A ROBUST METHOD OF EVALUATING COASTAL NUMERICAL MODELS: THE GLUE APPROACH

J Simmons[1], Lucy A. Marshall[1], Ian L. Turner[1], Kristen D. Splinter[1], Ron J. Cox[1], Mitchell D. Harley[1], David J. Hanslow[2] and Michael A. Kinsela[2]

[1] School of Civil and Environmental Engineering, UNSW Australia, Sydney, NSW

[2]Office of Environment and Heritage, NSW

Abstract

Sophisticated numerical models play an important role in forecasting beach erosion at high risk sites along NSW coastlines. These models contain free parameters that require calibration to available field data and little guidance (beyond the adoption of the default values provided) is presently available to inform the selection of best-fit parameter values. In practice, the means of calibrating erosion models through the optimisation of these parameters often lacks a sufficient consideration of parameter compensation for model error, the impacts of parameter interdependence and parameter-induced model uncertainty. The Generalised Likelihood Uncertainty Estimation (GLUE) method has been previously employed in the field of hydrology and has proven to be a conceptually simple and efficient method to evaluate model sensitivity to parameters, optimum calibrated parameter values and parameter uncertainty. This paper describes the application of the GLUE method to the XBeach storm erosion model, using data from a site in Italy where the XBeach model has been previously applied without such a rigorous calibration method. The results presented demonstrate the more generic effectiveness of GLUE in enhancing the performance of coastal numerical models. The sensitivity of XBeach to each trailed free parameter is determined in a rigorous and transparent manner, and parameter-induced uncertainty bounds are obtained. This enables the modeller to better quantify model skill in predicting observed and potential future erosion.

Introduction

A number of sites (‘hotspots’) have been identified along the NSW coastline which are susceptible to coastal erosion events and the ability to forecast the potential impact of storm events at these sites would greatly improve the effectiveness coastal protection and emergency response (NSW Office of Environment and Heritage, 2003). To assist coastal managers in the identification of marine storm risk, a number of international projects (Ciavola et al., 2011; Stockdon et al., 2012) have incorporated coastal erosion models into on-line tools which provide real-time beach erosion forecasts. These forecasts must be accompanied by information on prediction uncertainty to allow decision makers to act with confidence, especially when rapid forecasting requires the use of imperfect input data (Plant and Holland, 2011; Kinsela and Hanslow, 2013). To give accurate representations of uncertainty bounds, modellers must be able to reliably calibrate models, recognise model limitations and identify areas in which the physics of the real-world system are not adequately resolved.

Several studies have attempted to examine the sources of model uncertainty present in coastal engineering forecasts. Errors propagating through morphological response forecasts have been examined (Baart et al., 2011), with a focus on quantifying uncertainty in input data. A large body of work to-date on the quantification of uncertainty bounds for numerical modelling used in coastal engineering has focused on providing probabilistic estimates of storm erosion (Callaghan et al., 2008; Pender and Karunarathna, 2013). These studies often incorporate Monte Carlo sampling of storm characteristics combined with bootstrapping methods to produce estimates of storm erosion demand spanning a range of time scales (Callaghan et al., 2013; Li et al., 2014). Importantly, such an approach assumes that uncertainties stemming from the model itself are effectively negligible, once the initial calibration process has been completed (Li et al., 2013).

It is important in uncertainty estimation to be aware of model structural errors and limit these to ensure that the model is performing optimally. As such, calibration of model free parameters is vital before quantifying model performance. In reality, this crucial calibration process is often approached in a somewhat ‘ad hoc’ fashion, often due to the complex and non-linear interaction of parameters. Algorithms have been applied to simpler models with fewer free parameters, searching through the parameter space to find optimal parameter values by minimizing discrepancies between model results and observational data (Plant, 2004; Ruessink et al., 2007; Dubarbier et al., 2015).

Models like XBeach (eXtreme Beach behaviour model) have tended to be analysed in a less rigorous way due to the large number of free model parameters and longer run times. XBeach, a process-based model, is the current `state-of-the-art' model used to predict changes in coastal morphology arising due to storms (Roelvink et al., 2009). Most documented calibrations of XBeach have used modeller experience and one-at-a-time variation to determine sensitive parameters, and have varied these selected parameters by only a few values in order to determine the ‘optimal’ parameter set using a skill measure (Roelvink et al., 2009; Harley et al., 2011; Splinter and Palmsten, 2012; Callaghan et al., 2013; Pender and Karunarathna, 2013; Stockdon et al., 2014). A brief examination of these 6 example studies reveals that in each a different parameter subset was found to be sensitive. Few studies (Vousdoukas et al., 2012) have rigorously trailed a broad range of XBeach parameter combinations and values, and no generalised techniques have been applied to examine the impact of model uncertainty due to parameter selection.

The Generalised Likelihood Uncertainty Estimation (GLUE) method (formally presented in (Beven and Binley, 1992)) has the potential to address a number of the shortcomings identified above. It is a Monte Carlo based method that provides parameter sensitivity information, reliable parameter values and uncertainty estimation. The GLUE method is based upon the concept of ‘equifinality’, the idea that multiple parameter combinations may produce model runs equally skilful as estimators of observed morphological change in the system (Freer et al., 1996; Beven, 2006). This can be due to the complex and often non-linear interaction between model parameters (Yates et al., 2009), over-parameterization without sufficient observational data to inform parameter selection (Beven, 2006) and, errors in the observational data and model structure. These errors can be termed ‘epistemic’, referring to a lack of knowledge of the system within the model (Efstratiadis and Koutsoyiannis, 2010; Beven and Binley, 2014). For instance, the complex three-dimensional morphology may not be captured in the two-dimensional beach profile surveys, the observational data may not be collected immediately before/after the event allowing sediment transport to occur outside of the model simulation time, or the model may be formulated in such a way that overly simplifies the physics in the system (Dubarbier et al., 2015). Recognizing this, the modeller’s focus is then shifted to finding the most reliable parameter combination across studied sites and storms, whilst also assessing the validity of the model in truly representing the physics responsible for the observed change (Candela et al., 2005; Beven and Binley, 2014).

While the GLUE method has been employed extensively in the field of hydrology (Beven and Freer, 2001; Brazier et al., 2000; Candela et al., 2005; Freer et al., 1996; He et al., 2010; Jin et al., 2010), its application to coastal numerical model studies has so far been limited (Ruessink, 2005; Ruessink, 2006).

The purpose of this work is to present a suggested methodology for the application of the GLUE method to deterministic coastal numerical models, allowing coastal engineers to rigorously calibrate models and identify the sources and magnitudes of errors in the modelling process. The first section outlines this methodology and the second section presents an application of the GLUE method to a storm event using the model XBeach.

GLUE Methodology

Parameter ranges and sampling

A basic GLUE analysis begins with the assumption that all combinations of parameter values have an equal likelihood of producing the most reliable predictions, given a set of observations (Beven and Binley, 1992). Uniform Monte Carlo sampling of the parameter space is then performed and the model run to determine the skill of each individual parameter set. Values are subsequently assigned reflecting the likelihood of each parameter set as the best model realization simulating the observed event. These likelihood values assigned to each sampled parameter value are then used to form parameter posterior distributions which, despite being linked to the entire parameter set and therefore incorporating some noise (Freer et al., 1996; Candela et al., 2005), show parameter values generally resulting in skilful model realizations.

The first step in applying the GLUE methodology is to define the model parameters of interest. Given that the number of Monte Carlo runs required to conduct a GLUE analysis is dependent on the number of parameters included, careful attention must be paid in ensuring that appropriate tuning parameters are chosen. It is appropriate at this initial stage to combine the experience of the modeller and simple one-at-a-time parameter sensitivity testing (Morris, 1991). Alternatively, a GLUE analysis with a restricted number of Monte Carlo runs can also be undertaken to estimate the most sensitive parameters.

It is important that a wide range of values are sampled for each selected parameter in order to adequately define the parameter posterior distributions. Where modeller knowledge exists of the physical constraints or likely values a parameter should have, the range should be restricted accordingly. For example, the breaking parameter (gamma) in XBeach should be within the physically reasonable range of 0.3 to 0.6 (Short Ed., 1999). While these estimates can be used to narrow the range significantly, parameters must still be treated as free in order to investigate model performance. Model parameters can compensate for the presence of epistemic errors and therefore cannot be regarded as the equivalent of their physical counterparts (Ruessink et al., 2007; Hsu et al., 2006; Beven, 2006). Due to this effect, allowing parameters to deviate from ‘default’ and measured values provides the modeller information regarding processes in the model which may be missing or describe the system in an incomplete manner.

Likelihood measure

The likelihood measure in the GLUE method serves to rank the individual model runs as potential simulators of the system and is typically based on a measure of skill. A key advantage of the GLUE method lies in the fact that it does not require a formal and detailed knowledge of the sources and magnitudes of errors present in the model. The errors are treated implicitly in GLUE analyses and in this way we expect that if a model consistently underestimates beach erosion for observed data, it will continue to display similar errors during forecasting (Beven and Binley, 2014). Therefore the likelihood measure can be informal and simple, the only requirements being that it must have a value of zero when the model runs are deemed 'non-behavioural' and increase monotonically as the skill of the model in replicating the observed data increases (Beven and Binley, 1992).

Determining if a model run is 'non-behavioural' requires the modeller to impose a behavioural threshold. Parameter sets which do not meet the behavioural threshold criteria are deemed to have no skill in modelling the observed beach response and are therefore 'non-behavioural'. The criteria chosen will change depending on the purpose intended for the model and this is left to the modeller's discretion.

Within the context of coastal morphological modelling, a Brier Skill Score (BSS) is a suitable definition of model skill for GLUE and will be used here. The BSS compares model performance to a baseline profile (most commonly the initial profile) with a value of 1 representing perfect agreement of the model predictions with observational data and a value of 0 reflecting the performance of the baseline profile (Sutherland et al., 2004). It can be presented as:

$$BSS = 1 - \frac{MSE(m,o)}{MSE(b,o)} = 1 - \frac{\sum\left(\left|z_o - z_m\right|\right)^2}{\sum\left(\left|z_o - z_b\right|\right)^2} \tag{1}$$

where z_0 is a series of bed elevations observed post-storm, z_m the final modelled bed levels and z_b the selected baseline series of bed elevations. The BSS is particularly useful as meaningful thresholds have already been established (Sutherland et al., 2004). For example, for initially assessing parameter posterior distributions a threshold of $BSS > 0$ may be used, indicating that the model has provided a better estimate of post-storm beach morphology than the baseline (i.e. initial) bed elevation. However, if the modeller wished to keep only 'good' estimates, a threshold of $BSS > 0.2$ could be used, following the guidance provided by Sutherland et al. (2004). If a model run does not achieve a BSS greater than the pre-determined threshold value, the likelihood for that run is set to 0.

The likelihood measure (rescaled to sum to 1) is then simply defined as:

$$L_{BSS} = \frac{BSS_i}{\sum_{i=1}^{n} BSS_i} \tag{2}$$

where n is the total number of behavioural model runs and the likelihood is calculated from the skill of each individual behavioural model run (BSS_i). This likelihood measure applies provided the BSS threshold is 0 or greater.

Many more likelihood measures with much greater complexity can be found in the literature (Freer et al., 1996; Beven and Freer, 2001; He et al., 2010), however the simplistic measures detailed above are able to give adequate definition to the parameter posterior distributions and therefore provide reliable measures with which to differentiate between the skill of each parameter set.

Application of GLUE

Data and model setup

Study site

This section details an example application of the GLUE methodology using data from a site located along the Emilia-Romagna coastline (Northern Italy) collected before and after the 2012 'Halloween' storm. The Lido di Classe site is situated on the Adriatic Sea with a semidiurnal and micro-tidal regime (neap tidal range of ±0.15m, spring tidal range of ±0.4m), and low wave energy conditions dominating (Harley et al., 2015). When the area is impacted by SE ('Sirocco') winds, storm surge elevation can be twice that of the maximum tide (Armaroli et al., 2012). The storm surge risk coupled with the relatively low-lying dune crests (with elevations at the site ranging between 2.1m and 3.9m above mean sea level), leaves the dunes at this site vulnerable to overtopping. For the purpose of illustration, the GLUE method is applied at one profile line located at the Northern end of the site (Profile 'classe02' (Harley et al., 2015)) which was surveyed every two months using RTK-GPS.

Table 1. A brief description of the trialled XBeach parameters and the sampled ranges for the 'Halloween' dataset.

Parameter	Description	Default	Sampled Range
eps	threshold depth for differentiating between wet and dry cells	0.005	0.001-0.1
facua	the degree to which wave skewness and asymmetry influence the direction of sediment transport	0.1	0-1
gamma	the breaker index in the wave dissipation model	0.55	0.4-0.9
gammax	the maximum allowed wave height to water depth ratio	2	0.4-5
smax	maximum Shields parameter value before sheet flow conditions occur	-1	-1-3
wetslp	maximum beach slope of wet cells before avalanching occurs	0.3	0.1-1

Storm event

On 31 October 2012, a storm event drove SE winds along the Adriatic Sea, causing the peak water level to rise to 1.16m above mean sea level at the site (measured at 23:30 GMT). Significant wave height peaked at 2.43m at 3:00 GMT on 1 November (measured by a wave buoy in 10m water depth), with storm conditions subsiding within 24 hours (Harley et al., 2015). This combination of surge and relatively large waves resulted in the event being classified as a 1-in-20 to 1-in-50 year event and caused significant damage as widespread erosion and inundation occurred (Harley et al., 2015).

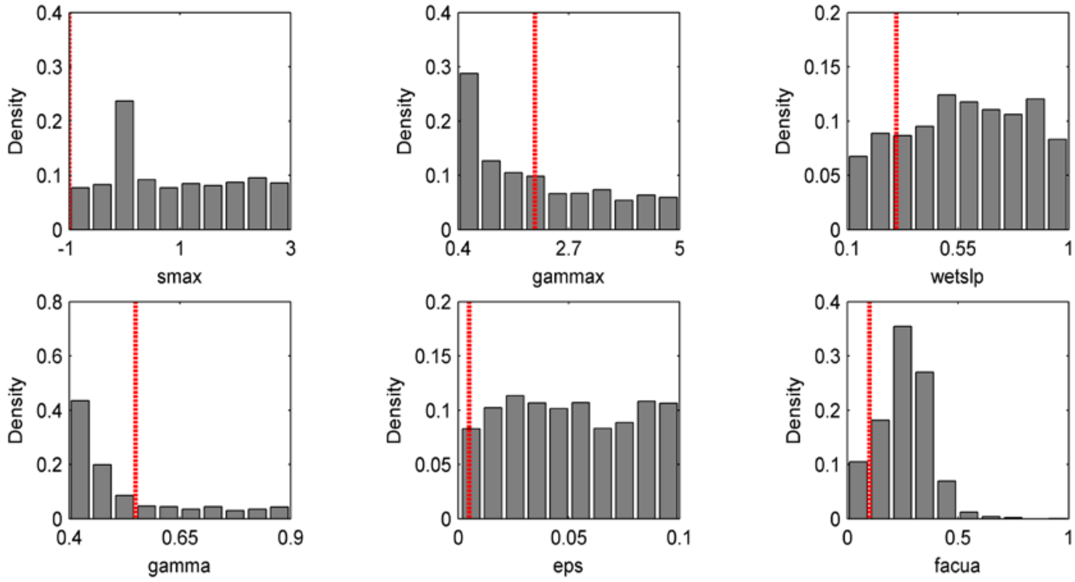

Figure 1. Weighted histograms showing the parameter posterior distributions across the sampled parameter ranges. The red dashed lines represent the XBeach default values for each parameter. The values of the higher likelihood bins of *gamma*, *gammax* and *facua* compared to the default XBeach values suggests erosion in the model must be minimised to ensure good model skill for this observational data.

In this study, XBeach was run in one-dimensional mode 15,000 times (to ensure convergence of the parameter posterior distributions) at the selected profile using the 'Halloween' storm dataset. A behavioural threshold of $BSS > 0$ has been selected for this study. Monte Carlo sampling of six free parameters (described in Table 1) was conducted in accordance with the GLUE method detailed above and these were chosen to be parameters deemed sensitive in the previous modelling study by Harley et al. (2015). The sampled ranges for each of the parameters were taken directly from the XBeach user's manual (www.xbeach.org).

Results

Parameter optimisation

The GLUE method provides a simple tool for choosing the most reliable of the trialled parameter sets, as the modeller can select the set with the highest likelihood value. In this study, the highest likelihood model realisation achieved a BSS (above mean sea level) of 0.906. In addition to this, it can be extremely useful to visualise the individual parameter posterior distributions by plotting a Probability Density Function (PDF) in order to find regions of the parameter space with high likelihood. Ideally the modeller should be able to see the full development and subsidence of the likelihood peak(s) within the sampled range and if not, the selected range should be re-evaluated to ensure that it is appropriate and incorporates all physically reasonable values.

Figure 1 displays the posterior distributions of the six trialled XBeach parameters for the Italy data. Although most parameters in XBeach represent measurable physical parameters, the GLUE method treats these as free and varies these in a random way in order to provide insight into the way the model parameters compensate for epistemic errors. The data from the sensitive parameters (identified in the next section) in Figure 1 show that the best results for this site (indicated by bins with high likelihood) are obtained when the parameters are set to values which minimise erosion in the model.

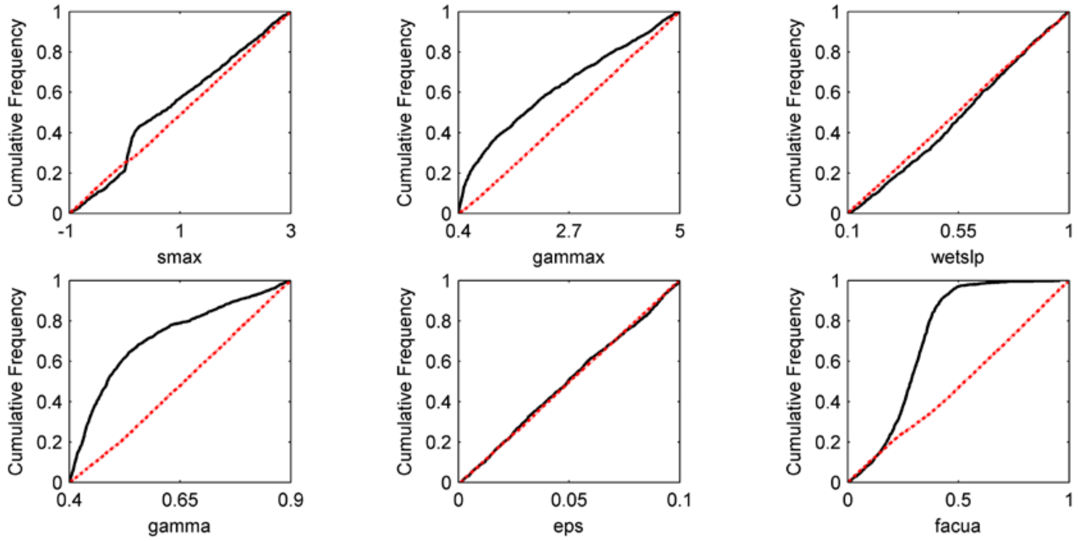

Figure 2. Generalised Sensitivity Analysis (GSA) plots displaying the relative sensitivity of XBeach to the 6 trialled parameters at Profile 'classe02'. A large difference between the behavioural (black) and non-behavioural (red dashed) curves indicates model sensitivity to a given parameter, while similarity indicates insensitivity.

Both gamma and gammax show higher likelihood at lower values (compared to the default values of 0.55 and 2 respectively) indicating that the model performs better when wave breaking is induced earlier, increasing energy dissipation before the wave reaches the shoreline. Higher likelihood values occur when the parameter facua is higher than the default value of 0.1, leading to a greater proportion of onshore sediment transport.

These results concur with the findings of (Harley et al., 2015), which reports a one-at-a-time sensitivity analysis of the same event and found that XBeach significantly overestimated sub-aerial beach erosion with default parameter settings . Similarly, when calibrating the model, (Harley et al., 2015) found the 'optimal' parameter combination to include lower values of gamma and gammax and higher values of facua compared to XBeach defaults.

Sensitivity testing

The GLUE method extends upon the Generalised Sensitivity Analysis (GSA) method developed by Spear and Hornberger (1980), providing a systematic way with which to rank the relative sensitivity of the model to trailed parameters. Using the GSA method, the cumulative frequency distributions of both the behavioural ($BSS > 0$) and non-behavioural ($BSS \leq 0$) parameter values are plotted on the same axes. A large difference between these two distributions is indicative of model sensitivity to that parameter (as the ability of the model to exceed the behavioural threshold shows a clear dependence upon the parameter), whereas a small difference indicates insensitivity (Beven and Binley, 1992; Freer et al., 1996; Beven and Freer, 2001). In this way, the modeller can rank the sensitivity of each trailed parameter using a simple visual method.

Figure 2 displays the insensitivity of XBeach to the parameters eps and wetslp when applied to the 'Halloween' storm dataset. The sensitivity of the model to the remaining parameters can be visually ranked from highest to lowest as facua, gamma, gammax and smax. This suggests that eps and wetslp could be discarded from future analyses

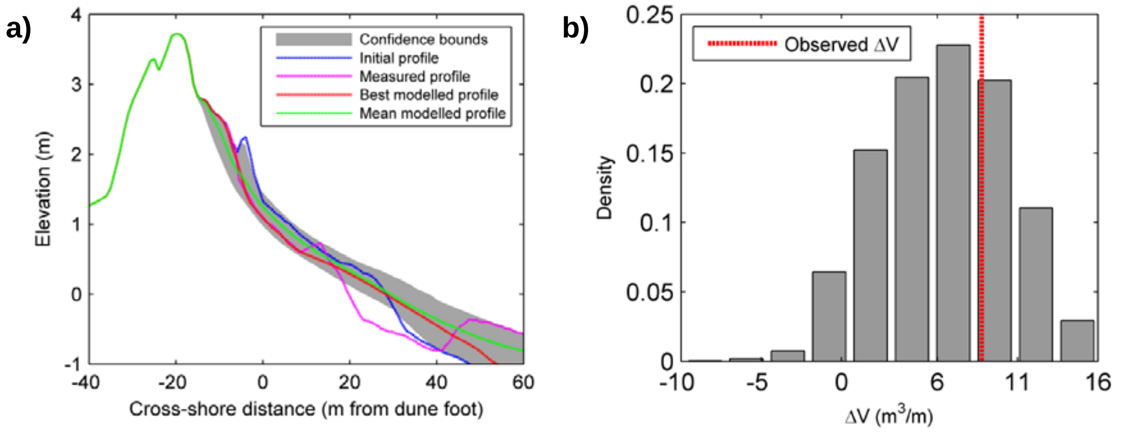

Figure 3. a) Uncertainty bounds (90% confidence interval) showing the range of behavioural model predictions with a threshold of BSS>0. b) Likelihood weighted distribution of the modelled behavioural erosion volumes (ΔV) above mean sea level.

(or as an iterative step in the GLUE process) in order to gain further detailed information about the other four parameter posterior distributions.

Uncertainty analysis

The GLUE method enables the modeller to assess the total model and parameter uncertainty, and compare performance between models. By combining the final bed elevations from all of the behavioural model realizations, upper and lower bounds of predicted bed elevation can be derived for each cross-shore position. In this way, uncertainty bounds can be plotted (as in Figure 3a) which represent errors stemming from model formulation, the selection of free parameter values and the collection of observational data (Brazier et al., 2000).

To give another indication of uncertainty, in Figure 3b the modelled behavioural erosion volumes (change in volume above mean sea level) have been plotted against their likelihood values. This plot shows that within the behavioural model runs, there is a slight bias to under-predict erosion at the site. The 90% confidence interval shows that XBeach expected erosion of between -1 and 12 (m3/m), however the higher likelihood model realizations tend to estimate erosion very close to the observed value of 8.8 m3/m.

The width of the prediction bounds will be closely linked to the choice of likelihood measure and behavioural threshold (Beven and Binley, 2014). For instance if the modeller were to choose a BSS behavioural threshold of 0.2 as opposed to 0, fewer model runs would be deemed behavioural and therefore the bounds would be narrower.

Rather than presenting an absolute uncertainty, the GLUE method is designed to produce uncertainty estimates that can be adapted to individual models and therefore allows the modeller to account for errors in the model structure and input data (Beven and Binley, 1992). For instance, it may be unreasonable to expect a one-dimensional model to perfectly predict the evolution of complex three-dimensional morphology and therefore the threshold must be lowered to a value deemed appropriate by the modeller so that enough behavioural runs are produced to allow study of the range of predicted final bed elevations. If the modeller is reasonable, explicit and consistent in the definition of the behavioural threshold, the GLUE method allows comparisons to be made about the certainty with which XBeach and other models make predictions at different sites (i.e. the uncertainty of XBeach models with $BSS > 0$ can be compared at two different profiles in order to gauge the ability of the model as a predictor of storm erosion at both).

If the uncertainty bounds do not contain the measured profile or an insufficient number of runs are regarded as behavioural, a clear indication is given to the modeller regarding the validity of the model structure (Beven and Binley, 2014). These results are not in any way disguised by an error model, allowing quick identification of problems in the model that need to be addressed in order for it to adequately simulate the system.

Summary

In order to apply deterministic coastal numerical models to the prediction of coastal change, it is important to appropriately optimise these whilst understanding the inherent limitations and uncertainties present in the model. Rather than providing an objective calibration method to search the parameter space for ‘optimal’ parameter values, the GLUE method provides a rigorous tool guiding the modeller through the calibration process whilst also identifying model deficiencies and providing uncertainty bounds for model predictions. This paper provides an example of the GLUE methodology used in a coastal engineering application. The method was applied to the model XBeach with a simple likelihood measure based on the BSS. Sensitive parameters and higher likelihood regions of the parameter space were identified, allowing insight into the compensation required in the model to limit the influence of epistemic errors for this dataset. Uncertainty bounds were provided using a behavioural threshold of $BSS > 0$, showing that the model provided good predictions of observed erosion above mean sea level.

The performance of GLUE when applied to a coastal numerical model suggests that future work could extend the method to multiple sites in order to examine the universality of sensitive parameters and ‘optimal’ parameter values in XBeach.

Acknowledgements

This study is supported by the Coastal Processes and Responses Node of the OEH Climate Change Research Hub and the Australian Research Council (DP150101339). Wave and tide data for the Halloween storm were kindly provided by ARPA-SIMC (Emilia-Romagna) and ISPRA. The first author is supported by an Australian Postgraduate Award.

References

Armaroli, C., Ciavola, P., Perini, L., Calabrese, L., Lorito, S., Valentini, A. and Masina, M., 2012. Critical storm thresholds for significant morphological changes and damage along the Emilia-Romagna coastline, Italy. *Geomorphology*, 143-144, pp.34–51.

Baart, F., van Gelder, P.H.A.J.M. and van Koningsveld, M., 2011. Confidence in real-time forecasting of morphological storm impacts. *Journal of Coastal Research*, (64), pp.1835–1839.

Beven, K., 2006. A manifesto for the equifinality thesis. *Journal of Hydrology*, 320(1-2), pp.18–36.

Beven, K. and Binley, A., 2014. GLUE: 20 years on. *Hydrological Processes*, 28(24), pp.5897–5918.

Beven, K. and Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, 6(May 1991), pp.279–298.

Beven, K. and Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249(1-4), pp.11–29.

Brazier, R.E., Beven, K.J., Freer, J. and Rowan, J.S., 2000. Equifinality and uncertainty in physically based soil erosion models: Application of the glue methodology to wepp-the water erosion prediction project-for sites in the UK and USA. *Earth Surface Processes and Landforms*, 25(8), pp.825–845.

Callaghan, D.P., Nielsen, P., Short, A. and Ranasinghe, R., 2008. Statistical simulation of wave climate and extreme beach erosion. *Coastal Engineering*, 55(5), pp.375–390.

Callaghan, D.P., Ranasinghe, R. and Roelvink, D., 2013. Probabilistic estimation of storm erosion using analytical, semi-empirical, and process based storm erosion models. *Coastal Engineering*, 82, pp.64–75.

Candela, A., Noto, L. V. and Aronica, G., 2005. Influence of surface roughness in hydrological response of semiarid catchments. *Journal of Hydrology*, 313(3-4), pp.119–131.

Ciavola, P., Ferreira, O., Haerens, P., Van Koningsveld, M., Armaroli, C. and Lequeux, Q., 2011. Storm impacts along European coastlines. Part 1: The joint effort of the MICORE and ConHaz Projects. *Environmental Science & Policy*, 14(7), pp.912–923.

Dubarbier, B., Castelle, B., Marieu, V. and Ruessink, G., 2015. Process-based modeling of cross-shore sandbar behavior. *Coastal Engineering*, 95, pp.35–50.

Efstratiadis, A. and Koutsoyiannis, D., 2010. One decade of multi-objective calibration approaches in hydrological modelling: a review. *Hydrological Sciences Journal*, 55(1), pp.58–78.

Freer, J., Beven, K. and Ambroise, B., 1996. Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resources Research*, 32(7), pp.2161–2173.

Harley, M., Armaroli, C. and Ciavola, P., 2011. Evaluation of XBeach predictions for a real-time warning system in Emilia-Romagna, Northern Italy. *Journal of Coastal Research*, (64), pp.1861–1865.

Harley, M.D., Valentini, A., Armaroli, C., Perini, L., Calabrese, L. and Ciavola, P., 2015. Can an early warning system help minimize the impacts of coastal storms? A case study of the 2012 Halloween storm, Northern Italy. *Natural Hazards and Earth System Sciences Discussions*, 3(5), pp.3409–3448.

He, J., Jones, J.W., Graham, W.D. and Dukes, M.D., 2010. Influence of likelihood function choice for estimating crop model parameters using the generalized likelihood uncertainty estimation method. *Agricultural Systems*, 103(5), pp.256–264.

Hsu, T.J., Elgar, S. and Guza, R.T., 2006. Wave-induced sediment transport and onshore sandbar migration. *Coastal Engineering*, 53(10), pp.817–824.

Jin, X., Xu, C.Y., Zhang, Q. and Singh, V.P., 2010. Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. *Journal of Hydrology*, 383(3-4), pp.147–155.

Kinsela, M. and Hanslow, D., 2013. Coastal Erosion Risk Assessment in New South Wales: Limitations and Future Directions. In *Proceedings of the NSW Coastal Conference, Port Macquarie*.

Li, F., Van Gelder, P.H. a J.M., Callaghan, D.P., Jongejan, R.B., Den Heijer, C. and Ranasinghe, R., 2013. Probabilistic modeling of wave climate and predicting dune erosion. *Journal of Coastal Research*, (65), pp.760–765.

Li, F., Gelder, P.H. a J.M. Van, Vrijling, J.K., Callaghan, D.P., Jongejan, R.B. and Ranasinghe, R., 2014. Probabilistic estimation of coastal dune erosion and recession by statistical simulation of storm events. *Applied Ocean Research*, 47, pp.53–62.

Morris, M.D., 1991. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, 33(2), pp.161–174.

NSW Office of Environment and Heritage, 2003. *Code of Practice under the Coastal Protection Act 1979*,

Pender, D. and Karunarathna, H., 2013. A statistical-process based approach for modelling beach profile variability. *Coastal Engineering*, 81, pp.19–29.

Plant, N.G., 2004. Prediction skill of nearshore profile evolution models. *Journal of Geophysical Research*, 109(C1), pp.1–19.

Plant, N.G. and Holland, K.T., 2011. Prediction and assimilation of surf-zone processes using a Bayesian network. Part I: Forward models. *Coastal Engineering*, 58(1), pp.119–130.

Roelvink, D., Reniers, A., van Dongeren, A., van Thiel de Vries, J., McCall, R. and Lescinski, J., 2009. Modelling storm impacts on beaches, dunes and barrier islands. *Coastal Engineering*, 56(11-12), pp.1133–1152.

Ruessink, B.G., 2006. A Bayesian estimation of parameter-induced uncertainty in a nearshore alongshore current model. *Journal of Hydroinformatics*, 8(1), pp.37–49.

Ruessink, B.G., 2005. Predictive uncertainty of a nearshore bed evolution model. *Continental Shelf Research*, 25(9), pp.1053–1069.

Ruessink, B.G., Kuriyama, Y., Reniers, a. J.H.M., Roelvink, J. a. and Walstra, D.J.R., 2007. Modeling cross-shore sandbar behavior on the timescale of weeks. *Journal of Geophysical Research: Earth Surface*, 112(3), pp.1–15.

Short Ed., A.D., 1999. *Handbook of Beach and Shoreface Morphodynamics* A. D. Short, ed., London: John Wiley & Sons Ltd.

Spear, R.C. and Hornberger, G.M., 1980. Eutrophication in Peel Inlet. II. Identification of critical uncertainties via generalized sensitivity analysis. *Water Research*, 14(1), pp.43–49.

Splinter, K.D. and Palmsten, M.L., 2012. Modeling dune response to an East Coast Low. *Marine Geology*, 329-331, pp.46–57.

Stockdon, H.F., Doran, K.J., Thompson, D.M., Sopkin, K.L., Plant, N.G. and Sallenger, A.H., 2012. *National Assessment of Hurricane-Induced Coastal Erosion Hazards : Gulf of Mexico*, Virginia.

Stockdon, H.F., Thompson, D.M., Plant, N.G. and Long, J.W., 2014. Evaluation of wave runup predictions from numerical and parametric models. *Coastal Engineering*, 92, pp.1–11.

Sutherland, J., Peet, A.H. and Soulsby, R.L., 2004. Evaluating the performance of morphological models. *Coastal Engineering*, 51(8-9), pp.917–939.

Vousdoukas, M.I., Ferreira, Ó., Almeida, L.P. and Pacheco, A., 2012. Toward reliable storm-hazard forecasts: XBeach calibration and its potential application in an operational early-warning system. *Ocean Dynamics*, 62(7), pp.1001–1015.

Yates, M.L., Guza, R.T. and O'Reilly, W.C., 2009. Equilibrium shoreline response: Observations and modeling. *Journal of Geophysical Research*, 114(C9), p.C09014.