Simulating Sediment Discharge at Water Treatment Plants Under Different Land Use Scenarios Using Cascade Modelling with an Expert-Based GIS Model and a Deep Neural Network

Edouard Patault[1], Valentin Landemaine[2], Jérôme Ledun[3], Arnaud Soulignac[4], Matthieu Fournier[1], Jean-François Ouvry[3], Olivier Cerdan[2], Benoit Laignel[1]

[1]Normandie UNIV, UNIROUEN, UNICAEN, CNRS, M2C, FED-SCALE, Rouen, France
[2]BRGM, 3 avenue Claude Guillemin, BP6009, F-45060 Orléans Cedex 2, France
[3]AREAS, 2 avenue Foch, F-76460 Saint-Valéry-en-Caux, France
[4]BRGM, 1039 rue de Pinville, F-34000 Montpellier, France

Correspondence to: Edouard Patault (edouardpatault@gmail.com) and Matthieu Fournier (matthieu.fournier@univ-rouen.fr)

Abstract. Excessive sediment discharge in karstic regions can be highly disruptive to water treatment plants. It is essential for catchment stakeholders and drinking water suppliers to limit the impact of high sediment loads on potable water supply, but their strategic choices must be based on simulations, integrating surface and groundwater transfers, and taking into account possible changes in land use. Karstic environments are particularly challenging as they face a lack of accurate physical description for the modelling process, and they can be particularly complex to predict due to the non-linearity of the processes generating sediment discharge. The aim of the study was to assess the sediment discharge variability at a water treatment plant according to multiple realistic land use scenarios. To reach that goal, we developed a new cascade modelling approach with an erosion-runoff GIS model (WaterSed) and a deep neural network. The model was used in the Radicatel hydrogeological catchment (106 km² in Normandy, France) where karstic spring water is extracted to a water treatment plant. The sediment discharge was simulated for five design storms under current land use and compared to four land use scenarios (baseline, ploughing up of grassland, eco-engineering, best farming practices, and coupling eco-engineering/best farming practices). Daily rainfall time series and WaterSed modelling outputs extracted at connected sinkholes (positive dye-tracing) were used as input data for the deep neural network model. The model structure was found by a classical trial and error procedure, and the model was trained on two significant hydrologic years. Evaluation on a test set showed a good performance of the model (NSE = 0.82), and the application of a monthly backward chaining nested cross validation revealed that the model is able to generalize on new datasets. Simulations made for the four land use scenarios suggested that ploughing up 33 % of grasslands would increase sediment discharge at the water treatment plant by 5% on average. On the contrary, eco-engineering and best farming practices will significantly reduce sediment discharge at the water treatment plant (respectively in the range of 10-44 % and 24-61 %). The coupling of these two strategies is the most efficient since it affects the hydro-sedimentary production and transfer processes (decreasing sediment discharge from 40 to 80 %). The cascade modelling approach developed in this study offers interesting opportunities for sediment discharge prediction at karstic springs or water treatment plants under multiple land use scenarios. It also provides robust decision-making tools for land use planning and drinking water suppliers.

Key words: erosion; runoff; karst; deep neural network; WaterSed model; land use scenarios

1 Introduction

In karstic environments, erosion and runoff can lead to a high load of sediments in surface and underground streams. Sediment discharge (SD) can occur through rapid and direct transfer via sinkholes, and/or via re-suspension of sediment in the karst network itself (Masséi et al., 2003). For suppliers of drinking water, excessive sediment input can be highly disruptive, requiring additional treatment or, in the worst cases, temporary shutdowns of a Water Treatment Plants (WTP; Stevenson et al., 2019). Impacts can be significant, including restrictions on the use of drinking water. Upper-Normandy (France) is particularly affected, and the economic cost linked to the restrictions on the use of drinking water due to excessive sediment inputs in raw water was estimated at €5 million during the period 1992-2018 (Patault et al., 2021a). 10000 to 20000 people are still affected every year by restrictions on the use of drinking water in the region (ARS, 2013). Reducing sediment delivery to sinkholes is therefore essential for catchment managers in order to reduce the impact on potable water supply. One way to achieve this would be to build a complex decision-making process modelling chain, integrating surface and groundwater transfers. Several approaches have been proposed to model erosion/runoff and the karstic response induced by rainfall, but these approaches are often treated separately which does not make it possible to evaluate the impact of a change in land use on the sediment load delivered to a WTP. Many studies in the literature have focused on hillslope erosion processes using different types of erosion models (Merritt et al., 2003; de Vente et al., 2013). Empirical models, such as RUSLE (Renard, 1994), are frequently used because of limited data requirements but are not able to fully represent spatial and temporal dynamic of erosion processes at the catchment scale (Verstraeten et al., 2007). Physical models, such as WEPP or LISEM (Laflen et al., 1991; Takken et al., 1999), can more accurately describe processes but may require many input parameters that are not available for application on large areas. Expert-based models (e.g. STREAM, WaterSed (Cerdan et al., 2001; Landemaine, 2016)) offer an interesting compromise focusing on the main driving factors of runoff and erosion. These models have been designed with cultivated areas of the European loess belt in mind and are particularly efficient where hortonian overland flow dominates (Souchère et al., 2005; Evrard et al., 2010; Landemaine, 2016). Some of hillslope erosion studies conducted in similar karstic environment, do not account for the transmissivity loss of water and sediment in sinkholes. Other studies that have focused on modelling karst processes, have mainly examined karstic floods but, for the most part, overlooking sedimentary fluxes (see review of Hartmann et al., 2014). Due to the non-linearity of the processes generating sediment at karstic springs, and the lack of accurate physical description of karstic environments, modelling surface-subsurface interactions with physical models can therefore be a difficult task (Savary et al., 2017; Jourde et al., 2018). Based on systemic approaches, as initiated by Mangin (1984), karst can be considered a system able to both transform an input (rainfall) into an output (discharge) and the input-output relation can be evaluated using mathematical functions. This approach can be considered a 'black-box' model to some extent, and recent research emphasized the advantages of using data-driven techniques, such as Deep Neural Networks (DNN) in similar situations (Yaseen et al., 2015; Kratzert et al., 2018; Kratzert et al., 2019b). DNN are

advanced Artificial Neural Networks (ANN) which have been gaining momentum since 2012 in computer sciences and whose adoption has been gradual in hydrology (Shen, 2018). DNN are now commonly used for modelling real world problems due to their ability to represent and generalize complex non-linear relationships between inputs and outputs (Meyers et al., 2016; 2017; Hafeez et al., 2019). DNN can help by providing both stronger predictive capabilities and a complementary avenue toward knowledge discovery in hydrologic sciences (Shen et al., 2018). Limited but conclusive applications were made to predict inflows to reservoirs, water levels of combined sewage outflow structures, turbidity, flood forecasting in karst regions, and integrated in rainfall-runoff modelling to better predict stream flows (Siou et al., 2011; Bai et al., 2016; Savary et al., 2017; Hu et al., 2018; Kratzert et al., 2018). The main objectives of this study were: (i) to develop a cascade modelling approach able to simulate hydro-sedimentary transfer at a WTP for specific daily rainfall events, and (ii) to evaluate the impact of different land use scenarios on the SD variability. This study was conducted in the Radicatel hydrogeological catchment (Normandy, France) where spring water is extracted to a WTP. We benefited from the use of an existing expert-based GIS model (WaterSed), developed and successfully applied on the studied area, to simulate the impact of land use management on hydro-sedimentary transfers to connected sinkholes. Rainfall events characteristics and WaterSed outputs were then used as input for a data-driven model (i.e. DNN) to simulate SD at the WTP. The cascade modelling approach was applied to multiple design storms (DS) under different scenarios in order to simulate hydro-sedimentary transfer in the hydrogeological catchment and evaluate the efficiency of different land use management strategies.

2 Study Site

The study site is in the Pays de Caux (Normandy, France) on the right bank of the Seine River about 30 km from the Seine estuary (Fig. 1a). The climate is temperate with an average temperature of 11°C. Annual precipitation ranges between 600 and 1100 mm/year with a rainy season occurring between October and May. The karst is typical of the geological setting of the lower Seine Valley. The karstic chalk plateaus of the North-Eastern side of the Normandy region are part of the western edge of the Paris Basin. The elevation ranges from 138 to 0 m above sea level and the median slope is 5.9%. The geology consists of Cenomanian to Campanian chalk overlaid by thick surficial formations. The major formation is composed of clay-with-flints which results from the weathering of the chalk (Laignel, 2003), loess, and tertiary sands (Lautridou, 1985). The thickness of the formation overlying the chalk ranges from 5 to 10 m. Water infiltrates from the uplands of the chalk aquifer to the valleys via rapid transfer through a highly developed karstic network, and via slower drainage through the thick surficial formations.

The hydrogeological catchment of Radicatel covers 106 km² and the WTP is located at the interface of the Seine alluvium and the karstic chalk (Fig. 1b). Water is pumped from four different springs and seven pumping wells near the WTP (Chédeville, 2015). According to the information system for groundwater management in Seine-Normandy (http://sigessn.brgm.fr/), hydrogeological investigations reported seven sinkholes positively connected to the springs (in-situ dye tracing already performed and confirming the connection). The WTP of Radicatel is exploited by the Le Havre-Seine Métropole (LHSM) to

supply Le Havre inhabitants. Turbidity (NTU) is measured with a nephelometer at the inlet of the pumping station and recorded since 1987 (Fig. 2). The maximum turbidity value is recorded every day (Δt = 1 d) and LHSM provided access to the entire turbidity time series (1987-2017). Incomplete periods or hydrologic years with missing data were discarded. Twenty-two hydrologic years with a complete time-series of turbidity were kept for the rest of the study.

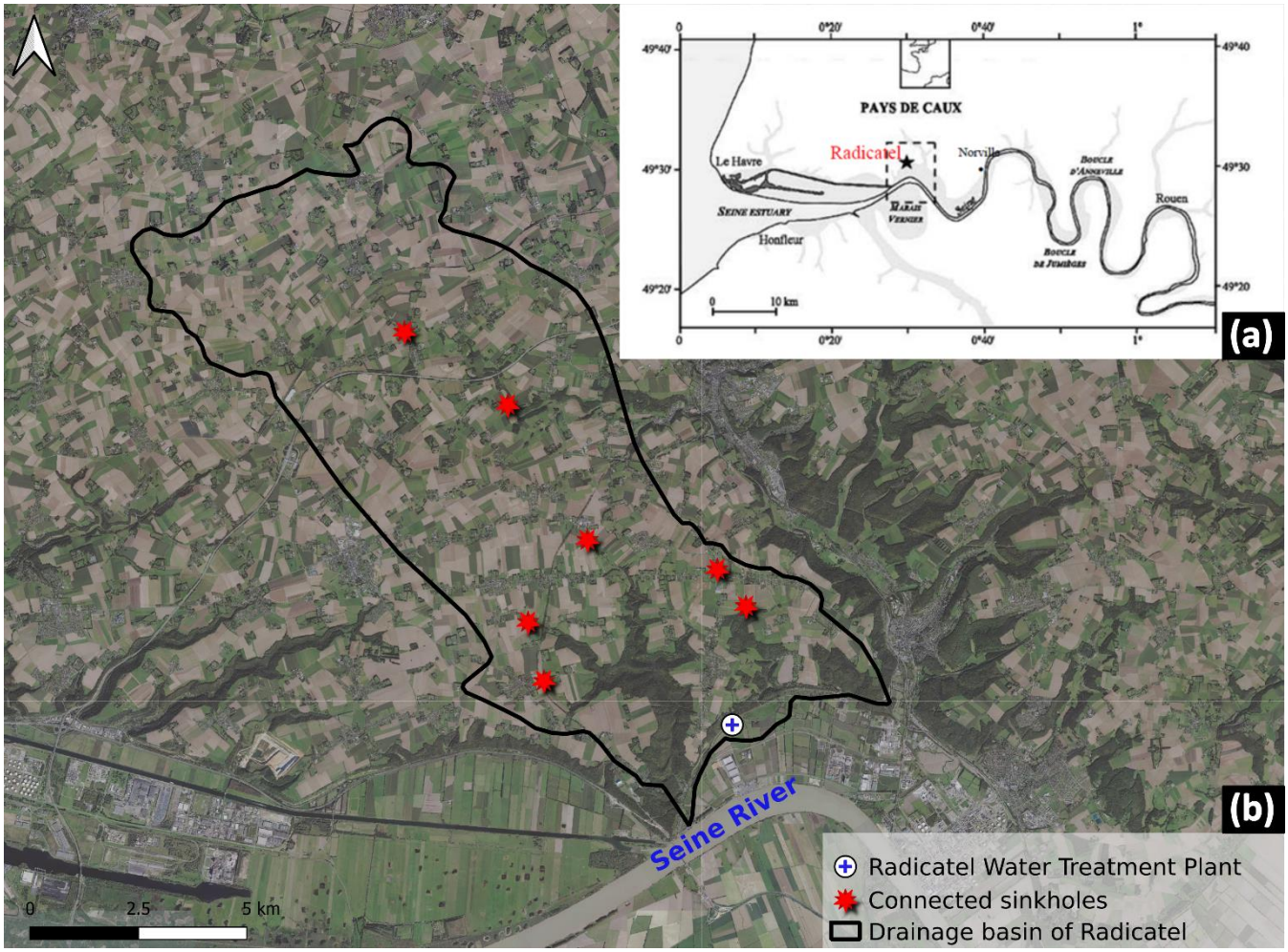
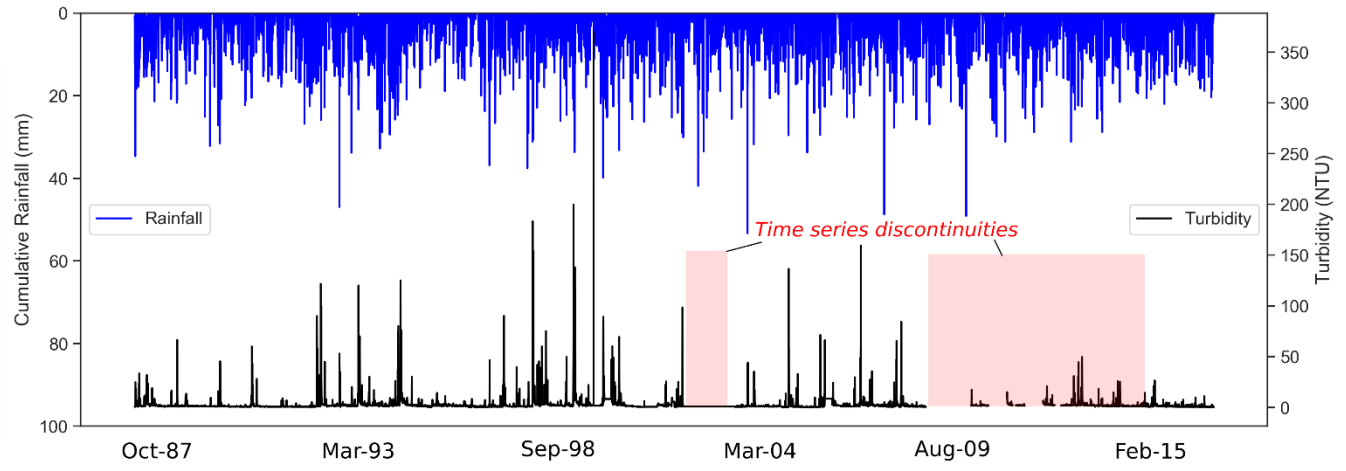


Figure 1: (a) Location of the study site in the lower Seine Valley (Normandy, France) on the right bank of the Seine River (Hanin, 2011), and (b) Location of the water treatment plant of Radicatel and the seven connected sinkholes. Background map was retrieved from the BD Ortho© – IGN.



105

Figure 2: Cumulative daily rainfall (mm) extracted from SAFRAN database over the Radicatel hydrogeological catchment and maximum daily turbidity (NTU) observed at the water treatment plant from 1987 to 2017.

3 Methodology

The proposed cascade modelling for the simulation of SD at WTP incorporates 2 different models that can be used separately
110 or as part of an integrated modelling framework. These components are described in detail in the following sections.

3.1 Expert-based GIS model

The WaterSed model uses a raster-based distributed approach to model the spatial distribution of runoff and soil erosion within a catchment for a given rainfall event. The WaterSed model is an upgrade of the STREAM model (Souchère et al., 1998; Cerdan et al., 2001) simulating hydrological processes by conceptualizing each raster grid cell as a reservoir whose properties
115 are calculated at the event-scale and by routing both water and sediment according to the surface flow network. The WaterSed model requires several datasets to compute runoff and erosion for any location in the catchment: (i) Digital Elevation Model (DEM) to extract slope and the runoff flow network, (ii) stream network (used to modify the DEM to ensure a steady downstream gradient according to the observed hydrographic network) and river width, (iii) land cover and soil texture map to associate each land cover in the catchment with the appropriate soil surface characteristics, (iv) decision tables, adapted for
120 the local conditions, to associate each soil surface characteristic (soil surface crusting, surface roughness and crop cover) observed in the land cover or field with a steady-state infiltration rate, a Manning's roughness coefficient, a single potential sediment concentration and a soil erodibility value (Cerdan et al., 2002a), and (v) rainfall events including total precipitation, antecedent moisture (48h) and effective rainfall duration.

3.3 Deep Neural Network

In this study, a multi-layer feed forward network (DNN) was built under Python version 3.6 using the high-level API Keras (Chollet, 2015) and Tensorflow (Abadi et al., 2016) as a background engine. A multilayer feedforward neural network is an interconnection of perceptrons in which data and calculations flow in a single direction, from the input data to the outputs. The number of layers in a neural network is the number of layers of perceptrons. The number of hidden layers and perceptrons depends on the characteristics of the input data, and there is no specific rule for selecting these parameters (Le et al., 2019). Neural networks are adjusted during a training stage when the parameters are calculated iteratively a gradient descent that seeks to minimize a mean squared error. For efficient learning, input variables were rescaled between 0 and 1 using normalization and re-transformed for the simulations using the normalization parameters. The mean squared error (MSE) was chosen as loss function and the adaptive moment estimation (Adam) was adopted as the model optimization algorithm. After training, the quality of the model is evaluated on a test set. The ability of the model on the test set is known as generalization, and it can be influenced by overfitting during the training stage. To avoid overfitting and increase model robustness, we used two regularization methods in this study; (i) early stopping and (ii) cross-validation. Early stopping is an efficient regularization method that prevents the model from overfitting (Sjörberg and Ljung, 1992). We visually monitored the learning curves and stopped the training as soon as the validation error reached a minimum. Then, we rolled back the model parameters to the point where the validation error was at the minimum. The choice of the training/test set remains arbitrary and may need to be evaluated to compute a robust estimate of model error. With time series data, care must be taken when splitting the data in order to prevent data leakage (Cochrane, 2018). To address this issue, we adapted a month-backward chaining nested cross-validation procedure, which provides an almost unbiased estimate of the true error (Varma and Simon, 2006). The procedure contains an inner loop for parameter tuning, and an outer loop for error estimation (see Fig. B1). The parameters that minimize error are chosen on the inner loop and we add an outer loop, which splits the data into multiple training/test sets. Then, the error is averaged on each split to evaluate the overall performance of the model. The optimum structure and configuration of the network (model design) was found by a classical trial and error procedure (training-evaluation process through optimization of errors; Ortiz-Rodriguez et al., 2013).

3.4 Performance evaluation

The performance of the model was evaluated through the Nash-Sutcliffe Efficiency (NSE) and the root mean square error (RMSE). The root mean square error (RMSE) corresponds to the standard deviation of the residuals (prediction errors). The residuals are a measure of the distance from the data points of the regression line. It evaluates how the predictions match to the observations, and values may range from no fit ($+\infty$) to perfect fit (0) based on the relative change of the data. The Nash-Sutcliffe Efficiency (Nash and Sutcliffe, 1970), indicates the model's ability to predict variables different from the mean, and gives the proportion of the initial variance accounted by the model. NSE values vary between $-\infty$ (poor model) and 1, indicating a perfect fit between observed and predicted values. Finally, to ensure that the model does not suffer from a weakness when

making simulations during extreme events, we performed an additional evaluation using the Generalized Extreme Value distribution (GEV) that is broadly applied to extreme events such as rainfall or river discharges (Carreau et al., 2013; De Michele and Avanzi, 2018). The GEV distribution was fitted to SD time series at the WTP of Radicatel. The maximum annual SD observed at the WTP was extracted for all complete hydrologic years (i.e. annual maximal blocks; n = 22) and the distribution was fitted using the package 'extRemes' of the R software (R Development Core Team, 2008; Gilleland and Katz, 2016; see Fig. C1). We compared the SD simulated at the WTP by the DNN to the values calculated by the GEV.

4 Data

Appropriate data handling can help address various concerns in DNN modelling, such as their ability of generalization beyond training limits (Kourgialas et al., 2015). Moreover, cascade modelling with GIS and DNN can require important computational effort; therefore, the input data must be carefully selected.

4.1 WaterSed data

To compute erosion and runoff, the WaterSed model need a DEM. The DEM (5m resolution) was retrieved from the BD Alti®. Depressions in the DEM were filled according to the algorithm developed by Wang and Liu (2006). The stream network location and width were provided by the BD TOPO®. To define the soil surface characteristics needed by the WaterSed model, we computed a land cover map and a soil texture map. The land cover map was developed for 2016 by combining two national databases: the French Land Parcel Identification System (RPG) and the Soil Observatory of Upper Normandy (OSCOM). The soil texture map (three classes: clay, silt, and sand) was derived from the Regional Soil Referential (RRP) with a precision of 1:250000. The parametrization consists of a characterization of the main parameters influencing runoff and infiltration in the studied context: soil surface crusting, surface roughness and crop cover. These soil surface characteristics were defined for each month and for each crop class according to the cropping calendar developed by Evrard et al. (2010) and Delmas et al. (2012). Steady-state infiltration and potential sediment concentration were assigned to each soil surface characteristics according to Cerdan et al. (2001) and Cerdan et al. (2002b). The WaterSed model requires Manning's roughness coefficient in order to compute flow velocity. Based on soil surface characteristics, Manning's values were derived from surface roughness (Morgan, 2013) and the percentage of crop cover (Gilley et al., 1991). Last, soil surface characteristics were also used to define the soil erodibility factor, varying in the [0-1] interval, by adapting table developed by Souchère et al. (2003a). Calibration parameters were extracted from a previous study (Landemaine, 2016; Landemaine et al., 2020b) conducted in the Austreberthe catchment, located 30 km east of the Radicatel catchment. For the training phase of the DNN, erosion and runoff were calculated with WaterSed model for 269 events to reduce computational efforts, considering that erosion and runoff occurs only for significant rainfall events (P > 2.5 mm/day; threshold below which no effective rainfall is generated; and runoff and sediment discharge entering connected sinkholes was set to 0).

4.2 DNN data

Cumulative daily rainfall (mm) from 1987 to 2017 was extracted from the SAFRAN database over the hydrogeological catchment of Radicatel. SAFRAN data are hydroclimatic data covering France at a resolution of 8 km on an extended Lambert-II projection and produced by MétéoFrance (Quintana-Segui et al., 2008; Vidal et al., 2010). We used the rainfall time series (i.e. daily cumulative rainfall (P); and 48h-antecedent rainfall (P_{48})) retrieved from the SAFRAN database, and WaterSed modelling outputs (runoff (R_{WS}) and sediment discharge (SD_{WS})) at connected sinkholes as input data for the DNN model. The SD time series at the Radicatel WTP was retrieved from turbidity data and considered as the output for the DNN. In this study, turbidity data distribution was explored using scatter plots (Fig. A1). The turbidity data was chosen so that it includes the best dispersion of values (0-370 NTU) and a strong recurrence of extreme turbidity values (> 150 NTU). Thus, two significant hydrologic years (H12-H13; from 10/1/1998 to 9/30/2000) were selected, accounting for 731 daily events. These two hydrologic years were selected as the training/test set of the model. Data were split as a training set (70 %) and a test set (30 %), while respecting the chronology of daily rainfall amounts. The training set was also split into training and stop set to fine-tune the hyperparameters (70-30 %). At the Radicatel WTP, the mean pumping flow rate is estimated to 19733 $m^3\ d^{-1}$, and the volume of water pumped in 2018 was approximately 7.2 million m^3. SD (kg d^{-1}) was assessed considering turbidity time series, mean pumping flow rate at the Radicatel WTP and the relation between turbidity and sediment concentration resulting from field investigations ([mg L^{-1}] = 0.96*[NTU]; R^2 = 0.97; Hanin, 2011). In accordance with previous studies by Masséi et al. (2006) and Hanin (2011) in karstic environment in Normandy, a lag of 1 day was applied to the SD time series to properly match the rainfall input.

4.3 Design storms and land use scenarios

Five design storms were constructed for the cascade modelling approach based on expert knowledge and depth-duration-frequency curves of the French Meteorological Survey (Météo-France) available from 1996-2006 on the studied area (Table 1). We considered winter events for low return period rainfall (0.5 and 2 years). In the studied region, even low daily rainfall depths can lead to severe erosion by water runoff due to saturated soils induced by heavy cumulated rainfall on antecedent days (Le Bissonnais et al., 2002; Evrard et al., 2010). For higher return periods (i.e. 10, 50, and 100 years) we considered spring events that are characterized by stronger rainfall intensities and can be particularly damaging in this region (Evrard et al., 2007).

Table 1: Design storms defined for the Radicatel catchment and considered for the simulations.

Return period (year)	Daily cumulative rainfall (mm)	6-min maximum intensity (mm h^{-1})	48h-antecedent rainfall (mm)	Season
0.5	19	25	45	Winter
2	31.4	25	45	Winter
10	51.9	45	0	Spring

50	74.7	45	0	Spring
100	87	45	0	Spring

Four land use change scenarios for year 2050 were investigated and compared to a baseline land use scenario (2018). The scenarios were incorporated in the model in order to simulate SD variability at the WTP and evaluate their impacts. All scenarios are described below:

- Baseline scenario (S_base): This scenario served as a reference and was built considering the latest available land use data on the catchment (see section 3.3). Existing erosion control measures in 2018 were considered and extracted from a regional database (BD CASTOR; http://bdcastor.fr/). The database contained 45 dams/retention ponds, 16 ponds, 1 fascine, and 4 hedges for the actual land use scenario on the Radicatel catchment, which have been included in the WaterSed model.
- Ploughing up of grassland (S_grass): Based on regional benchmarks (DRAAF; https://draaf.normandie.agriculture.gouv.fr/) for the 1970-2010 period in the studied region (Pays de Caux), we observed an average conversion rate of grasslands up to 900 ha yr^{-1}. Extrapolation by 2050 lead to the conversion of 33% of existing grasslands. The extrapolated rate was applied on the Radicatel catchment. Grasslands were ploughed up based on a slope criterion, taking into account the working conditions of farmers and prioritizing those with mild slopes (<12 %), therefore mainly located on the plateau upstream of the catchment.
- Eco-engineering (S_engi): based on expert knowledge, 181 fascines and 13.1 ha of grass strips were implemented in addition to existing erosion control measures to mitigate runoff/erosion on the catchment and reduce rapid transfer via the connected sinkholes. Grass strips were deployed on the flow paths in the vicinity of the sinkholes. Fascines were deployed throughout the catchment, also on flow paths and along plot boundaries. This scenario allows for a shift from a 0.19 per 30 ha erosion control measure density to nearly 1 per 30 ha which is advised to promote sedimentation and landscape restructuring (Ouvry et al., 2019). The localization was optimized according to the baseline scenario simulations.
- Best farming practices (S_farm): This scenario promotes the adoption of farming practices improving infiltration on the catchment (increasing crop cover or delaying the formation of the slaking crust). 50 % of the plots were randomly selected and applied a 15 % increase in infiltration capacity, respecting the actual proportions of winter and spring crops on the catchment. The applied value was set based on the study of Maetens et al. (2012), who synthetized the reduction in erosion and runoff following different agricultural practices across Europe. According to their results, 15 % increase in infiltration capacity can be regarded as a conservative assumption that can be easily achieved through simplified agricultural techniques (e.g. minimum tillage, no till, direct seeding, crop cover, etc.).

- Coupling eco-engineering and best farming practices (S_farm+engi): Both scenarios, S_farm and S_engi were combined. Experiments carried out in the study area suggested that combining both approaches is necessary to reduce the impact on sensitive or vulnerable areas (Ouvry et al., 2012).

5 Results

5.1 Modelling the inputs

In order to feed the DNN, erosion and runoff were simulated with the WaterSed model for 269 rainfall events over the entire Radicatel catchment. We used the WaterSed parametrization that was carried out on the adjacent catchment (La Lézarde; Landemaine et al., 2020a) and that already proved to be valid on the same area of the Pays de Caux (Landemaine, 2016; Landemaine et al., 2020b) and Belgium (Baartman et al., 2020). For each event, WaterSed outputs (i.e. runoff (R_{WS}; m^3) and sediment discharge (SD_{WS}; kg d^{-1}) values) were independently extracted over the connected sinkholes and summed to consider a unique contribution from the 7 sinkholes to the spring. Figure 3 shows that from October 1998 to September 2000, rainfall events led to a significant cumulative runoff and sediment discharge to the sinkholes (R_{WS} = 213225 m^3; SD_{WS} = 262806 kg). Most of the events occurred during spring (SD_{WS_spr} = 131672 kg; R_{WS_spr} = 94732 m^3) and winter (SD_{WS_win} = 88413 kg; R_{WS_win} = 62590 m^3). Sediment discharge and runoff occurred on the catchment for only 126 events over the 269 simulated. 61 % of the SD_{WS} (i.e. 162168 kg) and 52 % of the R_{WS} (i.e. 110877 m^3) were transported during seven major events (i.e. 0.95 % of the time). Over the 126 events, the maximum R_{WS} and SD_{WS} reached 38471 m^3 and 38380 kg (μSD_{WS} = 2102 kg; μR_{WS} = 1705 m^3; σSD_{WS} = 5826 kg; σR_{WS} = 4437 m^3).

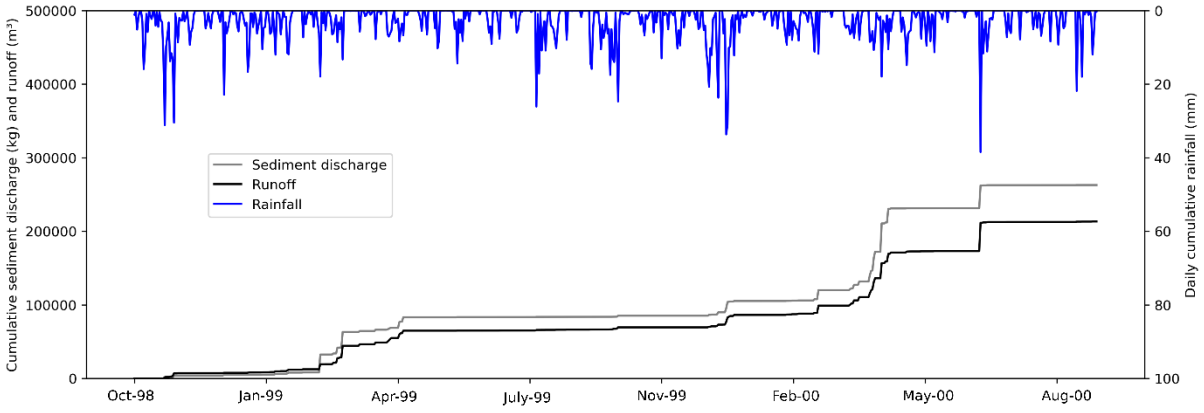


Figure 3: Cumulative sediment discharge (kg) and runoff (m^3) simulated by the WaterSed model, extracted and summed over all positively connected sinkholes from October 1998 to September 2000.

5.2 DNN: Calibration and Generalization

The final structure of the DNN was composed of one input layer with 4 variables, two hidden layers, and one output layer with the targeted variable. We set 40 neurons in the hidden layers as follows: 30-10. We used the rectified linear (ReLU) activation

265 function for the two hidden layers. The optimal number of iterations was set to 15, and the batch size to 1. We used the runoff and sediment discharge simulated by the WaterSed model for the two selected hydrologic years as inputs for the DNN. The rainfall time series (P, P_{48}) available on the catchment were also considered as inputs. Turbidity time series recorded at the WTP was transformed into sediment discharge knowing the mean pumping flow rate at the Radicatel WTP, and considering a mean sediment concentration of 1 mg L^{-1} from field investigations made by Hanin (2011). Thus, sediment discharge at the
270 WTP was considered as output of the DNN. The DNN was trained from October 1998 to February 2000 (n = 511 daily events). The remaining period from March 2000 to September 2000 was used as the test set (n = 220 daily events). Modelling results (Fig. 4) suggested a good agreement between observed and simulated SD at the WTP over the test set, with a NSE criterion reaching 0.82 and a RMSE around 383 kg. The temporal variability was also well reproduced. The results over the training set were slightly worse and reached a NSE of 0.6 and a RMSE of 420 kg. Over the investigated period, the cumulative SD observed
275 at the WTP reached 158611 kg whereas the simulated cumulative SD reached 129253 kg (underestimation of 19 %). Comparing SD_{WS} and observed SD at the WTP resulted in a 60 % recovery rate from WaterSed outputs which was consistent with previous hydrogeological tracing results on connected sinkholes. Hanin (2011) suggested the existence of fast karstic connections between sinkholes and springs on Radicatel catchment, with a 62 % recovery rate. For the 126 events for which erosion and runoff occurred on the catchment, the cumulative amount for both observed and simulated SD was estimated to
280 88114 kg and 78870 kg respectively, suggesting that direct transfers accounts for 55-61 %. These results were consistent with previous published results of Masséi et al. (2003) in same karstic environment (Norville catchment, Normandy, France), which evaluated the proportion of direct transfers at 73 % during erosive events.

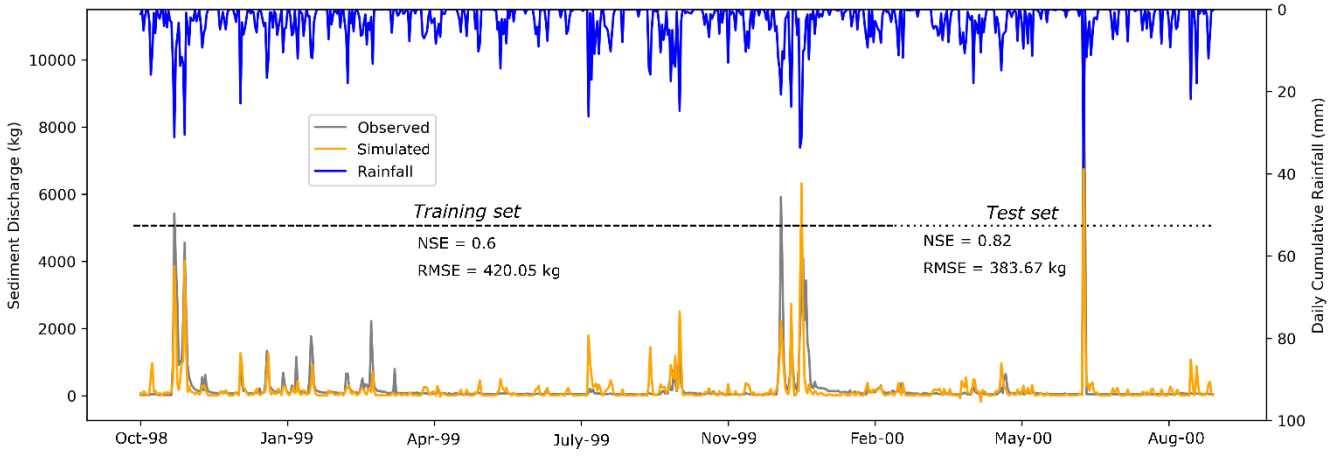


Figure 4: Observed and simulated sediment discharge (kg) at the water treatment plant using the DNN model.

285 The monthly-backward chaining nested cross-validation procedure made it possible to assess the generalization capacity of the DNN on a new data set. This procedure removed one month from the initial dataset (i.e. 731 events) and was repeated twelve times, while keeping at least one full hydrologic year as input for the modelling. The modelling results were more efficient for the inner loop (Fig. 4) than for the outer loops (Fig. 5 and Fig. B1). Overall, the median NSE value for the training

and the test sets for the outer loops was above 0.5. The median RMSE value on the training and the test sets was in the same order of magnitude as for the complete dataset (300 – 500 kg). The difficulties of the model to generalize suggested a classical problem in machine learning, the bias-variance dilemma (Geman et al., 1992). Ideally, the model should accurately reflect irregularities in the training data while generalizing to data testing, but it was not possible to achieve both at the same time. The higher NSE and lower RMSE values on training sets suggest a complex model with high variance and a reliable bias that represents the training data fairly well but presents a learning risk on the test data. The model can thus represent part of the random noise of the learning dataset.

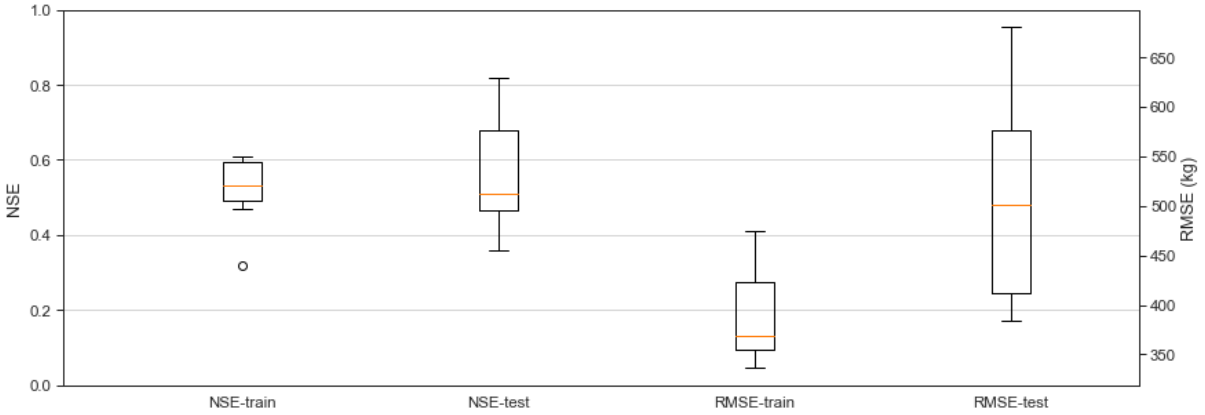


Figure 5: Boxplots of the performance metrics over the training and the test set using the month backward-chaining nested cross-validation.

To validate the approach for extreme values, we simulated overall SD for the baseline scenario at the connected sinkholes and for the five DS using the WaterSed model. Then, we used these modelling results as inputs for the DNN model and applied it to the five DS. Secondly, we compared the results with the calculated GEV distribution (Fig. 6). The SD simulated at the WTP by the DNN model can be considered to be in good agreement with the values calculated by the GEV, even if the latter evidences a high dispersion for higher return periods.

In parallel, we selected the 1% highest sediment discharge on records from October 1998 to September 2000 which represented 28% of the total sediment discharge observed at the WTP, and compared the results with simulated sediment discharge in a scatter plot. We observed a good relation between those two variables ($R^2 = 0.72$; Fig. D1), which strengthens our confidence in the model for simulating extreme events. While the performance of deep learning-based methods to model extreme events is discussed (Zhang et al., 2019), the results obtained here provide confidence in the model's ability to simulate them due to a careful selection of input data that allows the model to learn patterns of extreme events in historical time series.

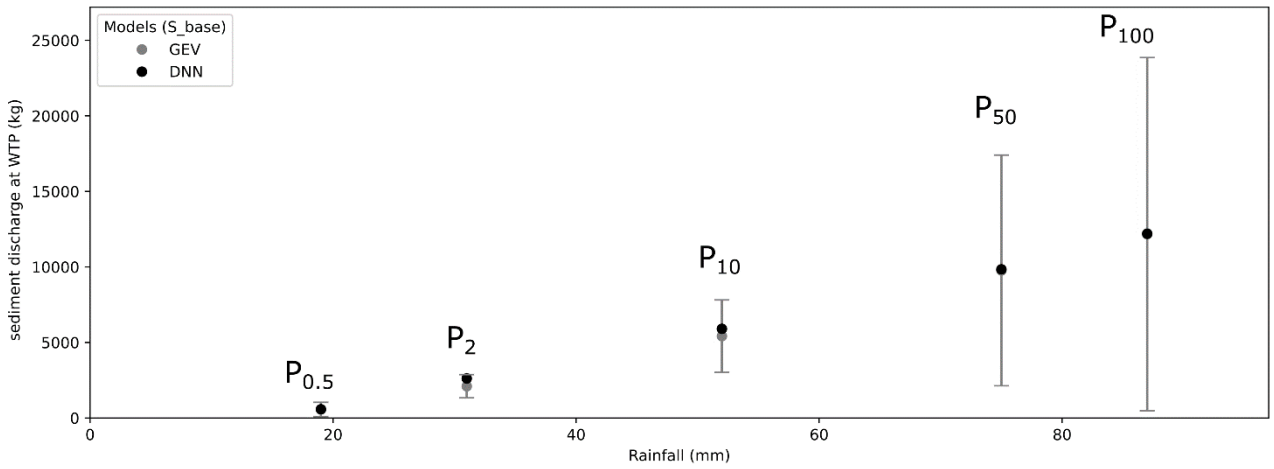


Figure 6: Simulated sediment discharge (kg) at the water treatment plant of Radicatel for the five design storms and the land use baseline scenario by (1) Generalized Extreme Values distribution, and (2) DNN modelling. The grey bars represent the 95% confidence intervals.

5.3 Prospective analysis

Once the calibration was completed and the architecture of the DNN defined, we applied the cascade modelling for all predefined scenarios. As a first step, and using the WaterSed model, we simulated SD_{WS} and R_{WS} at the connected sinkholes for the three additional scenarios and the five DS, comparing them to the baseline scenarios (Fig. 7a-b). For the first scenario (S_grass), 33 % of grasslands were ploughed up, which led to an increase of the spatial extent of runoff generation on the catchment. SD_{S_grass} ranged from 8468 to 188264 kg for the five DS with an average increase of 4.74 % compared to the results of the baseline scenario. The effect on runoff was higher, with an average increase of 8.4 % for the five DS, ranging from 2211 to 67097 m^3. The second scenario (S_engi) considered the implementation of erosion control measures by 2050 (i.e. 181 fascines and 13.1 ha of grass strips), which led to a global decrease in sediment flux rates. Simulated SD_{S_engi} to connected sinkholes ranged from 1659 to 133458 kg, resulting in an average decrease of 44 %. This scenario was more effective on small return periods (i.e. 0.5 and 2 years), with a SD decrease between 79 and 55 %, respectively. Erosion control measures were less effective on the three other DS, with decreases ranging from 25 to 34 %. The effects on runoff were not as important, with an average decrease of 7 %. This was not surprising because the considered management plan only included 13.1 ha of grass strips, and the main interest of the fascines lies in their ability to trap suspended sediments, especially its coarsest elements (AREAS, 2012). The third scenario (S_farm) represented the adoption of good farming practices (i.e. +15 % infiltration capacity) on 50 % of the plots. It can be observed that this scenario significantly reduced sediment production on the hillslopes (Fig. 8a-b). The simulated SD_{S_farm} ranged from 741 to 134001 kg, leading to an average decrease of 49 % compared to the baseline scenario results. The simulated values on runoff ranged from 720 to 47442 m^3. This scenario was more effective in reducing SD and runoff, on average, than the eco-engineering scenario. The SD decrease induced by good farming practices ranged from 25 to 90 % at the connected sinkholes. The fourth scenario (S_farm+engi) combined both effects of the erosion

control measures and good farming practices. We observed the highest decrease in sediment discharge and runoff at connected sinkholes. The simulated values on SD ranged from 547 to 97739 kg and from 598 to 45051 m^3 for runoff.

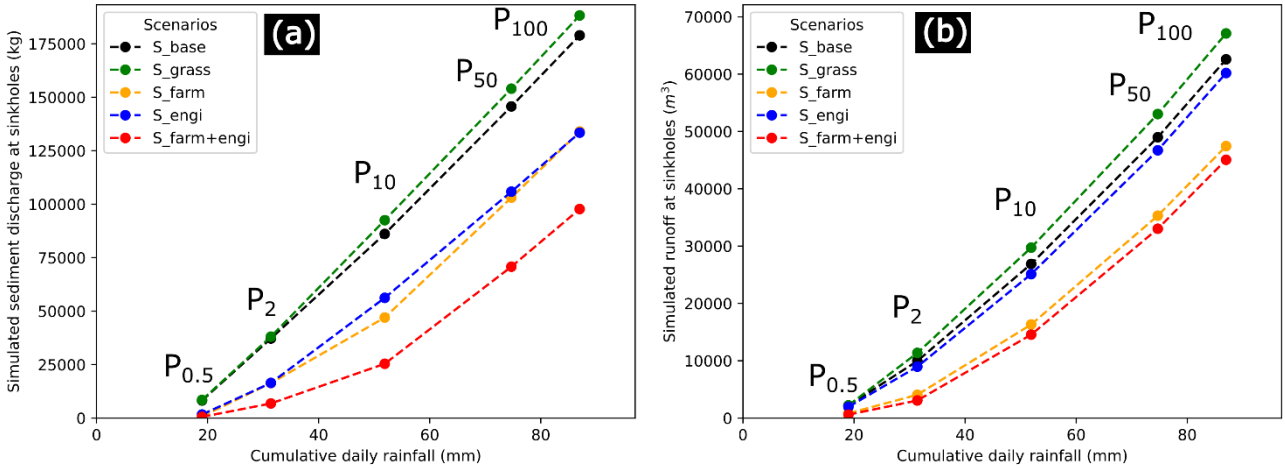


Figure 7: WaterSed modelling results according to the five designed storms for the four scenarios: Simulated (a) sediment discharge (kg) and (b) runoff (m^3), extracted and summed over the 7 connected sinkholes on the Radicatel catchment. (S_base = baseline scenario in 2018; S_grass = 33 % of grasslands ploughed up; S_farm = +15 % infiltration capacity on 50 % of the plots; S_engi = implementation of 181 fascines and 13.1 ha of grass strips, S_farm+engi = combination of S_farm and S_engi).

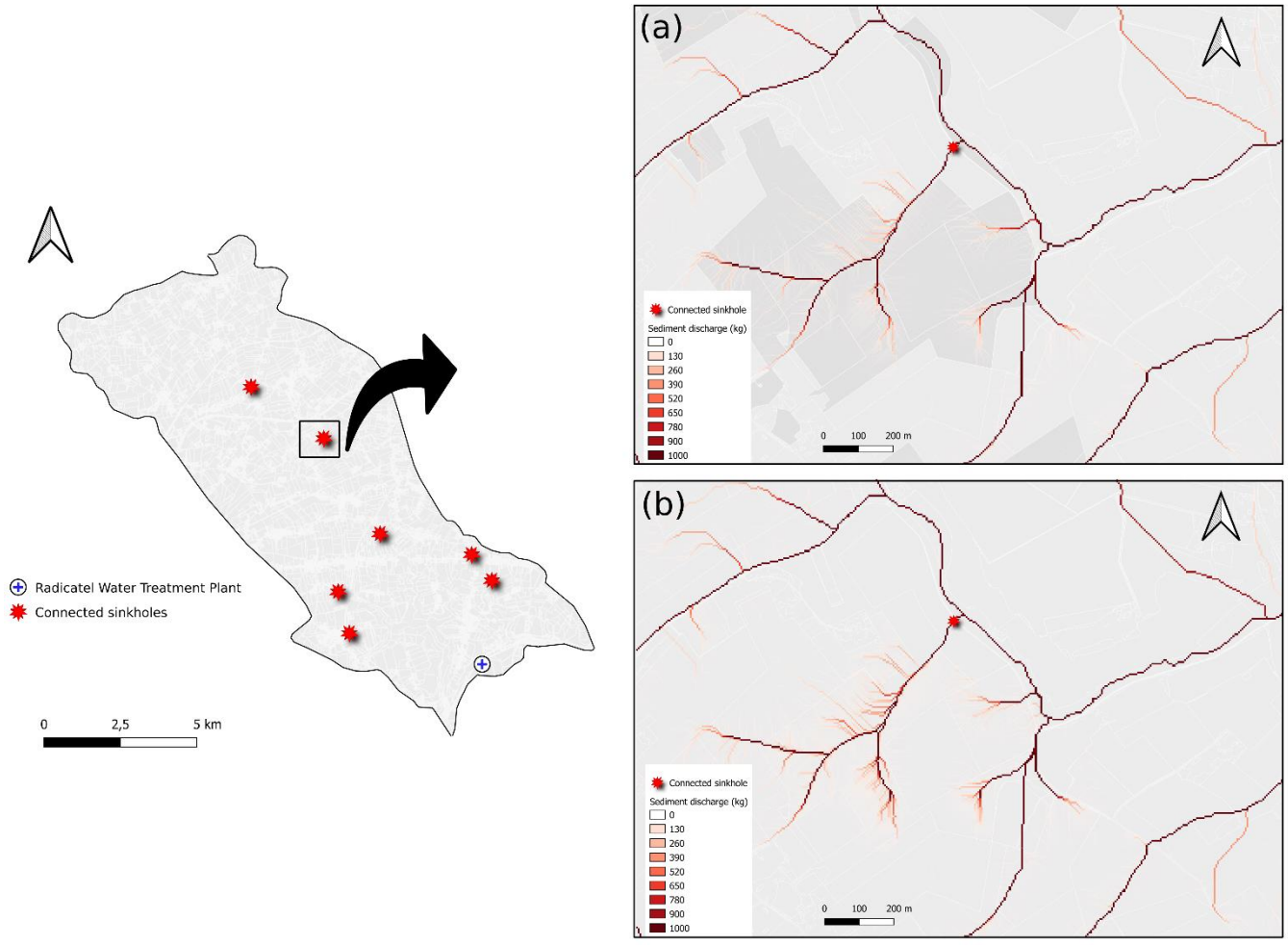


Figure 8: Mapping of flow path and sediment discharge for (a) the scenario including an increase of 15% of the infiltration capacity on selected plots (stronger grey colour) and (b) the baseline scenario, for the 10-year return period designed storm on the Radicatel catchment.

345 DS characteristics and simulated SD/runoff by the WaterSed model at connected sinkholes, and for all scenarios, were injected as inputs in the DNN model. On the baseline scenario, simulated SD at the WTP ranged from 576 to 12200 kg (see section 4.2), and Fig. 9 shows the global trends of the scenarios modelling results. For the S_grass scenario, simulated SD at the WTP were slightly higher than the baseline scenario, reaching 12894 kg for the 100 years DS. This scenario resulted in an average increase of 4.5 % for the SD at the WTP. The SD increase was less significant on small return periods (+0.42-3.29 %) than 350 longer return periods (+5.68-6.82 %). On the second scenario (S_engi), the modelling results suggested a significant SD decrease at the WTP. The simulated SD ranged from 515 to 9810 kg that led to an average decrease of 25.4 %. The SD reduction was particularly effective on the 2 years DS reaching a decrease of 44.7 %. The smallest SD reduction was observed for the 0.5 years DS (10 %) and ranged from 19 to 30 % on longer return periods. These results were consistent with those

found by Fournier et al. (2008), who showed by a parametric interpolation model that erosion control measures at connected sinkholes on the Caumont aquifer led to a 36 % decrease of the turbidity in the Varras WTP (Normandy, France). The third scenario (S_farm) was more efficient with simulated SD ranging from 223 to 9165 kg and an average decrease of 43 %. The SD reduction ranged from 28.9 to 61.3 % and was the more efficient for the 0.5 and 2 years return periods (55.6-61.3 %). For the last scenario (S_farm+engi), we observed a global decrease of 59.6%. The simulated SD values at the WTP ranged from 167 to 7298 kg. This scenario was more efficient for the 0.5, 2, and 10 years return periods (70% in average).

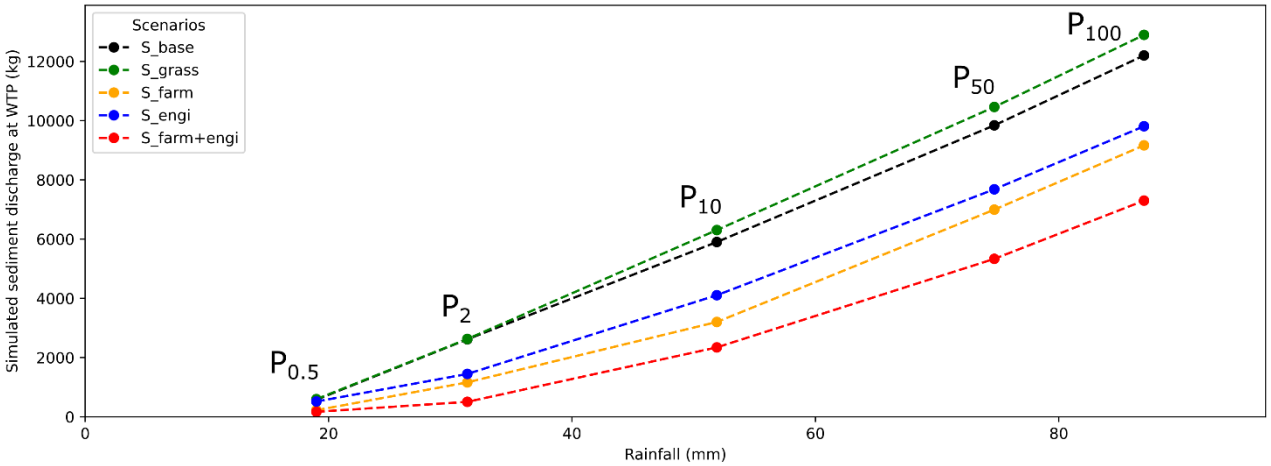


Figure 9: Simulated sediment discharge (kg) for designed storms and all scenarios at the Radicatel water treatment plant (WTP) using DNN model (S_base = baseline scenario in 2018; S_grass = 33 % of grasslands ploughed up; S_farm = +15 % infiltration capacity on 50 % of the plots; S_engi = implementation of 181 fascines and 13.1 ha of grass strips, S_farm+engi = combination of S_farm and S_engi).

6 Discussion

5.1 Cascade modelling approach: strength and uncertainties

The original cascade modelling approach developed in this study allowed the evaluation of surface and subsurface sediment transfer processes. As karstic processes are complex and difficult to model due to the lack of knowledge of their geometry, the cascade modelling with a process-based GIS model and a data-driven model (i.e. DNN) proved to be a powerful tool. This approach is efficient to assess the impact of different land-use scenarios on drinking water quality at drinking water treatment plants. Despite very encouraging results, one may notice that the cascade model was trained on only two years of complete selected measurements and therefore may not have captured the full distribution of possible cumulative daily rainfall on the catchment and/or turbidity values at the drinking water treatment plant. The degree of confidence in the model's output could be further improved with longer time series. One limitation of the presented cascade approach is the data availability and its ability to be validated at other study sites, as the turbidity data backup are not always properly done in all WTP. One specific issue for DL applications in hydrological sciences, as mentioned by Sit et al. (2020) is that data provided by the authorities are dispersed and they occasionally have mismatches in temporal or spatial coverage. We thus encourage the drinking water

supplier to keep turbidity records to allow the application of these data-driven models and to ensure that the time series do not have interruptions.

6.2 Implications for Future Land Use Strategies

The modelling results in this study suggest that two different land use strategies (i.e. increase of infiltration capacity and eco-engineering) can significantly reduce the SD incoming to connected sinkholes during extreme rainfall events (up to 90 %), and therefore, decrease the SD at the WTP (from 10 to 61%). The first strategy leads to a decrease in sediment production through simplified cultural techniques, while the second affects transfer processes. The adoption of better farming practices (e.g increase of crop cover, no-tillage, reduced tillage) inducing an increase of 15 % of the infiltration capacity on 50 % of the agricultural plots of the catchment, appears to be slightly more efficient than the implementation of eco-engineering infrastructures (181 fascines + 13.1 ha of grass strips) for reducing SD at the WTP. These results thus show that it is more interesting to adopt a land use strategy aimed at reducing hydro-sedimentary production directly on cultivated plots than transfer processes on the catchment, but these strategies could be combined. Additional simulations integrating both effects of best farming practices on cultivated plots and the implementation of eco-engineering infrastructures show a tenfold effect. As illustrated in Fig. 9, the coupling of the two strategies makes it possible to reduce significantly the SD at the WTP. The combined effect will not add up, but we can expect an improved decrease in sediment discharge at the WTP from 40 % to 80 %. Both land use strategies can also provide interesting effects on biodiversity and ecosystem services (Posthumus et al., 2015) and/or can improve soil resilience and promote sustainable agriculture (Lal, 2014). Public policies leading to the implementation of erosion control measures can be economically viable and efficient (Patault et al., 2021b). Even if some simplified cultural techniques may imply negative returns for farmers, they may be eligible for agri-environment payments (Posthumus et al., 2015). The implementation of these land use strategies may also require specific maintenance to keep their initial performance (Frankl et al., 2018) or specific training and machines for farmers.

According to the simulations, the ploughing up of 33 % of grasslands for the benefit of agricultural plots on the catchment by 2050, will not increase significantly the SD at the WTP (less than 5 % in average). Our simulations just extended the current trend observed in the studied region. Despite all, some precautions must be taken regarding the results of this scenario, which could be higher or lower depending on the localization of the ploughed-up plots. The same observations were made by Souchère et al. (2005) who suggested, according to their results, that the efficiency of all developments in reducing erosion and runoff is linked to their location within the catchment. Indeed, it is well known that hydrologic connectivity may lead to increments in runoff (Appels et al., 2011).

7 Conclusions

In this study, a new cascade modelling approach was developed in order to help decision-makers choosing an adapted erosion and runoff management strategy to reduce the impact of sediment discharge on drinking water supply. The expert-based GIS model (WaterSed) was used to simulate erosion and runoff at connected sinkholes (positive dye tracing) on the Radicatel catchment (Normandy, France) and used to feed a data-driven model (i.e. Deep Neural Network) to simulate karstic transfers. This new approach does not require a knowledge of the geometry of the karstic system studied and demonstrated the value of understanding hillslope erosion and runoff processes to model underground hydro-sedimentary transfers in karstic systems. Our modelling results suggest that the cascade model performed well during the calibrating phase. The cascade model was able to generalize on unknown datasets through an adapted monthly backward chaining nested cross-validation procedure and the cascade model was efficient for simulating extreme events. The results also suggest that a land use change scenario considering the adoption of simplified cultural techniques can significantly reduce sediment discharge at the water treatment plant (up to 61%). This scenario also outperformed the one which considered the implementation of eco-engineering control measures reducing erosion and runoff on the catchment (up to 45% at the water treatment plant). The coupling of the two previous land use strategies can be even more effective since it operates both on the hydro-sedimentary production and transfer processes (decreasing SD at WTP from 40 to 80 %). Finally, ploughing up 33% of actual grasslands on the catchment will not significantly increase sediment discharge at the water treatment plant. However, the results might be influenced by the spatial organization of grasslands on the catchment that is a key parameter for hydro-sedimentary connectivity and hydro-sedimentary transfers to the sinkholes. In the framework of this study, we suggest conducting a specific study on this hypothesis.

Appendix A: Scatter plot of turbidity time series

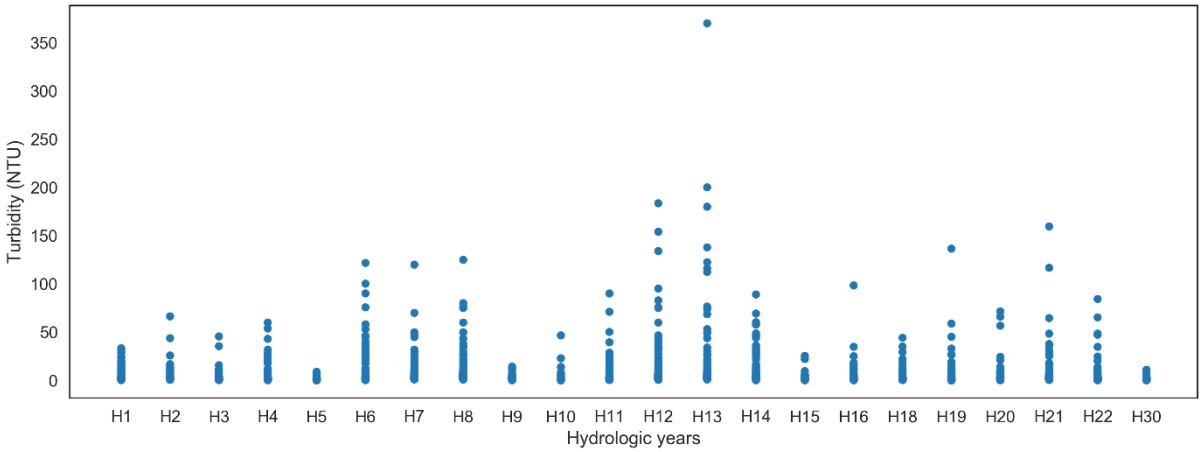


Figure A1: Scatter plot of turbidity time series recorded at the water treatment plant in the Radicatel catchment, Normandy, France.

Appendix B: Month-backward chaining nested cross-validation

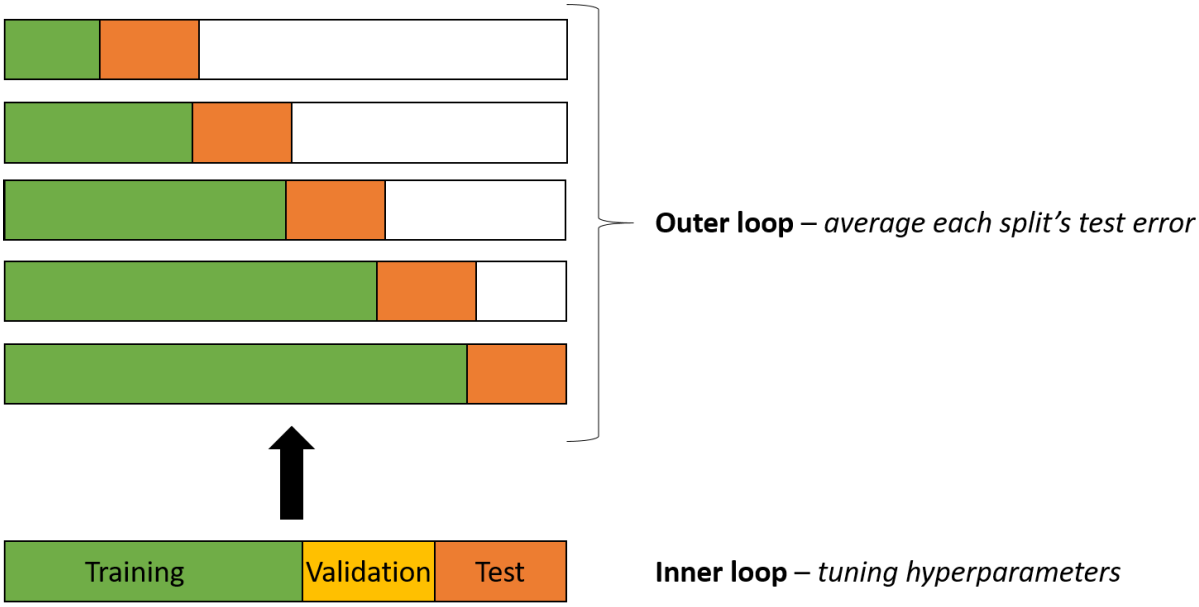
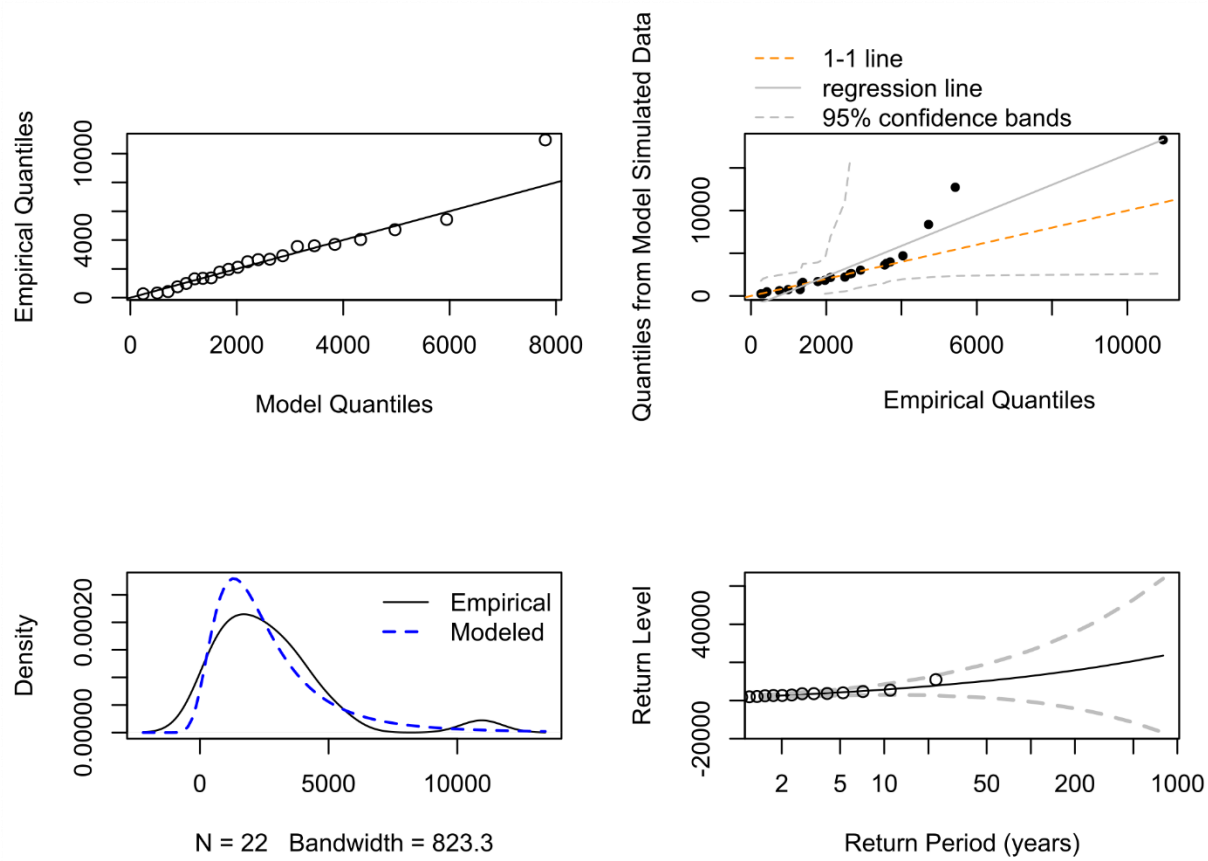


Figure B1: Month-backward chaining nested cross-validation procedure developed to test the generalization capacity of the model (modified after Cochrane (2018)).

Appendix C: Generalized Extreme Values Distribution



435 **Figure C1: Generalized Extreme Value distribution modeled with the 'exTremes' R package on the sediment discharge time series (22 hydrologic years) at the water treatment plant of the Radicatel catchment.**

Appendix D: Scatter plot of extreme events

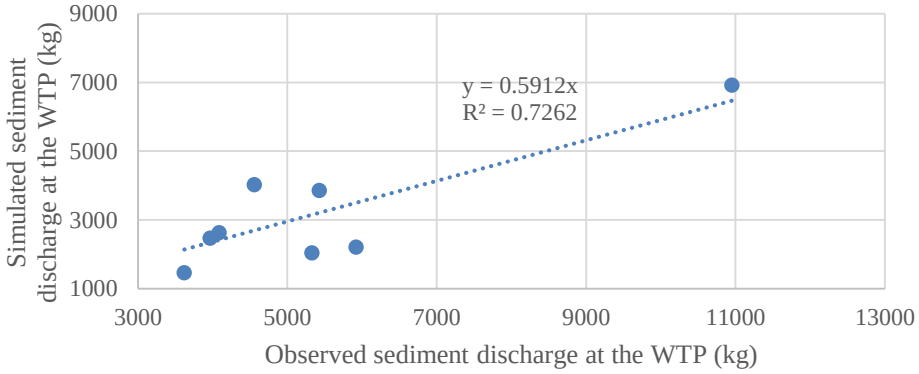


Figure D1: Scatter plot of the observed versus simulated sediment discharge by the DNN at the WTP for the 1% highest sediment discharge on records.

Author contribution

Conceptualization EP, MF, BL, VL, OC, JFO, JL, AS; Formal analysis EP, VL; Funding acquisition BL, OC, JFO; Investigation EP, VL, JL, AS; Methodology EP, MF, VL; Project administration MF, BL, OC, JFO; Software EP, VL; Supervision EP, VL; Validation EP, VL, MF, OC; Visualization EP; Writing – original draft EP, VL, MF, OC; Writing – review and editing

Data & Code availability

The code that supports the findings of this study are available on request from the corresponding author (MF).

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

This study is based on research undertaken as part of EVAPORE project “EVAluation de l’efficacité des POlitiques publiques pour les actions visant à REduire les impacts du ruissellement” to evaluate efficiency of public policy to reduce erosion and runoff impacts in Normandy, France. Authors are thankful to the University of Rouen, AREAS, BRGM, and Seine-Normandy Water Agency who co-founded this project. Authors are also thankful to the LHSM ‘Le Havre Seine Métropole’ who provided access to their data.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro,C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, available at: https://www.tensorflow.org/

Appels, W.M., bogaart, P.W., van der Zee, S.E.A.T.M. (2011). Influence of spatial variations of microtopography and infiltration on surface runoff and field scale hydrological connectivity. Advances in Water Resources, 34(2), 303-313 pp. https://doi.org/10.1016/j.advwatres.2010.12.003

AREAS (2012). Fascines & Haies pour réduire les effets du ruissellement érosif. Caractérisation de l'efficacité et conditions d'utilisation. 68p.

ARS (2013). Bilan 2013 de la qualité des eaux destinées à la consommation humaine et la protection des captages en Seine-Maritime. https://www.normandie.ars.sante.fr/sites/default/files/2017-03/bilan_76_qualite_eau_2013.pdf

Bai, Y., Chen, Z., Xie, J., Li, C. (2016). Daily reservoir inflow forecasting using multiscale deep feature learning with hybrid models. Journal of Hydrology, 532, 193-206 pp.

Baartman, J.E.M., Nunes, J.P., Masselink, R., Darboux, F., Bielders, C., Degre, A., Cantreul, V., Cerdan, O., Grangeon, T., Fiener, P., Wilken, F., Schindewolf, M., Wainwright, J. (2020). What do models tell us about water and sediment connectivity?, Geomorphology. https://doi.org/10.1016/j.geomorph.2020.107300

BRGM (2017). Analyse de l'érosion des sols sur le vignoble champenois de la Marne. BRGM/RP-66688-FR. Février 2017.

Carreau, J., Neppel, L., Arnaud, P., Cantet, P. (2013). Extreme rainfall analysis at ungauged sites in the south of France: Comparison of three approaches. Journal de la Société Française de Statistique, 154(2), 119-138 pp.

Cerdan, O., Souchère, V., Lecomte, V., Couturier, A., Le Bissonnais, Y. (2001). Incorporating soil surface crusting processes in an expert-based runoff model: Sealing and transfer by runoff and erosion related to agricultural management. Catena, 46, 189-205 pp. https://doi.org/10.1016/S0341-8162(01)00166-7

Cerdan, O., Le Bissonnais, Y., Souchère, V., Martin, P., Lecomte, V. (2002a). Sediment concentration in interrill flow : interactions between soil surface conditions, vegetation and rainfall. Earth Surf. Process. Landforms, 27, 193-205. https://doi.org/10.1002/esp.314

Cerdan, O., Le Bissonnais, Y., Couturier, A., Saby, N. (2002b). Modelling interrill erosion in small cultivated catchments. Hydrol. Process., 3226, 3215-3226 pp. https://doi.org/10.1002/hyp.1098

Chédeville, S. (2015). Etude de la variabilité du fonctionnement hydro-sédimentaire des karsts de l'Ouest du Bassin de Paris à partir de la comparaison des remplissages sédimentaires karstiques anciens, actuels et du signal turbide des eaux souterraines. Thèse de doctorat, Université de Rouen-Normandie, France, 458p.

Chollet, F. (2015). Keras, available at https://github.com/fchollet/keras

Cochrane, C. (2018). Time Series Nested Cross-Validation. https://towardsdatascience.com/time-series-nested-cross-validation-76adba623eb9

Delmas, M., Pak, L.T., Cerdan, O., Souchère, V., Le Bissonnais, Y., Couturier, A., Sorel, L. (2012). Erosion and sediment budget across scale : a case study in a catchment of the European loess belt. J. Hydrol. 420, 255-263 pp. https://doi.org/10.1016/j.jhydrol.2011.12.008

De Michele, C. & Avanzi, F. (2018). Superstatistical distribution of daily precipitation extremes: A worldwide assessment. Scientific Reports, 8:14204; doi: 10.1038/s41598018-31838-z

De Vente, J., Poesen, J., Verstraeten, G., Govers, G., Vanmaercke, M., Van Rompaey, A., Arabkhedri, M., Boix-Fayos, C. (2013). Predicting soil erosion and sediment yield at regional scales: Where do we stand ? Earth-Sciences reviews, 127, 16-29 pp. https://doi.org/10.1016/j.earscirev.2013.08.014

Evrard, O., Bielders, L.C., Vandaele, K., van Wesemael, B. (2007). Spatial and temporal variation of muddy floods in central Belgium, off-site impacts and potential control measures. Catena, 70(3), 443-454 pp. https://doi.org/10.1016/j.catena.2006.11.011

Evrard, O., Nord, G., Cerdan, O., Souchère, V., Le Bissonnais, Y., Bonté, P. (2010). Modelling the impact of land use change and rainfall seasonality on sediment export from an agricultural catchment of the northwestern European loess belt. Agriculture, Ecosystems and Environment, Elsevier Masson, 138(1-2), 83-94 pp. https://doi.org/10.1016/j.agee.2010.04.003

Fournier, M., Mouhri, A., Ratajczak, M., Rossi, A., Slimani, S., Mesquita, J. (2008). Fonctionnement hydrogéologique de l'aquifère de Caumont et incidence des aménagements de bassin versant sur la qualité des eaux du forage des Varras. https://www.unicaen.fr/m2c/IMG/pdf/rapportsersaep.pdf?262/447a7b605f6dbc8a8df025b439812a0bc73d39bf

Frankl A., Prêtre, V., Nyssen, J., Salvador, P-G. (2018). The success of recent land management efforts to reduce soil erosion in northern France. Geomorphology, 303, 84-93 pp. https://doi.org/10.1016/j.geomorph.2017.11.018

Geman, S., Bienenstock, E., Doursat, R. (1992). Neural networks and the bias/variance dilemma. Neural Computation, 4, 1-58 pp.

Gilley, J.E., Kottwitz, E.R., Wieman, G.A. (1991). Roughness coefficients for selected residue materials. J. Irrig. Drain. Eng. 117, 503-514 pp.

Gilleland, E. & Katz, R.W. (2016). extRemes 2.0: An extreme Value Analysis Package in R. Journal of Statistic Software, 72(8), 1-39 pp; doi:10.18637/jss.v072.i08

Hafeez, S., Sing Wong, M., Chak Ho, H., Nazeer, M., Nichol, J., Abbas, S., Tang, D., Ho Lee, K., Pun, L. (2019). Comparison of machine learning algorithms for retrieval of water quality indicators in case-II waters: a case study of Hong-Kong. Remote Sensing, 11, 617, https://doi.org/10.3390/rs11060617

Hanin, G. (2011). Contrôles structural et hydrogéologique sur la dynamique d'un champ captant en contexte crayeux karstique et sa sensibilité aux variations du signal climatique: Implications en matière de vulnérabilité de la ressource. PhD thesis, University of Rouen-Normandy, France, 320p.

Hartmann, A., Goldscheider, N., Wagener, T., Lange, J. and Weiler, M. (2014). Karst water resources in a changing world: Review of hydrological modeling approaches. Rev. Geophys., 52, 218-242, https://doi.org/10.1002/201RG000443

Hu, C., Wu, Q., Jian, S., Li, N., Lou, Z. (2018). Deep learning with a long short-term memory networks approach for rainfall-runoff simulation. Water, 10, 1543; doi:10.3390/w10111543

Jourde, H., Masséi, N., Mazzilli, N., Binet, S., Batiot-Guilhe, C., Labat, D., Steinmann, M., Bailly-Comte, V., Seidel, J-L., Arfib, B., Charlier, J-B., Guinot, V., Jardani, A., Fournier, M., Aliouache, M., Babic, M., Bertrand, C., Brunet, P., Boyer, J-F., Bricquet, J-P., Camboulive, T., Carrière, S-D., Celle-Jeanton, H., Chalikakis, K., Chen, N., Cholet, C., Clauzon, V., Dal Soglio, L., Danquigny, C., Défargue, C., Denimal, S., Emblanch, C., Hernandez, F., Gillon, M., Gutierrez, A., Hidalgo Sanchez, L., Hery, M., Houillon, N., Johannet, A., Jouves, J., Jozja, N., Ladouche, B., Leonardi, V., Lorette, G., Loup, C., Marchand, P., de Montety, V., Muller, R., Ollivier, C., Sivelle, V., Lastennet, R., Lecoq, N., Maréchal, J-C., Perotin, L., Perrin, J., Petre, M-A., Peyraube, N., Pistre, S., Plagnes, V., Probst, J-L., Simler, R., Stefani, V., Valdes-Lao, D., Viseur, S., Wang, X. (2018). SNO KARST: A French network of observatories for the multidisciplinary study of critical zone processes in karst watersheds and aquifers. Vadose Zone J. 17:180094. https://doi.org/10.2136/vzj2018.04.0094

Kourgialas, N.N., Dokou, Z., Karatzas, G.P. (2015). Statistical analysis and ANN modeling for predicting hydrological extremes under climate change scenarios: The example of a small Mediterranean agro-watershed. Journal of Environmental Management, 154, 86-101 pp.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., Herregger, M. (2018). Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. Hydrol. Earth Syst. Sci., 22, 6005-6022 pp.

Kratzert, F., Klotz, D., Herrneger, M., Sampson, A.K., Hochreiter, S., Nearing, G.S. (2019b). Toward Improved Predictions in Ungauged Basins: Exploiting the Power of Machine Learning. Water Resources Research. 10.1029/2019WR0260065

Laflen, J.M., Lane, L.J., Foster, G.R. (1991). WEPP: a new generation of erosion prediction technology. Journal of Soil and Water Conservation, 46, 34-38 pp.

Laignel, B. (2003). Caractérisation et dynamique érosive de systèmes géomorphologiques continentaux sur substrat crayeux. Exemple de l'Ouest du Bassin de Paris dans le contexte Nord-Ouest Européen, mémoire HDR, 138p.

Lal, R. (2014). Soil conservation and ecosystem services. International Soil and Water Conservation Research. 2(3), 36-47 pp. https://doi.org/10.1016/S2095-6339(15)30021-6

Landemaine, V. (2016). Erosion des sols et transferts sédimentaires sur les bassins versants de l'Ouest du bassin de Paris: analyse, quantification et modélisation à l'échelle pluriannuelle. PhD thesis. University of Rouen-Normandy, France, 236p.

Landemaine, V., Soulignac, A., Cerdan, O. (2020a). Analyse coût-bénéfice des actions de lutte contre le ruissellement et l'érosion des sols sur le bassin de la Lézarde. Rapport final. BRGM/RP-69650-FR, 142p.

Landemaine, V., Cerdan, O., Grangeon, T., Vandromme, R., Laignel, B., Evrard, O., Salvador-Blanes, S., Laceby, P. (2020b). Saturation-excess overland flow in the European loess belt: An underestimated process? In prep.

Lautridou, J-P. (1985). Le cycle périglaciaire pléistocène en Europe du Nord-Ouest et plus particulièrement en Normandie. PhD thesis, University of Caen, France, 908p.

Le, X-H., Ho, H.V., Lee, G. & Jung, S. (2019). Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. Water, 11, 1387 ; https://doi.org/10.3390/w11071387

Le Bissonnais, Y., Thorette, J., Bardet, C., Daroussin, J. (2002). L'érosion hydrique des sols en France. Rapport INRA-IFEN, 106p. http://eduterre.ens-lyon.fr/thematiques/sol/degradation-du-sol/erosion-hydrique-2002-br.pdf

Maetens, W., Poesen, J., Vanmaercke, M. (2012). How effective are soil conservation techniques in reducing plot runoff and soil loss in Europe and the Mediterranean ? Earth-Science Rev. 115, 21-36 pp. https://doi.org/10.1016/j.earscirev.2012.08.003

Mangin, A. (1984). Pour une meilleure connaissance des systèmes hydrologiques à partir des analyses corrélatoire et spectrale. Journal of Hydrology, 67 (1-4), 25-43 pp.

Masséi, N., Wang, H.Q., Dupont, J.P., Rodet, J., Laignel, B. (2003). Assessment of direct transfer and resuspension of particles during turbid floods at a karstic spring. Journal of Hydrology, 275, 109-121 pp.

Masséi, N., Dupont, J.P., Mahler, B.J., Laignel, B., Fournier, M., Valdes, D., Ogier, S. (2006). Investigating transport properties and turbidity dynamics of a karst aquifer using correlation, spectral, and wavelet analyses. Journal of Hydrology, 329, 244-255 pp.

Merritt, W.S., Letcher, R.A., Jakeman, A.J. (2003). A review of erosion and sediment transport models. Environmental modelling & software, 18(8-9), 761-799 pp.

Meyers, G., Kapelan, Z., Keedwell, E., Randall-Smith, M. (2016). Short-term forecasting of turbidity in a UK water distribution system. Procedia Engineering, 154, 1140-1147 pp.

Meyers, G., Kapelan, Z., Keedwell, E. (2017). Short-term forecasting of turbidity in trunk main networks. Water research, 124, 67-76 pp.

Morgan, R.P.C. (2013). Soil Erosion and Conservation. https://doi.org/10.1002/9781118351475.ch22

Nash, J.E. & Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I-a Discussion of Principles*. Journal of Hydrology, 10, 282-290 pp.

Ortiz-Rodriguez, J.M., Martinez-Blanco, M.R, Cervantes-Viramontes, J.M., Vega-Carrillo, H.R. (2013). Robust design of artificial neural networks methodology in neutron spectrometry. In Artificial Neural Networks – Architectures and Applications – Edition 1. Chapter 4, INTECH. https://dx.doi.org/10.5772/51274

Ouvry, J-F., Coufourier, N., Richet, J-B., Lhériteau, M., Pivain, S. (2012). Maîtrise du ruissellement et de l'érosion des sols : expérimentations sur les pratiques culturales : synthèse des résultats du ruissellement et d'érosion. Expérimentations sur les pratiques culturales, 2001-2010. 2021. Hal-02811122

Ouvry, J-F., Richet, J-B., Saunier, M. (2019). « Le rebocagement » : une réponse pertinente face aux enjeux érosifs ? Retour d'expérience du Pays de Caux, Revue Science Eaux & Territoires, Ressources en eau, ressources bocagères, 30, 54-59 pp. doi : 10.14758/SET-REVUE.2019.4.11

Patault, E., Ledun, J., Landemaine, V., Soulignac, A., Richet, J-B., Fournier, M., Ouvry, J-F., Cerdan, O., Laignel, B. (2021a). Analysis of off-site economic costs induced by runoff and soil erosion: Example of two areas in the northwestern European loess belt for the last two decades (Normandy, France). Land Use Policy, 108, 105541. https://doi.org/10.1016/j.landusepol.2021.105541

Patault, E., Soulignac, A., Landemaine, V., Ledun, J., Allard, E., Fournier, M., Ouvry, J-F., Cerdan, O., Laignel, B. (2021b). Analyse coût-bénéfice du programme d'actions visant à réduire les impacts du ruissellement et de l'érosion en Haute-Normandie : évaluation des actions passées et projections futures sur le bassin versant de la Lézarde, LHB, 107:1, 1-10, https://doi.org/10.108/00186368.2021.1912963.

Posthumus, H., Deeks, L.K., Rickson, R.J., Quinton, J.N. (2015). Costs and benefits of erosion control measures in the UK. Soil Use and Management, 31, 16-33 pp. https://doi.org/10.1111/sum.12057

Quintana-Segui, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchistéguy, L., Morel, S. (2008). Analysis of near surface atmospheric variables: validation of the SAFRAN analysis over France. Journal of Applied Meteorology and Climatology, 47, 92-107 pp. https://doi.org/10.1175/2007JAMC1636.1

R Development Core Team (2008). R: A language and environment for statistical computing.R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Renard, K.G & Freidmund J.R. (1994). Using monthly precipitation data to estimate the R factor in the revised USLE. Journal of Hydrology, 157, 287-306 pp.

Savary, M., Johannet, A., Masséi, N., Dupont, J-P., Hauchard, E. (2017). Operational turbidity forecast using both recurrent and feed-forward based multilayer perceptrons. In: Rojas, I., Pomares, H., Valenzuela, O. (eds) Advances in time series analysis and forecasting. ITISE 2016. Contributions to Statistics. Springer, Cham.

Shen, C. (2018). A transdisciplinary review of deep learning research and its relevance for water resources scientists. Water Resources Research, 54, 8558-8593 pp. https://doi.org/10.1029/2018WR022643

Shen, C., Laloy, E., Elshorbagy, A., Albert, A., Bales, J., Chang, F-J., Ganguly, S., Hsu, K-L., Kifer, D., Fang, Z., Fang, K., Li, D., Li, X., Tsai, W-P. (2018). HESS Opinions: Incubating deep-learning-powered hydrologic science advances as a community. Hydrol. Earth Syst. Sci., 22, 5639-5656 pp.

Siou, L.K.A., Johannet, A., Borrell, V., Pistre, S. (2011). Complexity selection of a neural network model for karst flood forecasting: The case of the Lez Basin (southern France). Journal of Hydrology, 403, 367-380 pp.

Sit, M., Demiray, B.Z., Xiang, Z., Ewing, G.J., Sermet, Y., Demir, I. (2020). A comprehensive review of deep learning applications in hydrology and water resources. Water Science and Technology. Doi: 10.2166/wst.2020.369

Sjörberg, J. and Ljung, L. (1992). Overtraining, regularization, and searching for minimum in neural networks. IFAC Adaptative Systems in Control and Signal Processing, Grenoble, France, 1992.

Souchère, V., King, D., Daroussin, J., Papy, F., Capillon, A. (1998). Effects of tillage on runoff directions: Consequences on runoff contributing area within agricultural catchments. J. Hydrol. 206, 256-267 pp. https://doi.org/10.1016/S0022-1694(98)00103-6

Souchère, V., Cerdan, O., Ludwig, B., Le Bissonnais, Y., Couturier, A., Papy, F. (2003a). Modelling ephemeral gully erosion in small cultivated catchments. Catena, 50, 489-505 pp. https://doi.org/10.1016/S0341-8162(02)00124-8

Souchère, V., Cerdan, O., Dubreuil, N., Le Bissonnais, Y., King, C. (2005). Modelling the impact of agri-environmental scenarios on runoff in a cultivated catchment (Normandy, France). Catena, 61, 229-240 pp.

Stevenson, M. & Bravo, C. (2019). Advanced turbidity prediction for operational water supply planning. Decision Support Systems, 119, 72-84 pp.

Takken, I., Beuselinck, L., Nachtergaele, J., Govers, G., Poesen, J., Degraer, G. Spatial Evaluation of physically-based distributed erosion model (LISEM). Catena, 37(3-4), 431-447 pp.

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics, 7, 91; https://doi.org/10.1186/1471-2105-7-91

Verstraeten, G., Prosser, I.P., Fogarty, P. (2007). Predicting the spatial patterns of hillslope sediment delivery to river channels in the Murrumbidgee catchment, Australia. Journal of Hydrology, 334 (3-4), 440-454 pp. https://doi.org/10.1016/j.hydrol.2006.10.025

Vidal, J.P., Martin, E., Franchisteguy, L., Baillon, M. & Soubeyroux, J.M. (2010). A 50-year high-resolution atmospheric reanalysis over France with the Safran system. International Journal of Climatology, 30, 1627-1644 pp.

Wang, L. & Liu, H. (2006). An efficient method for identifying and filling surface depressions in digital elevation model for hydrologic analysis and modelling. Int. J. Geogr. Inf. Sci. 20, 193-213 pp. https://doi:10.1080/13658810500433453

Yaseen, Z.M., El-shafie, A., Jaafar, O., Afan, H.A., Sayl, N.K. (2015). Artificial intelligence based models for stream-flow forecasting: 2000-2015. Journal of Hydrology, 530, 829-844 pp. http://dx.doi.org/10.1016/j.jhydrol.2015.10.038

Zhang, M.D.D., Yan, M.X.P., He, X. (2019). Modeling extreme events on time series prediction. KDD'19, august 4-8, 2019, anchorage, AK, USA. https://doi.org/10.1145/3292500.3330896