measurements. In the case of model-derived boundary data, knowledge of how the boundary data are produced by the larger model is required to define the accuracy and reliability (e.g., the number of constituents used and the spatial and temporal resolution). However, a water-level boundary may not be applicable in areas with little or no tidal height variation.

Noting that a flux (discharge) carries momentum and water-level variations across the model domain have to generate momentum, it can be argued that a flux (discharge) boundary condition is a more robust option for model calibration purposes. However, it is usually much more difficult to describe and apply. Reliance on water levels only can lead to serious model underperformance, and wherever possible, attempts must be made to use water level and flux at the model boundaries.

Since the appropriate model calibration accuracy can be obtained, the following boundary condition issues also need to be fully understood: (a) spikes in modelled boundary data attributable to instabilities in the original boundary data and (b) the selection of boundary data from larger model domain in unsuitable locations (e.g., close to land domain or elements that dry).

A flux (discharge) can be applied at any model boundary (e.g., the point of freshwater input into an estuarine model). Generally, these data are provided by measurements or derived from a coarser-scale model. The quality of these data depends on the accuracy of the measuring device or model. It is recommended that in areas of small tidal variation, or where multiple boundaries are included, at least one boundary is of a flux (discharge) type. It may be noted that using flexible grids (or nested models), the model domain can be extended to provide more robust boundary conditions owing to the large phase differences and gradient effects across the model domain. However, in practice, many model setups use only water levels as the primary driving force, and in many applications, this proves to be successful.

It is recommended that if measured tidal levels are used as model boundary conditions, then these are checked to ensure consistent phase and amplitude with values obtained from the harmonic constituents. While there will be small differences in amplitude attributable to meteorological effects, the phase should be very similar.

4. Assessing Model Performance

In engineering and environmental modelling studies, the use of quantitative model evaluation methods is perceived as providing more objective, consistent, and reproducible model validation and assessment. However, it is also self-evident that the identification of systematic or random errors in model results can also be detected quickly by the human eye. In practice, the assessment of model output is most effective when both qualitative and quantitative approaches are

employed. For example, in most shelf sea or estuarine applications, a combined visual and quantitative evaluation may be achieved by presenting the spatial distribution of current vectors for visual examination together with statistics that quantify differences between measured and predicted current speeds from several locations. These statistics can provide useful additional information about spatial coherence, correlations, and consistency and will often indicate explanations and origins of the possible differences between the model results and the measurements.

It is also important that model results receive expert assessment, ideally against a conceptual understanding of processes in each model domain established using a range of data resources. This might include some obvious checks on current speed, phase, and direction as well as more detailed investigations of sedimentation patterns. It is recommended that the initial assessment of model performance by whatever means should be undertaken before running models for extended periods. However, the period chosen for this preliminary examination depends on the processes being modelled. For example, a model of tidal currents run over one or two tidal cycles should be sufficient to determine how well the model is performing and which adjustments might be necessary. On the other hand, a model of sediment transport may require considerable time before the effects of net sediment movement are evident through changes in the bathymetry.

4.1. Error, Accuracy, and Uncertainty of Model Calibration Data. As it defines the metric against which model performance will be judged, the assessment of error, accuracy, and uncertainty in the data used for model calibration is an important step in the modelling process. Indeed, the accuracy of a numerical model is governed in part by the degree of error present in the model calibration data. It is essential therefore to quantify error, accuracy, and uncertainty through understanding of the instrumentation, the instrument deployment method, and its location as well as any data postprocessing issue.

It is necessary to distinguish between systematic and random measurement errors. All measurements are prone to systematic errors resulting, for example, from imperfect instrument calibration (zero error) and changes in the environmental conditions. Similarly, random errors are usually present in a measurement or other observations and result from inherently unpredictable fluctuations in the readings of a measurement apparatus or in the experimenter's interpretation of an instrumental reading or the environment. Different results for ostensibly the same repeated measurement are a clear indication of a random error. The error can be quantified by comparing multiple measurements and reduced by averaging multiple measurements. Systematic errors cannot be detected this way because they always "push" the results in the same direction. However, when identified, they are easier to eliminate from a data set using trend removal techniques (e.g., regression analysis).

Instruments collecting data from different spatial locations may also apply range-dependant spatial averaging to

the recorded data, leading to variable spatial resolution. For example, the horizontal averaging across spreading ADCP beams results in a measurement footprint that increases in size with distance from the instrument. Taking as an example the calibration of a 2D depth-averaged hydrodynamic model using ADCP data, it is first necessary to derive the depth-mean current from the ADCP measurement. This requires making assumptions about the vertical structure of the marine boundary layer (often occupying the region from the bed to the air-water interface) before time and spatially averaging the ADCP data to obtain a depth-mean current speed. These data processing steps introduce errors which are difficult to quantify. These problems are further compounded when attempting to extract a meaningful depth-mean representation of the current direction, especially in areas subject to significant current veering (e.g., adjacent to sand banks). Furthermore, if a given measurement footprint is within a highly turbulent flow field, then the accuracy of the mean flow measurement will be governed by the sampling time and can lead to significant errors if the flow is not sampled correctly at that location. With this example in mind when comparing predictions from a grid point in a model with measurements from single or multiple locations, attention must be given to spatial and temporal inconsistencies that might lead to calibration error and/or bias.

4.2. Sensitivity Analyses. Sensitivity analyses are used to study how the uncertainty in the output from a model can be apportioned to different sources of uncertainty in its inputs. Sensitivity analyses are undertaken by varying input parameters (within a range, i.e., physically realistic) and examining the model response. Sensitivity analyses can be useful for a range of purposes including (a) testing the robustness of model resulting in the presence of uncertainty, (b) increasing the understanding of the relationships between input and output variables in a model, (c) identifying errors in the model by encountering unexpected relationships between inputs and outputs, and (d) simplifying models by identifying model inputs that have no effect on the output, or identifying and removing redundant parts of the model structure. Sensitivity analyses can also help to reduce uncertainty by identifying the model inputs that cause the greatest uncertainty in the output, thereby allowing adjustments to increase the robustness of the model. Importantly, by making model results more understandable, compelling, or persuasive, sensitivity analyses can enhance interactions between modellers and the end users of modelling output. Sensitivity analyses are therefore a vital part of evaluating if a model is fit for purpose, and time must be set aside in any modelling study to undertake a credible model sensitivity study.

One area of sensitivity analysis that requires special consideration concerns the sensitivity of a given model to errors in the input data (e.g., bathymetry, water level, and depth-mean current speed). This is especially important when there are errors and/or uncertainty in more than one input data set which can result in compounded errors in the model output. For example, in an estuarine sediment

transport model, errors in the water depth and/or current speed at a given location will result in an over- or underestimation of the bed shear stress. Since sediment transport is related to a power of the bed shear stress (typically quadratic for bed load and cubic for suspended load), small errors in predicted bed shear stress can result in large errors in predicted gross and net sediment transport.

4.3. Time Series and Statistical Output. In many cases, the presentation of data in time-series format helps to reveal the goodness of fit between model and observation data, with gaps between observed and predicted data indicating visually discrepancies between the model predictions and the calibration data. Calibration should aim to minimise these discrepancies, and statistical analysis should be used to quantify the goodness of fit. Additionally, it is also informative to compare like-with-like values using a scatter plot showing observed versus modelled values. Some examples are provided below.

To quantify the temporal aspect of the model calibration further, statistical approaches are used to demonstrate that confidence can be placed in the model performance over temporal time scales in a clear and understandable way. The Danish Hydraulics Institute, DHI, Quality Indices Matrix calculating several goodness-of-fit statistics for comparison between observations and simulated results is an appropriate methodology to adopt. When necessary, and when data quality permits, additional types of analysis may be appropriate, such as Brier skill score analysis [38] or indices of agreement (e.g., [39]).

Simple statistics that demonstrate the level of agreement between measured/observed data and model prediction at a chosen location in the model domain include the mean and peak differences (often expressed as a percentage) and the standard deviation. In addition, there are several quality indices that can be used to demonstrate the statistical agreement between model predictions and observations (Table 4). In the table, O_i and S_i are the measured and predicted values of a given parameter at time t_i , respectively, and N_i is the total number of data points. The statistics are now defined.

Accuracy expresses the difference between the measured and modelled data which is defined as $dif_i = S_i - O_i$. In all cases, the aim should be to reduce the value of dif_i to the smallest value practicable. Ideally, a minimum dif_i should not exceed 10%, although this will be highly variable depending on the parameter being considered and the accuracy of the calibration data used in the model. The accuracy of the modelled data can also be quantified using the root mean square error (RMSE) statistic (Table 4). The RMSE value is often expressed as a percentage, where lower values indicate less residual variance and thus better model performance.

The bias expresses the difference between an estimator's expectation and the true value of the parameter being estimated and can be defined as being equal to the mean error in the data. Systematic bias reflects external influences that may affect the accuracy of statistical measurements. Detection bias

TABLE 4: Example statistics to demonstrate the level of agreement between measured/observed data and model prediction.

Quality index	Example	Formulae
Accuracy	Root mean square error	$RMSE = \sqrt{\frac{1}{N_i} \sum_{i=1}^{N_i} (S_i - O_i)^2}$
Bias	Average bias	$Bias = \sum_{i=1}^{N_i} \frac{1}{N_i} (S_i - O_i)$
Correlation	Pearson product-moment coefficient: $\bar{O}_i = \frac{1}{N_i} \sum_{i=1}^{N_i} O_i$ and $\bar{S}_i = \frac{1}{N_i} \sum_{i=1}^{N_i} S_i$	$R = \frac{\sum_{i=1}^{N_i} (S_i - \bar{S}_i)(O_i - \bar{O}_i)}{\sqrt{\sum_{i=1}^{N_i} (S_i - \bar{S}_i)^2} \sqrt{\sum_{i=1}^{N_i} (O_i - \bar{O}_i)^2}}$
Skill	Brier skill score (BSS): X_p is the postevent condition predicted by the model, X_m is the measured postevent condition, and X_b is the preevent condition. Skill: index of agreement [44], where X and \bar{X} are time series and time average of model and observed values	$BSS = \left[\frac{\% \langle X_p - X_m ^2 \rangle}{\% \langle X_b - X_m ^2 \rangle} \right]$ $Skill = 1 - \frac{\sum X_{mod} - X_{obs} ^2}{\sum (X_{mod} - \bar{X}_{obs} + X_{obs} - \bar{X}_{obs})^2}$
Agreement	Scatter index	$SI = \frac{\sqrt{\frac{1}{N_i} \sum_{i=1}^{N_i} (S_i - O_i)^2}}{\frac{1}{N_i} \sum_{i=1}^{N_i} O_i} \times 100$

is where a phenomenon is more likely to be observed and/or reported for a set of study subjects. Reporting bias involves a skew in the availability of data, such that observations of a certain kind may be more likely to be reported and consequently used in research.

The agreement or otherwise between measured/observed data and model prediction time series is frequently quantified using the Pearson product-moment coefficient, R (Table 4). It is essential to test the statistical significance of the correlation coefficient. In most cases, the Pearson method (one- or two-tailed) is appropriate. In statistical significance testing, if the null hypothesis is true, the p value is the probability of obtaining a test statistic at least as extreme as the one that was observed. One often “rejects the null hypothesis” when the p value is less than 0.05 or 0.01, corresponding, respectively, to a 5% or 1% chance of rejecting the null hypothesis when it is true. When the null hypothesis is rejected, the result is said to be statistically significant. In estuarine and shelf sea modelling studies, statistical significance at around the 95% confidence level is judged to be acceptable for most practical applications.

A range of statistical indices of model performance has been developed (e.g., [40–43]). The widely used Brier skill score, BSS [38], and Willmott’s dimensionless index of agreement [44] compare the mean square difference between the prediction and observation with the mean square difference between baseline prediction and observation. For example, perfect agreement gives a BSS score of 1, and negative values indicate that predictions are worse than the baseline value. van Rijn et al. [45] provides an interpretation of BSS values where $0 < BSS < 0.3$, $0.3 < BSS < 0.6$, $0.6 < BSS < 0.8$, and $BBB > 8$ indicated poor, reasonable/fair, good, and excellent, respectively. However, it has been recognised that the larger errors, when squared, overweight the influence of those errors on the sum of squared errors. This issue has recently been addressed by Willmott et al. [46]

who present a nontrivial improvement to the earlier index of agreement recommended for a wide range of model performance applications. Examples of model skill assessments for estuarine models are given by Sheng and Kim [47] and Warner et al. [48].

The scatter index, SI, is the RMSE normalised with the mean value. In most cases, the scatter index provides a useful indication of the model performance. However, taking wave model results as an example, the scatter index may appear to understate the skill of the model, as it tends to be large in shelf sea applications. The reason is that the RMSE of the significant wave height is normalised with the average significant wave height, which is usually rather small in shelf sea regions. For example, an RMSE of 0.25 m in the significant wave height in complex field conditions seems reasonable, but if the mean value is only 0.5 m, the scatter index attains the rather high value of 50%. The diagnostic model performance index MPI indicates the degree to which the model reproduces the observed changes of the waves. Like the scatter index, it is defined in terms of RMSE values in the form $MPI = 1 - (RMSE/RMSC)$. Here, the definition of RMSC is identical to that of RMSE, except that all S_i values are replaced by the incident O_i values. For a perfect model ($RMSE = 0$), the value of the MPI would obviously be 1, whereas it would be 0 for a model that (erroneously) predicts no changes ($RMSE = RMSC$) (cf. [49]).

5. Hydrodynamic Models

5.1. Data Sources. Water-level gauges and pressure sensors typically provide information on the water level relative to a defined datum at a suitable temporal resolution (typically no more than 30-minute intervals). Ideally, water level and current information should be obtained from as many key locations within the model domain as possible, and specifically, in areas of interest and areas of significant variation. Typically, errors

associated with these kinds of data include (a) incorrect time references (e.g., GMT/BST), (b) errors in datum corrections (see below), (c) errors in correctly defining the measured data locations in the model domain, and (d) instrument calibration error. Problems with the measuring device often appear as offsets and/or spikes in the measured data. Spikes should be either removed or substituted with artificial data. Interpolation over large gaps in the data should not be attempted, and alternative data sources with better temporal cover should be sought.

To define the confidence limits of the measured data, a quality review is required. This may result in the rejection of some data, or the adoption of other data with stated caveats. The more data that are available (depending on quality, format, and spatial and temporal resolution), the more reliable the model calibration is likely to be. To minimise potential uncertainties in model performance and to optimise model calibration, common misunderstandings and typical errors and uncertainties in hydrodynamic model input data are described below along with some suggested approaches which can aid model setup and calibration.

5.2. Performance Guidelines

5.2.1. Water Level. A model calibration for water level should include examination of amplitude, phase, and asymmetry. Specifically, the test should look at (a) differences in maximum and minimum surface elevations; differences in tidal phase, at high and low water; and RMSE (noting that this is not corrected for bias, and unless the bias is insignificant, this parameter can be difficult to interpret), (b) bias, and (c) scatter index (SI). It is recommended that the minimum-level model performance required for shelf sea areas is (a) water levels to within ± 0.10 m (or to within 10% and 15% of spring and neap tidal ranges, resp.) and (b) timing of high water to within ± 15 minutes. For estuaries, it is recommended that the minimum-level model performance required is (a) water levels to within ± 0.10 m at the mouth, ± 0.30 m at the head (or to within 10% and 15% of spring and neap tidal ranges, resp.) and (b) timing of high water at the mouth to within ± 15 min, ± 25 min at the head.

5.2.2. Current Speed. In 2D depth-average hydrodynamic models, current speed predictions should be examined with respect to amplitude, phase, direction, and asymmetry. Specifically, the test should look at (a) differences in peak flow speeds (ebb and flood tides), (b) mean flow direction, (c) RMSE, (d) bias, and (e) SI. However, appropriate depth-average current speed values must normally be derived from either point measurements at some reference height in the water column or measured vertical current profile data (e.g., ADCP data). In both cases, depth-average current speed can be calculated using the 1/7 power law (e.g., [50], p. 49) or similar. Normally for 3D hydrodynamic models, ADCP data can be used directly for calibration at one or more levels in the model. However, if the model layers are large in vertical extent, they may span one or more ADCP measurement bins, and the 1/7 power law or similar must be applied to interpolate an appropriate current speed value for the model layer.

It is recommended here that predicted current speeds from 2D and 3D hydrodynamic models in shelf sea areas and estuaries be less than ± 0.20 m/s (or $\pm 10\%$ to 20%) of the measured speed. To express the accuracy of tidal current speed predictions by models, Cefas (www.cefas.defra.gov.uk/media/.../report-on-first-asmo-workshop.pdf, accessed March 2014) expresses performance in terms of error in the maximum predicted velocity so that errors of < 0.05 m/s, < 0.1 m/s, < 0.2 m/s, and > 0.2 m/s express very good, good, moderate, and poor performance, respectively.

Results from statistical analyses of model performance need to be interpreted with care. The RMSE value provides a quantitative measure of how good the model fits the data based on the mean of the data. However, if there is significant bias in the data, then the goodness of this fit is not an appropriate statistic to use. It is recommended here that bias < 0.2 , SI < 0.5 , and RMSE < 0.2 demonstrate a statistically significant fit.

5.2.3. Current Direction. Since current direction is derived from vector quantities, it cannot be treated in the same way as other parameters (e.g., speed). However, the accuracy of predicted current direction can be examined using time-series plots and quantified, for example, using bias and SI statistics. To remove ambiguity from current direction data, the following steps are recommended: (a) detect whether the absolute difference between the directions is greater than 180° ; (b) if it is greater than 180° , then add 360° to the lesser direction before subtracting the greater direction; or (c) if it is less than 180° , then calculate the absolute difference between the directions. This method returns an absolute (positive) value describing the difference in directions which will be always less than 180° . For practical applications, it is suggested that preserving the sign (negative or positive) of the direction difference is not necessary, and it prevents a meaningful mean bias to be calculated from those differences. Once the absolute difference between the directions has been calculated, it is possible to calculate the bias. For shelf sea areas and estuaries, the minimum-level model performance is recommended here to be $\pm 10^\circ$ and $\pm 15^\circ$, respectively.

5.2.4. Bed Shear Stress. Except for some specialist research instruments, for example, a field-deployable shear plate prototype reported by Oebius [51] and laboratory-based shear plates reported by Grass et al. [52] and Rankin and Hires [53], no reliable direct way of measuring the bed shear stress is yet available. For most practical applications, the use of measurements to calibrate/validate bed shear stress values predicted by a model is therefore not possible.

When considering bed shear stress in the context of hydrodynamic and/or sediment transport, it is critically important to distinguish between the skin friction component of total bed shear stress responsible for sediment mobilisation and transport and the form drag imposed on the flow by pressure losses in the wake of bed obstacles such as bedforms. Most models predict the total bed shear stress using the quadratic stress law. This relates a depth-average flow speed to stress via a drag coefficient that characterises the hydrodynamic “roughness” of the bed. For skin friction bed shear stress, the roughness parameter expresses the drag

attributable only to sediment grains. Form drag (in part responsible for maintaining suspended sediment status) is then obtained through a partitioning approach. It is very important to understand how a model deals with drag partitioning and to use any resulting estimate of bed shear stress correctly. Soulsby [50] provides a clear account of bed shear stress components and their calculation and application.

When dealing with subgrid-scale bedforms, it is normal to parameterise bed roughness using a friction coefficient or an equivalent grain roughness. This can vary spatially (and in some cases temporally) and provides a means of moderating or enhancing the local bed shear stress and thereby “tune” the model against observational flow data. This needs to be undertaken with care to avoid implementation of unrealistic friction coefficient values, and guidance on appropriate friction coefficients should be sought (e.g., software Guides for the model being used; Soulsby [50]). To avoid significant over- or underestimation of sediment transport, it is recommended that bed shear stress requires estimation to within $\pm 0.05 \text{ N/m}^2$ for shelf sea and estuarine models. However, small errors in bed shear stress can be compounded over time, especially in morphological models. It is also noted that bed shear stress data can be usefully postprocessed to obtain estimates of bedforms and bed load and suspended transport using a range of empirical formulae [50]. However, these estimates are constrained by the data used to generate them and the accuracy of the algorithms used to estimate hydrodynamic stresses.

A simple illustration of a depth-average hydrodynamic model calibration using bed roughness is shown in Figure 4. Figure 4(a) shows time series of measured and predicted current speed at locations P_1 and P_2 in the mouth of a small tidal inlet. In this initial model run, a drag coefficient, C_d , value of 0.035 is assumed, leading to an underestimation of the current speed by the model. Reference to available bed sediment data suggests that a C_d value of 0.02 is more appropriate resulting in much better agreement between the measured and predicted current speed values (Figure 4(b)). However, it is also noted recently that erosion had occurred in the inlet since the last bathymetric survey. Iterative adjustments to the water depth in a subsequent series of model runs finally resulted in very good agreement between the measured and predicted current speed (Figure 4(c)). The lowering of the bed of the inlet channel by 0.45 m was subsequently confirmed by a repeat bathymetric survey undertaken after the modelling was completed.

As a further example of calibration targets to achieve, required hydrodynamic model performance statistics for estuarine flooding models from Defra/EA [3] are shown in Table 5. The statistics include RMSE for storm surge elevation (h_{surge}), RMSE for high water levels (h_{max}), the tolerances for predicted peak water levels, RMSE for flow velocity (U), the tolerances for predicted fluvial inputs (Q), the flood area required to be predicted correctly for two or more historical floods (A), and the predicted flood depth error (d_{err}). While these statistical tests are specific to the flooding application and are exacting since flood predictions must be accurate, they are typical of the model performance criteria that should be used for all shelf sea and estuarine models.

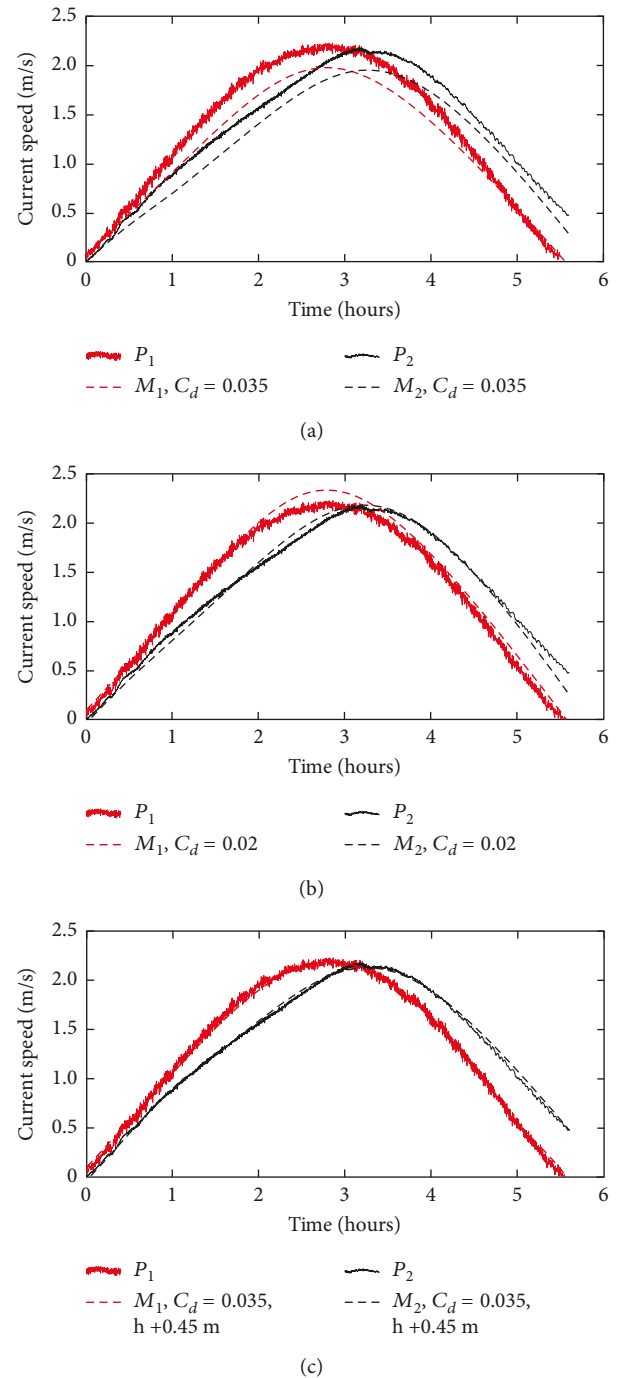


FIGURE 4: Simple examples of a hydrodynamic model calibration: (a) measured (P_1 and P_2) and predicted (M_1 and M_2) water levels at two locations with C_d set to 0.0035 in the model; (b) the same plot as (a) with $C_d = 0.002$; and (c) the same plot as (a) with the bathymetry corrected by +0.45 m.

6. Wave Models

6.1. Data Sources. Typically, there are four sources of data available for use in wave model calibration: visual observations; buoys and platforms; satellites; and numerical models driven by surface wind fields. Each has a different level of confidence and uncertainty, and good accounts of wave measurement and data are given by Steele et al. [54],

TABLE 5: Model performance statistics for flood studies in estuaries including storm surge elevation (h_{surge}), high water levels (h_{max}), predicted peak water levels, flow velocity (U), predicted fluvial inputs (Q), flood area predicted correctly for two or more historical floods (A), and predicted flood depth error (d_{err}) (A = design; B = appraisal; C = strategy; and U = unsatisfactory); from Defra/EA [3].

Type	RMSE for h_{surge} (m)	RMSE for h_{max} (m)	Predicted peak water level (mm)	RMSE for U (m/s)	Q (% of measured flows)	A (%)	d_{err} (m)
A	<0.20	<0.20	± 150	≤ 0.1	± 5	>95	≤ 0.10
B	0.20 to 0.35	0.20 to 0.25	± 250	≤ 0.2	± 10	>90	≤ 0.20
C	0.35 to 0.50	<0.30	± 350	≤ 0.3	± 15	>80	≤ 0.30
U	>0.50	>0.30	>350	>0.3	>15	<80	>0.30

Komen et al. [55], Krogstad et al. [56, 57], and Lindroth and Leijon [58].

6.1.1. Visual Observations. Visual observations of wave parameters (height, period, and direction) taken from ships of opportunity are sometimes available for long periods (decades). This data source has clear limitations and many potential errors, particularly in stormy weather (cf. [59]). Other significant limitations include the number of observations and extent of data coverage which is typically limited to shipping lanes. However, in the past, there are many areas of the world where wave measurements by other means are absent and visual observations may be the only source of information.

6.1.2. Buoy and Platforms. Surface-following buoys are the most common instrument used to measure waves, with deployment depths between 10 m and a few hundred metres (cf. [56, 57, 60]). There is a large variation in the quality of the data available from these devices depending on their age and type. Typically, the latest devices can capture an estimate of the main wave parameters (significant wave height, H_s , mean and peak period, T_m and T_p , and the related directional information such as mean direction, mean directional spread, kurtosis, and skewness for the full 2D spectrum). Typically, the measurements are taken at 1-hour intervals. Wave-measuring buoys are accurate instruments, and the related error for H_s is usually only a few percent. Uncertainty occurs due to sampling variability and resolution of the frequency distribution (peak periods). In the high H_s range, the buoys tend to “slip” around the highest crests. In doing so, this introduces a negative bias in estimation of the higher wave height values.

6.1.3. Remote Sensing. Satellite altimeter and scatterometer data are now available from a range of platforms (e.g., ERS2, ASCAT METOP-A, Krogstad and Barstow [61]). The altimeter provides information on wind speed and wave height, and the scatterometer provides a wider band of information on wave and wind parameters. However, in areas of complex geometry (typically shelf sea and estuarine areas), satellite data usually provide a poor estimate of sea state due to the strong spatial gradients which cannot be well resolved by the satellite sensors. Other limitations include poor temporal coverage due to satellite overpass frequency which can prevent acquisition of high-frequency time series

for a chosen location. However, developments reported by Young et al. [62] demonstrate clearly that useful global wave data sets can now be assembled using data from a range of remote sensing platforms and that these data have high utility in regions of the world where wave data are scarce. Wave data can also be obtained using HF (e.g., [63]) and X-band (e.g., [64]) radar systems and through the use of video (cf. Argus Video [65]). Although these approaches require some calibration, each has a capability of measuring nearshore waves and can help calibrate and validate wave models in complex regions where reflection, refraction, and diffusion processes may be present.

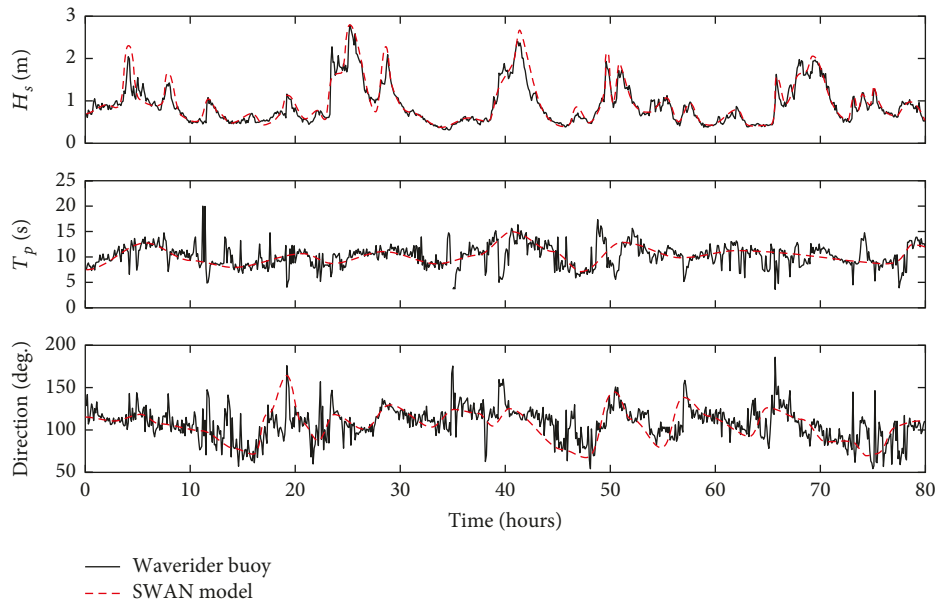
6.2. Performance Guidelines. Typically, for waves, the required model performance at the calibration and validation stage is judged to be acceptable if the wave model outputs are biased to within (a) $\pm 10\%$ of the mean observed height, (b) $\pm 20\%$ of the mean observed period, and (c) $\pm 15^\circ$ of the mean observed direction. Considering design, appraisal, and strategy applications, Table 6 provides practical wave model performance guidelines concerned with model resolution, minimum record (or hindcast) lengths required to define extreme wave statistics, and RMSE values for H_{m0} and average peak H_{m0} [3].

These wave model performance statistics are intended only a guide, and often more stringent agreement between observed and modelled data may be required. Equally, these criteria might be too exacting for all regions of the modelled area. Meeting these criteria for at least 90% of positions/time combinations is likely to be a less stringent and acceptable criterion in most circumstances. In cases where waves from more than one direction are present simultaneously (e.g., swell and wind sea), mean wave direction is meaningless and reference must be made to the directional wave spectra to characterise the observed and modelled wave field. Scatter plots and correlation statistics are also useful to demonstrate agreement between measured and modelled wave direction for multidirectional sea states. It is also helpful in some circumstances to examine directional wave spreading since many third-generation spectral wave models tend to underestimate this parameter (cf. [66, 67]).

Examples of useful plots that help assess wave model performance are shown in Figure 5 which shows measured H_s , T_p , and direction data from the SWAN model [68] and measured values from a Directional Waverider buoy. Good agreement between the model and the observations is demonstrated. Figure 6(a) shows a scatter diagram of

TABLE 6: Model performance statistics for wave models (A = design; B = appraisal; C = strategy; and U = unsatisfactory); from Defra/EA [3].

Type	Model resolution (km)	Record length (years)	RMSE, H_{m0} (m)	RMSE average peak, H_{m0} (m)
A	≤ 12	≥ 30	≤ 0.4	≤ 0.10
B	≤ 12	≥ 20	≤ 0.5	≤ 0.12
C	≤ 12	≥ 10	≤ 0.6	≤ 0.15
U	> 12	< 10	> 0.6	> 0.15

FIGURE 5: Comparisons between measured and predicted wave time series showing H_s , T_p , and direction.

measured and modelled H_s . Here, the line of unity indicates that the model slightly overestimates H_s . In Figure 6(b), a peak over threshold (POT) analysis for modelled and measured H_s data is shown. These diagrams both help identify agreement or otherwise between measured and modelled wave data, and it is recommended that these visual checks are used when evaluating wave predictions.

In the case of spectral wave models, it is also helpful to examine the frequency domain differences in the energy distribution between model and measured spectra. The same approach applies to directional wave spectra, noting however, that it is unusual to have measured wave spectra from more than one location in the model domain.

7. Sediment Models

Accurately simulating the behaviour of sediments in numerical models presents one of the greatest challenges. The principal aim of sediment models is to reproduce the observed spatial and temporal variations in observed erosion and accretion. Here, guidance is provided for the calibration of numerical models for sand (sediment coarser than sand (e.g., shingle) cannot be represented reliably through 2-dimensional modelling) (median grain diameter, $D_{50} > 63 \mu\text{m}$) and silt/mud ($D_{50} < 63 \mu\text{m}$). Attention is first given to the essential data required to successfully calibrate sediment models of estuaries and open shelf sea

environments. The methods used to measure bed load, suspended load, and net sediment transport are reviewed briefly, drawing attention to potential errors and uncertainties that must be considered in the model calibration process [69]. The issues associated with the calibration of cohesive, noncohesive, and mixed grain-size sediment models are then discussed.

It is emphasised from the outset that a primary requirement of all sediment modelling is accurate information about the physical properties of the sediment (grain-size distribution, bulk density, porosity, etc.), bedforms (active and moribund), and the spatial distribution and thickness of the sediments. Also, sight should not be lost of the fact that although the physical characteristics of sediments may be well expressed, biological mediation and the behaviour of some cohesive sediment remain difficult to parameterise in models [70], and although some of these problems can be overcome using in situ measurements (e.g., field flumes [71]), these are costly to deploy and are not normally undertaken in practice.

7.1. Data Sources. Obtaining suitable data of sufficient quality for the calibration of a sediment model is a widely recognised challenge [72]. In addition to sediment data obtained directly from in situ water and bed samples (grabs), typically, there are two primary types of data which are

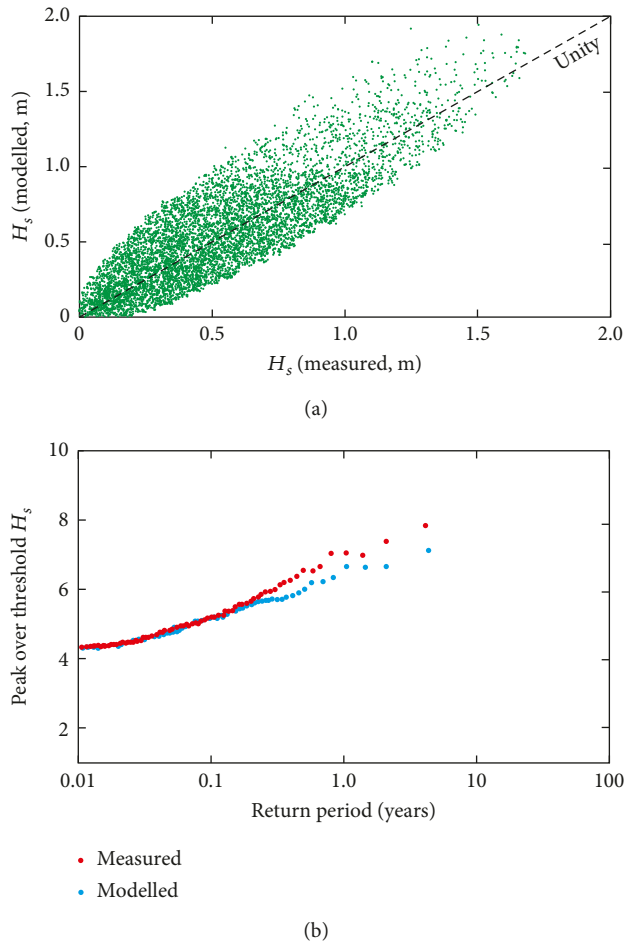


FIGURE 6: (a) Scatter diagram showing the relationship between measured and modelled H_s and (b) a peak over threshold (POT) analysis of H_s between modelled and measured wave data.

required for the calibration of sediment models: (a) measurements of the sediment transport flux (bed load and suspended load) over time scales of a few tidal cycles and (b) measurements of bed-level changes attributable to local erosion and accretion to provide information on net sediment transport over a period of weeks and months. A comprehensive review of instrumentation used to measure sediment transport is given by Williams [73].

7.1.1. Measuring Bed Load and Bedforms. In estuaries and shelf sea environments, bed load is the dominant mode of sediment transport for sand. Sediment traps, frequently used to measure bed load in rivers, have been deployed in estuarine and shelf sea environments with mixed success (cf. [74]). The Arnhem, Helley–Smith, and Delft Nile samplers are the most commonly used devices owing to their robustness and ease of handling in the field. However, their accuracy depends on the number of samples collected which may be restricted by high analysis costs.

As an alternative to traps, bed load is sometimes estimated using fluorescent tracers (e.g., [32]) and less frequently using magnetic tracers (e.g., [75]). The method

involves deploying a quantity of a material at a known location and subsequent sampling campaigns on a grid of sample positions to determine the dispersion of the material. Both approaches use materials with the same dynamic behaviour as the natural sediments and with sufficiently distinct characteristics to make it easily detectable in very low concentrations (<http://www.partrac.com/> (accessed on 1 August 2014)). Sediment tracing techniques have values in studies examining the sediment flux and have been used effectively to study dredging impacts and disruptions to sediment supply attributable to structures. A comprehensive review of tracing techniques and options is given by Black et al. [76].

Passive acoustic techniques using hydrophones to record the sediment-generated noise, SGN, arising during bed load transport of coarse sediments have been used (e.g., [77]) but to be effective, they require objective calibration which can be very costly. However, improvements to processing software and computing power now allow automated analysis of video images to detect particle displacements at subsecond temporal resolution for the entire field of view and the visual analysis of bed load images in providing useful data. Attempts to quantify bed load have also exploited the bottom tracking feature of ADCPs in combination with conventional pressure difference samplers. Together, these instruments can be used to determine the bed load transport velocity and bed load transport rate, respectively (e.g., [78, 79]). Rates of bed load transport have also been inferred from rates of bed form migration measured using rotary sonar devices (e.g., [80, 81]). At a much larger scale, remote sensing techniques have been applied to link large bedform migrations with bed load sediment transport rates (e.g., [82]).

7.1.2. Measuring Suspended Load. Only very fine sand, silt, and mud are transported in suspension in estuarine and shelf sea environments. In the simplest approach used to quantify the suspended load, water samples are collected in situ to determine the concentration of suspended particulate matter (SPM) and the grain-size distribution either at the surface or at a specified depth in the water column using, for example, triggered water bottles or pump sampling (cf. [83, 84]). Samples can be collected either at discrete times or at set times throughout a tidal cycle. There is a low-to-moderate level of certainty in the resultant SPM data owing to potential errors in the way the water samples are collected, the short temporal sampling period, and the presence of varying quantities of organic particles. The deployment of collocated current meters enables the sediment flux to be determined, which in turn can provide useful information on sediment resuspension and settling velocity.

Turbidity meters can provide continuous or discrete measure of SPM concentrations by detecting the attenuation of light passing through the instrument's sampling volume. They are best suited to suspensions of silt and clay-size particles. Self-logging turbidity meters are capable of recording accurately turbidity at a single depth within the water column for long periods [85]. Turbidity data are also obtained using a CTD probe (conductivity, temperature, and

depth) equipped with a turbidity sensor. In estuaries and shelf sea locations, CTDs are normally lowered and raised through the water column for a period of 12.5 hours to provide information on the temporal changes in the SSC profile during a tidal cycle. However, there is the potential for the sensors to become fouled over time giving erroneous data, especially if the sensor becomes exposed during low water and the optical systems are compromised by sediment and/or biological films, leading to unrealistically high turbidity values. In addition, the gain setting (sensitivity of the instrument) must be correctly adjusted to accommodate the range of SPM concentrations in the area. For example, turbidity measurements could reach the upper limit of the instrument if the gain setting is too low or barely register turbidity if the gain setting is too high. This problem can be overcome using an instrument with a logarithmic response which allows measurements of SPM spanning several orders of magnitude. Although overcoming the problems associated with saturation and aliasing, the overall instrument precision is reduced as a result.

Optical turbidity instruments are calibrated using primary solutions such as formazin, and turbidity is expressed in formazin turbidity units (FTU) or nephelometric turbidity units (NTU). The main problem with this approach arises when conversions are made between FTU (or NTU) and in situ water samples where differences can be as large as $\pm 200\%$ [86]. Furthermore, the material in suspension may be a complex mixture of organic and inorganic particles which adds further complexity to the conversion between FTU and SSC. Calibration is therefore required to get the concentration of SPM into meaningful units for sediment modelling purposes (e.g., mg/l). This can be achieved either by collecting water samples at specific times to calibrate the measurements or by calibrating the instrument in laboratory conditions for a range of concentrations before and after deployment. It is important that calibration is performed over an applicable range of SPM concentration values. However, there are additional problems attributable to flocculation when measuring muddy sediments. In these cases, water sampling can destroy the delicate structures, and changes in temperature/salinity can enhance or reduce flocculation potential. Both factors can lead to errors, and thus, these kinds of data must be treated with caution in the context of model calibration [87].

Optical backscatter sensors (OBSs) measure turbidity and suspended solids concentrations by detecting infrared light scattered from SPM (cf. [88, 89]). OBS instruments are best suited to suspensions of silt and clay-size particles. The response of the OBS sensors strongly depends on the size, composition, and shape of the suspended particles, and calibration like that used for turbidity sensors is required to obtain SPM concentration data. OBS instruments are subject to the same problems with biofouling and other optical contamination as turbidity sensors (e.g., [90]). SSC profiles can be obtained using vertical arrays of OBS. Further information is given by, for example, Kineke and Sternberg [91], Hoitink and Hoekstra [92], and Boss et al. [93].

Aerial and satellite remote sensing imagery can be used in some circumstances to indicate the advection rate and

direction of suspended sediment plumes in the surface and near-surface layers of the water column. Remote sensing algorithms have been widely used to extract information on suspended sediment concentrations from multispectral sensor data (e.g., [94, 95]).

For sand-size particles, the use of multiple-frequency acoustic backscatter (ABS) to measure the concentration of suspended sediment is becoming more widespread. Inversion techniques can be applied to obtain suspended sediment concentration (SSC) profiles directly (e.g., [96]), and information about the grain size in suspension can also be extracted (e.g., [97]). Typically, these instruments measure SSC at intervals of 1 cm up to a few metres above the bed where the bulk of suspended sediment is present. SSC profiles can also be derived from ADCP data, albeit with less spatial resolution, using a similar acoustic inversion technique (e.g., [98]). While one or more samples are required for calibration and measurements are spatially averaged, the instrument can provide useful SSC profile information over extended periods.

7.1.3. Estimations of Net Sediment Transport. In many modelling studies, the required outcome concerns the prediction of net sediment transport over periods of days, weeks, or months. There are several useful data that can assist the model calibration process for this aim. In areas where frequent (e.g., annual) maintenance dredging is undertaken, information is likely to be available to describe the frequency and volumes of sediment removed. These data can be used to define changes in bed levels (e.g., accretion amounts and rates over known periods), and through comparisons between predicted accretion rates and rates derived from dredging data, it may be possible to calibrate a sediment model, albeit with limited accuracy (e.g., [99]). In addition, dredging volume data can also be used to provide an indication of the interannual variability of accretion and guide the modelling process. However, owing to sediment loss during the dredging activities and to uncertainty about the bulk density of the material removed, these measurements may not be as accurate as might be desired and should only be used to provide an indication of the volumes of sediment accreting in the area. It is not possible to attribute accretion to a mode of transport, and thus, sediment formulae that predict total sediment transport must be used.

Several acoustic systems have been developed to image the bed at a large scale including echo sounding devices and side-scan and multibeam sonar (e.g., [100–102]). These data can be used to determine net sediment budgets and transport pathways and assist model calibration (e.g., [103, 104]). Repeat subtidal bathymetric surveys can provide valuable information on bedform mobility from which net sediment responses can be determined (e.g., [105]). At the scale of estuaries, Mason et al. [77] illustrate how areas and volumes of sediment accretion and erosion can be estimated using the waterline method employing remote sensing and hydrodynamic modelling. Recent advances in LiDAR now make it possible to penetrate water to depths exceeding 10 m provided water clarity is good enough and thus allow

subtidal survey opportunities (e.g., [106]). Monitoring of large-scale changes in morphology and/or bathymetry in coastal and estuarine environments brought about by sediment mobilisation, transport, and accretion can also now be measured routinely with systems such as ARGUS (<http://www.planetargus.com/> (accessed on 1 August 2014)) and X-band radar (<http://www.oceanwaves.de/> (accessed on 1 August 2014)) (e.g., [107]). Although remote sensing would never be selected to generate a primary bathymetric data set, it has been used in situations where monitoring of rapid bathymetric changes may be required (e.g., following beach nourishments, breaching, etc). Bathymetric and topographic survey data are obtained at irregularly spaced locations. To make the data usable in a numerical model, it is necessary to use an interpolation routine to transform the data onto a grid of regularly spaced data points. Care must be taken to select the most appropriate interpolation method as any error will impact on calculations of change in bathymetry and topography.

LiDAR data are especially useful for intertidal areas. Although spatial positioning is accurate (typically ± 5 cm), the vertical accuracy is at best ± 20 cm. Furthermore, standing water on the beach can result in spurious data, and significant postprocessing may be required. Although the use of LiDAR to determine accurate accretion and erosion rates is not recommended, it does provide extensive (and rapid) spatial cover which may prove to be useful in several applications. In some instances, an assessment of changes in beach topography might be enhanced through reference to fixed, identifiable structures (e.g., quay walls and engineering structures) which can be used to calibrate repeat surveys.

7.2. Sediment Transport Models. Shelf sea and estuarine models normally provide output defining the predicted cumulative erosion/sedimentation for a stated bulk density giving the cumulative change in bed level over the model period. Total sand transport is usually expressed as a net value over a specified period, allowing transport vectors to be plotted which may be comparable with information directly available from the literature.

A wide range of sediment transport formulae are available to predict bed load transport, suspended load transport, and total load transport of noncohesive and cohesive sediments (e.g., [108]). All are derived to represent the best fit to empirical data sets derived in the laboratory or in some cases from the field. The sediment calibration data available will determine the accuracy of the model and limit how much validation is possible. Typically, a sediment model will be calibrated using SSC data, with validation utilising measured sedimentation rates (e.g., [109]). It is important to keep in mind that the sediment transport model is driven by modelled hydrodynamics and that highly nonlinear relationships exist between bed shear stress, flow turbulence, and sediment mobilisation, transport, and accretion. Thus, any limitation with the initial hydrodynamic calibration could impact significantly on the sediment model. It is therefore critically important to obtain the best hydrodynamic calibration possible.

When modelling sediment transport, it is important to recognise the heterogeneity of the seabed and the homogeneity of most sediment transport models. It is therefore essential that roughness maps previously described are used to characterise as accurately as possible the physical properties of the sediments (grain roughness) and the morphology of the bed (form drag). It should be remembered when interpreting model outputs that, since sediment transport formulae are empirical and are based on a limited amount of calibration data from laboratory and/or field studies, the prediction of sediment responses to hydrodynamic forcing is at best limited to accuracy of no more than a factor of two ([110–112]).

7.2.1. Sediment Properties. As the physical and dynamic properties of noncohesive sediments are less complex, the amount of information needed to setup and calibrate sand transport models is less than that required for mud. In the absence of measurements, the specific density, porosity, and bulk density for quartz sand are assumed to be 2650 kg/m^3 , 0.45 kg/m^3 , and 1460 kg/m^3 , respectively. The median grain diameter (D_{50}) is normally measured using grain-size analysis of samples or published data (e.g., BGS (<http://www.bgs.ac.uk/discoverymetadata/13605549.html>)). While the spatial distribution of sand-sized sediment and information on the depth of any deposit is helpful, it is rarely available, leading to ambiguity about sediment source limitations. Information about bedforms is available either from observations described above or generated through theoretical equations linking bedform dimensions with the sediment grain size and the hydrodynamic regime [50]. The threshold bed shear stress, a critical parameter in sediment models defining the bed shear stress required to mobilise the sediments, is normally calculated using a selected empirical formula (e.g., [50]) and expressed as a Shields parameter.

For cohesive sediment transport models, the following sediment data are normally required: sediment density; grain size; settling velocity; and the threshold bed shear stress for erosion and deposition. If this information cannot be obtained from in situ measurements and/or analysis of samples, Whitehouse et al. [113] provide a good account of formulae for deriving some useful properties of cohesive sediments. It is common practice to measure “wet” sediment density in situ using a density probe which can then be converted to “dry” density (e.g., [113]). To obtain the correct dry density required by some models, it is recommended that the porosity factor is changed until the wet density is the same as the measured density. Typically, porosity values between 0.75 and 0.98 should be applied for sediments consolidated for less than 1 year and then 0.25 to 0.75 for longer periods of consolidation.

Typical settling velocities for mud range from 0.003 m/s to 0.0001 m/s and can be calculated based on the grain size (if known) using empirical formulae (cf. [113]). However, caution must be exercised when using this approach due to flocculation, which can increase significantly the size (and hence settling velocity) of particles in suspension. These may also incorporate organic matter in their matrix, thereby

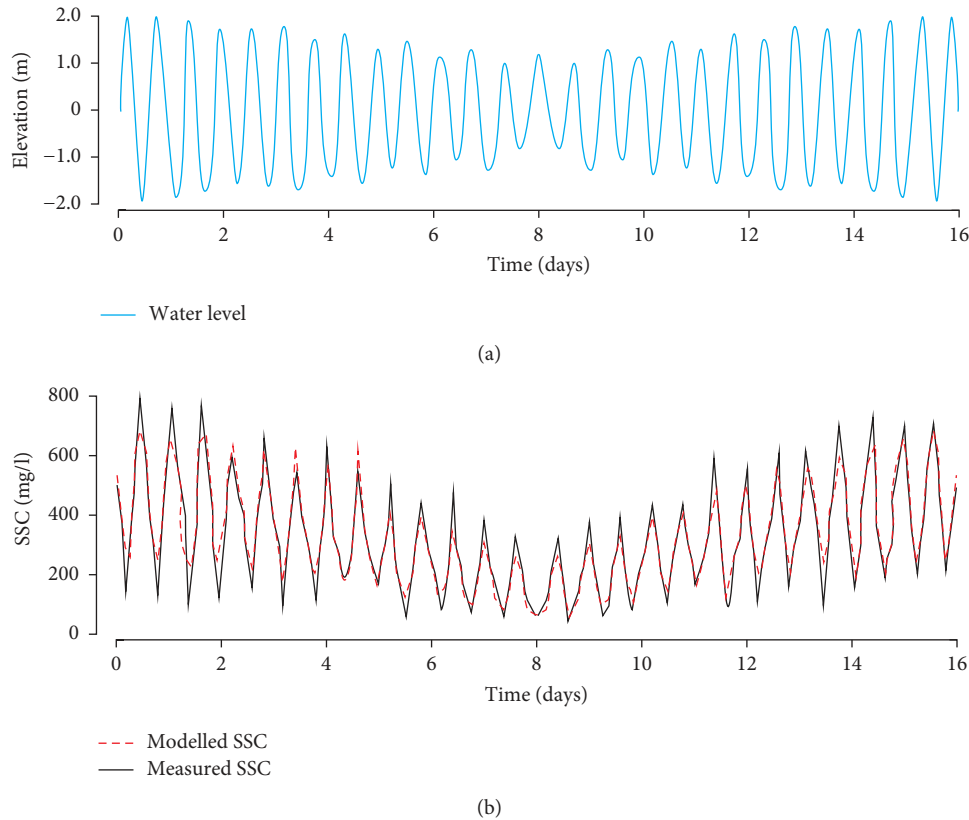


FIGURE 7: Time series of (a) measured water level and (b) SSC (continuous) and predicted SSC from a calibrated cohesive sediment model.

affecting the density (and possibly reducing the settling velocity). Furthermore, sampling of suspended sediments in situ frequently destroys the flocs or alters significantly their physical properties. There is no simple solution to this problem, and modelling assumptions and limitations must be stated clearly. Although knowledge of mineralogy, salinity, turbulent kinetic energy, and water temperature makes it possible to calculate the potential floc size, this is further complicated by temporal and spatial variations in these parameters. It is also noted that many mud models use the SSC as a parameter for defining the settling velocity, not grain size.

The critical bed shear stress for erosion, τ_{crit_E} , can be estimated using the Mitchener et al. [114] formula which accounts for sediment density. Several methods to measure τ_{crit_E} exist and comprise laboratory devices to analyse samples from the field and carousel flumes for field deployment [71]. These can be very effective and allow investigation of how τ_{crit_E} changes as erosion of a given sample proceeds (normally increasing). As with most sediment dynamics, extreme care should be exercised when attempting to parameterise physical properties and processes using empirical approaches. The critical bed shear stress for deposition, τ_{crit_D} , is frequently used as a calibration parameter. However, it is highly dependent on the local conditions. Generally, values between 0.1 N/m^2 and 0.3 N/m^2 provide effective calibration settings for mud models. It is important to note that the default value in some models may not be appropriate for a case (e.g., in Delft3D

$\tau_{\text{crit}_D} = 1000 \text{ N/m}^2$ and must be changed prior to any model runs). Examples to guide the use and calibration of cohesive sediment models are provided by van Kessel et al. [115] and Carniello et al. [116].

7.3. Performance Guidelines

7.3.1. Noncohesive and Cohesive Sediments: Suspended Sediments. Sediment model calibration success can be assessed visually by comparing measured and predicted average concentrations over a set period, typically a spring-neap cycle. Example time series of measured water level and measured and predicted SSC over a period of 8 days are shown in Figure 7. In this case, SSC is measured continuously using a turbidity instrument. A second example of model output and SSC calibration data is shown in Figure 8. In this case, SSC data were obtained from water samples. Both Figures 7 and 8 demonstrate that the general pattern of SSC is similar for the modelled and measured data. For most applications, the aim should be to achieve a model calibration of $\pm 20\%$ of the measured average concentrations. In areas where time series of SSC measurements are available from multiple sites, a calibration level of $\pm 30\%$ for average SSC at most of the sites would be deemed as a good level of calibration. If there are only discrete values of SSC from water samples (or a handheld turbidity probe), experience shows that calibration of only $\pm 40\%$ is achievable since the discrete measurements are subject to higher levels of uncertainty.

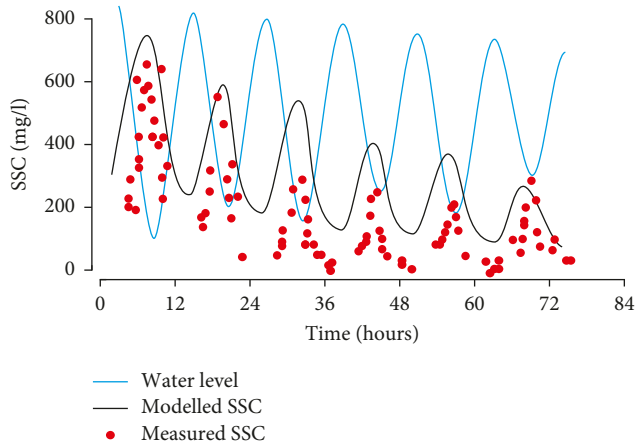


FIGURE 8: Details of measured water level and SSC (intermittent samples) and predicted SSC from a calibrated cohesive sediment model.

7.3.2. Noncohesive and Cohesive Sediments: Sedimentation. Provided sufficient good quality data are available, sedimentation rates provide one of the best means of validating the longer-term performance of sediment models and provide an integrated view of the net result of modelled suspended sediments and bed load. Given the complexity of sediment transport and the errors associated with measurements and empirical sediment transport formulae alluded to above, it is normal practice to apply a scaling factor to the modelled sediment transport rates. In effect, this is a global correction factor to the sediment transport rates predicted by the model that provides an effective means of matching the model predictions of sedimentation with the observations. The scaling factor has no physical meaning and simply represents many complex physical processes not present within the model including, for example, biological and sediment consolidation factors which can significantly alter the physical properties of the sediment with respect to mobilisation and transport.

As a general rule, the scaling factor should be less than 5. Higher values indicate a more significant issue with the accuracy of the modelling, or with site-specific complexity, that may require a nonstandard approach. In such cases, the model approximations cannot be relied upon to describe sedimentation/accretion, and field monitoring is recommended to supplement the model deficiencies.

Dredging data are frequently used as a measure of long-term sedimentation and are normally expressed as the volume of sediment removed from an area per year. Although such data are useful, they are frequently complicated by a poorly defined relationship between the dredged volume and the rarely provided bulk density value which can give rise to significant errors. Recourse must be made to estimated bulk density values, and sensitivity analyses should be used to quantify sedimentation for a plausible range of values. When validating a model using dredging data, the volume of sediment accumulation predicted by the model (normally over a 15-day spring-neap cycle) would normally be scaled to match as closely as possible the measurements. In recognition of the many sources of errors

and uncertainty, a model predicting the dredged volumes to within 50% of the measured rates is normally deemed to be satisfactory for most practical applications. For example, in study of the Humber, the modelled sedimentation volume was 2,180,000 m³/yr, while the average volume of sediment dredged was 1,830,000 m³/yr, with the values ranging from 790,000 m³/yr to 3,915,000 m³/yr over a 5-year period (Mott MacDonald, *per. com.*).

8. Morphological Models

Morphological modelling in estuarine and coastal environments is challenging, and useful description of the range of approaches employed is provided by Roelvink [117]. The primary limitation to the accuracy of morphological models concerns the length of time over which the model is run, with results from long runs (e.g., monthly–decadal) likely to deviate significantly from reality [118, 119].

From the outset, it is very important to establish a conceptual understanding of sediment transport and historical morphological changes in each study area before attempting a morphological model. This must draw together existing evidence and provide a qualitative description of the process controls and how the morphology of the system responds to these drivers. For long-term assessments of morphology, this also requires consideration of climate change factors. A conceptual understanding can provide the hypothesis (e.g., sources and sinks) with which to test the performance of the model and to provide some guidance on expected magnitudes and directions of sediment transport and the associated morphological changes.

The largest constraint to calibrating morphological models is the availability of high-quality data sets that adequately describe the model parameters over a sufficient length of time. In an assessment of data requirements by Splinter et al. [120], it was concluded that (a) calibration of a seasonally dominated site required longer data sets but was less sensitive to sampling interval and (b) calibration of a storm-dominated site required shorter and more frequently sampled data sets. Most studies show that morphological calibrations that are based on short observational records (i.e., < one year) are not robust. To determine initial estimates of calibration coefficients and to hindcast short-term (1–5 years) shoreline variability, Splinter et al. [120] recommend monthly monitoring programs for at least two years. For longer-term predictions of morphology, longer data sets are required to improve the performance of the models.

Morphodynamic models of shelf sea and estuarine environments usually comprise a controlling programme that invokes sequentially subroutines predicting hydrodynamics (e.g., [68, 121, 122]), sediment transport (e.g., [123]), and bed level (via the sediment continuity equation). These are all then linked via the well-established morphodynamic feedback loop (e.g., [124, 125]). Examples of morphological models include (a) the deterministic process-based model Delft3d-MOR (e.g., [126]), suitable for short-term morphology predictions, (b) ASMITA (e.g., [127]), a semi-empirical model using large process-based units which can

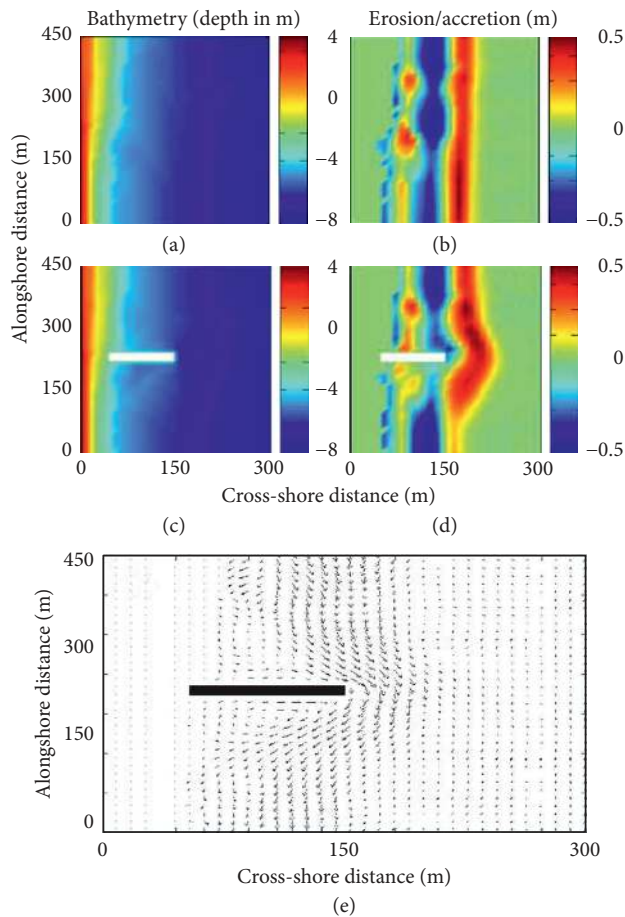


FIGURE 9: Examples of XBeach model output: (a) baseline bathymetry; (b) baseline erosion/accretion after 8 tidal cycles with oblique waves; (c) scheme starting bathymetry with a shore-normal groyne; (d) scheme erosion/accretion after 8 tidal cycles with oblique waves; and (e) snapshot of tidal and wave-induced flows around the groyne.

iterate towards an equilibrium condition and predict large-scale changes in sediment balances (sediment budget) over medium- to long-term periods, and (c) XBeach, a deterministic process-based model suitable for predicting morphological changes resulting from storm impacts (e.g., [128, 129]). An example of outputs from an XBeach model setup to predict the impact of a shore-normal groyne is shown in Figure 9.

However, bed-level changes occur over long time scales compared with the hydrodynamic forcing, and thus until recently, owing to computational limitations, shelf sea morphodynamic models have been unable to predict very far into the future using traditional morphodynamic upscaling techniques such as the “continuity correction” method. To overcome this limitation, Lesser et al. [122] and Roelvink [130] have developed the morphological acceleration factor (MORFAC) concept to enable morphological predictions to extend over decadal (e.g., [131]) and centennial [132] time scales.

Whilst the certainty in predicting morphology cannot be proven, it may be possible to bound the uncertainty using

sensitivity analysis for key process drivers and to determine a range of possible outcomes. Where possible, a range of different morphological modelling approaches should be applied, and where there is general agreement between approaches, then it may be possible to draw additional confidence from the results using an ensemble of model outputs.

Defining what is and what is not a good morphological model performance depends on the spatial and temporal scales considered. At a minimum: (a) the observed/measured sedimentation-erosion patterns must be broadly in agreement with the model outputs; (b) contour plots of measured and computed sedimentation and erosion need to agree as closely as practicable; (c) predicted volume changes over control areas must agree as closely as practicable with soundings or dredging figures; and (d) the shape, migration, and area change of measured and computed cross sections are required to agree.

The incomplete description of the physics underpinning morphological processes and an imperfect knowledge of the initial conditions and parameters will always lead to increasing errors in the model predictions and limit the ability of shelf sea and estuarine morphodynamic models to accurately predict the future true state of the environment.

9. Improving Predictions and Reducing Uncertainty

An alternative emerging approach to address the problem of model prediction uncertainties involves the application of data assimilation techniques. These techniques keep model parameters fixed and produce an updated model state that matches as closely as possible the true state by combining observational data with model predictions. This updated model state is then used to initiate the next model forecast. However, even if the initial system state can be described flawlessly, model parameters simplify the physical processes and, by doing so, will result in the growth of prediction errors. At present, assimilation methods being developed to improve morphological forecast reliability are producing encouraging results (e.g., [133–137]). For example, ad hoc data assimilation schemes and techniques using more refined heuristic tuning of model state variables are being used to improve the performance of suspended sediment transport models (e.g., [138, 139]).

Two main types of uncertainty pervade morphological models: (a) *scenario uncertainty* stems from uncertainty about the nature of the future weather and weather events (magnitude and frequency) responsible for driving morphological change and (b) *response uncertainty* relates to the uncertainty in predicting how the morphological system will respond to given forcing conditions. To reduce uncertainty in morphological model predictions, the ensemble approach widely adopted by a climate change scientist (e.g., [140]) may prove to be helpful. The ensemble modelling approach aims to address uncertainty arising from two main sources: (a) incomplete description of the physical processes bringing

TABLE 7: Statistical guidelines to establish calibration standards for a minimum level of performance for coastal and estuarine hydrodynamic and sediment models. The table is based in part on the recommendations from Evans [1] and Bartlett [2].

Model predictions	RMSE	Bias	R	SI
Bathymetry	$\pm 2.5\%$ of the mean water depth	<0.10	>0.95	$<10\%$
Water level (coast)	± 0.1 m or 10% of the measured level (spring tide); $\pm 15\%$ (neap tide)	<0.10	>0.95	$<10\%$
Water level (estuary)	± 0.1 m (mouth); ± 0.3 m (head) or 10% of the measured level (spring tide); $\pm 15\%$ (neap tide)	<0.20	>0.95	$<15\%$
Water-level phase (coast)	± 15 minutes	<0.20	>0.90	$<20\%$
Water-level phase (estuary)	± 15 minutes (mouth); ± 25 minutes (head)	<0.25	>0.90	$<20\%$
Average current speed	± 0.1 m/s or $\pm 10\%$ to 20% of the measured speed	<0.10	>0.95	$<10\%$
Peak current speed	Within <0.05 m/s (very good), <0.1 m/s (good), <0.2 m/s (moderate), and <0.3 m/s (poor) of the measured peak speed	<0.15	>0.90	$<15\%$
Current direction (coastal)	$\pm 10^\circ$ of the measured direction	<0.25	>0.90	$<20\%$
Current direction (estuary)	$\pm 15^\circ$ of the measured direction	<0.30	>0.90	$<20\%$
Bed shear stress	$\pm 10\%$ N/m ² of the measured mean stress	<0.10	>0.95	$<10\%$
Wave height	$\pm 10\%$ of the measured mean observed height	<0.15	>0.95	$<10\%$
Wave period	$\pm 20\%$ of the measured mean observed period	<0.20	>0.90	$<20\%$
Wave direction	$\pm 30\%$ of the measured mean observed direction	<0.25	>0.90	$<25\%$
Mean SPM concentration	$\pm 20\%$ of the mean measured SPM concentration	<0.20	>0.90	$<20\%$
Accretion based on dredge volumes	$\pm 50\%$ of the measured dredge volume	N/A	N/A	N/A
Temperature	$\pm 0.5^\circ$	—	—	—
Salinity	± 1 psu	—	—	—

about morphological changes and (b) limited computing power that constrains how accurately processes can be parameterised in models. For example, in models of estuaries or shelf sea systems, subgrid scale processes such as turbulence can only be represented in a simple way. There are two possible routes to take in ensemble modelling: (a) *perturbed-physics* studies investigate how model predictions are affected by the choice of input parameters through running systematically a single model with different parameter values and (b) *multimodel studies* investigate how predictions differ between different models. The effect of the initial conditions of the model can also be tested using both approaches.

10. Summary and Conclusions

The modelling guidance presented in the paper has drawn on published guidelines and on the extensive practical experience of the authors and their colleagues using a range of model types in modelling projects concerned with, for example, managed realignment, environmental impact assessments for offshore wind farms, tidal energy, coastal defences, dredging/disposal sites, beach and estuarine morphodynamics, barrages, and cooling water discharges. Statistical guidelines to establish calibration standards for a minimum level of performance for coastal and estuarine hydrodynamic and sediment models are summarised in Table 7 and are based in part on the recommendations from Evans [1] and Bartlett [2].

While naturally these guidelines remain open to challenges from modellers requiring more exacting model performance, they have been found to deliver models with a good prognostic performance across a broad range of metrics and recognise the

practical limitations imposed on model calibration processes by the accuracy and the temporal and spatial resolutions of the available calibration data. Their use in coastal and estuarine modelling studies is therefore recommended.

Nomenclature

ADCP:	Acoustic Doppler current profiler
ADV:	Acoustic Doppler velocimeter
AWAC:	Acoustic waves and currents
BST:	British Summer Time
CD:	Chart datum
CTD:	Conductivity, temperature, and depth
DHI:	Danish hydraulics institute
DGPS:	Differential global positioning system
DEM:	Digital elevation model
D_{50} :	Median grain diameter
D_{90} :	90% finer than
FTU:	Formazin turbidity units
GMT:	Greenwich Mean Time
HAT:	Highest astronomical tide
HF:	High frequency
h_{\max} :	Maximum wave height
H_{m0} :	H_s computed using spectral analysis
H_s :	Significant wave height
JONSWAP:	Joint North Sea Wave Project
LAR:	Largest astronomical range
LAT:	Lowest astronomical tide
LiDAR:	Light detection and ranging
MHWN:	Mean high water of neap tides
MHWS:	Mean high water of spring tides

MLWS:	Mean low water of spring tides
MLWN:	Mean low water of neap tides
MSL:	Mean sea level
N:	Newton
N_i :	Total number of data
NTU:	Nephelometric turbidity units
OBS:	Optical backscatter sensors
ODN:	Ordnance Datum Newlyn
OSGB:	Ordnance Survey Great Britain
O_i :	Observed wave data
POT:	Peak over threshold
PT:	Pressure transducer
Q:	Sediment transport
QA:	Quality assurance
QMS:	Quality management system
R:	Pearson product-moment correlation coefficient
RMS:	Root mean square
SET:	Sedimentation erosion table
SI:	Scatter index
S_i :	Simulated or model data
SSC:	Suspended sediment concentration
T_m :	Mean wave period
T_p :	Spectral peak wave period
T_z :	Mean wave period
WGS84:	The new world geodetic system
WW3:	WaveWatch 3
G:	Gram
kg:	Kilogram
k_s :	Equivalent sand roughness
l:	Litre
mg:	Milligram
s:	Second
ϵ :	Rate of turbulent energy dissipation
κ :	Rate of turbulent kinetic energy production
μ :	Micron (10^{-6} m)
θ :	Wave direction (degrees)
θ_m :	Mean wave direction (degrees)
τ :	Bed shear stress (N/m^2).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

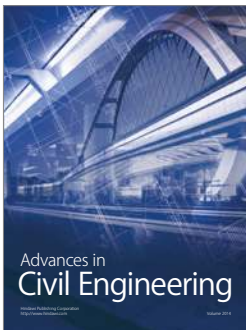
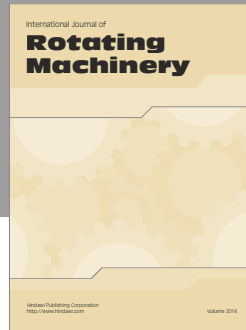
- [1] G. P. Evans, "A framework for marine and estuary model specification in the UK," Tech. Rep. FR0374, p. 58, Foundation for Water Research, 1993.
- [2] J. M. Bartlett, "Quality control manual for computational estuarine modelling, Binnie, Black & Veatch, R&D," Tech. Rep. W113, p. 80, Environment Agency, 1998.
- [3] Defra/Environment Agency, "Protocols for minimum standards in modelling-flood warning management system phase 2a," Tech. Rep. W5C-021/2a, Report prepared for Defra/Environment Agency FCERM management R&D Programme, 2005.
- [4] Defra/Environment Agency, "Benchmarking the latest generation of 2D hydraulic modelling packages," Tech. Rep. SC120002, Joint Defra/EA FCERM R&D programme project SC120002, 2013.
- [5] M.-H. Hsu, A. Y. Kuo, J.-T. Kuo, and W.-C. Liu, "Procedure to calibrate and verify numerical models of estuarine hydrodynamics," *Journal of Hydraulic Engineering*, vol. 125, no. 2, pp. 166–182, 1999.
- [6] B. Yan, Y. Jia, and J. B. Hinwood, "Use of one-dimensional modelling in estuary management: entrance depth: model calibration," *Journal of Coastal Conservation*, vol. 17, no. 1, pp. 191–196, 2013.
- [7] Y. P. Sheng, "On modeling three-dimensional estuarine and marine hydrodynamics," *Elsevier Oceanography Series*, vol. 45, pp. 35–54, 1987.
- [8] H. V. Gupta, K. J. Beven, and T. Wagener, "Model calibration and uncertainty estimation," in *Encyclopedia of Hydrological Sciences*, John Wiley & Sons, Ltd., Hoboken, USA, 2005.
- [9] M. Sousa and J. M. Dias, "Hydrodynamic model calibration for a mesotidal lagoon: the case of Ria de Aveiro (Portugal)," *Journal of Coastal Research*, vol. 6, pp. 1075–1080, 2007.
- [10] D. O. Lambkin, J. M. Harris, W. S. Cooper, and T. Coates, *Coastal Process Modelling for Offshore Wind Farm Environmental Impact Assessment: Best Practice Guide*, COWRIE Limited, London, UK, 2009.
- [11] D. P. Dee, "A pragmatic approach to model validation," in *Quantitative Skill Assessment for Coastal Ocean Models*, vol. 47 of Coastal and Estuarine Studies, pp. 1–13, 1995.
- [12] R. H. van Waveren, S. Groot, H. Scholten et al., *Good Modelling Practice Handbook*, STOWA Report 99-05, RWS-RIZA, Utrecht, Netherlands, <http://waterland.net/riza/aquest>.
- [13] L. Ma, J. C. Ascough II, L. R. Ahuja, M. J. Shaffer, J. D. Hanson, and K. W. Rojas, "Root zone water quality model sensitivity analysis using Monte Carlo simulation," *Transactions of the American Society of Agricultural Engineers*, vol. 43, no. 4, pp. 883–895, 2000.
- [14] J. G. Arnold, D. N. Moriasi, P. W. Gassman et al., "SWAT: model use, calibration, and validation," *Transactions of the American Society of Agricultural and Biological Engineers*, vol. 55, no. 4, pp. 1491–1508, 2012.
- [15] M. W. van Liew, J. G. Arnold, and D. D. Bosch, "Problems and potential of autocalibrating a hydrologic model," *Transactions of the American Society of Agricultural Engineers*, vol. 48, no. 3, pp. 1025–1040, 2005.
- [16] Q. D. Duan, V. K. Gupta, and S. Sorooshian, "Effective and efficient global optimization for conceptual rainfall-runoff models," *Water Resources Research*, vol. 28, no. 4, pp. 1015–1031, 1992.
- [17] R. M. Errico, "What is an adjoint model?" *Bulletin American Meteorological Society*, vol. 78, no. 11, pp. 2577–2591, 1997.
- [18] M. Guarracino, M. Dowd, J. Sheng, and J. J. Cullen, "Calibration of a coupled biological-physical model for prediction in a coastal inlet," *Continental Shelf Research*, vol. 31, no. 16, pp. 1713–1727, 2011.
- [19] D. M. Mulcare, *NGS Toolkit, Part 9: The National Geodetic Survey VERTCON Tool*, Professional Surveyor, vol. 24, no. 3, 2004.
- [20] L. Cea and J. R. French, "Bathymetric error estimation for the calibration and validation of estuarine hydrodynamic models," *Estuarine, Coastal and Shelf Science*, vol. 100, pp. 124–132, 2012.
- [21] N. G. Plant, K. T. Holland, and J. A. Puleo, "Analysis of the scale of errors in nearshore bathymetric data," *Marine Geology*, vol. 191, no. 1–2, pp. 71–86, 2002.
- [22] N. G. Plant, K. L. Edwards, J. M. Kaihatu, J. Veeramony, L. Hsu, and K. T. Holland, "The effect of bathymetric filtering

- on nearshore process model results,” *Coastal Engineering*, vol. 56, no. 4, pp. 484–493, 2009.
- [23] B. Mourre, P. De Mey, F. Lyard, and C. Le Provost, “Assimilation of sea level data over continental shelves: an ensemble method for the exploration of model errors due to uncertainties in bathymetry,” *Dynamics of Atmospheres and Oceans*, vol. 38, no. 2, pp. 93–121, 2004.
- [24] J. Nikuradse, *Laws of Flow in Rough Pipes*, National Advisory Committee for Aeronautics, WA, USA, 1950.
- [25] R. Manning, “On the flow of water in open channels and pipes,” *Transactions of the Institution of Civil Engineers of Ireland*, vol. 20, pp. 161–207, 1891.
- [26] A. Chézy, “Formula to find the uniform velocity that the water will have in a ditch or in a canal of which the slope is known,” *Journal Association of Engineering Societies*, 1776. File No. 847, vol. 18, pp. 363–369, Ms. 1915 in the Library of Ecole des Ponts et Chaussées, English translation in H. Clemens, On the origin of the Chézy formula, 1897.
- [27] A. Lefebvre and A. P. Lyons, “Quantification of roughness for seabed characterization,” in *Proceedings of the 4th International Conference on Underwater Acoustic Measurements: Technologies and Results*, Kos Island, Greece, June 2011.
- [28] M. Guerrero, V. Di Federico, and A. Lamberti, “Calibration of a 2-D morphodynamic model using water–sediment flux maps derived from an ADCP recording,” *Journal of Hydroinformatics*, vol. 15, no. 3, pp. 813–828, 2013.
- [29] N. Huybrechts, H. Smaoui, A. Ouahsine et al., “Prediction of the bed friction coefficient using either high resolution bathymetric data or granulometry samples,” in *Proceedings of the XXth TELEMAC-MASCARET User Conference (TUC 2013)*, p. 109, Karlsruhe, Germany, October 2013.
- [30] A. J. Cooper, “TELEMAC-2D version 3.0-validation document,” Report EDF-LNH HE-43/96/006/A, 74 pages, 1993.
- [31] J. M. Hervouet and P. Bates, “The TELEMAC modelling system—special issue,” *Hydrological Processes*, vol. 14, no. 13, pp. 2207–2208, 2000.
- [32] X. Bertin, A. Oliveira, and A. B. Fortunato, “Simulating morphodynamics with unstructured grids: description and validation of a modeling system for coastal applications,” *Ocean Modelling*, vol. 28, no. 1, pp. 75–87, 2009.
- [33] M. H. Hsu, C. J. Huang, A. S. Chen, and Y. T. Chou, “Inundation simulation model with local grid refinement,” in *Proceedings of the 34th World Congress of the International Association for Hydro-Environment Research and Engineering: 33rd Hydrology and Water Resources Symposium and 10th Conference on Hydraulics in Water Engineering*, p. 4360, Engineers Australia, Brisbane, Australia, June–July 2011.
- [34] H. W. Kernkamp, A. van Dam, G. S. Stelling, and E. D. de Goede, “Efficient scheme for the shallow water equations on unstructured grids with application to the Continental Shelf,” *Ocean Dynamics*, vol. 61, no. 8, pp. 1175–1188, 2011.
- [35] X. Liu and L. Ren, “Optimal selection of grid-scale for land surface model in the numerical simulation of spring soil moisture,” *Journal of China Agricultural University*, vol. 5, p. 030, 2012.
- [36] J. J. Maynard and M. G. Johnson, “Scale-dependency of LiDAR derived terrain attributes in quantitative soil-landscape modeling: effects of grid resolution vs. neighborhood extent,” *Geoderma*, vol. 230, pp. 29–40, 2014.
- [37] K. El Kheishy, J. McCorquodale, I. Georgiou, and E. Meselhe, “Bed forms resistance dependency on numerical model grid size spatial resolution,” *Journal of Spatial Science*, vol. 59, no. 2, pp. 363–372, 2014.
- [38] G. W. Brier, “Verification of forecasts expressed in terms of probabilities,” *Monthly Weather Review*, vol. 78, no. 1, pp. 1–3, 1950.
- [39] K. A. Haas, I. A. Svendsen, M. C. Haller, and Q. Zhao, “Quasi-three-dimensional modeling of rip current systems,” *Journal of Geophysical Research: Oceans*, vol. 108, no. C7, pp. 1978–2012, 2003.
- [40] J. E. Nash and J. V. Sutcliffe, “River flow forecasting through conceptual models part I—A discussion of principles,” *Journal of Hydrology*, vol. 10, no. 3, pp. 282–290, 1970.
- [41] I. G. Watterson, “Non-dimensional measures of climate model performance,” *International Journal of Climatology*, vol. 16, no. 4, pp. 379–391, 1996.
- [42] D. R. Legates and G. J. McCabe Jr, “Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation,” *Water Resources Research*, vol. 35, no. 1, pp. 233–241, 1999.
- [43] P. W. Mielke Jr. and K. J. Berry, *Permutation Methods: A Distance Function Approach*, Springer-Verlag, New York, USA, p. 352, 2001.
- [44] C. J. Willmott, “On the validation of models,” *Physical Geography*, vol. 2, no. 2, pp. 184–194, 1981.
- [45] L. C. van Rijn, D. J. R. Walstra, B. Grasmeijer, J. Sutherland, S. Pan, and J. P. Sierrae, “The predictability of cross-shore bed evolution of sandy beaches at the time scale of storms and seasons using process-based profile models,” *Coastal Engineering*, vol. 47, no. 3, pp. 295–327, 2003.
- [46] C. J. Willmott, S. M. Robeson, and K. Matsuura, “A refined index of model performance,” *International Journal of Climatology*, vol. 32, no. 13, pp. 2088–2094, 2012.
- [47] Y. P. Sheng and T. Kim, “Skill assessment of an integrated modeling system for shallow coastal and estuarine ecosystems,” *Journal of Marine Systems*, vol. 76, no. 1–2, pp. 212–243, 2009.
- [48] J. C. Warner, W. R. Geyer, and J. A. Lerczak, “Numerical modeling of an estuary: a comprehensive skill assessment,” *Journal of Geophysical Research: Oceans*, vol. 110, no. C5, pp. 1978–2012, 2005.
- [49] R. C. Ris, L. H. Holthuijsen, and N. Booij, “A third-generation wave model for coastal regions 2. Verification,” *Journal of Geophysical Research*, vol. 104, no. C4, pp. 7667–7681, 1999.
- [50] R. L. Soulsby, *Dynamics of Marine Sands. A Manual for Practical Applications*, Thomas Telford Publications, London, UK, 1997.
- [51] H. U. Oebius, *Circulation and Sediment Transport on Sand-Banks in European Shelf Seas. Part: Development of a Shear Stress Meter*, Report No. 1208/92, VWS, Berlin, Germany, p. 100, 1992.
- [52] A. J. Grass, R. R. Simons, R. D. MacIver, M. Mansour-Tehrani, and A. Kalopedis, “Shear cell for direct measurement of fluctuating bed shear stress vector in combined wave current flow,” in *Proceedings of the 26th IAHR Congress “HYDRA 2000”*, pp. 415–421, London, UK, September 1995.
- [53] K. L. Rankin and R. I. Hires, “Laboratory measurement of bottom shear stress on a movable bed,” *Journal of Geophysical Research: Oceans*, vol. 105, no. C7, pp. 17011–17019, 2000.
- [54] K. E. Steele, C. C. Teng, and D. W. C. Wang, “Wave direction measurements using pitch-roll buoys,” *Ocean Engineering*, vol. 19, no. 4, pp. 349–375, 1992.
- [55] G. J. Komen, L. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann, and P. A. E. M. Janssen, *Dynamics and Modelling of Ocean Waves*, Cambridge University Press, New York, NY, USA, 1996.
- [56] H. E. Krogstad, S. F. Barstow, S. E. Aasen, and I. Rodriguez, “Some recent developments in wave buoy measurement

- technology," *Coastal Engineering*, vol. 37, no. 3-4, pp. 309–329, 1999.
- [57] H. E. Krogstad, J. Wolf, S. P. Thompson, and L. R. Wyatt, "Methods for intercomparison of wave measurements," *Coastal Engineering*, vol. 37, no. 3-4, pp. 235–257, 1999.
- [58] S. Lindroth and M. Leijon, "Offshore wave power measurements—a review," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 9, pp. 4274–4285, 2011.
- [59] T. P. Jardine, "The reliability of visually observed wave heights," *Coastal Engineering*, vol. 3, pp. 33–38, 1980.
- [60] J. J. De Vries, J. Waldron, and V. Cunningham, "Field tests of the new Datawell DWR-G GPA wave buoy," *Sea Technology*, vol. 44, no. 12, pp. 50–55, 2003.
- [61] H. E. Krogstad and S. F. Barstow, "Satellite wave measurements for coastal engineering applications," *Coastal Engineering*, vol. 37, no. 3-4, pp. 283–307, 1999.
- [62] I. R. Young, S. Zieger, and A. V. Babanin, "Development and application of a global satellite database of wind and wave conditions," in *Proceeding of the 34th Conference on Ocean, Offshore, and Arctic Engineering*, vol. 7, pp. 62–71, St. John's, Canada, May–June 2015.
- [63] L. R. Wyatt, J. B. D. Jaffrés, and M. L. Heron, "Spatial averaging of HF radar data for wave measurement applications," *Journal of Atmospheric and Oceanic Technology*, vol. 30, no. 9, pp. 2216–2224, 2013.
- [64] J. Wolf and P. S. Bell, "Waves at Holderness from X-band radar," *Coastal Engineering*, vol. 43, no. 3-4, pp. 247–263, 2001.
- [65] R. A. Holman and J. Stanley, "The history and technical capabilities of Argus," *Coastal Engineering*, vol. 54, no. 6-7, pp. 477–491, 2007.
- [66] G. Z. Forristall and K. C. Ewans, "Worldwide measurements of directional wave spreading," *Journal of Atmospheric and Oceanic Technology*, vol. 15, no. 2, pp. 440–469, 1998.
- [67] R. M. Gorman and C. G. Neilson, "Modelling shallow water wave generation and transformation in an intertidal estuary," *Coastal Engineering*, vol. 36, no. 3, pp. 197–217, 1999.
- [68] N. Booij, L. H. Holthuijsen, and R. C. Ris, "The "SWAN" wave model for shallow water," in *Proceedings of the 25th International Conference on Coastal Engineering*, Orlando, FL, USA, September 1996.
- [69] D. G. Wren, B. D. Barkdoll, R. A. Kuhnle, and R. W. Darrow, "Field techniques for suspended-sediment measurement," *Journal of Hydraulic Engineering*, vol. 126, no. 2, pp. 97–104, 2000.
- [70] P. Le Hir, Y. Monbet, and F. Orvain, "Sediment erodability in sediment transport modelling: can we account for biota effects?," *Continental Shelf Research*, vol. 27, no. 8, pp. 1116–1142, 2007.
- [71] C. L. Amos, J. Grant, G. R. Daborn, and K. Black, "Sea carousel—a benthic, annular flume," *Estuarine, Coastal and Shelf Science*, vol. 34, no. 6, pp. 557–577, 1992.
- [72] J. J. Wang, X. X. Lu, and M. Kumm, "Sediment load estimates and variations in the Lower Mekong River," *River Research and Applications*, vol. 27, no. 1, pp. 33–46, 2011.
- [73] J. J. Williams, "Recent advancements in understanding of sediment processes," in *Sediments, Morphology, and Sedimentary Processes on Continental Shelves, Special Publication of the International Association of Sedimentologists (IAS)*, M. Li, C. Sherwood, and P. Hill, Eds., John Wiley & Sons, Hoboken, USA, 2012.
- [74] C. W. Emerson, "A method for the measurement of bedload sediment transport and passive faunal transport on intertidal sandflats," *Estuaries*, vol. 14, no. 4, pp. 361–371, 1991.
- [75] K. D. van der Post, F. Oldfield, and G. Voulgaris, "Magnetic tracing of beach sand: preliminary results," in *Proceedings of the Coastal Dynamics'94*, pp. 323–334, ASCE, Barcelona, Spain, February 1994.
- [76] K. S. Black, S. Athey, P. Wilson, and D. Evans, "The use of particle tracking in sediment transport studies: a review," *Geological Society, London, Special Publications*, vol. 274, no. 1, pp. 73–91, 2007.
- [77] D. C. Mason, M. Amin, I. J. Davenport, R. A. Flather, G. J. Robinson, and J. A. Smith, "Measurement of recent intertidal sediment transport in Morecambe Bay using the waterline method," *Estuarine, Coastal and Shelf Science*, vol. 49, no. 3, pp. 427–456, 1999.
- [78] J. W. Gartner, "Estimating suspended solids concentrations from backscatter intensity measured by acoustic Doppler current profiler in San Francisco Bay, California," *Marine Geology*, vol. 211, no. 3-4, pp. 169–187, 2004.
- [79] C. D. Rennie and P. V. Villard, "Site specificity of bed load measurement using an acoustic Doppler current profiler," *Journal of Geophysical Research: Earth Surface*, vol. 109, no. F3, 2004.
- [80] J. J. Williams, P. S. Bell, and P. D. Thorne, "Field measurements of flow fields and sediment transport above mobile bed forms," *Journal of Geophysical Research: Oceans*, vol. 108, no. C4, 2003.
- [81] J. G. Venditti, M. Church, and S. J. Bennett, "Morphodynamics of small-scale superimposed sand waves over migrating dune bed forms," *Water Resources Research*, vol. 41, no. 10, 2005.
- [82] C. R. Bates and D. J. Oakley, "Bathymetric sidescan investigation of sedimentary features in the Tay Estuary, Scotland," *International Journal of Remote Sensing*, vol. 25, no. 22, pp. 5089–5104, 2004.
- [83] A. Mudroch and S. D. MacKnight, *Handbook of Techniques for Aquatic Sediments Sampling*, CRC Press, Boca Raton, FL, USA, 1994.
- [84] L. Thomsen, G. Graf, V. Martens, and E. Steen, "An instrument for sampling water from the benthic boundary layer," *Continental Shelf Research*, vol. 14, no. 7, pp. 871–882, 1994.
- [85] J. P. Minella, G. H. Merten, J. M. Reichert, and R. T. Clarke, "Estimating suspended sediment concentrations from turbidity measurements and the calibration problem," *Hydrological Processes*, vol. 22, no. 12, pp. 1819–1830, 2008.
- [86] A. Ziegler, *Issues related to use of turbidity measurements as a surrogate for suspended sediment. Turbidity and Other Sediment Surrogates Workshop*, Reno, NV, USA, 2002.
- [87] F. Mietta, C. Chassagne, R. Verney, and J. C. Winterwerp, "On the behavior of mud floc size distribution: model calibration and model behavior," *Ocean Dynamics*, vol. 61, no. 2-3, pp. 257–271, 2011.
- [88] J. P. Downing, R.W. Sternberg, and C. R. B. Lister, "New instrument for the investigation of sediment suspension processes in the shallow marine environment," *Marine Geology*, vol. 42, no. 1-4, pp. 19–34, 1981.
- [89] M. O. Green and J. D. Boon, "The measurement of constituent concentrations in non-homogeneous sediment suspensions using optical backscatter sensors," *Marine Geology*, vol. 110, no. 1-2, pp. 73–81, 1993.
- [90] T. F. Sutherland, P. M. Lane, C. L. Amos, and J. Downing, "The calibration of optical backscatter sensors for suspended sediment of varying darkness levels," *Marine Geology*, vol. 162, no. 2-4, pp. 587–597, 2000.
- [91] G. C. Kineke and R. W. Sternberg, "Measurements of high concentration suspended sediments using the optical

- backscatterance sensor," *Marine Geology*, vol. 108, no. 3-4, pp. 253–258, 1992.
- [92] A. J. F. Hoitink and P. Hoekstra, "Observations of suspended sediment from ADCP and OBS measurements in a mud-dominated environment," *Coastal Engineering*, vol. 52, no. 2, pp. 103–118, 2005.
- [93] E. Boss, L. Taylor, S. Gilbert et al., "Comparison of inherent optical properties as a surrogate for particulate matter concentration in coastal waters," *Limnology & Oceanography: Methods*, vol. 7, no. 11, pp. 803–810, 2009.
- [94] R. P. Stump and J. R. Pennock, "Calibration of a general optical equation for remote sensing of suspended sediments in a moderately turbid estuary," *Journal of Geophysical Research: Oceans*, vol. 94, no. C10, pp. 14363–14371, 1989.
- [95] B. Nechad, K. G. Ruddick, and Y. Park, "Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters," *Remote Sensing of Environment*, vol. 114, no. 4, pp. 854–866, 2010.
- [96] P. D. Thorne and P. J. Hardcastle, "Acoustic measurements of suspended sediments in turbulent currents and comparison with in-situ samples," *Journal of the Acoustical Society of America*, vol. 101, no. 5, pp. 2603–2614, 1997.
- [97] A. M. Crawford and A. E. Hay, "Determining suspended sand size and concentration from multifrequency acoustic backscatter," *Journal of the Acoustical Society of America*, vol. 94, no. 6, pp. 3312–3324, 1993.
- [98] L. M. Merckelbach and H. Ridderinkhof, "Estimating suspended sediment concentration using backscatterance from an acoustic Doppler profiling current meter at a site with strong tidal currents," *Ocean Dynamics*, vol. 56, no. 3-4, pp. 153–168, 2006.
- [99] I. Townend and P. Whitehead, "A preliminary net sediment budget for the Humber Estuary," *Science of the Total Environment*, vol. 314, pp. 755–767, 2003.
- [100] J. E. Hughes Clarke, L. A. Mayer, and D. E. Wells, "Shallow-water imaging multibeam sonars: a new tool for investigating seafloor processes in the coastal zone and on the continental shelf," *Marine Geophysical Research*, vol. 18, no. 6, pp. 607–629, 1996.
- [101] J. V. Gardner, P. Dartnell, L. A. Mayer, and J. E. Hughes Clarke, "Geomorphology, acoustic backscatter, and processes in Santa Monica Bay from multibeam mapping," *Marine Environmental Research*, vol. 56, no. 1-2, pp. 15–46, 2003.
- [102] P. K. Pandian, S. Ramesh, M. V. R. Murthy, S. Ramachandran, and S. Thayumanavan, "Shoreline changes and near shore processes along Ennore coast, East coast of south India," *Journal of Coastal Research*, vol. 203, pp. 828–845, 2004.
- [103] P. L. Barnard, D. M. Hanes, and D. M. Rubin, "Giant sand waves at the mouth of San Francisco Bay," *Eos*, vol. 87, no. 29, pp. 285–289, 2006.
- [104] M. C. Buijsman and H. Ridderinkhof, "Long-term evolution of sand waves in the Marsdiep inlet. I: high-resolution observations," *Continental Shelf Research*, vol. 28, no. 9, pp. 1190–1201, 2008.
- [105] J. L. D. Woods and D. M. Kennedy, "The measurement of modern sedimentation in estuarine environments in New Zealand," *New Zealand Geographer*, vol. 67, no. 1, pp. 39–50, 2011.
- [106] J. Wozencraft and D. Millar, "Airborne Lidar and integrated technologies for coastal mapping and nautical charting," *Marine Technology Society Journal*, vol. 39, no. 3, pp. 27–35, 2005.
- [107] P. S. Bell, "Shallow water bathymetry derived from an analysis of X-band marine radar images of waves," *Coastal Engineering*, vol. 37, no. 3-4, pp. 513–527, 1999.
- [108] F. Cayocca, R. Verney, S. Petton et al., "Development and validation of a sediment dynamics model within a coastal operational oceanographic system," *Mercator Ocean Quarterly Newsletter*, vol. 49, no. 2, pp. 76–86, 2014.
- [109] N. K. Ganju and D. H. Schoellhamer, "Calibration of an estuarine sediment transport model to sediment fluxes as an intermediate step for robust simulation of geomorphic evolution," *Continental Shelf Research*, vol. 29, no. 1, pp. 148–158, 2009.
- [110] C. Teisson, "Cohesive suspended sediment transport: feasibility and limitations of numerical modelling," *Journal of Hydraulic Research*, vol. 29, no. 6, pp. 755–769, 1991.
- [111] W. S. Merritt, R. A. Letcher, and A. J. Jakeman, "A review of erosion and sediment transport models," *Environmental Modelling & Software*, vol. 18, no. 8-9, pp. 761–799, 2003.
- [112] L. C. van Rijn, "Unified view of sediment transport by currents and waves. I: initiation of motion, bed roughness, and bed-load transport," *Journal of Hydraulic Engineering*, vol. 133, no. 6, pp. 649–667, 2007.
- [113] R. J. Whitehouse, R. L. Soulsby, W. Roberts, and H. Mitchener, *Dynamics of Estuarine Muds*, Thomas Telford Publications, Heron Quay, London, UK, 2000.
- [114] H. J. Mitchener, H. Torfs, and R. J. S. Whitehouse, "Erosion of sand/mud mixtures," *Coastal Engineering*, vol. 29, no. 1-2, pp. 1–25, 1996.
- [115] T. van Kessel, J. Vanlede, and J. de Kok, "Development of a mud transport model for the Scheldt estuary," *Continental Shelf Research*, vol. 31, no. 10, pp. S165–S181, 2011.
- [116] L. Carniello, A. Defina, and L. D'Alpaos, "Modeling sand-mud transport induced by tidal currents and wind waves in shallow microtidal basins: application to the Venice Lagoon (Italy)," *Estuarine, Coastal and Shelf Science*, vol. 102, pp. 105–115, 2012.
- [117] D. Roelvink, J. A. Roelvink, and A. Reniers, *A Guide to Modeling Coastal Morphology*, Vol. 12, World Scientific, Singapore, 2011.
- [118] L. O. Amoudry and A. J. Souza, "Deterministic coastal morphological and sediment transport modelling: a review and discussion," *Reviews of Geophysics*, vol. 49, no. 2, 2011.
- [119] M. van der Wegen and B. E. Jaffe, "Towards a probabilistic assessment of process-based, morphodynamic models," *Coastal Engineering*, vol. 75, pp. 52–63, 2013.
- [120] K. D. Splinter, I. L. Turner, and M. A. Davidson, "How much data is enough? The importance of morphological sampling interval and duration for calibration of empirical shoreline models," *Coastal Engineering*, vol. 77, pp. 14–27, 2013.
- [121] J. Sutherland, A. H. Peet, and R. L. Soulsby, "Evaluating the performance of morphological models," *Coastal Engineering*, vol. 51, no. 8-9, pp. 917–939, 2004.
- [122] G. Lesser, J. A. Roelvink, J. A. T. M. Van Kester, and G. S. Stelling, "Development and validation of a three-dimensional morphological model," *Coastal Engineering*, vol. 51, no. 8-9, pp. 883–915, 2004.
- [123] L. C. van Rijn, *Principles of Sediment Transport in Rivers, Estuaries and Coastal Seas*, Vol. 1006, Aqua Publications, Amsterdam, Netherlands, 1993.
- [124] J. Holden, *An Introduction to Physical Geography and the Environment*, Pearson Education, London, UK, 2005.
- [125] R. Ranasinghe, C. Swinkels, A. Luijendijk et al., "Morphodynamic upscaling with the MORFAC approach: dependencies and sensitivities," *Coastal Engineering*, vol. 58, no. 8, pp. 806–811, 2011.

- [126] A. A. Németh, S. J. M. H. Hulscher, and R. M. J. van Damme, "Simulating offshore sand waves," *Coastal Engineering*, vol. 53, no. 2-3, pp. 265–275, 2006.
- [127] M. Knaapen, I. H. Townend, K. Rossington, C. A. Fletcher, and J. Spearman, "Modelling the dynamics of intertidal mudflat and salt marshes within estuaries," *Environmentalist*, vol. 84, pp. 12–15, 2009.
- [128] D. Roelvink, A. Reniers, A. van Dongeren, J. de Vries, R. McCall, and J. Lescinski, "Modeling storm impacts on beaches, dunes and barrier islands," *Coastal Engineering*, vol. 56, no. 11-12, pp. 1133–1152, 2009.
- [129] J. A. Roelvink, A. Reniers, A. van Dongeren, J. de Vries, J. Lescinski, and R. McCall, *XBeach Model Description and Manual*, Deltares/TU Delft, Netherlands, 2009.
- [130] J. A. Roelvink, "Coastal morphodynamic evolution techniques," *Coastal Engineering*, vol. 53, no. 2-3, pp. 277–287, 2006.
- [131] G. Lesser, *An Approach to Medium-Term Coastal Morphological Modeling*, Ph.D. thesis, Delft University of Technology, Delft, Netherlands, 2009, ISBN 978-0-415-55668-2.
- [132] D. M. P. K. Dissanayake, R. Ranasinghe, and J. A. Roelvink, "Effect of sea level rise on tidal inlet evolution," *Journal of Coastal Research*, vol. 56, no. 2, pp. 942–946, 2009.
- [133] T. R. Scott and D. C. Mason, "Data assimilation for a coastal area morphodynamic model: Morecambe Bay," *Coastal Engineering*, vol. 54, no. 2, pp. 91–109, 2007.
- [134] P. J. Smith, G. D. Thornhill, S. L. Dance, A. S. Lawless, D. C. Mason, and N. K. Nichols, "Data assimilation for state and parameter estimation: application to morphodynamic modelling," *Quarterly Journal of the Royal Meteorological Society*, vol. 139, no. 671, pp. 314–327, 2013.
- [135] K. Chu, C. Winter, D. Hebbeln, and M. Schulz, "Improvement of morphodynamic modeling of tidal channel migration by nudging," *Coastal Engineering*, vol. 77, pp. 1–13, 2013.
- [136] N. Margvelashvili, J. Andrewartha, M. Herzfeld, B. J. Robson, and V. E. Brando, "Satellite data assimilation and estimation of a 3D coastal sediment transport model using error-subspace emulators," *Environmental Modelling & Software*, vol. 40, pp. 191–201, 2013.
- [137] T. Mayo, T. Butler, C. Dawson, and I. Hoteit, "Data assimilation within the Advanced Circulation (ADCIRC) modeling framework for the estimation of Manning's friction coefficient," *Ocean Modelling*, vol. 76, pp. 43–58, 2014.
- [138] G. Y. El Serafy, M. A. Eleveld, M. Blaas, T. van Kessel, S. G. Aguilar, and H. J. van der Woerd, "Improving the description of the suspended particulate matter concentrations in the southern North Sea through assimilating remotely sensed data," *Ocean Science Journal*, vol. 46, no. 3, pp. 179–204, 2011.
- [139] N. Margvelashvili and E. Campbell, "Sequential data assimilation in fine resolution models using error subspace emulators: theory and preliminary evaluation," *Journal of Marine Systems*, vol. 90, no. 1, pp. 13–22, 2012.
- [140] W. S. Parker, "Ensemble modeling, uncertainty and robust predictions," *Wiley Interdisciplinary Reviews: Climate Change*, vol. 4, no. 3, pp. 213–223, 2013.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

