Computers & Geosciences 177 (2023) 105362

Contents lists available at ScienceDirect

Computers and Geosciences

journal homepage: www.elsevier.com/locate/cageo

Research paper

A fully coupled superficial runoff and soil erosion basin scale model with efficient time stepping

Federico Gatti [a,*], Luca Bonaventura [a], Alessandra Menafoglio [a], Monica Papini [b], Laura Longoni [b]

[a] MOX-Department of Mathematics, Politecnico di Milano, Milan, Italy
[b] Department of Civil and Environmental Engineering, Politecnico di Milano, Milan, Italy

ARTICLE INFO

Keywords:
Basin scale modeling
Soil erosion modeling
Semi-implicit time discretization
Finite differences
Finite volumes
Depth-integrated models

ABSTRACT

We present a numerical model of soil erosion at the basin scale that allows one to describe surface run-off without *a priori* identifying drainage zones, river beds and other water bodies. The model is based on robust semi-implicit numerical techniques and guarantees exact mass conservation and positivity of the surface and subsurface water layers. Furthermore, the method is equipped with a geostatistical preprocessor that can perform downscaling of data retrieved from digital databases at coarser resolutions. Numerical experiments on both idealized and realistic configurations demonstrate the effectiveness of the proposed method in reproducing transient high resolution features at a reduced computational cost and to reproduce correctly the main hydrographic features of the considered catchment. Furthermore, probabilistic forecasts can be carried out based on soil data maps automatically generated by the geostatistical preprocessor that are distributed among the available MPI ranks to carry out simulations independently thus reducing the total cost of the simulation. The numerical experiments show the ability of the model to provide robust estimates of water levels, discharge and of the order of magnitude of the total sediment yield.

1. Introduction

Soil erosion in mountain catchments is a widely studied phenomenon with major economical and societal impacts. A great number of models for the simulation of this phenomenon exists, see e.g. Dutta (2016), Fatichi et al. (2016), Fu et al. (2019), Ketema and Dwarakish (2019), Merritt et al. (2003), Pandey et al. (2016), Borrelli et al. (2021) and the references therein. Among erosion models, physically based models are the most complex ones and try to simulate directly most of the processes that take place in reality, like the impact of raindrops on the soil and the consequent detachment of soil particles. These models usually rely on the numerical solution of the mass and momentum conservation equations for the water flow and of the mass conservation equation for sediments. As already pointed out in Epple et al. (2022), these models require a large amount of input data, which are often scarce or inconsistent. This leads to poor input datasets that create errors in output, which can be seen as a drawback that generally applies to all the models of this group. Another problem is the risk of over-parametrization. More specifically, the large number of parameters (tens or even hundreds) that these models require makes it almost impossible to find optimal values (Merritt et al., 2003). Although the parameters have physical meaning and should be measured in the field, this is sometimes impossible in practice due to their temporal and spatial variability and to time and budget limitations. As a result, either calibration or some form of uncertainty quantification have to be performed, thus reducing the effectiveness of the model predictions. Among these models, we can mention LISEM (Roo and Jetten, 1999), WEPP (Soto and Fierros, 1998), EUROSEM (Morgan et al., 1998), SHESED (Wicks and Bathurst, 1996) (which is based on the previous SHE hydrological model (Abbott et al., 1986a,b)), DHSVM (Wigmosta et al., 1994; Doten et al., 2006), TOPKAPI (Ciarapica and Todini, 2002; Sinclair and Pegram, 2013) and GEOTOP (Rigon et al., 2006; Simoni et al., 2008; Endrizzi et al., 2014). While including rather sophisticated descriptions of soil erosion, all these models use simplified equations for the surface run-off, require *a priori* identification of rivers and drainage zones and are not capable of handling domains in which larger water bodies are present, such as lakes or estuaries. Moreover, they usually need for substantial pre-processing of the orography data.

In this work, we propose a multi-event, time adaptive basin-scale numerical model which couples the superficial runoff dynamics with the soil erosion modeling and is able to automatically detect drainage zones. With respect to other models already available in the literature, the proposed model requires few input parameters, in particular, is equipped with a geostatistical preprocessor, fully described in Gatti et al. (2021), which allows to downscale to fine resolution meshes the

* Corresponding author.
E-mail address: federico.gatti@polimi.it (F. Gatti).

https://doi.org/10.1016/j.cageo.2023.105362
Received 26 July 2022; Received in revised form 30 March 2023; Accepted 6 April 2023
Available online 5 May 2023

soil composition data available in global databases like SoilGrids (Hengl and al., 2017), thus enabling to perform estimations even in absence of detailed geological surveys.

The runoff model equations are approximated numerically using a very well tested, efficient and robust semi-implicit discretization method (Casulli, 1990; Casulli and Cheng, 1992). The proposed technique yields an effective spatially distributed model, able to handle automatically the wide range of transients that can arise in long term simulations and to run even at high resolution with time step values that are dictated by accuracy rather than stability reasons. Furthermore, the numerical discretization guarantees exact mass conservation, positivity of the water and sediment layers and consistency of the discretized equations for tracers with that for water mass, according to the prescriptions in Gross et al. (2002). The model has been implemented in an object oriented programming language and is equipped with a simple user interface. The results of the first verification and validation experiments confirm the model efficiency, robustness and flexibility and demonstrate its potential role as the basis for a more comprehensive simulation tool with uncertainty quantification capabilities.

The structure of the paper is the following. The governing equations of the proposed model are introduced in Section 2, while their discretization together with the adopted time stepping procedure is presented in Section 3. A first attempt at verification and validation of the model is presented in Section 4, finally some conclusions and some perspectives for future developments are discussed in Section 5.

2. Mathematical model

In this section, we describe the governing equations of the proposed basin scale model, which can be seen as an improvement and extension of the model proposed in Bemporad et al. (1997). In the following, to clarify the presentation, we express some modeling quantities in the International System of Units (SI) (Taylor, 1995).

We consider a cartesian domain $\Omega = [0, L_x] \times [0, L_y]$ which contains a sub-domain $\Omega_b \subset \Omega$ that represents the basin under study. This sub-domain is usually identified by geometric and hydrologic considerations. Consider now the drainage sub-domain $\Omega_d \subset \Omega_b$, whose extension varies in time and which is only implicitly defined as the portion of Ω_b where we have a strictly positive surface water layer H. Notice that $H = \eta - b$, where b denotes the orography profile defined on Ω and η is the height of water free surface. For $\mathbf{x} \in \Omega_d$, we model the motion of the surface water layer by the de Saint-Venant equations

$$\begin{aligned} &\partial_t H + \nabla \cdot (H\mathbf{u}) = (1-\mu)p - f, \\ &\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = -g\nabla\eta - \gamma(\mathbf{u})\mathbf{u}. \end{aligned} \tag{1}$$

Here $\mathbf{u}$ is the surface water velocity, g is the gravitational field and γ is a friction term to be defined later. The source term p denotes the precipitation intensity, which is corrected to account for snow by the non-dimensional parameter μ. This parameter takes the value of 1 if the ground temperature is lower or equal than the melting temperature T_m and 0 if it is higher. The infiltration term f represents the exchange of water mass between the surface layer and the gravitational layer and will be defined later in this section.

Note that the equation for the surface water layer thickness is in divergence form, which guarantees exact mass conservation if appropriate numerical methods are employed. Note also that we assume that the orography profile is not changing in time, so that $\partial_t H = \partial_t \eta$. This simplification is justified in the limit of thin sediment layers, while bed evolution can be taken into account, if necessary, by a decoupled approach, see e.g. Garegnani et al. (2013). Notice that, the equations above are valid also for $\mathbf{x} \in \Omega_b \backslash \Omega_d$, i.e. in dry conditions. They modify assuming zero horizontal velocities, so that in this region one has

$$\begin{aligned} &\partial_t H = (1-\mu)p - f, \\ &\mathbf{u} = \mathbf{0}. \end{aligned} \tag{2}$$

In this way, drainage zones are automatically identified while preserving mass conservation. Then, on the wet–dry interface $\partial\Omega_d$, the outflow or zero mass flux boundary conditions are automatically imposed depending on local orography profile. Concerning the friction coefficient γ, it is defined by the Manning-Strickler formula

$$\gamma(\mathbf{u}) = \frac{g\, n^2}{H^{\frac{4}{3}}} |\mathbf{u}|, \tag{3}$$

where n is the Manning friction coefficient, see e.g. Chanson (1999). The appropriate value of this coefficient for steep slopes and mountain torrents has been discussed in a series of papers Chiari and Rickenmann (2007), Rickenmann (1994), Rickenmann (2001), where a number of corrections to the values typically employed in river hydraulics have been proposed. Following Rickenmann (1994), we will consider the formula

$$\begin{aligned} \frac{1}{n_r} &= \frac{0.56\, g^{0.44} (H|\mathbf{u}|)^{0.11}}{|\nabla b|^{0.33} (Rd_{90})^{0.45}} \quad \text{for} \quad |\nabla b| > 0.6\%, \\ \frac{1}{n_r} &= \frac{2.73\, g^{0.49} (H|\mathbf{u}|)^{0.03}}{|\nabla b|^{0.08} (Rd_{90})^{0.24}}, \quad \text{for} \quad |\nabla b| \le 0.6\%, \end{aligned} \tag{4}$$

where d_{90} denotes the value such that 90% of the soil sediment has grain diameter $d \le d_{90}$. Notice that we have introduced a roughness coefficient R that multiplies d_{90} in the formulation above, to be used to account for model incompleteness and intrinsic model errors. This is the only empirical parameter that is not directly derived from data but left instead to be determined by model calibration experiments. Here we consider the roughness coefficient as a scalar coefficient, but in general, it could be assumed to be a spatially varying function to improve the accuracy of the results. In order to allow for regions with small values of the slope, the Manning coefficient actually used in the model is defined as $n = \max\{n_r, n_{\min}\}$, where a minimum value $n_{\min}$ is set by the user depending on the specific application. For all the simulations presented in this article, we use $g = 9.81$ m s^{-2} and $n_{\min} = 0.01$s m$^{-\frac{1}{3}}$.

The model is then completed by a number of equations for the time evolution of the equivalent depths of other two-dimensional, depth-averaged water and sediment layers, all of which are defined for $\mathbf{x} \in \Omega_b$. More specifically, we consider a snow layer with equivalent depth h_{sn}, a gravitational layer with equivalent depth h_g and a sediment layer with equivalent depth h_{sd}. For each of these layers, conservation of mass is assumed, so as to obtain

$$\begin{aligned} \partial_t h_{sn} &= \mu p - s, \\ \partial_t h_{sd} + \nabla \cdot \mathbf{f}_{sd}(h_{sd}, \mathbf{u}) &= w, \\ \partial_t h_g + \nabla \cdot \mathbf{f}_g(h_g) &= s + f - ev. \end{aligned} \tag{5}$$

Here $\mathbf{f}_{sd}, \mathbf{f}_g$ denote the sediment and gravitational layer fluxes, while p, s, f denote, as before, the precipitation, snowfall and infiltration rates, respectively, while ev represents the evapotraspiration rate and w is the sediment source rate. For the horizontal mass fluxes and the mass exchanges among layers, relatively simple models are employed in the present formulation, which will now be discussed in greater detail, starting from the topmost layer. Each of these could however be replaced by more sophisticated approaches, the only limitations being in practice the available data and the implied computational cost.

The atmospheric component is not modeled directly, but it is rather considered as a reservoir of infinite capacity. Water leaves this reservoir through precipitation, in form of snow or rain, which is characterized by intensity, duration and spatial distribution. On the other hand, water may enter back the atmospheric layer via evapotranspiration. Precipitation can take the form of rain or snow, depending on the surface temperature. Rain occurs if the temperature is higher then the melting threshold of $T_m = 2$ °C. In this case, water is assumed to end up in the surface run-off layer. In the opposite case, precipitation takes the form of snow and is being accumulated at the surface until temperature reaches values high enough to cause melting. We use a linear relationship between temperature and orography height, based

on the International Standard Atmosphere, see e.g. Davies (2003). The snow melting rate s, is computed according to the degree-day approach (Idso et al., 1977; Day et al., 2006)

$$s = \delta(T - T_m), \tag{6}$$

where δ is a parameter that determines the amount of snow that melts in one day at a given temperature T. The evapotranspiration rate ev, expressed in [m s^{-1}], is modeled via the Hargreaves equation (Hargreaves and Allen, 2003), given by

$$ev = 0.0023\ R_a(\bar{T} + 17.8)\ \sqrt{T_{\max} - T_{\min}}. \tag{7}$$

Here, R_a is the water equivalent extraterrestrial radiation, $\bar{T}$ is the mean air temperature, $T_{\max}, T_{\min}$ the daily maximum and minimum air temperatures in [°C], respectively. The extraterrestrial radiation, R_a, expressed in [MJ m^{-2} day^{-1}], for each day of the year and for different latitudes can be estimated from the solar constant, the solar declination and the time of the year by

$$R_a = \frac{24 \cdot 60}{\pi} G_{sc}\ d_r(\omega_s\ \sin\phi\ \sin\delta + \cos\phi\ \cos\delta\ \sin\omega_s), \tag{8}$$

where the solar constant is given by $G_{sc} = 0.0820$ MJ m^{-2} min^{-1}, d_r denotes the inverse of the Earth-Sun distance and the sunset hour angle is given by $\omega_s = \arccos(-\tan\phi\ \tan\delta)$, where δ is the solar declination and ϕ is the latitude.

The sediment flux is expressed as a function of the surface run-off velocity and of the local terrain slope. This dependency is modeled following the proposal in Zhang et al. (2009) as

$$\mathbf{f}_{sd} = \alpha\ |\nabla b|^{\beta}\ h_{sd}\mathbf{u}, \tag{9}$$

where α and β are empirical coefficients. Possible values for these coefficients are $\alpha = 2.5$ and $\beta = 1.6$, as suggested in Smart and Jäggi (1983). In this model, the critical shear stress is set to zero, so that the sediment movement begins simultaneously with the water movement. The sediment source term w, expressed in [m s^{-1}], is defined according to the Gavrilović method (Bemporad et al., 1997; Gavrilović, 1988; Globevnik et al., 2003). It corresponds to the rate of sediment production due to erosive processes as a result of precipitation and it is computed as

$$w = \pi\ (1 - \mu)\ p\ \tau_g\ Z^{3/2}. \tag{10}$$

Terms τ_g and Z are empirical coefficients of the Gavrilović method that depend on temperature and land use, respectively, while p is the precipitation intensity in [mm s^{-1}]. Although the Gavrilović method gives results on a yearly basis, it is assumed that it is also valid for shorter periods in which w will be interpreted as an intensity.

The gravitational layer is the soil portion in which water can move due to gravitational forces. This movement is governed mainly by the permeability of the soil. For the present version of the model, the same, extremely simplified description of the subsurface mass fluxes is adopted as in Bemporad et al. (1997). More specifically, the horizontal mass fluxes in the gravitational layer $\mathbf{f}_g$, expressed in [m^2 s^{-1}], are defined as

$$\mathbf{f}_g(h_g) = h_g\mathbf{u}_g, \tag{11}$$

where $\mathbf{u}_g$ represents the water velocity vertically averaged over the layer. This velocity is modeled as $\mathbf{u}_g = \beta_g(h_g, \mathbf{x})\mathbf{n}$, where β_g is the hydraulic conductivity, a function of the soil permeability and of the water level in the layer, while $\mathbf{n} = -\nabla b/|\nabla b|$. We estimate the hydraulic conductivity from the characteristic diameters of the soil particles, using the Hazen model, see Hazen (1892). This yields

$$\beta_g = C_H \frac{g}{\nu} d_{10}^2, \tag{12}$$

where C_H is a non dimensional coefficient with a reference value $6.54 \cdot 10^{-4}$, ν is the water kinematic viscosity ($\nu = 0.89 \cdot 10^{-6}$ m^2 s^{-1} at 25°C) and d_{10} denotes the value such that 10% of the soil sediment has grain diameter $d \le d_{10}$. This very crude model will be replaced in future implementations by a more accurate one based on vertically averaged Darcy equations coupled to the surface layer flow, along the lines proposed in several papers, see e.g. Casulli (2015, 2017), Discacciati et al. (2002), Miglio et al. (2003).

To describe the water infiltration f from the surface run-off layer to the gravitational layer we use the well known SCS-CN or curve number method, see e.g. Ponce and Hawkins (1996). This approach has been very widely applied in practice, also to areas similar to those that will constitute our main case study, see e.g. Rosso (2004). In particular, to enable the multi-event simulation, we propose a recasting of the classical SCS-CN model, i.e. a recasting in terms of rates rather than cumulative quantities. Refer to the Appendix for the derivation of the infiltration law used in this work.

3. Numerical framework

In this section, we describe first the space discretization employed in the solution of the model equations presented in the previous section, and then describe the time-adaptation procedure employed for the time step selection.

3.1. Space discretization

The de Saint-Venant equations (1) are discretized via a classical semi-implicit approach proposed first in Casulli (1990) and Casulli and Cheng (1992) and subsequently applied to a large number of coastal and river circulation models. In this way the Courant number, which characterizes the numerical stability of the scheme, depends only on the surface velocity $|\mathbf{u}|$ and not on the celerity $\sqrt{gH}$, which can be much larger than $|\mathbf{u}|$ if water bodies of significant depth are present in the simulation domain. This allows to use larger values of the time step compared to those required to classical explicit schemes, see e.g. LeVeque (1992). An application to two-dimensional sediment transport in rivers is presented in Rosatti et al. (2005), while a one-dimensional, section averaged model for river hydraulics based on the same approach is presented in Garegnani et al. (2013) and Rosatti et al. (2011). Even if higher order variants in space and time of this approach can be devised, in this work we will stick to the simplest, first order formulation, since high order formal accuracy is less relevant than robustness for the target applications.

We consider a Cartesian structured mesh with resolution Δx, Δy discretizing the domain $\Omega = [0, L_x] \times [0, L_y]$. In particular, we denote as computational domain the set formed by the cells providing a discretization of the basin domain $\Omega_b \subset \Omega$, here Δt is the given time step. At a generic time n and in a generic cell (i, j), we consider a staggered variable arrangement, with discrete velocity variables $u^n_{i+\frac{1}{2},j}$ and $v^n_{i,j+\frac{1}{2}}$ defined at half integer locations and discrete variables $\eta^n_{i,j}, H^n_{i,j}$ defined at integer locations. Wet cells are defined as those for which $H^n_{i,j} > 0$. These cells provide a discretization of the drainage domain Ω_d. The space and time discretization is given for all basin cells and corresponding edges by

$$\begin{aligned}\eta^{n+1}_{i,j} = \eta^n_{i,j} &- \frac{\Delta t}{\Delta x}\left[H^n_{i+\frac{1}{2},j}\, u^{n+1}_{i+\frac{1}{2},j} - H^n_{i-\frac{1}{2},j}\, u^{n+1}_{i-\frac{1}{2},j}\right] \\ &- \frac{\Delta t}{\Delta y}\left[H^n_{i,j+\frac{1}{2}}\, v^{n+1}_{i,j+\frac{1}{2}} - H^n_{i,j-\frac{1}{2}}\, v^{n+1}_{i,j-\frac{1}{2}}\right] \\ &+ \Delta t\,(1-\mu)\, p^n_{i,j} - \Delta t\, f^n_{i,j},\end{aligned} \tag{13}$$

$$u^{n+1}_{i+\frac{1}{2},j} = \mathcal{F}u^n_{i+\frac{1}{2},j} - g\frac{\Delta t}{\Delta x}(\eta^{n+1}_{i+1,j} - \eta^{n+1}_{i,j}) - \Delta t\gamma^n_{i+\frac{1}{2},j}\, u^{n+1}_{i+\frac{1}{2},j}, \tag{14}$$

$$v^{n+1}_{i,j+\frac{1}{2}} = \mathcal{F}v^n_{i,j+\frac{1}{2}} - g\frac{\Delta t}{\Delta y}(\eta^{n+1}_{i,j+1} - \eta^{n+1}_{i,j}) - \Delta t\gamma^n_{i,j+\frac{1}{2}}\, v^{n+1}_{i,j+\frac{1}{2}}. \tag{15}$$

Here $\mathcal{F}u^n_{i+\frac{1}{2},j}$, $\mathcal{F}v^n_{i,j+\frac{1}{2}}$ denote some explicit discretization of the momentum advection terms. In this work, a first order semi-Lagrangian

method with bilinear interpolation is employed. Furthermore, the water layer depths $H^n_{i+\frac{1}{2},j}$, $H^n_{i,j+\frac{1}{2}}$ are defined in an upwind fashion, so that

$$H^n_{i+\frac{1}{2},j} = H^n_{i+1,j} \quad \text{if} \quad u^{n+1}_{i+\frac{1}{2},j} < 0, \quad H^n_{i+\frac{1}{2},j} = H^n_{i,j} \quad \text{if} \quad u^{n+1}_{i+\frac{1}{2},j} > 0, \tag{16}$$

as suggested in Gross et al. (2002). While the method defined by (13) is only first order in time and space, the resulting discretization is very robust and stable and allows to employ relatively long time steps in most flow regimes. The practical solution of Eqs. (13) is achieved as follows. The equations for $u^{n+1}_{i+\frac{1}{2},j}$, $v^{n+1}_{i,j+\frac{1}{2}}$ are first rewritten as

$$u^{n+1}_{i+\frac{1}{2},j} = \alpha^n_{i+\frac{1}{2},j} \mathcal{F} u^n_{i+\frac{1}{2},j} - g\alpha^n_{i+\frac{1}{2},j} \frac{\Delta t}{\Delta x}(\eta^{n+1}_{i+1,j} - \eta^{n+1}_{i,j}), \tag{17}$$

$$v^{n+1}_{i,j+\frac{1}{2}} = \alpha^n_{i,j+\frac{1}{2}} \mathcal{F} v^n_{i,j+\frac{1}{2}} - g\alpha^n_{i,j+\frac{1}{2}} \frac{\Delta t}{\Delta y}(\eta^{n+1}_{i,j+1} - \eta^{n+1}_{i,j}), \tag{18}$$

where

$$\alpha^n_{i+\frac{1}{2},j} = \frac{1}{1+\Delta t\gamma^n_{i+\frac{1}{2},j}}, \quad \alpha^n_{i,j+\frac{1}{2}} = \frac{1}{1+\Delta t\gamma^n_{i,j+\frac{1}{2}}}, \tag{19}$$

respectively. These equation are then substituted into the equation for $\eta^{n+1}_{i,j}$, so as to obtain, for the interior nodes, the equations

$$\begin{aligned}\eta^{n+1}_{i,j} = \mathcal{G}^n_{i,j} &- g\frac{\Delta t^2}{\Delta x^2}\alpha^n_{i+\frac{1}{2},j} H^n_{i+\frac{1}{2},j}(\eta^{n+1}_{i+1,j} - \eta^{n+1}_{i,j}) + g\frac{\Delta t^2}{\Delta x^2}\alpha^n_{i+\frac{1}{2},j} H^n_{i-\frac{1}{2},j}(\eta^{n+1}_{i,j} - \eta^{n+1}_{i-1,j}) \\ &- g\frac{\Delta t^2}{\Delta y^2}\alpha^n_{i,j+\frac{1}{2}} H^n_{i,j+\frac{1}{2}}(\eta^{n+1}_{i,j+1} - \eta^{n+1}_{i,j}) + g\frac{\Delta t^2}{\Delta y^2}\alpha^n_{i,j-\frac{1}{2}} H^n_{i,j-\frac{1}{2}}(\eta^{n+1}_{i,j} - \eta^{n+1}_{i,j-1}),\end{aligned} \tag{20}$$

where

$$\begin{aligned}\mathcal{G}^n_{i,j} = \eta^n_{i,j} &+ \Delta t(1-\mu)p^n_{i,j} - \Delta t f^n_{i,j} - \frac{\Delta t}{\Delta x}\left[H^n_{i+\frac{1}{2},j}\alpha^n_{i+\frac{1}{2},j}\mathcal{F}u^n_{i+\frac{1}{2},j}\right. \\ &\left. - H^n_{i-\frac{1}{2},j}\alpha^n_{i-\frac{1}{2},j}\mathcal{F}u^n_{i-\frac{1}{2},j}\right] \\ &- \frac{\Delta t}{\Delta y}\left[H^n_{i,j+\frac{1}{2}}\alpha^n_{i,j+\frac{1}{2}}\mathcal{F}v^n_{i,j+\frac{1}{2}} - H^n_{i,j-\frac{1}{2}}\alpha^n_{i,j-\frac{1}{2}}\mathcal{F}v^n_{i,j-\frac{1}{2}}\right].\end{aligned} \tag{21}$$

Eq. (20) can be rewritten as

$$\begin{aligned}&\left[1 + g\frac{\Delta t^2}{\Delta x^2}\left(\alpha^n_{i+\frac{1}{2},j}H^n_{i+\frac{1}{2},j} + \alpha^n_{i-\frac{1}{2},j}H^n_{i-\frac{1}{2},j}\right)\right. \\ &\left. + g\frac{\Delta t^2}{\Delta y^2}\left(\alpha^n_{i,j+\frac{1}{2}}H^n_{i,j+\frac{1}{2}} + \alpha^n_{i,j-\frac{1}{2}}H^n_{i,j-\frac{1}{2}}\right)\right]\eta^{n+1}_{i,j} \\ &- g\frac{\Delta t^2}{\Delta x^2}\alpha^n_{i+\frac{1}{2},j}H^n_{i+\frac{1}{2},j}\eta^{n+1}_{i+1,j} - g\frac{\Delta t^2}{\Delta x^2}\alpha^n_{i-\frac{1}{2},j}H^n_{i-\frac{1}{2},j}\eta^{n+1}_{i-1,j} \\ &- g\frac{\Delta t^2}{\Delta y^2}\alpha^n_{i,j+\frac{1}{2}}H^n_{i,j+\frac{1}{2}}\eta^{n+1}_{i,j+1} - g\frac{\Delta t^2}{\Delta y^2}\alpha^n_{i,j-\frac{1}{2}}H^n_{i,j-\frac{1}{2}}\eta^{n+1}_{i,j-1} = \mathcal{G}^n_{i,j}.\end{aligned} \tag{22}$$

These equations, defined for all computational cells, define a linear system whose matrix is symmetric, positive definite and diagonally dominant, thus guaranteeing the possibility of a fast and accurate numerical solution via efficient numerical linear algebra solvers. We notice that, from a numerical point of view the positivity constraint on the water depth in most of the cases could be too restrictive, we then relax this condition by considering a tolerance $H_{\min}$ which can be user defined and corresponds to the maximum violation of the positivity constraint in each cell. In other words, we are numerically extending the dry state condition up to the threshold $H_{\min}$. All the other equations in (5) are discretized by a classical finite volume approach over each cell of the computational domain, with first order upwind definitions of the numerical fluxes $f_{i\pm\frac{1}{2},j}$, $f_{i,j\pm\frac{1}{2}}$ and explicit Euler time discretization, see e.g. LeVeque (1992). In particular, while the time scales of the gravitational layer is generally mush slower than the one of the superficial run-off, it is not always true for the sediment transport equation. This means that generally, for the sediment equation, more than one time step must be performed for a given step of the superficial run-off.

3.2. Time adaptation

Following the paper Porta et al. (2012), we adopt a recovery based adaptation procedure which provides an estimation of the time step based on an *a posteriori* error estimator. The main idea of the recovery approach is to consider a suitable projection of the numerical gradient onto richer spaces obtained by means of interpolations and averaging procedures, see Zienkiewicz and Zhu (1992a,b). We note that, this time adaptation procedure has already been successfully applied by some of the authors in the framework of fast landslides simulations (Gatti et al., 2023a,b).

To obtain a prediction of the time step, we need an estimator of the discretization error. Let us consider the water depth H, and let $H_{i,j}$ be the corresponding discretized counterpart in the given cell (i,j), where we assume to know its values $H^k_{i,j}$ in previous times $t^k, k \in \{0,\dots,n-1\}$. Considering a generic time t^n and a generic coordinate $\mathbf{x} \in \Omega$, we can write the H^1-seminorm of the discretization error as

$$|e^z_{i,j}(\mathbf{x})|^2_{H^1(\Delta t^{n-1})} = \int_{I_{n-1}} |\partial_t H - \partial_t H_{i,j}|^2 \, dt, \tag{23}$$

where $I_{n-1} = [t^{n-1}, t^n]$. We need to replace the function $H_{i,j}$ on the interval I_{n-1} with its linear interpolation, using the pointwise values $H^k_{i,j}$, with $k \in \{0,\dots,n\}$. This allows us to approximate the time derivative of the discrete solution $\partial_t H_{i,j}$ in the interval I_{n-1} as

$$\partial_t H_{i,j} \simeq \frac{H^n_{i,j} - H^{n-1}_{i,j}}{\Delta t^{n-1}}. \tag{24}$$

Once computed the above quantity, the time derivative of the exact solution $\partial_t H$, is replaced by the recovered solution H^* computed by quadratic interpolating three couples $(t^{n-2}, H^{n-2}_{i,j})$, $(t^{n-1}, H^{n-1}_{i,j})$ and $(t^n, H^n_{i,j})$. The cell-wise estimator then reads

$$(\eta^2_{I_{n-1}})_{i,j} = \tilde{T}\int_{I_{n-1}} \left|\partial_t H^* - \frac{H^n_{i,j} - H^{n-1}_{i,j}}{\Delta t^{n-1}}\right|^2 dt, \tag{25}$$

where $\tilde{T}$ is a scale factor, here we choose Δt^{n-1} as suggested in Porta et al. (2012). We notice that, the integral above can be computed exactly, since the functions $\partial_t H^*$ and $\partial_t H_{i,j}$ are respectively linear and constant polynomials of time in each interval I_{n-1}. Since the above estimator is a cell-wise contribution, we can get a global indicator by summing all of these contributions on the mesh. Hence, considering I_b the set of all indices of the cells providing a discretization of the basin domain, in each interval I_{n-1}, the estimator reads

$$(\eta^T_{I_{n-1}})^2 = \sum_{(i,j)\in I_b} (\eta^2_{I_{n-1}})_{i,j}. \tag{26}$$

Now we are able to compute the new time step Δt_n. For this end, we fix a tolerance $\tau^{\Delta t}$ for the local estimator, i.e. $\eta^T_{I_{n-1}} = \tau^{\Delta t}$, which depends on the problem at hand, and so we have

$$\Delta t_n = \frac{\tau^{\Delta t}\,\Delta t_{n-1}}{\eta^T_{I_{n-1}}}. \tag{27}$$

It is important to fix a limit value $\Delta t_{\max}$ and compare it with the Δt_n computed in (27). For the first two time steps we can simply set $\Delta t_{\max} = \Delta t_0 = \Delta t_1$. The $\Delta t_{\max}$ is chosen from numerical stability reasons, i.e. based on the Courant number and we set a $\Delta t_{\min}$ to be half the maximum time step $\Delta t_{\max}$. We point out that, in the implementation, we have included an a-posteriori control on the maximum violation of the positivity constraint $H_{\min}$, as defined previously, so to use a time step smaller than $\Delta t_{\min}$ in case is needed for conservation purposes. In this way, if for a given time step we have at least a cell under the threshold $\Delta t_{\min}$ we consider new time step one order of magnitude lower than the current one and we solve again the run-off linear system.

4. Numerical experiments

In this section, we present two sets of numerical experiments, aimed to provide a first verification and validation of the basin scale model presented in Section 2. We consider first two cases with idealized orographies, aimed at a first verification of the correctness of the model implementation. We then study a more realistic test case, corresponding to the Caldone catchment (Ivanov et al., 2020a). For all the tests, we assume $g = 9.81$ m s^{-2}, a minimum value of the Manning coefficient of $n_{\min} = 0.01$ s m$^{-\frac{1}{3}}$. Regarding the roughness coefficient if not said otherwise we consider $R = 1$. We then consider a tolerance on the positivity constraint equal to $H_{\min} = 10^{-5}$ m for the ideal tests and equal to $H_{\min} = 10^{-3}$ m for the simulations with realistic orography. We run the ideal tests on a laptop with an Intel i7 CPU having 2.60 GHz clock frequency and 16 GB of RAM, while we perform the realistic orography simulations on a HPC cluster of the Department of Mathematics of Politecnico di Milano with 5 nodes each one equipped with 2 Intel Xeon Gold 6238R at 2.20 GHz processors with 512 GB RAM. We use an isotropic space resolution for the following numerical tests, i.e. $\Delta x = \Delta y = \Delta$. For the solution of the linear system for the superficial run-off dynamics we use the *conjugate gradient* method with an *incomplete Cholesky factorization* preconditioner, see e.g. Kershaw (1978). The library used for the solution of the linear system is the Eigen library, see Guennebaud et al. (2010), considering a tolerance equal to 10^{-6}. For the time adaptation procedure we consider a tolerance $\tau^{\Delta t} = 10^{-5}$ m.

The model can compute values of quantities of interest such as the water height, the water discharge and solid discharge in given geographical locations. We represent the geographical location in the discrete model as a cell superimposed to the computational mesh with a user-defined resolution we call Δ_p. We define the set of all intersecting mesh cells with $\mathcal{I}$ and we indicate with N the number of elements of this set. In this way, the water height in the geographical location is defined as $\bar{H}(t) = \frac{1}{N}\sum_{Q\in\mathcal{I}} H_Q$, the water discharge as $\bar{Q}(t) = \Delta\frac{1}{N}\sum_{Q\in\mathcal{I}} H_Q|\mathbf{u}_Q|$ and the solid discharge as $\bar{Q}_s(t) = \Delta\alpha S^\beta \frac{1}{N}\sum_{Q\in\mathcal{I}} h_{sd,Q}|\mathbf{u}_Q|$, H_Q being the discrete superficial water value in the cell Q, $\mathbf{u}_Q$ being the discrete superficial velocity in the cell Q, S being the slope computed in the geographical point from the DTM, $h_{sd,Q}$ being the discrete sediment height in the cell Q.

4.1. Data preprocessing

For practical applications of the model described in the previous sections, several input datasets must be made available, including, in particular, orography data and soil composition data. The former is usually provided by easily accessible Digital Terrain Models (DTM). The preprocessing of soil composition data requires instead a special treatment, since the relevant data are rarely available at the scales of interest for the application of the proposed model. For this reason, a geostatistical preprocessor has been developed, which is described in more detail in Gatti et al. (2021), which allows to downscale to fine resolution meshes the soil composition data available in global databases such as SoilGrids, see Hengl and al. (2017). The key concepts underlying this preprocessor will be briefly reviewed here.

The potential maximum soil moisture retention is a function of soil texture, which can in turn be determined from particle-size fractions, i.e. the relative percentages of clay, silt and sand in the soil, see e.g. Matthews (2013). Since these are compositional data such as the sum particle-size fractions is always equal to one, well established statistical approaches have been applied to avoid spurious correlations among the recovered ratios (Aitchison, 1982; Kim, 1999). The particle-size fractions are transformed via an Isometric Log Ratio transform (Martín-Fernández et al., 2012), so as to be able to work within the Aitchison geometry without modifying the classical statistical techniques based on the Euclidean metric. Data coming from larger scale databases are then downscaled to match the desired resolution using the Isometric Log-Ratio Area-To-Point Regression Co-Kriging (ILR-ATPRCoK) as described in Gatti et al. (2021). To perform uncertainty quantification, we notice that, our preprocessor is able to produce stochastic soil maps through Block Sequential Gaussian Simulation (BSGS) (Boucher, 2003) since it relies on the same hypothesis as the ATPRCoK method. We provide an example of Monte Carlo simulation in the following of the present section.

We note that, in case of stochastic simulation, if more than one processor is available in the computer architecture, the current implementation relies on the OpenMPI protocol to split the computations among available cores. Fig. 1 illustrates the special case where we have N MPI ranks available and we need to perform N simulations each one initialized with a different Particle Size Fractions (psf) map generated with BSGS method.

4.2. Idealized orography

Here, we test the proposed model in a set of idealized orography simulations. For simplicity, we assume $\Omega_b = \Omega$ and start from null initial conditions.

We consider a square domain with $L_x = 10$ km, $L_y = 10$ km and with an orographic profile given by the plane $b(x, y) = 2 + 0.0005\,L_y + 0.0005x - 0.0005y$. We consider a computational mesh with $\Delta = 100$ m and a maximum time step $\Delta t = 72$ s, which corresponds to 50 time steps per hour. The soil properties are characterized here by a constant curve number $CN = 79$. In a first simulation, we simulate a total of 20 days, over which a constant and uniform rainfall of 5 mm h^{-1} and apply reflecting boundary conditions, on the water height H, at the boundary of the basin domain.

In Fig. 2 panel (a) we plot the time evolution of the mean surface run-off mass flux and in Fig. 2 panel (b) we plot time evolution of the mean gravitational water content. It can be observed that, as the soil reaches the saturated conditions corresponding to the maximum soil moisture retention, the infiltration rate approaches zero, see Fig. 2 panel (b). Finally, in Fig. 2 panels (c), (d) the spatial distribution of the state variables involved in the surface run-off process is plotted at the final time. Notice that, considering the final time simulation results, we have a celerity based Courant number approximately equal to 6 in the deepest portion of the formed lake. This highlights the ability of our framework to employ a time step much larger than that required by classical explicit time discretizations.

We have then performed a multi-event simulation on the same basin, in order to test the response of our implementation to time dependent rainfall rates. In this case, we apply non-reflecting boundary conditions, on the water height H, at the boundary of the basin domain and we consider the evapotranspiration. As in the previous case, we apply a homogeneous rainfall rate over the whole basin, which is however changing periodically in time with a period of 5 days and duty cycle of 20% as shown in Fig. 3 panel (a), over a total of 365 simulated days. Fig. 3 panel (b) shows the temperature profile input of the simulation necessary in this case to feed the evapotranspiration. Finally, to show the ability of the model to deal with transients, we show in Fig. 3 panels (c), (d) the time evolution of the average water content in the gravitational layer and the maximum soil moisture retention with infiltration rate.

4.3. Realistic orography: Caldone basin

Here we consider a domain corresponding to the Caldone catchment, which is located in the vicinity of the city of Lecco (Northern Italy). For more details on the river catchment structure, we refer to Ivanov et al. (2020a). The hydrographic catchment covers an area of 28 km^2. The land cover of the catchments mainly includes forests (67%), herbaceous vegetation (16%) and urban areas (13%), according to the CORINE Land Cover Classification (CLC) (Büttner et al., 2004;

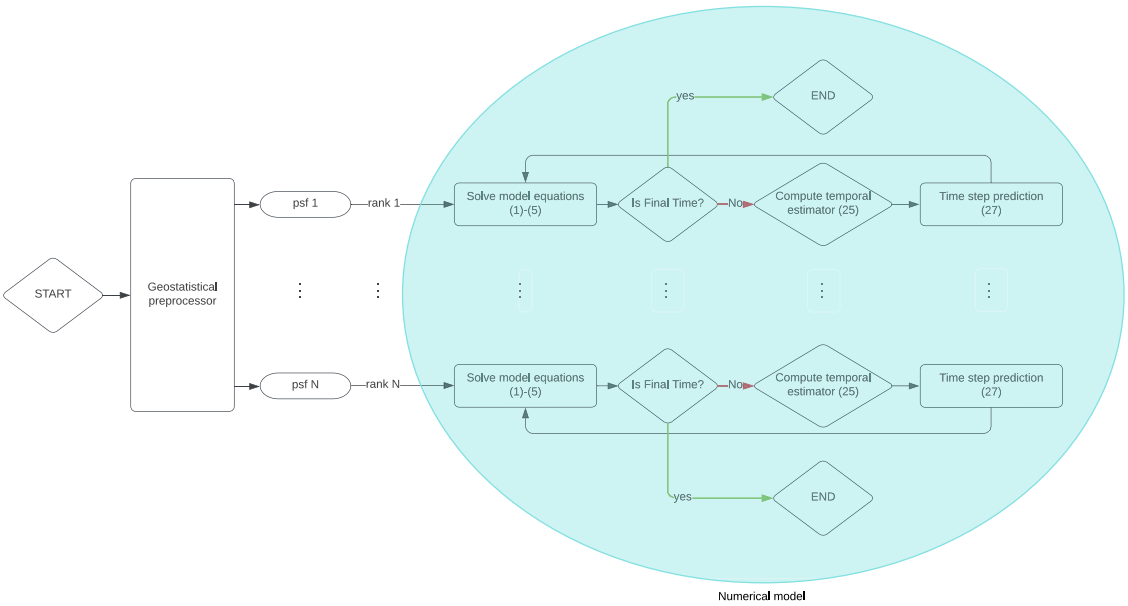

Fig. 1. Flowchart of the complete model workflow process. The numbers refer to the formulas in the manuscript.

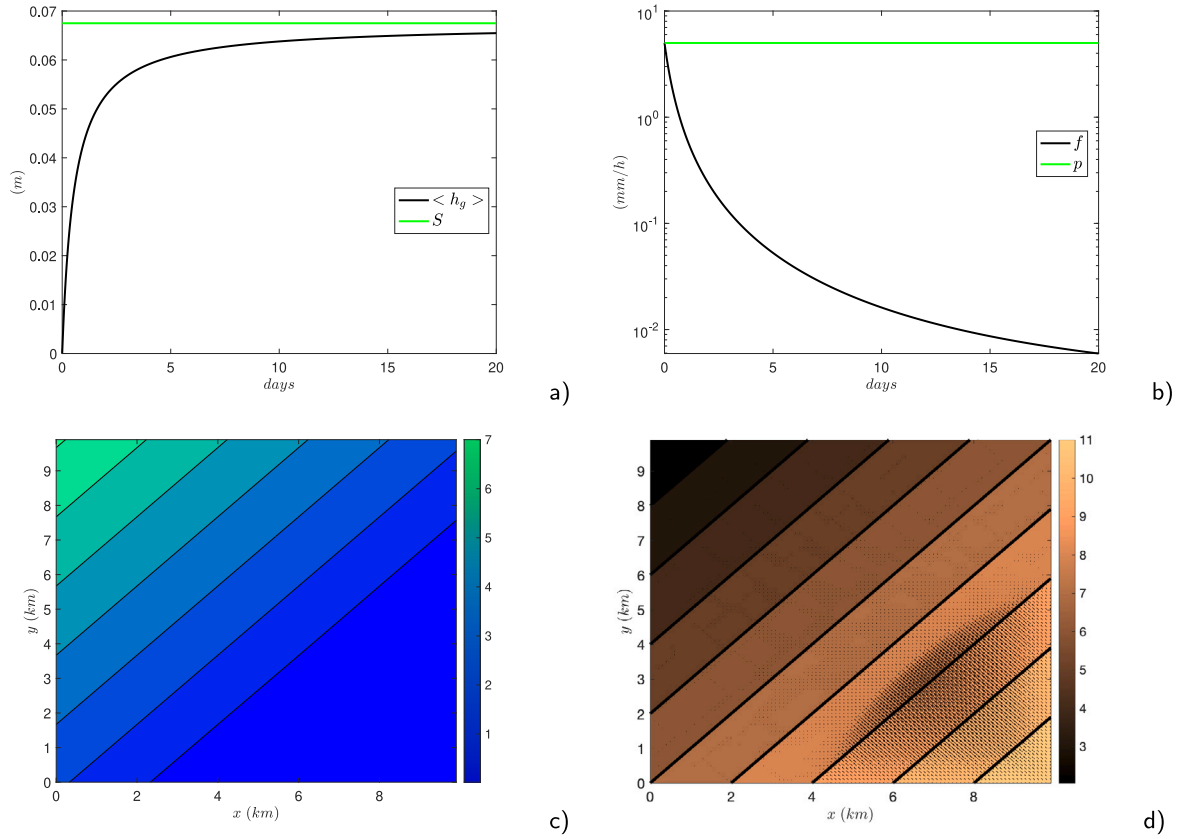

Fig. 2. Idealized test, single event simulation: (a) time evolution of average water content in the gravitational layer $\langle h_g \rangle$ (black) and maximum soil moisture retention (green) (b) time evolution of mean infiltration rate (black) and precipitation rate (green). (c) isolines at final time of water content in the surface layer in meters and (d) surface water velocity field at final time, superimposed onto the orography isolines.

Büttner, 2014). As in most pre-Alpine environments, active geomorphic processes include colluvial and fluvial transport responsible for the yield and further propagation of sediment downstream.

In Fig. 4 panels (a), (b), (c) we represent respectively the DTM, its slope field in degrees and the CLC map for the case study at a resolution of 5 m. In particular, Fig. 4 panel (a) represents the DTM raster. In the same figure, we report in blue the mean location of a tank, i.e. a control area we will use for the sediment transport and, in red, the control point we will use for the superficial water height, i.e. the location where we have hydrometer data. Being the tank located in the river flow path and upstream with respect to the hydrometer, and being the impossibility to add the discretized tank to the DTM (this would have constrained to

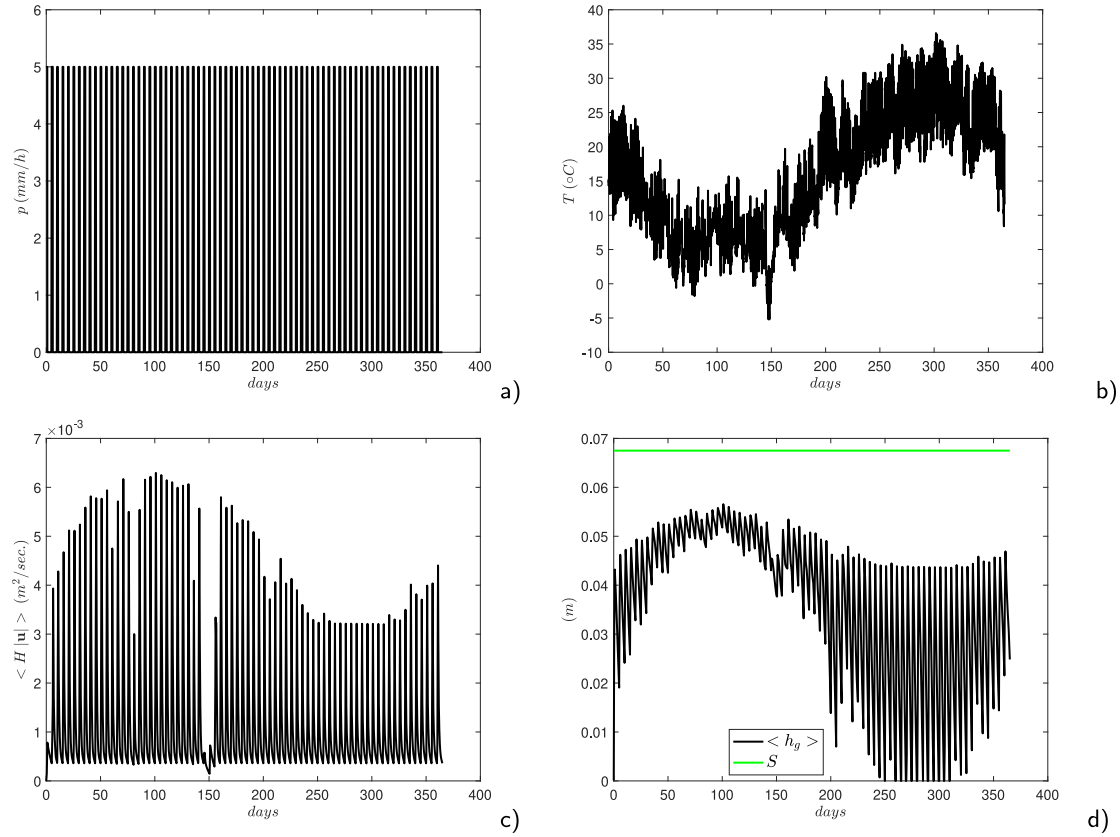

Fig. 3. Idealized test, multi event simulation: time evolution of (a) precipitation rate, it is a square wave; (b) temperature rate; (c) mean surface run-off mass flux $\langle H\,|\mathbf{u}|\rangle$; (d) average water content in the gravitational layer $\langle h_g \rangle$ (black) and maximum soil moisture retention (green).

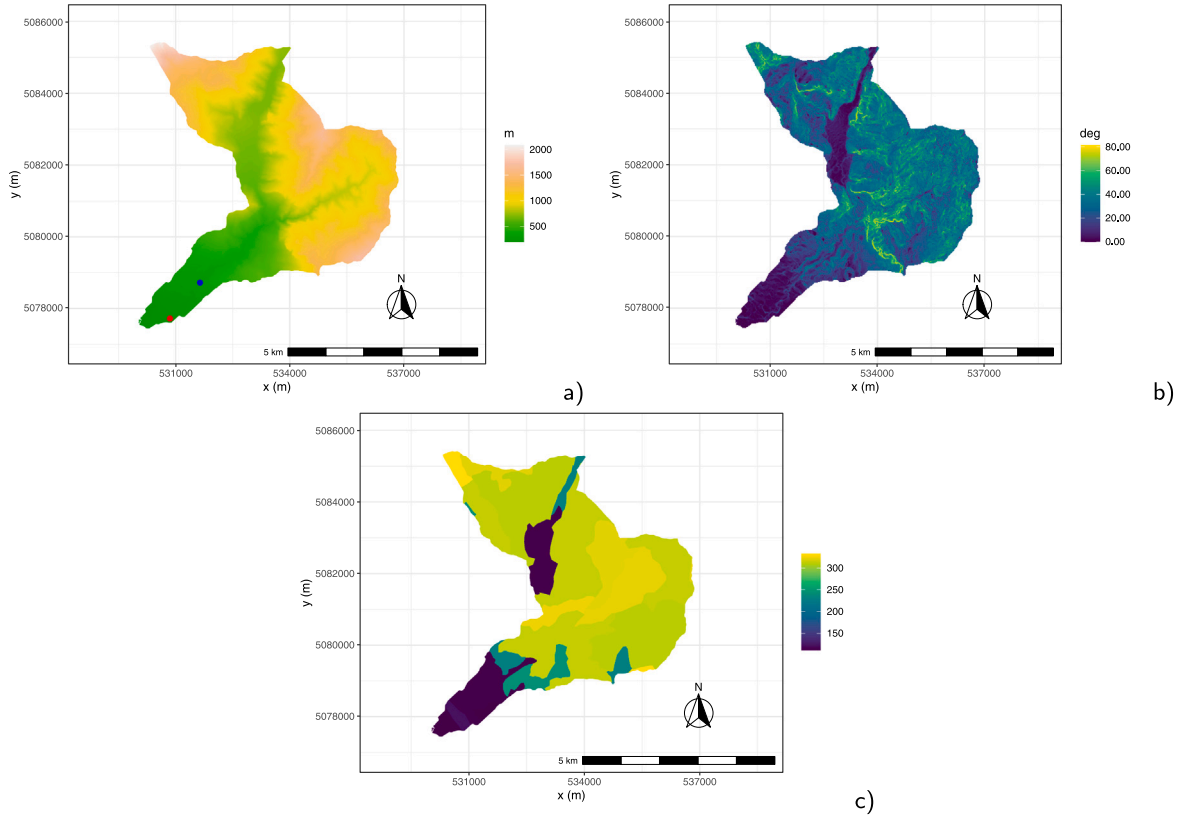

Fig. 4. Caldone basin. (a) 5 m-resolution DTM of the study area, blue point identifies the tank location while the red point is the control point for the superficial water height; (b) slope of 5 m-resolution DTM, i.e. isolines of $\arctan(|\nabla b|)$ expressed in degrees; (c) CLC map.

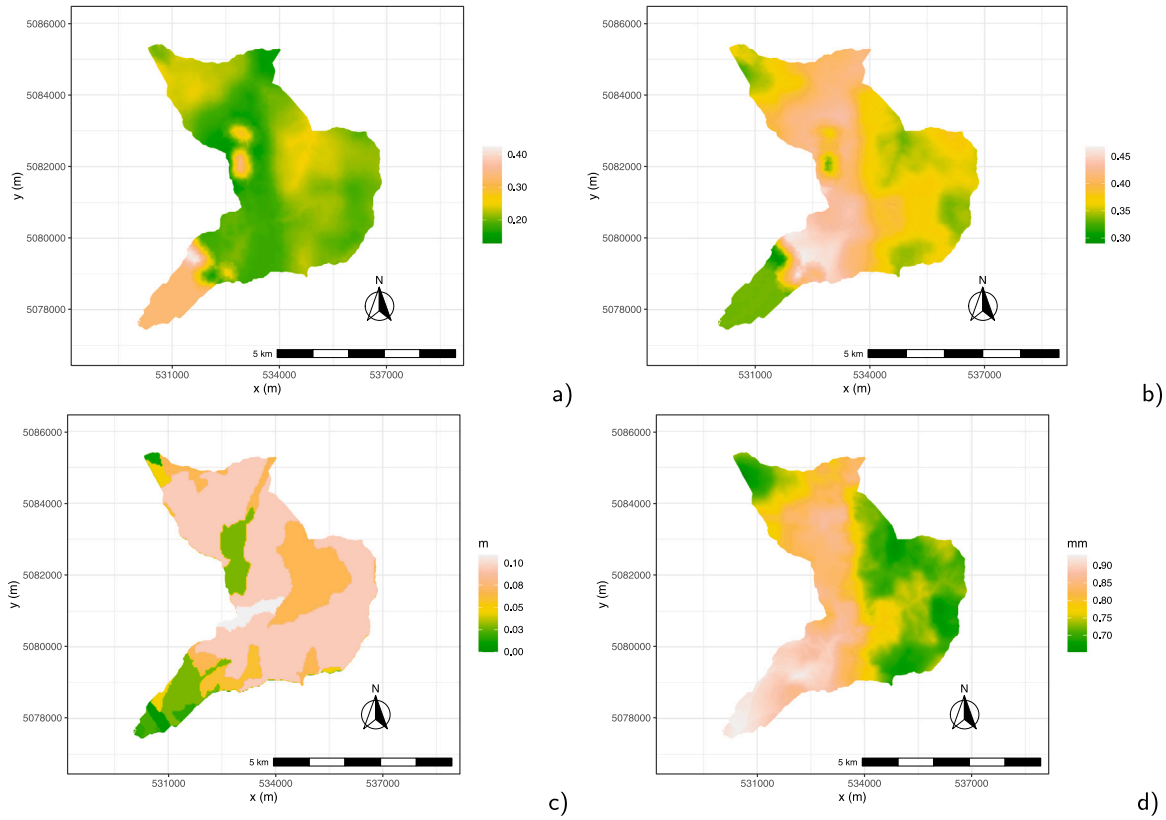

Fig. 5. Caldone basin. Panels (a), (b) Clay and Sand maps resulting from ILR-ATPRCoK; (c) Maximum soil moisture retention map in m in the basin domain; (d) d_{90} in mm. We do not report the silt map because it is complementary of clay and sand map via using the unit-sum constraint of compositional data.

adopt on the whole catchment at least the same minimum resolution to do not filter out the discretized tank and would result in too high computational cost), we consider the red point also in the computation of a solid flux to estimate the solid volume transported downstream by the river. For the red point we define a $\Delta_p = 45$ m.

In the following, we provide both deterministic and stochastic results in order to show the potentiality of the overall implementation and its effectiveness to provide reasonable results without any calibration process.

To avoid to limit the maximum time step, we consider non-null initial conditions for the superficial run-off dynamics coming from a fictitious uniform and constant rain for a duration of roughly one day where we put a maximum time step $\Delta t_{\max} = 7.2$ s. This is just to provide an initial state that is not completely dry and facilitates to bring the global system to the state of river formation. Note that, this however does not affect the final result due to the nature of the model equations that tend to loose the initial conditions with the passing of time.

For all the presented simulations we choose a maximum time step equal to $\Delta t_{\max} = 72$ s which corresponds to a minimum of 50 time steps per hour. The precipitation term p, is a space–time dependent field, in the numerical model the space dependency is obtained via Inverse Distance Weighting (IDW) method, see e.g. Chen and Liu (2012), considering data coming from nine rain gauges stations situated in nine different locations, data are available from the regional hydrological service (Lombardia, 2021).

4.3.1. Deterministic scenarios

The geostatistical preprocessor was applied to obtain maps for psf at a resolution of 5 m, here we consider as psf input maps the result of ILR-ATPRCoK not of the conditional simulation, see Fig. 5 panels (a), (b). The psf data are then combined with the CLC map in order to get the maximum soil moisture retention map, see Fig. 5 panel (c). Due to the computational limitation of performing the simulation at 5 m resolution, in the following we consider three upscaled input DTM with resolutions of $\Delta = 20, 50, 90$ m, as input to the simulation which are obtained by bilinearly interpolating the starting 5 m resolution DTM.

Here, we present two sets of simulations, "short" and "long" term, i.e. respectively one with a period of 35 days with intense rain (intense single event) and the other one is one year simulation, i.e. a multi-event simulation. The aim of these two sets of simulations is to show the ability of the method to deal with both multi and single-intense events in case of a real mountain orography scenario.

In Fig. 6 we show the temporal sequences, for both the event and multi-event simulations, of the superficial water height and water discharge against the observed data coming from a hydrometer located in the observation point as depicted in Fig. 4 panel (a). Note the ability of the implementation to capture the ongoing dynamics: the model is able to identify the timing of the "peaks" in the same temporal location as hydrometer, for both single- and multi-event simulation.

We point out that the results presented above are achieved without any calibration process or back analysis. In Fig. 7, for the case $\Delta = 50$ m and for the intense rain case, we show the sensitivity of the model to a perturbation of the d_{90}, by considering uniform roughness coefficients $R = 10^2, 10^3, 10^4$.

An example of the model capability to reproduce realistic run-off patterns without the need to prescribe run-off regions *a priori* is shown in Fig. 8 panels (a), (b), (c) where we represent the final time results of the intense event simulation for the case $\Delta = 20$ m. We also notice the lake formation in regions where actually small water bodies are usually present in case of intense rainfall periods. In Fig. 8 panel (d) we report the number of time steps per hour for the three space resolution considered, note that the peak values are in correspondence of the peak values experienced in Fig. 6 panels (b), (d).

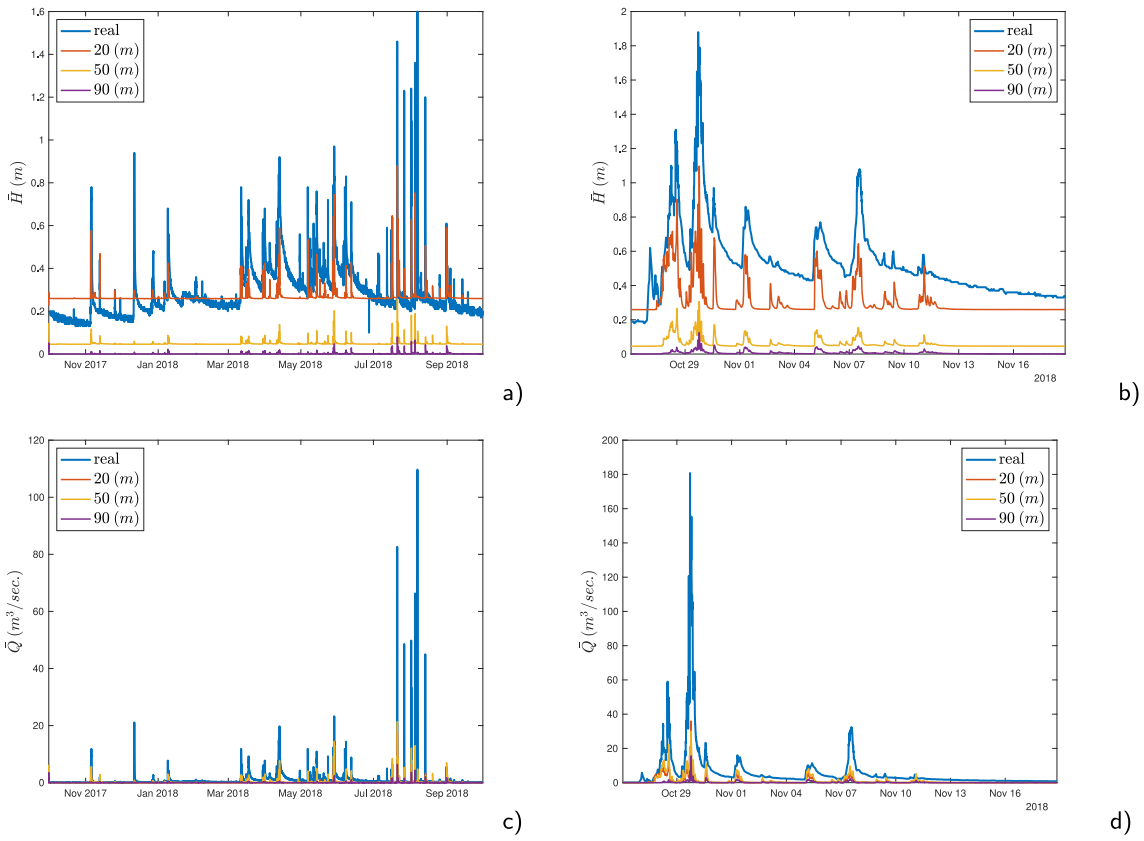

Fig. 6. Caldone basin. Superficial water height in meters and water flow rate in $m^3\ s^{-1}$, in case of year, panels (a) and (c), and intense event simulation, panel (b) and (d). In both cases, the blue line represents the hydrometer data.

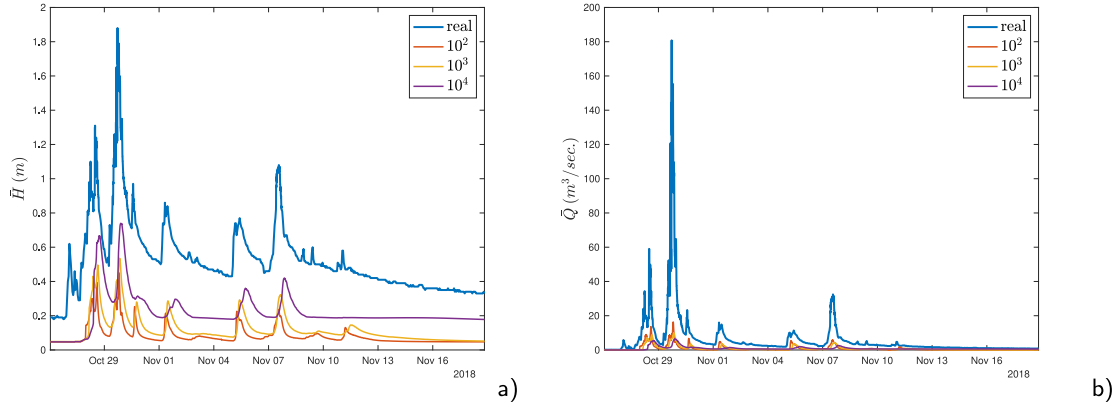

Fig. 7. Caldone basin. Sensitivity to a variation of d_{90} through coefficient R for the case $\Delta = 50$ m. Superficial water height in meters and water flow rate in $m^3\ s^{-1}$, in case of intense event simulation. The blue line represents the hydrometer data.

4.3.2. Stochastic scenarios

We perform two Monte Carlo simulations, for two different sets of Gavrilović coefficients, at a resolution $\Delta = 20$ m, for a period of 89 days on the same mountain catchment, specifically starting from Monday 20th July, 2020. Each Monte Carlo consists of 50 simulations each one initialized with a different soil composition map coming from the geostatistical preprocessor. Using 50 processors, the numerical model takes roughly 100 h to perform one Monte Carlo simulation on the time period considered. The input soil composition maps are determined via BSGS considering an uniform zero centered noise with standard deviation equal to 0.1.

We show the ability of the model to provide an estimation of the sediment discharge, as mentioned in the preamble of the present section, for each simulation we consider an estimation of the sediment flux and, for this point, we consider its temporal cumulative to provide an estimation of the total amount of sediments that are expected to sediment in the tank ($V_s(t) = \int_0^t \bar{Q}_s(\bar{t})d\bar{t}$), the blue dot in the map, see Fig. 4 panel (a). The expectations of the sedimentation from field data

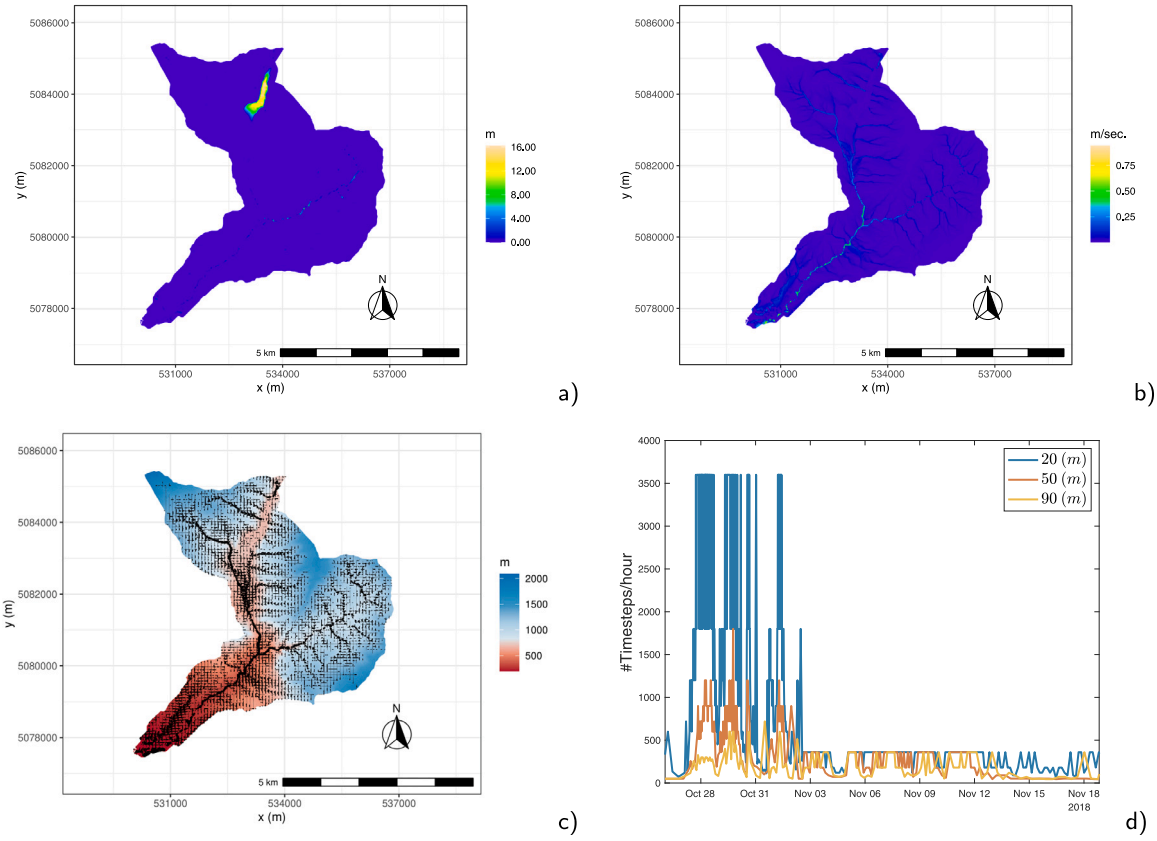

Fig. 8. Caldone basin. Intense event simulation. Final time results at resolution $\Delta = 20$ m: isolines in m of water height in panel (a), isolines of the velocity run-off in m s^{-1} in panel (b), while in panel (c) we report the velocity vectors of the superficial run-off velocity superimposed to the orography colormap. In panel (d) we report number of time steps per hour for the three space resolution considered.

is around ≈ 440 m^3 for this event, refer to Papini et al. (2017), Ivanov et al. (2020b), Brambilla et al. (2018). In Fig. 9(a), (b) we report two functional boxplots (Sun and Genton, 2011) derived from the two sets of 50 realizations of the cumulated sediment yields (as obtained using the python package statsmodels, see Seabold and Perktold (2010)). The probabilistic forecast allows one to identify the central functional quartiles of the simulated distribution and to achieve a robust estimate of a phenomenon well known for the strong associated uncertainties. As the considerable variation in the two sets suggests, a calibration could be performed choosing properly the Gavrilović coefficients, which control the sediment production, to fit the desired production on a given case study.

Finally, in Fig. 9 panel (c) we report, in blue, one realization of the $\bar{H}(t)$ against the observed water level in red. We do not provide a functional boxplot of this quantity since the d_{90} does not vary "considerably" among realizations in the numerical model. This results in an unchanged bias water level but in a change in the height of the peak values since the maximum soil moisture retention map results to be sensible for such kind of perturbation of the psf.

4.4. Realistic orography: Dragone basin

In this section we consider another case study with real orography. This is a test where real data are available and that has already been studied in Cuomo et al. (2015) with the LISEM software (Jetten, 2013).

We consider one mountain basin, the Dragone basin, located in the western part of the Lattari Mountains (Amalfi Coast, southern Italy). The Dragone basin has an area of 9.3 km^2 (and a perimeter of 15.7 km), with a linear main stream channel 6.5 km long. As stated in Cuomo et al. (2015), on 9 September 2010, cumulative rainfall of 126 mm was recorded in about 3 h, with a maximum hourly rainfall intensity equal to 92.2 mm. The total mobilized soil volume was estimated between 10^4 m^3 and $3 \cdot 10^4$ m^3 and peak discharge of water and debris at the basin outlet was estimated between 65 m^3 s^{-1} and 100 m^3 s^{-1}.

To reproduce this event, we consider the input DTM with resolution $\Delta = 10$ m and consider a set of uniform rain scenarios with a rainfall duration equal to 3 h and with a maximum rainfall intensity equal to 100 mm h^{-1} in 30 min, which corresponds to the rainfalls Ea, Eb, Ec as considered in Cuomo et al. (2015). The maximum rainfall intensity is set in the time intervals if $0 \text{ min} \le t \le 30$ min, if $75 \text{ min} \le t \le 105$ min, if $150 \text{ min} \le t \le 180$ min, for the three considered rainfalls respectively. To preserve the cumulative rain equal to 126 mm in the total time period, we set a rainfall intensity equal to 30.4 mm h^{-1} everywhere else. We consider also an uniform and constant rainfall in the given time period with intensity 42 mm h^{-1} to always to preserve the cumulative rain equal to 126 mm.

We report the results obtained with psf input maps produced with ILR-ATPRCoK for a set of Gavrilović coefficients. We mention that, we did a stochastic analysis with 50 psf maps generated via BSGS considering an uniform zero centered noise with standard deviation equal to 0.1 but we did not obtain "considerable" variations on the output among samples. This is probably due to the "short intense" event considered compared to the soil water content dynamic. Fig. 10 shows the results of such analysis. The time series are taken at the geographical location of the basin outlet computed with $\Delta_p = 10$ m. In particular, panel (a) shows the mean value of the water discharge $\bar{Q}$ while panel (b)

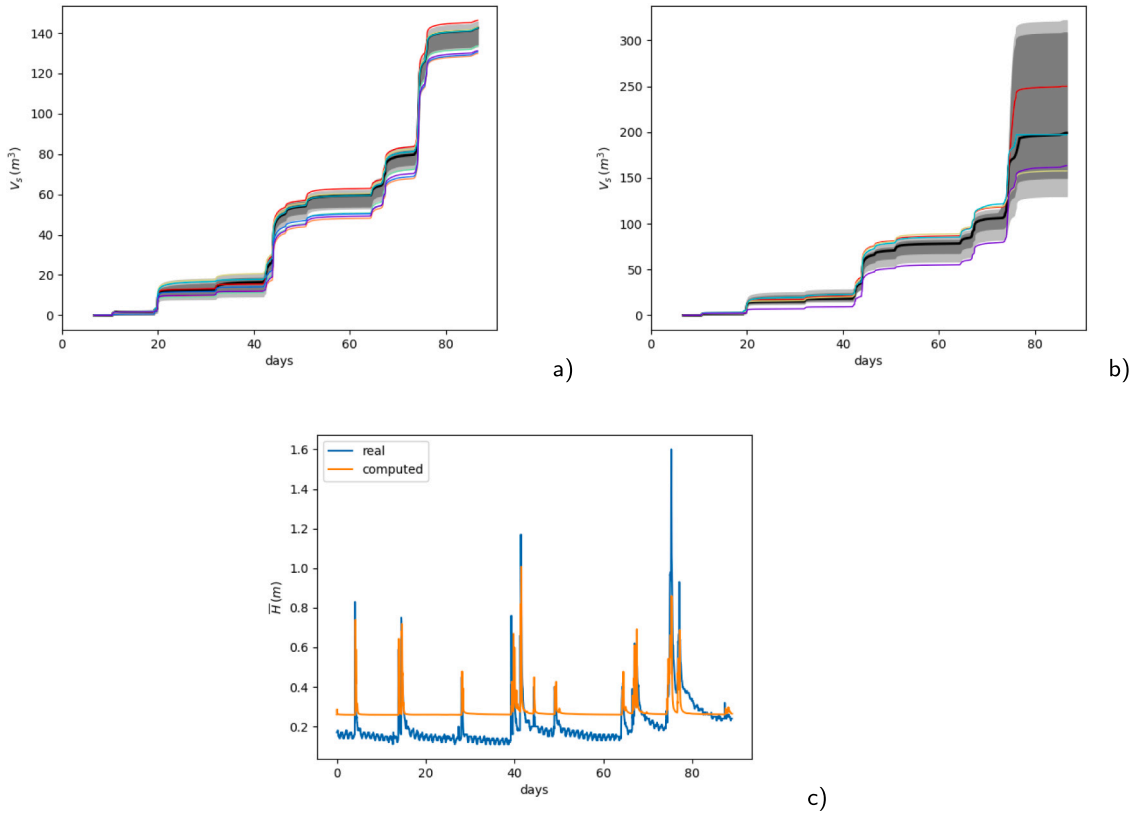

Fig. 9. Caldone basin. Stochastic scenarios, functional boxplots of the simulated cumulative sediment volume V_s in panels (a) and (b), while in panel (c) we report one simulated realization (in orange) and the hydrometer data (in blue). For the functional boxplots we have used as input parameter wfactor = 2.58 to identify the outlier curves. The black curves correspond to the median curve, the dark and light gray bands correspond to the regions where 50% and 75% of data lie respectively, the others colored curves denote the outliers.

shows the temporal evolution of the maximum value of the simulated sediment concentration C_v in the neighborhood $\Delta_p = 10$ m. Regarding the former panel, we experience a maximum value of the mean water discharge at the outlet point around 115 $m^3\ s^{-1}$ for events Ea, Eb, Ec, which fits rather well the water discharge measurements. Regarding panel (b), we experience a final time maximum sediment concentration ranging approximately between 17% and 19% for events Ea, Eb, Em, with a maximum value in the period equal to roughly 29% for event Eb. We particularly note the ability of the model to predict peaks both in the water discharge and sediment concentration due to rainfall peaks. Peaks in C_v that were not experienced in the set of simulations carried on in Cuomo et al. (2015) for some values of calibration coefficients. For event Ec we report a lower final time value probably due to the delay between the peak rainfall value and the sediment response at the basin outlet. This is suggested by the C_v behavior which is similar to the one experienced in the correspondence of the peak rainfall value for event Eb. Delay that would have result in a peak value of the quantity C_v after the 3 h observation. Finally, panel (c) shows the sediment production map, in meters, for the Dragone catchment for the set of the considered Gavrilović coefficients. From this map we compute a total eroded volume equal to $V_s \approx 1.2223 \cdot 10^4$ m^3 thus ranging in the expected interval. In the same panel we represent in red the outlet point of the basin. Panel (d) shows the final time isolines map of the absolute value of the superficial run-off velocity in m s^{-1} for the event Eb.

5. Conclusions

We have presented a numerical model for soil erosion at the basin scale, in which the de Saint-Venant equations are used to model surface water flow over the whole domain, without *a priori* identification of drainage zones. This allows to model basins in which strong variations of the surface run-off occur without ad hoc hypotheses, as well as to include naturally lakes and other water reservoirs, so to help in the identification of large water bodies, particularly useful in mountain catchments to identify potential flood-risk regions. The model is equipped with a geostatistical preprocessor, that allows to downscale to fine resolution meshes the soil composition data available in global databases. The model equations are approximated numerically using a very well tested, efficient and robust semi-implicit discretization method (Casulli, 1990; Casulli and Cheng, 1992). The proposed technique is able to handle automatically the wide range of transients that can arise in long term simulations and to run even at high resolution. Exact mass conservation and positivity of the water and sediment layers are guaranteed. The results obtained in the first verification and validation experiments confirm the model's efficiency, robustness and flexibility, as well as its ability to reproduce transient high resolution features at a low computational cost.

In future developments, the numerical model will be further integrated with its geostatistical counterpart, so as to allow for stochastic simulations and uncertainty quantification by multifidelity approaches, see e.g. Grujic et al. (2018). Less elementary choices will be employed for the description of the many physical processes involved, in particular for the description of the infiltration process. It is also planned to improve the description of the subsurface flow, employing vertically averaged Darcy equations and coupling the subsurface flow to the surface flows along the lines proposed in several papers, see e.g. Casulli (2015, 2017), Discacciati et al. (2002), Miglio et al. (2003). Finally, also the local refinement techniques proposed in Casulli (2019) for the

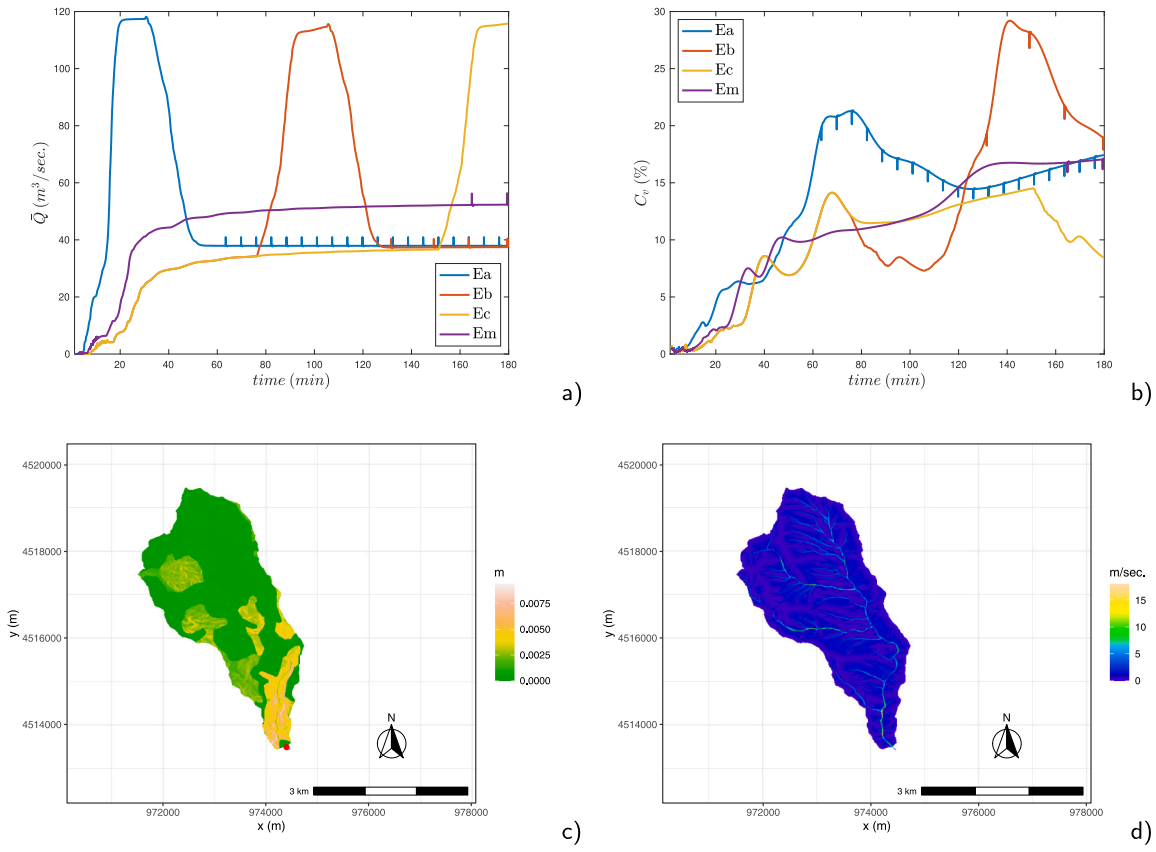

Fig. 10. Dragone basin. Panels (a), (b) show the temporal sequences of the mean water discharge and of the maximum value of the sediment concentration in the basin outlet point. Panels (c), (d) show in m the production map in the Dragone basin with in red the basin outlet point and the isolines of the magnitude of the superficial run-off velocity in m s^{-1}.

same kind of numerical method employed here can be incorporated in the model, further enhancing its ability to resolve fine spatial features at a low computational cost.

CRediT authorship contribution statement

Federico Gatti: Code development, Writing. **Luca Bonaventura:** Numerical methodology supervision. **Alessandra Menafoglio:** Geostatistical methodology supervision. **Monica Papini:** Review. **Laura Longoni:** Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request

Acknowledgments

The authors gratefully acknowledge the financial support of Fondazione Cariplo in the framework of the SMART-SED project, grant number 2017-0722. Many useful discussions with A. Radice and D. Brambilla are also kindly acknowledged, as well as the contributions by A. Abbate, J.K. Grudnicki, N. Togni, D. Rossi to the preliminary phases of the development of this model.

Code availability section SMART-SED can be accessed through its dedicated repository on GitHub (https://github.com/federicg/smartse d) under the GNU General Public License. The software is written in C++ and is equipped with a geostatistical preprocessor written in R.

Appendix. Infiltration law

In order to get the expression of the infiltration rate, we consider the mass conservation laws for surface and gravitational run-off in absence of boundary fluxes, together with the proportionality hypothesis of the SCS-CN method

$$\begin{aligned} &\partial_t H = p - f, \\ &\partial_t h_g = f, \\ &\frac{F}{S} = \frac{P - F}{P}, \end{aligned} \tag{28}$$

where S is the maximum soil moisture retention and F and P are the time integrals of f and p, respectively. For the maximum soil moisture retention, we will use the following expression (in mm)

$$S = 254 \left(\frac{100}{CN} - 1 \right), \tag{29}$$

where CN is the non dimensional parameter known as curve number. The curve number usually varies from 30 to 100 and depends on soil type, land cover and land use. Large values of such parameter are associated with impermeable surfaces, and subsequently to lower storage capability and higher run-off. The Soil Conservation Service provides a table in which each type of land cover and land use is associated to four values of CN, depending on the hydrologic soil group, which in turn defines the actual run-off capability of the soil underneath. Such soil groups are usually identified with a letter, from A to D, where A identifies the most impermeable soils and D the most permeable ones. For example, paved roads are associated with a CN value of 98 regardless of the soil group, woods in fair conditions

and favorable soil are associated with a value of 42 and woods with unfavorable soil 80. More details on the criteria used to determine the CN values will be given in Section 4. Finally, in order to derive an expression for f, we recast the last equation in (28) as

$$\frac{h_g}{S} = \frac{H}{H + h_g}. \tag{30}$$

It is then possible to solve the system (28) in order to get an expression for F, which denoting with H_0 and $h_{g,0}$ the initial conditions for surface and gravitational run-off, reads

$$F = \frac{S\,P}{P + S + H_0 + h_{g,0}}, \tag{31}$$

which is the common expression of the SCS-CN formulation. To get an expression of f, one can derive (30) with respect to time, under the hypothesis of constant S, so as to obtain

$$\partial_t H = \partial_t h_g \left[\frac{S^2}{(S - h_g)^2} - 1 \right]. \tag{32}$$

By substituting (32) in the last equation in (28), one gets for f

$$f = p \left(\frac{S - h_g}{S} \right)^2. \tag{33}$$

Notice that, since $h_g \leq S$, in Eq. (33) the precipitation rate is multiplied by a real quantity lying between 0 and 1, which ensures that $f \leq p$. An initial loss, I_a, can be added to this model to represent interception and depression storage. This loss occurs prior to the onset of run-off. The final infiltration rate model reads

$$f = \begin{cases} 0, & \text{if } H + h_g \leq I_a, \\ p \left(\frac{S - h_g}{S} \right)^2, & \text{if } H + h_g > I_a. \end{cases} \tag{34}$$

The initial loss is usually modeled to be proportional to the maximum soil moisture retention, i.e. $I_a = c\,S$, where c is areal constant usually set equal to 0.2 or 0.05. In this work, we choose $c = 0.2$.

References

Abbott, M., Bathurst, J., Cunge, J., O'Connell, P., Rasmussen, J., 1986a. An introduction to the European Hydrological System—Systeme Hydrologique Europeen,"SHE", 1: History and philosophy of a physically-based, distributed modelling system. J. Hydrol. 87, 45–59.

Abbott, M., Bathurst, J., Cunge, J., O'Connell, P., Rasmussen, J., 1986b. An introduction to the European Hydrological System—Systeme Hydrologique Europeen,"SHE", 2: Structure of a physically-based, distributed modelling system.. J. Hydrol. 87, 61–77.

Aitchison, J., 1982. The statistical analysis of compositional data. J. R. Stat. Soc. Ser. B Stat. Methodol. 44, 139–160.

Bemporad, G.A., Alterach, J., Amighetti, F.F., Peviani, M., Saccardo, I., 1997. A distributed approach for sediment yield evaluation in Alpine regions. J. Hydrol. 370–392.

Borrelli, P., Alewell, C., Alvarez, P., Anache, J.A.A., Baartman, J., Ballabio, C., Bezak, N., Biddoccu, M., Cerdà, A., Chalise, D., Chen, S., Chen, W., De Girolamo, A.M., Gessesse, G.D., Deumlich, D., Diodato, N., Efthimiou, N., Erpul, G., Fiener, P., Freppaz, M., Gentile, F., Gericke, A., Haregeweyn, N., Hu, B., Jeanneau, A., Kaffas, K., Kiani-Harchegani, M., Villuendas, I.L., Li, C., Lombardo, L., López-Vicente, M., Lucas-Borja, M.E., Märker, M., Matthews, F., Miao, C., Mikoš, M., Modugno, S., Möller, M., Naipal, V., Nearing, M., Owusu, S., Panday, D., Patault, E., Patriche, C.V., Poggio, L., Portes, R., Quijano, L., Rahdari, M.R., Renima, M., Ricci, G.F., Rodrigo-Comino, J., Saia, S., Samani, A.N., Schillaci, C., Syrris, V., Kim, H.S., Spinola, D.N., Oliveira, P.T., Teng, H., Thapa, R., Vantas, K., Vieira, D., Yang, J.E., Yin, S., Zema, D.A., Zhao, G., Panagos, P., 2021. Soil erosion modelling: A global review and statistical analysis. Sci. Total Environ. 780, 146494.

Boucher, A., 2003. Conditional joint simulation of random fields on block-support.

Brambilla, D., Papini, M., Longoni, L., 2018. Temporal and spatial variability of sediment transport in a mountain river: A preliminary investigation of the Caldone River, Italy. Geosciences 8 (5), 163.

Büttner, G., 2014. CORINE land cover and land cover change products. Land Use Land Cover Mapp. Eur.: Pract. Trends 55–74.

Büttner, G., Feranec, J., Jaffrain, G., Mari, L., Maucha, G., Soukup, T., 2004. The CORINE land cover 2000 project. EARSeL EProc. 3, 331–346.

Casulli, V., 1990. Semi-implicit finite difference methods for the two-dimensional shallow water equations. J. Comput. Phys. 86, 56–74.

Casulli, V., 2015. A conservative semi-implicit method for coupled surface–subsurface flows in regional scale. Int. J. Numer. Methods Fluids 79, 199–214.

Casulli, V., 2017. A coupled surface-subsurface model for hydrostatic flows under saturated and variably saturated conditions. Int. J. Numer. Methods Fluids 85, 449–464.

Casulli, V., 2019. Computational grid, subgrid, and pixels. Int. J. Numer. Methods Fluids 90, 140–155.

Casulli, V., Cheng, R., 1992. Semi-implicit finite difference methods for three-dimensional shallow water flow. Int. J. Numer. Methods Fluids 15, 629–648.

Chanson, H., 1999. The Hydraulics of Open Channel Flow: An Introduction. Physical Modelling of Hydraulics. Butterworth-Heinemann.

Chen, F., Liu, C., 2012. Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. Paddy Water Environ. 10 (3), 209–222.

Chiari, M., Rickenmann, D., 2007. The influence of form roughness on modelling sediment transport at steep slopes. In: Proceedings of the International Conference on: Erosion and Torrent Control As a Factor in Sustainable River Basin Management. pp. 25–28.

Ciarapica, L., Todini, E., 2002. TOPKAPI: A model for the representation of the rainfall-runoff process at different scales. Hydrol. Process. 16, 207–229.

Cuomo, S., Della Sala, M., Novità, A., 2015. Physically based modelling of soil erosion induced by rainfall in small mountain basins. Geomorphology 243, 106–115.

Davies, M., 2003. Standard Handbook for Aeronautical and Astronautical Engineers. McGraw-Hill Education.

Day, T., et al., 2006. Degree-Days: Theory and Application, Vol. 106. The Chartered Institution of Building Services Engineers, London.

Discacciati, M., Miglio, E., Quarteroni, A., 2002. Mathematical and numerical models for coupling surface and groundwater flows. Appl. Numer. Math. 43, 57–74.

Doten, C., Bowling, L., Lanini, J., Maurer, E., Lettenmaier, D., 2006. A spatially distributed model for the dynamic prediction of sediment erosion and transport in mountainous forested watersheds. Water Resour. Res. 42.

Dutta, S., 2016. Soil erosion, sediment yield and sedimentation of reservoir: A review.. Model. Earth Syst. Environ. 2, 123.

Endrizzi, S., Gruber, S., Dall'Amico, M., Rigon, R., 2014. GEOtop 2.0: Simulating the combined energy and water balance at and below the land surface accounting for soil freezing, snow cover and terrain effects. Geosci. Model Dev. 7, 2831–2857.

Epple, L., Kaiser, A., Schindewolf, M., Bienert, A., Lenz, J., Eltner, A., 2022. A review on the possibilities and challenges of today’s soil and soil surface assessment techniques in the context of process-based soil erosion models. Remote Sens. 14 (10).

Fatichi, S., Vivoni, E., Ogden, F.L., Ivanov, V., et al., 2016. An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. J. Hydrol. 537, 45–60.

Fu, B., Merritt, W., Croke, B., Weber, T., Jakeman, A., 2019. A review of catchment-scale water quality and erosion models and a synthesis of future prospects. Environ. Model. Softw. 114, 75–97.

Garegnani, G., Rosatti, G., Bonaventura, L., 2013. On the range of validity of the Exner-based models for mobile-bed river flow simulations. J. Hydraul. Res. 51, 380–391.

Gatti, F., Fois, M., de Falco, C., Perotto, S., Formaggia, L., 2023a. Parallel simulations for fast-moving landslides: Space-time mesh adaptation and sharp tracking of the wetting front. Internat. J. Numer. Methods Fluids http://dx.doi.org/10.1002/fld.5186.

Gatti, F., Menafoglio, A., Togni, N., Bonaventura, L., Brambilla, D., Papini, M., Longoni, L., 2021. A novel downscaling procedure for compositional data in the Aitchison geometry with application to soil texture data. Stoch. Environ. Res. Risk Assess. 35 (6), 1223–1241.

Gatti, F., Perotto, S., de Falco, C., Formaggia, L., 2023b. A positivity-preserving well-balanced numerical scheme for the simulation of fast landslides with efficient time stepping. MOX-Rep. 11.

Gavrilović, Z., 1988. Use of an empirical method(erosion potential method) for calculating sediment production and transportation in unstudied or torrential streams. In: International Conference on River Regime. Hydraulics Research Limited, Wallingford, pp. 411–422.

Globevnik, L., Holjević, D., Petkovsek, G., Rubinić, J., 2003. Applicability of the Gavrilović method in erosion calculation using spatial data manipulation techniques. In: Proceedings of the Symposium on: Erosion Prediction in Ungauged Basins, Integrating Methods and Techniques. IAHS, pp. 224–234.

Gross, E., Bonaventura, L., Rosatti, G., 2002. Consistency with continuity in conservative advection schemes for free-surface models. Int. J. Numer. Methods Fluids 38, 307–327.

Grujic, O., Menafoglio, A., Yang, G., Caers, J., 2018. Cokriging for multivariate Hilbert space valued random fields: Application to multi-fidelity computer code emulation. Stoch. Environ. Res. Risk Assess. 32, 1955–1971.

Guennebaud, G., Jacob, B., et al., 2010. Eigen v3. http://eigen.tuxfamily.org.

Hargreaves, G., Allen, R., 2003. History and evaluation of Hargreaves evapotranspiration equation. J. Irrig. Drain. Eng. 129, 53–63.

Hazen, A., 1892. Physical properties of sands and gravels with reference to use in filtration. Rep. Massachusetts State Board Health 539.

Hengl, T., al., J.D.J., 2017. SoilGrids250m: Global gridded soil information based on machine learning. PLoS One 12:e0169748.

Idso, S., Jackson, R., Reginato, R., 1977. Remote-sensing of crop yields. Science 196, 19–25.
Ivanov, V., Radice, A., Papini, M., Longoni, L., 2020a. Event-scale pebble mobility observed by RFID tracking in a pre-Alpine stream: A field laboratory. Earth Surf. Process. Landforms 45, 535–547.
Ivanov, V., Radice, A., Papini, M., Longoni, L., 2020b. Event-scale pebble mobility observed by RFID tracking in a pre-Alpine stream: A field laboratory. Earth Surf. Process. Landforms 45 (3), 535–547.
Jetten, V., 2013. OpenLISEM. http://sourceforge.net/projects/lisem/.
Kershaw, D., 1978. The incomplete Cholesky—conjugate gradient method for the iterative solution of systems of linear equations. J. Comput. Phys. 26 (1), 43–65.
Ketema, A., Dwarakish, G., 2019. Water erosion assessment methods: A review. ISH J. Hydraulic Eng. 1–8.
Kim, J., 1999. Spurious correlation between ratios with a common divisor. Statist. Probab. Lett. 44, 383–386.
LeVeque, R., 1992. Numerical Methods for Conservation Laws. Springer.
Lombardia, A., 2021. Portale dati idrologici Arpa Lombardia. URL http://idro.arpalombardia.it/pmapper-4.0/map.phtml.
Martín-Fernández, J., Hron, K., Templ, M., Filzmoser, P., Palarea-Albaladejo, J., 2012. Model-based replacement of rounded zeros in compositional data: Classical and robust approaches. Comput. Statist. Data Anal. 56, 2688–2704.
Matthews, J., 2013. Encyclopedia of Environmental Change. Sage.
Merritt, W., Letcher, R., Jakeman, A.J., 2003. A review of erosion and sediment transport models. Environ. Model. Softw. 18, 761–799.
Miglio, E., Quarteroni, A., Saleri, F., 2003. Coupling of free surface and groundwater flows. Comput. & Fluids 32, 73–83.
Morgan, R., Quinton, J., Smith, R., Govers, G., Poesen, J., Auerswald, K., Chisci, G., Torri, D., Styczen, M., 1998. The European Soil Erosion Model (EUROSEM): A dynamic approach for predicting sediment transport from fields and small catchments. Earth Surf. Process. Landforms 23, 527–544.
Pandey, A., Himanshu, S., Mishra, S., Singh, V., 2016. Physically based soil erosion and sediment yield models revisited. Catena 147, 595–620.
Papini, M., Ivanov, V.I., Brambilla, D., Arosio, D., Longoni, L., 2017. Monitoring bedload sediment transport in a pre-alpine river: An experimental method. Rendiconti Online Soc. Geol. Italiana 43, 57–63.
Ponce, V., Hawkins, R., 1996. Runoff curve number: Has it reached maturity? J. Hydrol. Eng. 1, 11–19.
Porta, G.M., Perotto, S., Ballio, F., 2012. A space-time adaptation scheme for unsteady shallow water problems. Math. Comput. Simul. 82 (12), 2929–2950.
Rickenmann, D., 1994. An alternative equation for the mean velocity in gravel-bed rivers and mountain torrents. Hydraulic Eng. 94, 672–676.
Rickenmann, D., 2001. Comparison of bed load transport in torrents and gravel bed streams. Water Resour. Res. 37, 3295–3305.
Rigon, R., Bertoldi, G., Over, T., 2006. GEOtop: A distributed hydrological model with coupled water and energy budgets. J. Hydrometeorol. 7, 371–388.
Roo, A.D., Jetten, V., 1999. Calibrating and validating the LISEM model for two data sets from the Netherlands and South Africa. CATENA 37, 477–493.
Rosatti, G., Bonaventura, L., Deponti, A., Garegnani, G., 2011. An accurate and efficient semi-implicit method for section-averaged free-surface flow modelling. Int. J. Numer. Methods Fluids 65, 448–473.
Rosatti, G., Chemotti, R., Bonaventura, L., 2005. High order interpolation methods for semi-Lagrangian models of mobile-bed hydrodynamics on Cartesian grids with cut cells. Int. J. Numer. Methods Fluids 47, 1269–1275.
Rosso, R., 2004. Mappatura Dell' Indice Di Assorbimento E Del Massimo Volume Specifico Di Ritenzione Potenziale Del Terreno. Relazione Finale Progetto SHAKEUP-2, ARPA Lombardia.
Seabold, S., Perktold, J., 2010. Statsmodels: Econometric and statistical modeling with Python. In: 9th Python in Science Conference.
Simoni, S., Zanotti, F., Bertoldi, G., Rigon, R., 2008. Modelling the probability of occurrence of shallow landslides and channelized debris flows using GEOtop-FS. Hydrol. Process. 22, 532–545.
Sinclair, S., Pegram, G., 2013. A sensitivity assessment of the TOPKAPI model with an added infiltration module. J. Hydrol. 479, 100–112.
Smart, G., Jäggi, M., 1983. Sediment transport on steep slopes. Technical Report, Mitteilung. 64. Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie, ETH Zurich.
Soto, B., Fierros, F.D., 1998. Runoff and soil erosion from areas of burnt scrub: Comparison of experimental results with those predicted by the WEPP model. CATENA 31, 257–270.
Sun, Y., Genton, M., 2011. Functional boxplots. J. Comput. Graph. Statist. 20, 316–334.
Taylor, B., 1995. Guide for the Use of the International System of Units (SI): The Metric System. DIANE Publishing.
Wicks, J., Bathurst, J., 1996. SHESED: A physically based, distributed erosion and sediment yield component for the SHE hydrological modelling system. J. Hydrol. 175, 213–238.
Wigmosta, M., Vail, L., Lettenmaier, D., 1994. A distributed hydrology-vegetation model for complex terrain. Water Resour. Res. 30, 1665–1679.
Zhang, G., Liu, Y., Han, Y., Zhang, X., 2009. Sediment transport and soil detachment on steep slopes: I. transport capacity estimation. Soil Sci. Am. J. 73, 1291–1297.
Zienkiewicz, O.C., Zhu, J., 1992a. The superconvergent patch recovery and a posteriory error estimates I. The recovery technique. Internat. J. Numer. Methods Engrg. 33 (7), 1331–1364.
Zienkiewicz, O.C., Zhu, J., 1992b. The superconvergent patch recovery and a posteriory error estimates II. Error estimates and adaptivity. Internat. J. Numer. Methods Engrg. 33 (7), 1365–1382.