Ensuring numerical stability of wave propagation by tuning model parameters using genetic algorithms and response surface methods

Riccardo Angelini Rota Roselli[1,5,][*] *Giuliano Vernengo*[2]*, Corrado Altomare*[3,4]*, Stefano Brizzolara*[5]*, Luca Bonfiglio*[6] *and Roberto Guercio*[1]

[1] Dept. of Civil, Construction and Environmental Engineering (DICEA), University of Rome La Sapienza, Via Eudossiana, 18, 00184 Roma, Italy

[2] Dept. of Electric, Electronical and Telecomunication Engineering and Naval Architecture (DITEN), University of Genova, Via Montallegro 1, 16145, Genova, Italy

[3] Flanders Hydraulic Research, Berchemlei 115, 2140 Antwerp, Belgium

[4] Dept. of Civil Engineering, Ghent University, Technologiepark 904, 9052 Ghent, Belgium

[5] Aerospace and Ocean Engineering, Virginia Tech, 460 Old Turner Street, Blacksburg, 24061, VA, USA

[6] MIT Sea Grant, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, USA

email: riccardo.angelinirota@uniroma1.it; giuliano.vernengo@unige.it; corrado.altomare@mow.vlaanderen.be; stebriz@vt.edu; bonfi@mit.edu; roberto.guercio@uniroma1.it

Abstract

The effectiveness of a Metamodel-Embedded Evolution Framework for model parameter identification of a Smoothed Particles Hydrodynamic (SPH) solver, called DualSPHysics, is demonstrated when applied to the generation and propagation of progressive ocean waves. DualSPHysics is an open-source code that provides GP-GPU acceleration, allowing for highly refined simulations. The automatic optimization framework combines the global-convergence capabilities of a Multi-Objective Genetic Algorithm (MOGA) with Response Surface Method (RSM) based on a Kriging approximation. The proposed Metamodel-Embedded Evolutionary framework is used to find the set of SPH model parameters that ensures an accurate reproduction of a 2nd order Stokes wave propagating in a numeric flume tank. In order to demonstrate the consistency of the obtained results, the optimum set of parameters found by the framework is finally used to reproduce other 2nd and 3rd order Stokes waves propagating over the same flume tank.

KEY WORDS

Smoothed Particle Hydrodynamics (SPH), Wave propagation, Non-dominated Sorting Genetic Algorithm (NSGA-II), Response Surface Method (RSM), Parameter Identification

1. Introduction

One of the major complexities in the study of coastal and ocean engineering process consists in the necessity of modeling small scale phenomena, such as boundary layer flows and turbulence,

for physical systems involving large scale problems such as wave run-up, wave breaking, fluid-structure interaction (FSI) and coastal flooding. As is typical in numerical problems, two conflicting tasks need to be tackled: on one hand prediction models must rely on accurate physics description; on the other hand the computational burden of the numerical model needs to suit engineering requirements.

The ability to correctly reproduce a certain wave field is one of the fundamental requirements for prediction tools. Many theories have been formulated over the past years to model different wave types, but the complexity of analytic models restricts their practical utility. In the past 20 years easier accessibility of high performance computing allowed application of Computational Fluid Dynamics (CFD) techniques to coastal engineering problems.

In this respect different numerical methods have been presented for the solution of wave propagation problems. One successful application has been presented by Zijlema et al. (2011), who developed a time-domain, depth-averaged, non-hydrostatic model for free surface waves based on Non-Linear Shallow Water (NLSW) equations for large domain applications. NLSW equations are a simplified form of the Navier-Stokes equations, assuming depth-integrated free-surface flows. This model has been widely used for the prediction of coastal engineering problems, such as wave transformation and propagation over small sloped beaches, wave breaking, wave run-up and wave overtopping of coastal defenses (see for instance Suzuki et al. (2011), Smit et al. (2014) and Suzuki et al. (2017)). Kennedy et al. (2000), Madsen et al. (2002) and Fuhrman and Madsen (2008) presented a Boussinesq model in which non-linearities and wave dispersion are retained in the solution of depth integrated N-S equations.

Limitations of Boussinesq and NLSW in computing vertical flow characteristics and accurately reproducing phenomena such as wave flow through porous structures, have been overcome by

---

1

using fully 3D, unsteady Navier-Stokes (N-S) models for the solution of pressure, velocity and turbulence characteristics. N-S models are applicable to a wide range of coastal structures of complex geometry, both permeable and impermeable. Among N-S techniques, the most popular models are based on an Eulerian flow description, in which fluid motion equations are written considering a specific fixed location in the fluid domain. VOFbreak (Austin and Schlueter, 1982), FLOW-3D (see e.g. Bayon et al., 2016) and IHFOAM (Higuera et al., 2013) are some examples of N-S Eulerian models. The most relevant flaw of these models is represented by time consuming mesh generation processes, often required for complex geometries. Moreover Eulerian approaches present severe technical challenges associated with implementing conservative multi-phase schemes able to capture the non-linearities within rapidly changing geometries. A second family of N-S techniques consists in meshless models in which the wave field is specified through a Lagrangian approach and the tracking of free surface is an intrinsic property of the fluid discretization in unit elements or particles. Among Lagrangian meshless models, Smoothed Particle Hydrodynamics method (SPH) has recently gained significant popularity in the coastal engineering field (Violeau, 2012). In SPH, the fluid is discretized into a set of particles. Each of these particles is a nodal point where physical quantities (e.g. position, velocity, density, pressure) are computed as an interpolation of the values of the neighboring particles.

In this paper we employ a SPH-based model (DualSPHysics, see Crespo et al., 2015) to reproduce different progressive waves on a flat-bottom 2D tank ending with a sloped beach. DualSPHysics is an open-source code based on the Smoothed Particle Hydrodynamics (SPH) method. It has been derived from the SPH formulation implemented in the open-source code SPHysics (Gomez-Gesteira et al. (2012a)). DualSPHysics has been mainly used to study wave transformation and breaking at detailed scale close to the shoreline. Many studies recently demonstrated the validity of this method. Barreiro et al. (2013), Vacondio et al. (2013), Rota et al. (2014), Altomare et al. (2014), Altomare et al. (2015a) and Altomare et al. (2017) are some of the many examples available in literature. The particle formulation of SPH methods is particularly suitable for modeling coastal processes especially in the surf and swash zones characterized by strong non-linearities such as run-up and wave breaking. The computational burden required for large domain simulations becomes unbearable, hence the *GPGPU* acceleration implemented in DualSPHysics, allows for feasible computational requirements as demonstrated by Valdez-Balderas et al. (2013) and Vacondio et al. (2014).

In general, any numerical model requires a preliminary validation study in which numerical predictions are compared to analytic results or experiments. This stage gives the chance to calibrate numerical models by opportunely tuning parameters in order to increase their fidelity in the description of each specific problem. A major complexity in properly tuning SPH model parameters is due to their dependence on the particular problem that needs to be solved. Considering the physical and mathematical meaning of each parameter, this paper proposes an innovative method, based on data mining, for the calibration of the model setting that better suits a specific problem. The problem under investigation is the propagation of a wave train in a tank. In particular the numerical calibration process is applied to minimize the error between the computed and the theoretic 2nd order Stokes wave profile. To achieve this target, a two-step computational framework for parameter calibration has been set up based on a combination of a Multi-Objective Optimization Algorithm and a Response Surface Method (RSM).

As outlined by Simpson et al. (2008) the improvements in computational resources are nowadays used to add complexities to the solution of particular problems. High-fidelity models are created to fill the gap between numerical simulations and physical observations. RSM (also referred to as *metamodels* or *surrogate models*) represent a less computationally expensive emulator of the original function under study. So they play a key role in reducing computational effort when a large number of function evaluations are needed. For a comprehensive review of surrogate-based technique applications in water resources problems and related issues the reader is referred to Razavi et al. (2012) and Brunetti et al. (2017).

The paper is organized as follows: theoretic background on Stokes waves and wavemaker theory is provided in Section 2. 3 describes the formulation of the DualSPHysics solver, highlighting the physical and mathematical meaning of each of the analyzed model parameters. The optimization-based calibration framework is described in Section 4 in all its parts: the Multi-Objective Optimization Algorithm (4.1), the RSM (4.2) and the formulation of the objective function (4.3). Results obtained by the proposed approach are presented and discussed in Section 5.

2. Physical and theoretical backgrounds of wave generation and propagation

The parameter identification is carried out on a second order Stokes wave. This wave theory is able to model a wide range of wave amplitudes $\frac{H}{gT^2}$ and depth ratios $\frac{d}{gT^2}$; it is a widely employed model for ocean engineering problems. Moreover, the non-linear theory at the basis of their mathematical formulation allows for steep waves modeling. The wave profile $\eta(x, t)$ is obtained considering a 2D potential flow assumption:

$$
\begin{aligned}
\eta(x,t) &= a\cos(kx - \omega t) + \\
&+ ka^2 \cdot \frac{\cosh(kd)\,[2 + \cosh(2kd)]}{4\sinh^3(kd)} \cdot \cos\left[2\,(kx - \omega t)\right]
\end{aligned}
\tag{1}
$$

Where d is the water depth, H the wave height and T the wave period. The non linear wave profile is obtained by superimposing two harmonics of frequency ω and 2ω, solutions of the first and second order problems respectively. Non linear Stokes wave theory is generally valid for deep water for $\frac{H}{d} \ll 1$ and $Ur \ll 10$, where Ursell number Ur is defined as follow:

$$
Ur = \frac{HL^2}{d^3}
\tag{2}
$$

In the context of a wave flume, progressive waves are generated using a piston-type wavemaker (see e.g. Dean and Dalrymple,

1991). The ratio between the expected wave height and the imposed stroke S of the wavemaker is theoretically provided by Eq. (3).

$$\frac{H}{S} = \frac{2\left(cosh(2kd) - 1\right)}{\sinh(2kd) + 2kd} \tag{3}$$

3. Smoothed Particle Hydrodynamics for Numerical Modeling of Progressive Waves

SPH method basically develops in two phases: a kernel approximation and a particle approximation. The former involves the representation of a field variable and its derivatives in a continuous integral form by a suitable kernel function; the latter refers to the discretization process of the computational domain that is redefined by an initial distribution of discrete particles. According to this discrete model, field variables on a particle are computed by approximation using the nearest neighbor particles. The contribution of the nearest particles is weighted according to distance between particles. The kernel function W is used to measure this contribution depending on the inter-particle distance that is defined using a smoothing length h. The smoothing length is a characteristic length used to define the area of influence of the kernel and the kernel presents compact support to avoid contributions with other particles beyond that distance.

3.1 Governing Equations

The basic mathematical formulation of the SPH relies on integral interpolation functions. Therefore any function F can be computed by the integral approximation according to Eq. (4).

$$F(r) = \int_\Omega F(r')W(r - r', h)dr' \tag{4}$$

This function F can also be expressed in discrete form based on particles. Thus, the approximation of the function is interpolated at particle a and the summation is performed over all the particles within the region defined by the compact support of the kernel as in Eq. (5).

$$F(r_a) \approx \sum_b F(r_b)W(r_a - r_b, h)\frac{m_b}{\rho_b} \tag{5}$$

where the volume associated to the neighboring particle b is $\frac{m_b}{\rho_b}$, with m and ρ being the mass and the density, respectively. The kernel function W must fulfill several properties (see Monaghan, 1992), such as being positive inside the area of interaction, having compact support, being normalized ($\int_\Omega W\left(x - x'\right)dx' = 1$) and monotonically decreasing with distance. DualSPHysics implements two different kernel functions, namely the cubic spline (Monaghan and Lattanzio, 1985) and the quintic kernel function by Wendland (1995). The former has been used in the present

work and can be expressed as follows:

$$W(r, h) = \frac{7}{4\pi h^2} \begin{cases} \frac{3}{2}q^2 + \frac{3}{4}q^3 & 0 \le q \le 1 \\ \frac{1}{4}(2 - q)^3 & 1 \le q \le 2 \\ 0 & q \ge 2 \end{cases} \tag{6}$$

In DualSPHysics, Navier-Stokes equations are solved considering weakly compressible fluid (Gomez-Gesteira et al. (2012b)). Conservation laws of continuum fluid dynamics, in the form of differential equations, are transformed into their particle forms by employing kernel functions. The momentum equation proposed by Monaghan (1992) has been used to determine the acceleration of a particle (a) as the result of the particle interaction with its neighbors (particles b) as in Eq. (7).

$$\frac{d\mathbf{v}_a}{dt} = -\sum_b m_b \left(\frac{P_b^2}{\rho_b} + \frac{P_a^2}{\rho_a} + \Pi_{ab}\right)\Delta_a W_{ab} + \mathbf{g} \tag{7}$$

Here, $\mathbf{v}$ is the velocity, P the pressure, ρ the density, m the mass, $\mathbf{g} = (0, 0, -9.81)\ ms^{-2}$ the gravitational acceleration and W_{ab} the kernel function that depends on the distance between particles a and b. Π_{ab} is the viscous term according to the artificial viscosity proposed by Monaghan (1992). This term is defined as follows:

$$\Pi = \begin{cases} \frac{\alpha\mu_{ab}\bar{c}_{ab}}{\rho_{ab}} & \mathbf{v}_{ab}\mathbf{r}_{ab} < 0 \\ 0 & \mathbf{v}_{ab}\mathbf{r}_{ab} > 0 \end{cases} \tag{8}$$

Here, $\mathbf{r}_{ab} = \mathbf{r}_a - \mathbf{r}_b$ and $\mathbf{v}_{ab} = \mathbf{v}_a - \mathbf{v}_b$ and $\mathbf{r}_k$ and $\mathbf{v}_k$ are the particle position and velocity, respectively. The mean speed of sound is defined as

$$c_{ab} = 0.5(c_a + c_b) \tag{9}$$

and the coefficient μ_{ab}, as:

$$\mu_{ab} = h\mathbf{v}_{ab} \cdot \frac{\mathbf{r}_{ab}^2}{(\mathbf{r}_{ab}^2 + \eta^2)} \tag{10}$$

where, $\eta^2 = 0.001h^2$ and α is a coefficient regulating wave dissipation if properly tuned. Based on experiments on a wave flume, Altomare et al. (2015b) proposed a reference value of $\alpha = 0.01$ for wave propagation and induced loading onto coastal structures.

The mass of each particle is constant, so that changes in fluid density are computed by solving the conservation of mass or continuity equation in SPH form as stated in Eq. (11):

$$\frac{d\rho_a}{dt} = \sum_b m_b\mathbf{v}_{ab} \cdot \Delta_a W_{ab} \tag{11}$$

According to the weakly compressible approach, pressure is computed from density values of the particle using a stiff equation of state. The compressibility is adjusted in order to artificially lower the the speed of sound. This means that the size of time step taken at any moment can be maintained at a reasonable value. The time step is determined according to a *Courant-Friedrichs-Lewy* (CFL) condition based on the currently calculated speed of sound for all particles. Time step adjustment,

however, restricts the sound speed to be at least ten times faster than the maximum fluid velocity. This implies density variations must be lower than 1% to avoid introducing relevant deviations with respect to an incompressible approach. Following Batchelor (1969), the relationship between pressure and density follows the expression in Eq. (12):

$$P = B\left[\left(\frac{\rho}{\rho_0}\right)^\gamma - 1\right] \tag{12}$$

with $B = \frac{c_0^2 \rho_0}{\gamma}$, $\gamma = 7$ and reference water density $\rho_0 = 1000 \left[kg\ m^{-3}\right]$. The speed of sound c_0 is defined at the reference density as:

$$c_0 = c(\rho_0) = \sqrt{\frac{\partial P}{\partial \rho}|_{\rho_0}} \tag{13}$$

In DualSPHysics c_0 from Eq. (13) is computed to be at least equal to 10 times the maximum velocity in the system. The latter is determined as the wave-front velocity of a dam-break case:

$$c_0 = coeff_{sound}\sqrt{g\ h_{swl}} \tag{14}$$

where h_{swl} is the still water level and $coef_{sound}$ is usually taken in the range [10;30].

The *Symplectic* time integration algorithm (Leimkuhler et al. (1996)) has been used in the present work. This scheme is an explicit second-order scheme with an accuracy in time of $O(\Delta t^2)$ and involves both predictor and corrector stages. During the predictor stage the values of acceleration and density are estimated at the middle of the time step according to the following formulation:

$$\begin{cases} \mathbf{r}_a^{n+\frac{1}{2}} = \mathbf{r}_a^n + \frac{\Delta t}{2}\mathbf{v}_a^n \\ \rho_a^{n+\frac{1}{2}} = \rho_a^n + \frac{\Delta t}{2}\frac{d\rho_a^n}{dt} \end{cases} \tag{15}$$

where the superscript n denotes the time step and $t = n\Delta t$. During the corrector stage, the term $\dfrac{d\mathbf{v}_a^{n+\frac{1}{2}}}{dt}$ is used to calculate the corrected velocity, and therefore the position, of the particles at the end of the time step according to the following Eq. (16):

$$\begin{cases} \mathbf{v}_a^{n+1} = \mathbf{v}_a^{n+\frac{1}{2}} + \frac{\Delta t}{2}\frac{dv_a^{n+\frac{1}{2}}}{dt} \\ \mathbf{r}_a^{n+1} = \mathbf{r}_a^{n+\frac{1}{2}} + \frac{\Delta t}{2}\mathbf{v}_a^{n+1} \end{cases} \tag{16}$$

The corrected value of density $\dfrac{d\rho_a^{n+1}}{dt}$ is finally calculated using the updated values of $\mathbf{v}_a^{n+1}$ and ρ_a^{n+1}. A variable time step is calculated based on the CFL condition, the force terms and the viscous diffusion term. According to Monaghan (1999) the variable time step Δt is computed according to Eq. (17).

$$\begin{cases} \Delta t_f = CFL \cdot \min \Delta t_f, \Delta t_{cv} \\ \Delta t_f = \min_a \sqrt{\frac{h}{|f_a|}} \\ \Delta t_{cv} = \min_a \frac{h}{c_s + \max_b \left|\frac{h\mathbf{v}_{ab}\cdot\mathbf{r}_{ab}}{(r_{ab}^2 + \eta^2)}\right|} \end{cases} \tag{17}$$

where Δt_f is based on the force per unit mass ($|f_a|$), and Δt_{cv} combines the Courant and the viscous time step controls. In DualSphysics, boundaries are described using a discrete set of boundary particles that exert a repulsive force on the fluid particles as they are approached. A dynamic boundary condition where the boundary particles satisfy the same equations as the fluid particles is used (Crespo et al., 2007). However, they do not move according to the forces exerted on them. Instead, they remain fixed (*fixed boundary*) or they move according to some externally imposed movement (e.g. gates, flaps, etc.). Using such a boundary treatment, when a fluid particle approaches a boundary particle, as the relative distance decreases beyond the kernel range, the density of the boundary particles increases giving rise to an increase of the pressure. This results in a repulsive force being exerted on the fluid particle due to the pressure term in the momentum equation.

3.2 Discussion on the selected model parameters for wave propagation

Five model parameters have been selected for the parameter calibration of DualSPHysics: the initial particle inter-space, the artificial viscosity coefficient, the coefficient for the speed of sound, the smoothing length and the *Courant-Friedrichs-Lewy* coefficient. Even though SPH-based models have been already largely applied for engineering purposes (e.g. Rogers et al., 2010, Barreiro et al., 2013, St-Germain et al., 2013, Altomare et al., 2015a), sensitivity analysis on the main model parameters has been carried out manually for a limited range of values. The choice of the ranges to be analyzed, far from being incorrect, was often based on the authors' expertise. To attain satisfactory simulation results, the user's selection of the model parameter can require several attempts depending on the expertise in the particular numerical method and in the specific phenomenon that is the object of study. The selection of proper values for the model parameter might also be case-dependent and can be accompanied with higher or lower degree of accuracy, which will inevitably affect the model results. This process is actually a sort of model manipulation, defined by Chau (2004) as a process that is used to improve modeling results for representation of real phenomena. Therefore, the mathematical and physical meaning of each model parameter needs to and must be analyzed in setting up a numerical simulations. This is of course also valid for DualSPHysics. The proposed automatic optimization based approach is intended to be a useful method to better support the expertise of the users, both allowing exploration of wider ranges of parameter values and accelerating convergence to the most suitable set for a specific problem. In this context, the capability of the SPH method to deal with the propagation of long-crested sea wave in a flat-bottom horizontal tank is studied. Wave propagation is characterized by slow dynamics, with little energy dissipation. This phenomenon differs from the classical application of free-surface SPH to fluid-structure impact flows or violent flows involving breaking waves and interface fragmentation. Rather than implementing new numerical schemes or improving existing ones, the

present work aims to explore the capabilities for wave propagation of DualSPHysics model.

3.2.1 Initial particle interspace, dp

The initial particle interspace, dp, is the grid size that is initially used in SPH to locate the particles inside the domain at $t = t_0$. When the simulation starts, the fluid particles are free to move, however the selected dp determines the total amount of fluid and boundary particles, hence being an expression of the model resolution. Convergence studies are usually carried out to analyze the influence of the model resolution on the results. This has been done also in SPH and DualSPHysics (see for instance Altomare et al., 2015a) however the optimal value of the initial particle relative distance is case-dependent. Experience suggests to use a value of dp that, for wave propagation purposes, would be at least $\frac{1}{10}$ of the wave height H. Several authors, such as Altomare et al. (2015a) and De Padova et al. (2014), show that increasing the spatial resolution (i.e. using smaller values of dp) improves the numerical model accuracy. However, convergence has not yet been completely proved in SPH and the choice of this parameter influences the contribution of artificial viscosity terms in Eq. (7) and Eq. (8) (see De Padova et al., 2014).

3.2.2 Artificial viscosity coefficient, α

Recently De Padova et al. (2014) proposed a systematic analysis focused on the effects of the artificial viscosity on the propagation and breaking of regular waves in a Weakly Compressible SPH (WCSPH) model. Based on their results, they concluded that the value of the artificial viscosity coefficient α can drastically modify the results of the simulation (see Eq. (8)). However, in absence of any other numerical diffusion scheme (e.g. as in Molteni and Colagrossi (2009), Antuono et al. (2010), Marrone et al. (2011)), the pressure field obtained with the WCSPH formulation contains non-physical high frequency noise (see Gotoh et al. (2005) and Meringolo et al. (2017)). This noise is partially damped out by introducing an artificial viscosity such the one described by Monaghan (1992). De Padova et al. (2014) show that using excessive values of viscosity coefficient lead to high diffusivity and wave height under-prediction (particularly at breaking). On the contrary, if the empirical coefficient α falls below a certain threshold, the model becomes unstable and the particles begin to move chaotically, even out of the computational domain. Furthermore, the choice of the optimal value of α should be related to the particular smoothing length h and initial particle resolution dp.

3.2.3 Coefficient for the speed of sound, $coeff_{sound}$

In DualSPHysics the liquid phase is treated as weakly-compressible media (see Eq. (12)-(15)). $coeff_{sound}$ in Eq. (15) is an empirical coefficient usually selected in the range [10;30]. Besides the fact that the use of a WCSPH approach involves high

computational costs because of the needed time-space resolutions, such a method is often affected by acoustic perturbations related to the speed of sound (Meringolo et al., 2017). Even if diffusive correction schemes have been largely introduced in SPH to obtain more reliable pressure results, acoustic perturbations still persist due to the weakly compressible approach. Furthermore, low values of compressibility will lead to particle penetration if density or pressure of boundary particles are too low. Therefore, the selection of the $coeff_{sound}$ not only will influence the model efficiency but also its accuracy, causing possibly unreliable results for engineering applications.

3.2.4 Smoothing length, h

The smoothing length h is computed in DualSPHysics as $h = coefh\sqrt{dx^2 + dy^2 + dz^2}$. In 2D, assuming $dx = dz = dp$ and $dy = 0$, it results in $h = coefh\sqrt{2\,dp^2} = coefh \cdot dp\sqrt{2}$. $coefh$ is an empirical coefficient, typically case-dependent. Based on the authors' experience, the smoothing length might be an even more important parameter than diffusive schemes for reliable wave propagation. The choice for $coefh$ determines the value of the dimensionless smoothing length ratio $\frac{h}{dp}$. According to De Padova et al. (2014) the latter quantity should be $\frac{h}{dp} \geq 1.4$ for cases of regular wave breaking on plane slope. Experience suggests that values ranging from 1.7 up to 2.5 provide more accurate results with low wave decay for wave propagation in large numerical domains. However, this has not been systematically analyzed yet.

3.2.5 *Courant-Friedrichs-Lewy* coefficient, CFL

Using explicit time integration schemes, in DualSPHysics, the time step is calculated based on the *Courant-Friedrichs-Lewy* condition, the forcing term and the viscous diffusion term. The CFL coefficient (see Eq. (17)) is a factor to correct minimum criteria derived by the forcing term and the viscous diffusion term. Usually, $CFL = 0.1$ or 0.2 should be used to keep the mean water level stable.

4. Computational Approach to Model Parameter Identification

The problem of identifying the parameters of the hydrodynamic method is solved by the three-step approach shown in the flowchart of Figure 1. Phase 1 (operations bounded by a red dotted rectangle) involves the use of an automatic optimization algorithm (orange boxes) that, coupled to the SPH solver (blue box), is used to find an initial set of feasible candidate designs. This already includes a preliminary search toward the desired solution. In Phase 2 (highlighted by a blue dotted rectangle) the preliminary set of results are interpolated in order to formulate a surrogate model (red box) that allows for a continuous description of the investigated domain. The same optimization process is conceptually applied again in Phase 2 but the hydrodynamic solution

is obtained using the surrogate model estimates. The results of this optimization process are the so-called *virtual designs*.

In Phase 3 (highlighted by a green dotted rectangle), the optimal *virtual design* found in Phase 2 (or eventually a set of possible optimal designs) is verified through new SPH simulations, hence generating a new *real design* set. In case the difference between the *virtual design* and the corresponding *real design* is within a given small tolerance ϵ, such a design can be chosen as the *real optimal design* concluding the calibration of the SPH model parameters. On the contrary, if those two designs differ by a certain amount greater than ϵ, the surrogate surface will be updated including the new *real design*. The second optimization process is repeated until convergence.

The architecture of the proposed approach, i.e. a Metamodel-Embedded Evolution framework (Razavi et al., 2012), is in principle comparable to other global convergence multi-objective optimization algorithms (e.g. Particle Swarm Optimizer) or other exact (interpolating) emulators (e.g. Radial Basis Functions).

4.1 Population-based multi-objective optimization algorithm

The Non-dominated Sorting Genetic Algorithm (NSGA-II) proposed by Deb et al. (2002) is selected for the first phase of the process. This is a well established Multi-Objective Optimization Algorithm (MOGA) able to solve complex multi-modal constrained problems (see for instance Zitzler et al., 2000). Many successful examples of the application of this optimization algorithm to a variety of problems belonging to different disciplines can be found in literature. It has been used for example in shape optimization of unconventional ships (Vernengo and Brizzolara, 2017) and propellers (Gaggero et al., 2016) or in calibrating the Soil and Water Assessment Tool (Zhang et al., 2013; Ercan and Goodall, 2016).

This algorithm adopts different levels of classification of the generated designs, creating sub-groups that are used during the evolution strategy. Besides the classic genetic operators (selection, crossover and mutation), the optimization process is also driven by the so called *crowding distance* assigned to each design in order to improve the exploration of the design space.

4.2 Surrogate Model for Data Interpolation

The interpolation of the reduced set of designs obtained at the end of the first optimization process is performed using a Kriging method. This is a technique first developed for interpolation of geostatistical spatial data (Krige, 1951) and then formalized for general interpolation problems (Matheron, 1963). Examples of application of such a data interpolation method can be found in many fields: for example in environmental-related problems (see for instance Li et al., 2011; Rohmer and Idier, 2012; Tsoukalas and Makropoulos, 2015; Barca et al., 2017)) or in fluid dynamics applications (see e.g. Simpson et al., 2001; Duvigneau and Chandrashekar, 2012). Further details on this method and on its
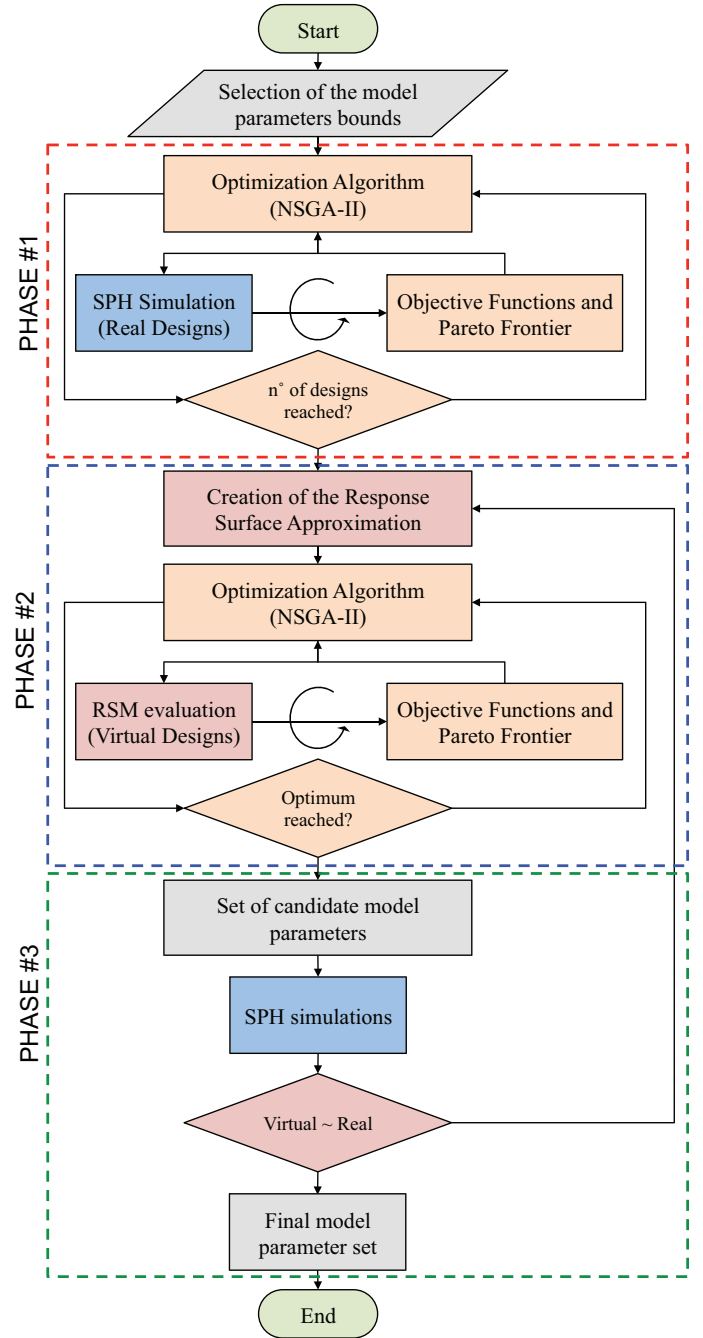
Figure 1: High-level flow chart of the computational framework for SPH model parameter identification. The three computational phases are highlighted by colored dashed rectangles.

formulation can be found among the other in Jones et al. (1998), Santner et al. (2003) or Forrester et al. (2008). There also exist specific formulations based on the Kriging model that aim at improving the prediction of the solution in unknown points of the domain under investigation. For example by using anisotropic semivariograms as in Friedland et al. (2016), by merging information from sources at different level of fidelity as proposed

by Kennedy and O'Hagan (2000) and applied e.g. in Bonfiglio et al. (2017) or by using particular rules on data co-variance as in Clark Jr et al. (2016).

Kriging is mainly a linear interpolation method able to give a prediction of the value of an objective function at an unknown location of the domain under investigation using the observations at known locations. The prediction is provided by weighting the known values in the neighbors of the point of interest. Due to the linearity of the Kriging estimator, the covariance can be used to express the variance error. The covariograms of the objective function, identifying the correlations among the data, are used to define the weighting coefficients. The *unbiasedness* property according to which the real and the predicted (virtual) value in a given location should be the same is enforced by specific requirements on these weigh coefficients.

The computational domain is denoted by $\Omega \subset \mathbf{R}^{n_p}$, being n_p the number of free parameters of the problem. Considering a point $\mathbf{v} = \left[v_1, v_2, ..., v_{n_p}\right]$ of the domain where the value of the objective function $\chi(\mathbf{v})$ is unknown, and a given number n_v of neighbor points in the proximity of the desired estimation point, the Kriging estimation $\tilde{\chi}(\mathbf{v})$ of $\chi(\mathbf{v})$ in $\mathbf{v} \subset S$ is given by Eq. (18):

$$\tilde{\chi}(\mathbf{v}) - E(\mathbf{v}) = \sum_{i=1}^{n_v} \lambda_i(\mathbf{v}) \cdot [\chi(\mathbf{v}_i) - E(\mathbf{v}_i)] \qquad (18)$$

where $E(\mathbf{v})$ and $E(\mathbf{v_i})$ are the expected values of $\tilde{\chi}(\mathbf{v})$ and $\chi(\mathbf{v}_i)$ respectively, and $\lambda_i(\mathbf{v})$ are the weights coefficients for the estimation point $\mathbf{v}$.

Assuming an isotropic stationary model, the covariance C of the function χ of two points $(\mathbf{v}_i, \mathbf{v}_j)$ depends only on the distance between the two. Hence the covariance and the covariance matrix are defined by the following Eq. (19) and Eq. (20), respectively:

$$C(\chi(\mathbf{v}_i), \chi(\mathbf{v}_j)) = E[(\chi(\mathbf{v}_i) - E(\chi(\mathbf{v}_i))) \cdot \\ \cdot (\chi(\mathbf{v}_j) - E(\chi(\mathbf{v}_j)))] \qquad (19)$$

$$\mathbf{C} = \begin{vmatrix} \sigma^2 & C(||\mathbf{v}_1 - \mathbf{v}_2||) & \cdots & C(||\mathbf{v}_1 - \mathbf{v}_{n_v}||) \\ C(||\mathbf{v}_2 - \mathbf{v}_1||) & \sigma^2 & \cdots & C(||\mathbf{v}_2 - \mathbf{v}_{n_v}||) \\ \vdots & \vdots & \ddots & \vdots \\ C(||\mathbf{v}_{n_v} - \mathbf{v}_1||) & C(||\mathbf{v}_{n_v} - \mathbf{v}_2||) & \cdots & \sigma^2 \end{vmatrix} \qquad (20)$$

The difference between the predicted function value at a point $\tilde{\chi}(\mathbf{v})$ (virtual design) and its real value $\chi(\mathbf{v})$ is called the estimation error, as defined in Eq. (21):

$$\varepsilon(\mathbf{v}) = \frac{\tilde{\chi}(\mathbf{v}) - \chi(\mathbf{v})}{\chi(\mathbf{v})} \qquad (21)$$

To build an optimal function estimation method the weight coefficients $\lambda_i(\mathbf{v})$ are derived in order to ensure the minimization of the variance of the estimation error. This condition is written as follows:

$$\frac{\partial E[(\tilde{\chi}(\mathbf{v}) - \chi(\mathbf{v}))^2]}{\partial \lambda_i(\mathbf{v})} = 0 \qquad (22)$$

Developing Eq. (22) and imposing the unbiasedness condition on the weight coefficients result in a constrained optimization problem that can be solved by Lagrangian multiplier technique, leading to the definition of a closed form for the Kriging predictor model.

4.3 Objective functions for progressive wave propagation

The aim of the automatic calibration framework is to obtain the set of model parameters that ensures a stable propagation of a given wave. Here a 2nd order Stokes wave is chosen but the method is general and is in principle applicable to every kind of wave.

Both space and time wave characteristics are monitored. As a consequence, two objective functions are defined in terms of spatial and time errors of the numerical wave with respect to the theoretical wave, giving rise to a multi-objective optimization problem.

The computational domain is defined in a fixed Cartesian reference system centered at the intersection between the bottom of the tank and the wavemaker at rest. The tank has a flat bottom of length $L_{Tank} = 15\ m$ ending on a sloped beach used to damp the incoming waves. The initial water depth is $d = 1.2\ m$. The computational domain with a developed wave is shown in Figure 2. Wave profile $\eta^{SPH}(x, t)$ is marked by blue dots.

According to Eq. (23) the objective function in space F_S is defined as the root mean square of the relative errors between the theoretical wave profile $\eta(x, t)$ and the computed one at a given time $t = \bar{t}$. $\bar{t}$ is chosen to avoid the first transient phase of the simulation. The time objective function F_T has the same definition but considering a wave profile measured at a fixed location $\frac{\bar{x}}{L} = \frac{1}{3}$ over a suitable range of time steps (see Eq. (24)). The simulation time has been choose in order to avoid reflections of the first waves that have reached the sloped beach (which is dissipative too).

$$F_S = \frac{1}{\bar{\eta}(x, \bar{t})} \sqrt{\sum_{i=1}^{n_p} (\eta_i^{SPH}(x, \bar{t}) - \eta_i(x, \bar{t}))^2} \qquad (23)$$

$$F_T = \frac{1}{\bar{\eta}(\bar{x}, t)} \sqrt{\sum_{i=1}^{n_{\Delta t}} (\eta_i^{SPH}(\bar{x}, t) - \eta_i(\bar{x}, t))^2} \qquad (24)$$

5. Results of the Identification Method for a 2nd order Stokes Wave

The goal of the optimization-based framework is to calibrate the SPH parameters with the goal of accurately reproducing a 2nd order Stokes wave of amplitude $H = 0.1\ m$ and length $\lambda = 2\ m$. The five-dimensional space of parameters is described in Table 1 where we indicate lower and upper bounds of each variable together with five relevant SPH parameters selected for the calibration of the wave generation model. A uniform distribution has been assigned to each of them. The dp parameter is directly
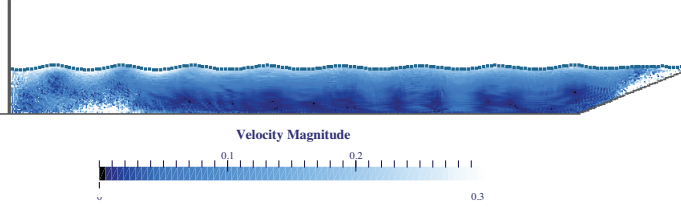
Figure 2: Example of computational domain used to resolve the wave generation. The free surface profile is highlighted by blue dots while particle velocities are shown using colors from blue to white.

Table 1: SPH model parameter involved in the calibration process.

| SPH Model Parameter	Lower bound	Upper bound
$\dfrac{H}{dp}$	3.5	12.5
$coeff_{sound}$	10	30
α	0.001	0.03
h	0.5	2.0
CFL	0.100	0.50

related to the number of particles used for the simulation. Therefore it strongly affects the required computational time. Its value is selected so that the ratio $\frac{H}{dp}$ results in the range $(3.5, 12.5)$. The suggested value of $\frac{H}{dp} = 10$ is then included in these bounds. The lower bound is chosen to achieve a reasonable limit on the overall computational time. As described in Section 3, $coeff_{sound}$ is a coefficient modeling the speed of sound at the reference density while α is a coefficient used in the artificial viscosity computation. The *Courant-Friedrichs-Lewy* (CFL) condition, together with a forcing and a viscous diffusion terms, regulates the time step for the explicit time integration scheme. These parameters have a direct impact on numerical issues related to the solver (such as possible flow instability, particle elimination and over-damped waves).

Figure 3 presents examples of possible inaccurate predictions of the wave profiles for the 2nd order Stokes wave over the selected flume tank, occurring during the automatic optimization process. All three simulations have been carried out with $dp = 0.01$, $coeff_{sound} = 15.8$. Simulation $SPH - A$, considered as a reference in this example, has $CFL = 0.5$, $h = 1.8$ and $\alpha = 0.0035$. With respect to the reference, simulation $SPH - B$ has $CFL = 0.176$ and $\alpha = 0.05$, while simulation $SPH - C$ has $CFL = 0.176$ and $h = 0.86$. All the values have been automatically selected by the identification framework. The predicted wave shapes strongly differ from the theoretic solution, ultimately demonstrating the sensitivity of the solution to the model parameters.

The NSGA-II optimization algorithm has been run for about 7 generations in Phase 1. Based on previous experiences, according to the rule $n_{Designs} = 10 * n_{Parameters}$ (also used e.g. in Jones et al., 1998) each generation is made of 50 candidate de-
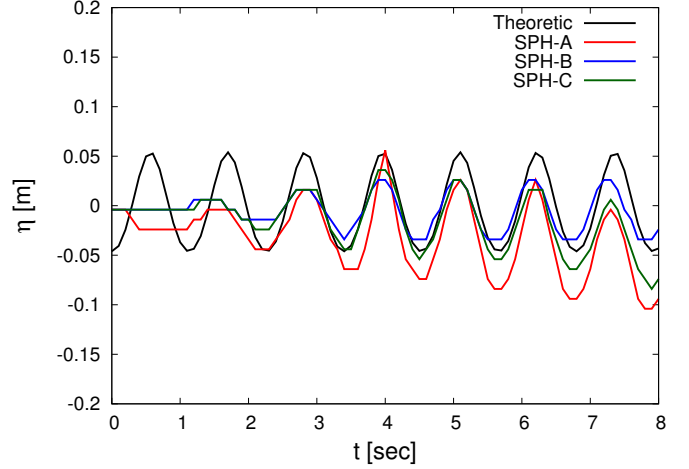
Figure 3: Comparison of theoretic (solid black curve) and (failed) numerical time series of surface elevation of the 2nd order Stokes wave measured at $x = 5\ m$ from the wave maker.
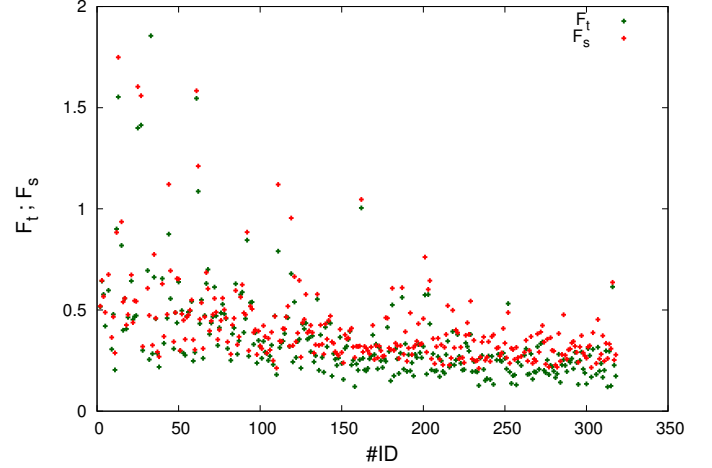
Figure 4: Evolution of the two objective functions during the optimization process. Red dots for F_s and green dots for F_t.

signs resulting in about 320 total designs evaluated. The initial Design of Experiments (DoE), i.e. the first generation of the algorithm has been sampled in the five-dimensional parameter space using a Sobol algorithm. The crossover probability has been set to $C_R = 0.9\%$ and the mutation probability to $C_M = 0.6\%$. The two optimization histories for the space and time objective functions are shown in Figure 4. This plot presents the evolution of the two objective functions during the optimization computation. The algorithm is able to reduce the initial spread among the solutions by minimizing the values of both objective functions. The genetic algorithm optimization is interrupted before complete convergence is reached, with objective functions showing a trend toward the optimum.

Considering this first set of results, the effect of each parameter on the objective functions has been preliminary analyzed by us-

ing the t-Student test, originally developed by Gossett (Student, 1908). This statistical test might be used even on a very small set of sampled data (De Winter, 2013) without need of bootstrapping, as happens e.g. in other variance based methods for sensitivity analysis (Pianosi et al., 2015). The value of the t parameter is obtained according to Eq. (25).

$$t = \frac{|M_l - M_u|}{\sqrt{\frac{S_G^2}{n_u} + \frac{S_G^2}{n_l}}} \quad (25)$$

Where S_G^2 is the general variance computed by Eq. (26), S_u^2 and S_l^2 are the variances of the designs for the output variable x in the upper and lower parts of the domain respectively as in Eq. (27) and (28).

$$S_G^2 = \frac{(n_l - 1)S_l^2 + (n_u - 1)S_u^2}{(n_l + n_u - 2)} \quad (26)$$

$$S_u^2 = \frac{\sum (x_u - M_u)^2}{(n_u - 1)} \quad (27)$$

$$S_l^2 = \frac{\sum (x_l - M_l)^2}{(n_l - 1)} \quad (28)$$

M_u and M_l are the means of the values of the output variable x in the upper and lower parts respectively of domain of the input variables while n_u and n_l are the number of values in the upper and lower parts of that domain. According to this definition, the significance represents the probability that the difference between upper and lower ranges of the response variable is due to chance. Then, low values of significance mean a stronger relation between factor and response variable while high values stand for weaker factor-response relation. As already applied in Vernengo et al. (2016), the influence of each parameter is evaluated as the opposite of the Student t-test significance t. Results of this analysis are represented in Figure 5 in a normalized form (i.e. each significance is divided by the overall summation on all the parameters). Both the objective functions are affected by the selected model parameters in the same extent. Smoothing length coefficient h and the CFL cover in both cases more than the 60% of the total percentage of the significance. $coeff_{sound}$ is the third most influencing variable followed by the artificial viscosity coefficient α ranging from 12% on F_S up to 13.6% on F_T. In this case, the number of particles governed by dp does not seem to have a relevant impact on the simulation.

This analysis concludes the computations performed in Phase 1. At the end of Phase 1 the number of objective function evaluations do not allow to clearly identify the so-called Pareto Frontier, i.e. the extreme boundary of the non-dominated solutions. This is due to a certain dispersion of the solutions that the optimization algorithm has not been able to tighten because of the relatively small number of generations.

In Phase 2 we construct a Kriging surrogate model for each objective function, using a subset of 100 designs from the optimization analysis of Phase 1 computation. These 100 designs are selected from the tails of the optimization histories, hence excluding the first 220 designs that show relatively higher dispersion. Both surrogate models are preliminary analyzed by using a

cross-validation procedure according to the *Leave-one-out* technique: each ith design is removed from the design space and a surrogate model is generated based on the $n_P - 1$ designs. We use the surrogate model to predict the objective function for the ith design and we compare Kriging prediction with the value obtained by CFD simulation. The estimation error, defined in Eq. (21), is on average equal to $\epsilon_{F_S} = 10.39\%$ and $\epsilon_{F_t} = 8.86\%$ F_S and F_t surrogate models. Despite the non-negligible average error, these models are able to properly predict the trends of the responses as shown in the following discussion of the results. We could reduce the average error by constructing surrogate models using different data sets, but in the context of the present study we employ the Kriging models to improve and confirm the prediction given by the genetic optimization algorithm.

Figure 6 presents the solutions found using this innovative optimization process in the space of the two objective functions where we can visually identify the Pareto Frontier. Figure 6 also shows the initial set of designs and the final optimum (both from Kriging response surface and from the SPH computation). Using this second optimization process we were able to analyze about 60000 cases in few minutes on a single *cpu*. Each new wave evaluation can be cheaply sampled from the response surface constructed using the Kriging surrogate model. The whole set of designs evaluated during the surrogate-based optimization is presented in Figure 6 by means of blue dots which finally identify the *Pareto Frontier*.

When using such a kind of Metamodel-Embedded Evolution Framework there are choices and relevant issues to be aware of (see Razavi et al., 2012). Examples are the size of the initial DoE, possible increases of inaccuracy due to the use of multiple surrogate models, *overfitting* of the surrogates due to the *conformability* factor. Most of these issues are difficult to predict in advance and many choices are based on users' experience and preferences. So beyond the validation of the single models of the framework (i.e. the SPH solver and the kriging interpolation) a further verification of the obtained results is required to assess the validity of the process.

The best Pareto solution found using the Kriging approximation has then been directly computed using the SPH solver in order to verify the real improvements achieved in Phase 2. Overall results of this computation are summarized in Table 2. The surrogate-based optimization is ultimately able to reduce both objectives by $\Delta F_S = -15.7\%$ and $\Delta F_T = -9.5\%$ compared to Phase 1. The optimal set of model parameters is reported in Table 3.

Figure 7 demonstrates the good agreement between the time evolution at $\frac{x}{L} = \frac{1}{3}$ of the simulated 2nd order Stokes wave obtained with the optimal set of model parameter (Best Pareto (Real)-Phase 2) and the theoretic wave profile.

Figure 8 displays as a bubble-chart the SPH solutions used to generate the surrogate model in the plane of the two objective functions. The $\frac{H}{dp}$ parameter is shown using both bubble size and color. It is not trivial to understand the trends of both objective functions with respect to e.g. the selected parameter. The ratio $\frac{H}{dp}$ has a strong impact on the computational time but its effect on the objective functions depends on the combinations of the

(a) Model parameter influence on F_T

(b) Model parameter influence on F_S
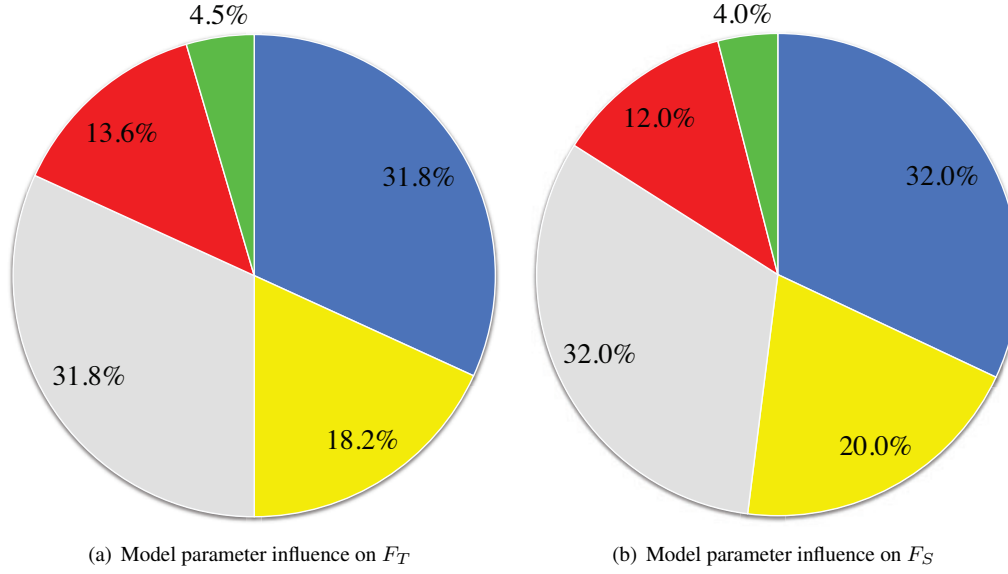
Figure 5: Influence of each model parameter with respect to F_t (left side) and F_s (right side). CFL in blue. $coeff_{sound}$ in yellow. h in grey. α in red. dp in green

other parameters. This is highlighted in Figure 9 for a subset of five SPH simulations. Increasing that ratio from a reference value of $\frac{H}{dp} = 4.5$ (blue circle) might also lead to double the computational time while the objective functions do not show a clear monotonic trend. This exactly depends on the combination of all the model parameters involved in the simulation. This clarifies the need for an automatic procedure for parameter calibration. The optimum set of parameters does not have the maximum allowed value of $\frac{H}{dp}$, theoretically related to the highest possible accuracy (i.e. the finest particle resolution), and the use of the framework avoids screening all the possible solutions and allows selection of those that can be *a priori* considered the most accurate.

Finally, the effectiveness of the results obtained by the calibration model has been proven on four different waves, i.e. two 2nd order Stokes waves and two 3rd order ones. Simulated waves have been compared with the corresponding theoretic profiles and the two objective functions used during the calibration process have been computed according to Eq. (24) and Eq. (23). Results of this additional verification study are included in Table 4. Results of the comparison between the time evolution of the numeric wave profile and the theoretic one of a 3rd order Stokes wave are shown in Figure 10 and 11. In the former the wave profiles are taken at $\frac{x}{L} = \frac{1}{3}$. In the latter two other slices at $\frac{x}{L} = \frac{2}{3}$ and $\frac{x}{L} = 1$ have been considered. In these last cases fewer wave periods are shown because, as already mentioned, the simulation time has been set to avoid the reflection of the first wave reaching the sloped beach. The good results achieved by these final comparisons on both 2nd and 3rd order Stokes waves different from the one used for calibration confirmed the validity of the proposed method based on different optimization techniques for this specific case.

Table 2: Results of the surrogate-based optimization.

	Best-Phase 1	Pareto (Kriging)-Phase 2	Pareto (SPH)-Phase 2
F_S	$2.11E^{-01}$	$1.84E^{-01}$	$1.91E^{-01}$
F_T	$1.27E^{-01}$	$9.51E^{-02}$	$1.07E^{-01}$

Table 3: Optimal set of the SPH model parameter resulting from the calibration process on the selected theoretic 2nd order Stokes wave.

SPH Model Parameter	Optimal value
$\dfrac{H}{dp}$	10.204
$coeff_{Sound}$	15.80
α	$3.54E^{-03}$
h	1.80
CFL	$1.76E^{-01}$

Table 4: Verification of the optimal set of SPH model parameters on other waves.

Wave type	$T\ [sec]$	$\frac{H}{L}$	F_T	F_S
Stokes 2nd order	1.14	0.033	$1.02E^{-01}$	$0.83E^{-01}$
Stokes 2nd order	0.76	0.033	$1.86E^{-01}$	$2.05E^{-01}$
Stokes 3rd order	1.27	0.070	$0.61E^{-01}$	$1.93E^{-01}$
Stokes 3rd order	0.76	0.070	$1.64E^{-01}$	$1.10E^{-01}$
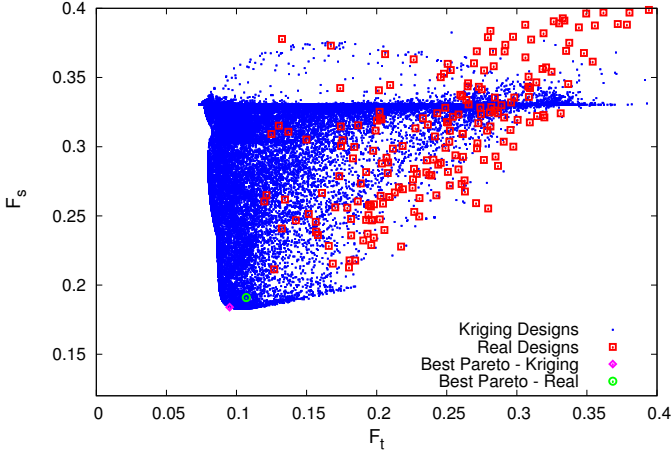
Figure 6: F_S versus F_T. Designs resulting from the optimization on the surrogate Kriging model are shown by blue dots. Real designs resulting by Phase 1 are shown by red empty rectangles. The best-Pareto design according to Kriging model is highlighted by a pink rhombus and the same design re-evaluated using the SPH solver is highlighted by a green empty circle.
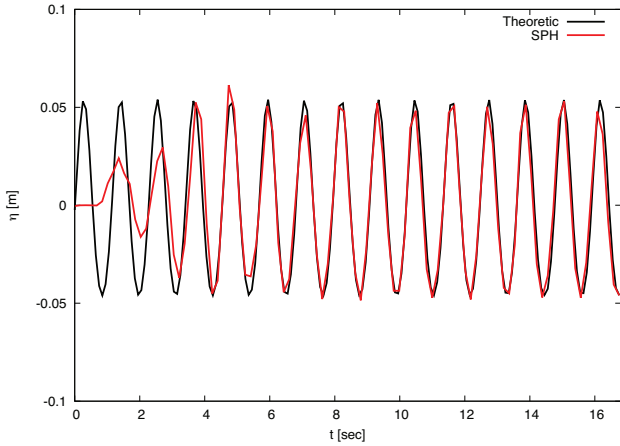

Figure 7: Comparison of theoretic (black solid curve) and computed (red solid curve) 2nd order Stokes' wave at $\frac{x}{L} = \frac{1}{3}$.

6. Conclusions

The time-consuming calibration of the model parameters of a Smoothed Particle Hydrodynamics (SPH) needed to obtain accurate physical results for a specific problem typically relies on tedious one-at-the-time tunings which strongly rely on user experience. In this paper, we solve this task using an innovative multi-objective optimization framework specifically developed to minimize the overall computational time needed to find the optimal parameter set. The computational process involves three main phases:

1. the initial exploration and sampling of the domain, aimed at tightening parameter bounds, and performed by a multi-
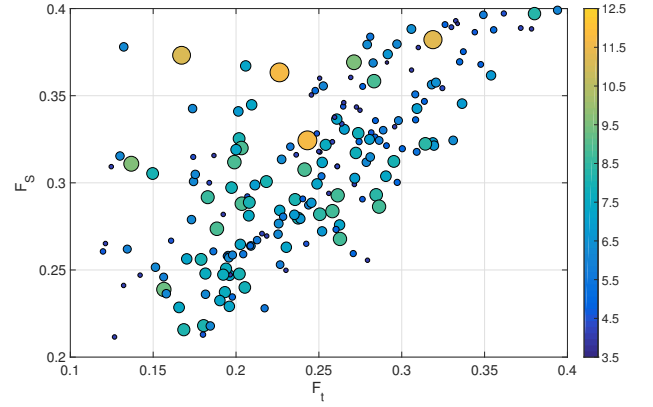
Figure 8: F_S versus F_T. Bubble-chart of the selected SPH solutions for surrogate creation. $\frac{H}{dp}$ is represented both as the bubble size and by colors.
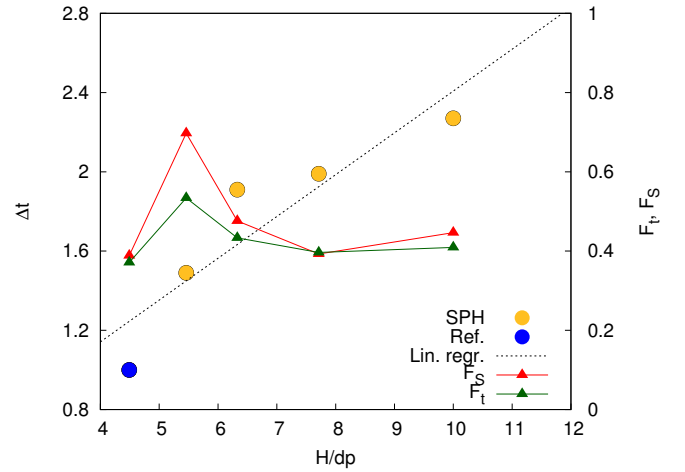

Figure 9: Example of simulation time versus particle resolution. Blue point is taken as the reference for Δt. Linear regression is shown too.

objective optimization algorithm.

2. interpolation of resulting data set by a Response Surface Method (RSM) in order to generate a surrogate model for each objective function

3. optimization exploiting the newly generated surrogate models to perform a candidate optimum search able to minimize both space and time errors with respect to a known solution.

The capabilities of the proposed framework are shown in an example case, i.e. the simulation of a 2nd order Stokes wave propagating in a numeric flume tank ending with a sloped beach. At the end of Phase 1 an appreciable improvement of the errors between the theoretic and numeric waves is achieved but at the cost of few hundreds SPH simulations (ranging from less than 1 hour up to about 2 hours). The use of surrogate models gives a marked
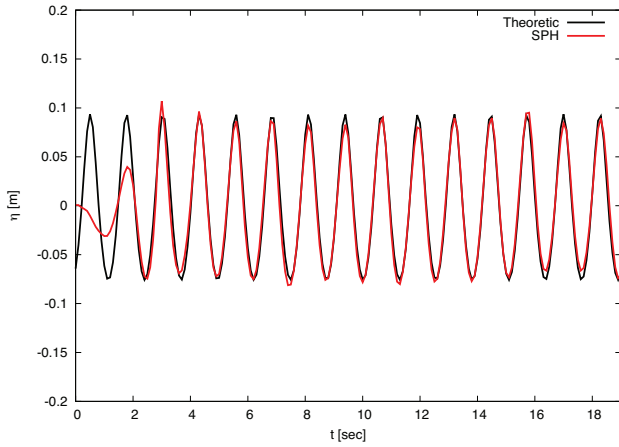
Figure 10: Comparison of theoretic (black solid curve) and computed (red solid curve) 3rd order Stokes' wave at $\frac{x}{L} = \frac{1}{3}$.
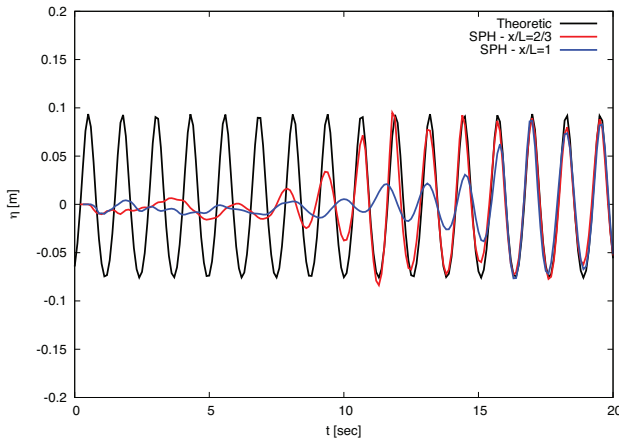
Figure 11: Comparison of theoretic (black solid curve) and computed 3rd order Stokes' wave at $\frac{x}{L} = \frac{2}{3}$ (red solid curve) and $\frac{x}{L} = 1$ (blue solid curve).

speed-up to the optimization process, allowing to compute about 60,000 extra designs in few minutes, leading to a final improvement of $\Delta F_t = -15.7\%$ and $\Delta F_s = -9.5\%$ with respect to the lowest error found from Phase 1. The accuracy of the SPH model, properly tuned with the optimum numerical and physical parameters identified for the selected theoretic 2nd order Stokes wave, is finally verified on a set of four new Stokes waves of both 2nd and 3rd order. In all these cases the computed error functions F_T and F_S provide consistent results of the same order of those obtained on the wave used in the calibration process. This confirms that the results obtained by the proposed framework have a fairly general validity for such a class of problems. Such results also encourage the application of the proposed calibration framework to other types of fluid related problems that can be solved by this SPH method, such as wave-structure interaction, solitary wave propagation or fluid-driven object dynamics.

REFERENCES

Altomare, C., Crespo, A. J., Domínguez, J. M., Gómez-Gesteira, M., Suzuki, T., and Verwaest, T. (2015a). Applicability of smoothed particle hydrodynamics for estimation of sea wave impact on coastal structures. *Coastal Engineering*, 96:1–12.

Altomare, C., Crespo, A. J., Rogers, B., Domínguez, J., Gironella, X., and Gómez-Gesteira, M. (2014). Numerical modelling of armour block sea breakwater with smoothed particle hydrodynamics. *Computers & Structures*, 130:34–45.

Altomare, C., Domínguez, J., Crespo, A., González-Cao, J., Suzuki, T., Gómez-Gesteira, M., and Troch, P. (2017). Long-crested wave generation and absorption for sph-based dual-sphysics model. *Coastal Engineering Journal*, 127:37–54.

Altomare, C., Domínguez, J., Crespo, A., Suzuki, T., Caceres, I., and Gómez-Gesteira, M. (2015b). Hybridization of the wave propagation model swash and the meshfree particle method sph for real coastal applications. *Coastal Engineering Journal*, 57(04):1550024.

Antuono, M., Colagrossi, A., Marrone, S., and Molteni, D. (2010). Free-surface flows solved by means of sph schemes with numerical diffusive terms. *Computer Physics Communications*, 181(3):532–549.

Austin, D. I. and Schlueter, R. S. (1982). A numerical model of wave/breakwater interactions. In *Coastal Engineering 1982*, pages 2079–2096.

Barca, E., Porcu, E., Bruno, D., and Passarella, G. (2017). An automated decision support system for aided assessment of variogram models. *Environmental Modelling & Software*, 87:72–83.

Barreiro, A., Crespo, A., Domínguez, J., and Gómez-Gesteira, M. (2013). Smoothed particle hydrodynamics applied in fluid structure interactions. *Fluid Structure Interaction VII*, 129:75.

Batchelor, G. (1969). An introduction to fluid dynamics.

Bayon, A., Valero, D., García-Bartual, R., López-Jiménez, P. A., et al. (2016). Performance assessment of openfoam and flow-3d in the numerical modeling of a low reynolds number hydraulic jump. *Environmental Modelling & Software*, 80:322–335.

Bonfiglio, L., Perdikaris, P., Brizzolara, S., and Karniadakis, G. E. (2017). A multi-fidelity framework for investigating the performance of super-cavitating hydrofoils under uncertain flow conditions. In *19th AIAA Non-Deterministic Approaches Conference*, page 1328. AIAA.

Brunetti, G., Simunek, J., Turco, M., and Piro, P. (2017). On the use of surrogate-based modeling for the numerical analysis of low impact development techniques. *Journal of Hydrology*, 548:263–277.

Chau, K. (2004). Selection and calibration of numerical modeling in flow and water quality. *Environmental Modeling & Assessment*, 9(3):169–178.

Clark Jr, D. L., Bae, H.-R., Gobal, K., and Penmetsa, R. (2016). Engineering design exploration using locally optimized covariance kriging. *AIAA Journal*, pages 3160–3175.

Crespo, A., Gómez-Gesteira, M., and Dalrymple, R. A. (2007). 3d sph simulation of large waves mitigation with a dike. *Journal of Hydraulic Research*, 45(5):631–642.

Crespo, A. J., Domínguez, J. M., Rogers, B. D., Gómez-Gesteira, M., Longshaw, S., Canelas, R., Vacondio, R., Barreiro, A., and García-Feal, O. (2015). Dualsphysics: Open-source parallel cfd solver based on smoothed particle hydrodynamics (sph). *Computer Physics Communications*, 187:204–216.

De Padova, D., Dalrymple, R. A., and Mossa, M. (2014). Analysis of the artificial viscosity in the smoothed particle hydrodynamics modelling of regular waves. *Journal of Hydraulic Research*, 52(6):836–848.

De Winter, J. C. (2013). Using the student's t-test with extremely small sample sizes. *Practical Assessment, Research & Evaluation*, 18(10).

Dean, R. G. and Dalrymple, R. A. (1991). *Water wave mechanics for engineers and scientists*. World Scientific Publishing, NJ, USA.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197.

Duvigneau, R. and Chandrashekar, P. (2012). Kriging-based optimization applied to flow control. *International Journal for Numerical Methods in Fluids*, 69(11):1701–1714.

Ercan, M. B. and Goodall, J. L. (2016). Design and implementation of a general software library for using nsga-ii with swat for multi-objective model calibration. *Environmental Modelling & Software*, 84:112–120.

Forrester, A., Keane, A., et al. (2008). *Engineering design via surrogate modelling: a practical guide*. John Wiley & Sons.

Friedland, C. J., Joyner, T. A., Massarra, C., Rohli, R. V., Treviño, A. M., Ghosh, S., Huyck, C., and Weatherhead, M. (2016). Isotropic and anisotropic kriging approaches for interpolating surface-level wind speeds across large, geographically diverse regions. *Geomatics, Natural Hazards and Risk*, pages 1–18.

Fuhrman, D. R. and Madsen, P. A. (2008). Simulation of nonlinear wave run-up with a high-order boussinesq model. *Coastal Engineering*, 55(2):139–154.

Gaggero, S., Gonzalez-Adalid, J., and Sobrino, M. P. (2016). Design of contracted and tip loaded propellers by using boundary element methods and optimization algorithms. *Applied Ocean Research*, 55:102–129.

Gomez-Gesteira, M., Rogers, B. D., Crespo, A. J., Dalrymple, R. A., Narayanaswamy, M., and Dominguez, J. M. (2012a). Sphysics–development of a free-surface fluid solver–part 1: Theory and formulations. *Computers & Geosciences*, 48:289–299.

Gomez-Gesteira, M., Rogers, B. D., Crespo, A. J., Dalrymple, R. A., Narayanaswamy, M., and Dominguez, J. M. (2012b). Sphysics–development of a free-surface fluid solver–part 1: Theory and formulations. *Computers & Geosciences*, 48:289–299.

Gotoh, H., Ikari, H., Memita, T., and Sakai, T. (2005). Lagrangian particle method for simulation of wave overtopping on a vertical seawall. *Coastal Engineering Journal*, 47(02n03):157–181.

Higuera, P., Lara, J. L., and Losada, I. J. (2013). Simulating coastal engineering processes with openfoam®. *Coastal Engineering*, 71:119–134.

Jones, D. R., Schonlau, M., and Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492.

Kennedy, A. B., Chen, Q., Kirby, J. T., and Dalrymple, R. A. (2000). Boussinesq modeling of wave transformation, breaking, and runup. i: 1d. *Journal of waterway, port, coastal, and ocean engineering*, 126(1):39–47.

Kennedy, M. C. and O'Hagan, A. (2000). Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1):1–13.

Krige, D. (1951). A statistical approach to some mine valuations and allied problems. *PhD thesis. University of Witwatersrand.*

Leimkuhler, B. J., Reich, S., and Skeel, R. D. (1996). Integration methods for molecular dynamics. In *Mathematical Approaches to biomolecular structure and dynamics*, pages 161–185. Springer.

Li, J., Heap, A. D., Potter, A., and Daniell, J. J. (2011). Application of machine learning methods to spatial interpolation of environmental variables. *Environmental Modelling & Software*, 26(12):1647–1659.

Madsen, P. A., Bingham, H., and Liu, H. (2002). A new boussinesq method for fully nonlinear waves from shallow to deep water. *Journal of Fluid Mechanics*, 462:1–30.

Marrone, S., Antuono, M., Colagrossi, A., Colicchio, G., Le Touzé, D., and Graziani, G. (2011). δ-sph model for simulating violent impact flows. *Computer Methods in Applied Mechanics and Engineering*, 200(13):1526–1542.

Matheron, G. (1963). Principles of geostatistics. *Economic geology*, 58(8):1246–1266.

Meringolo, D., Colagrossi, A., Marrone, S., and Aristodemo, F. (2017). On the filtering of acoustic components in weakly-compressible sph simulations. *Journal of Fluids and Structures*, 70:1–23.

Molteni, D. and Colagrossi, A. (2009). A simple procedure to improve the pressure evaluation in hydrodynamic context using the sph. *Computer Physics Communications*, 180(6):861–872.

Monaghan, J. J. (1992). Smoothed particle hydrodynamics. *Annual review of astronomy and astrophysics*, 30:543–574.

Monaghan, J. J. and Lattanzio, J. C. (1985). A refined particle method for astrophysical problems. *Astronomy and astrophysics*, 149:135–143.

Monaghan, J.J., K. A. (1999). Solitary waves on a cretan beach. *Journal of Waterway, Port, Coastal and Ocean Engineering*, 125(3):145,154.

Pianosi, F., Sarrazin, F., and Wagener, T. (2015). A matlab toolbox for global sensitivity analysis. *Environmental Modelling & Software*, 70:80–85.

Razavi, S., Tolson, B. A., and Burn, D. H. (2012). Review of surrogate modeling in water resources. *Water Resources Research*, 48(7).

Rogers, B. D., Dalrymple, R. A., and Stansby, P. K. (2010). Simulation of caisson breakwater movement using 2-d sph. *Journal of Hydraulic Research*, 48(S1):135–141.

Rohmer, J. and Idier, D. (2012). A meta-modelling strategy to identify the critical offshore conditions for coastal flooding. *Natural Hazards and Earth System Sciences*, 12(9):2943.

Rota, R., Angelini, R., Brizzolara, S., Chryssostomidis, C., Guercio, R., et al. (2014). Numerical modeling of breaking periodical waves on a sloped beach profile by sph. In *The Twenty-fourth International Ocean and Polar Engineering Conference*. International Society of Offshore and Polar Engineers.

Santner, T., Williams, B., and Notz, W. (2003). The design and analysis of computer experiments springer-verlag. *New York. 283pp*.

Simpson, T. W., Mauery, T. M., Korte, J. J., and Mistree, F. (2001). Kriging models for global approximation in simulation-based multidisciplinary design optimization. *AIAA journal*, 39(12):2233–2241.

Simpson, T. W., Toropov, V., Balabanov, V., and Viana, F. A. (2008). Design and analysis of computer experiments in multidisciplinary design optimization: a review of how far we have come or not. volume 5.

Smit, P., Janssen, T., Holthuijsen, L., and Smith, J. (2014). Non-hydrostatic modeling of surf zone wave dynamics. *Coastal Engineering*, 83:36–48.

St-Germain, P., Nistor, I., Townsend, R., and Shibayama, T. (2013). Smoothed-particle hydrodynamics numerical modeling of structures impacted by tsunami bores. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 140(1):66–81.

Student (1908). The probable error of a mean. *Biometrika*, pages 1–25.

Suzuki, T., Altomare, C., Veale, W., Verwaest, T., Trouw, K., Troch, P., and Zijlema, M. (2017). Efficient and robust wave overtopping estimation for impermeable coastal structures in shallow foreshores using swash. *Coastal Engineering*, 122:108–123.

Suzuki, T., Verwaest, T., Hassan, W., Veale, W., Reyns, J., Trouw, K., Troch, P., and Zijlema, M. (2011). The applicability of swash model for wave transformation and wave overtopping: A case study for the flemish coast. In *Proc. 5th Int. Conf. Advanced Computational Methods Engineering (ACOMEN 2011), Liège, Belgium*, pages 14–17.

Tsoukalas, I. and Makropoulos, C. (2015). Multiobjective optimisation on a budget: exploring surrogate modelling for robust multi-reservoir rules generation under hydrological uncertainty. *Environmental Modelling & Software*, 69:396–413.

Vacondio, R., Dal Palù, A., and Mignosa, P. (2014). Gpu-enhanced finite volume shallow water solver for fast flood simulations. *Environmental Modelling & Software*, 57:60–75.

Vacondio, R., Mignosa, P., and Pagani, S. (2013). 3d sph numerical simulation of the wave generated by the vajont rockslide. *Advances in Water Resources*, 59:146–156.

Valdez-Balderas, D., Domínguez, J. M., Rogers, B. D., and Crespo, A. J. (2013). Towards accelerating smoothed particle hydrodynamics simulations for free-surface flows on multi-gpu clusters. *Journal of Parallel and Distributed Computing*, 73(11):1483–1493.

Vernengo, G. and Brizzolara, S. (2017). Numerical investigation on the hydrodynamic performance of fast swaths with optimum canted struts arrangements. *Applied Ocean Research*, 63:76–89.

Vernengo, G., Gaggero, T., and Rizzuto, E. (2016). Simulation based design of a fleet of ships under power and capacity variations. *Applied Ocean Research*, 61:1–15.

Violeau, D. (2012). *Fluid Mechanics and the SPH method: theory and applications*. Oxford University Press.

Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in computational Mathematics*, 4(1):389–396.

Zhang, X., Beeson, P., Link, R., Manowitz, D., Izaurralde, R. C., Sadeghi, A., Thomson, A. M., Sahajpal, R., Srinivasan, R., and Arnold, J. G. (2013). Efficient multi-objective calibration of a computationally intensive hydrologic model with parallel computing software in python. *Environmental modelling & software*, 46:208–218.

Zijlema, M., Stelling, G., and Smit, P. (2011). Swash: An operational public domain code for simulating wave fields and rapidly varied flows in coastal waters. *Coastal Engineering*, 58(10):992–1012.

Zitzler, E., Deb, K., and Thiele, L. (2000). Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolutionary computation*, 8(2):173–195.