Catena 212 (2022) 105953

Contents lists available at ScienceDirect

Catena

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

A comparison of performance of SWAT and machine learning models for predicting sediment load in a forested Basin, Northern Spain

Patricia Jimeno-Sáez [a,*], Raquel Martínez-España [b,c], Javier Casalí [d], Julio Pérez-Sánchez [a,e], Javier Senent-Aparicio [a]

[a] Department of Civil Engineering, Universidad Católica San Antonio de Murcia, Campus de Los Jerónimos s/n, 30107 Guadalupe, Murcia, Spain
[b] Department of Computer Engineering, Universidad Católica San Antonio de Murcia, Campus de Los Jerónimos s/n, 30107 Guadalupe, Murcia, Spain
[c] Department of Information and Communication Engineering, Universidad de Murcia, Espinardo 30100, Murcia, Spain
[d] Department of Engineering, Public University of Navarra, Campus de Arrosadía s/n, 31006 Pamplona, Navarra, Spain
[e] Department of Civil Engineering, Universidad de Las Palmas de Gran Canaria, Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain

ARTICLE INFO

Keywords:
SWAT
Machine learning
M5P
Random forest
Suspended sediment load
Oskotz

ABSTRACT

In water bodies, sediment transport is a potential source of numerous negative effects on water resource projects and can damage environmental services. Two machine learning (ML) algorithms, the M5P and random forest (RF) models, have been explored for the first time as alternatives to the Soil and Water Assessment Tool (SWAT) model to estimate suspended sediment load (SSL) in the Oskotz river basin, a forested experimental basin in Navarra, northern Spain. In the ML models, streamflow and precipitation data were used to estimate daily SSL, testing different combinations of these inputs. The ML models were more accurate than the physically based hydrological SWAT model for all input scenarios tested at the daily scale. Moreover, although the SWAT results improved considerably at the monthly scale, the statistics obtained were generally inferior compared to the ML models. For the best combination of inputs, M5P demonstrated a superior ability to estimate SSL (R2 = 0.73, MAE = 135.04, RSR = 0.54, NSE = 0.71 and PBIAS = 5.19), compared to RF (R2 = 0.72, MAE = 143.39, RSR = 0.57, NSE = 0.67 and PBIAS = 11.60) and SWAT (R2 = 0.57, MAE = 181.24, RSR = 0.65, NSE = 0.57 and PBIAS = -1.27). The average sediment loads in winter, the season with the highest sediment generation in the Oskotz basin, were 2,094.04, 1,831.08 and 2,242.67 tonnes for M5P, RF and SWAT, respectively, compared to an observed SSL of 1,878.16 tonnes. These results indicate that M5P and RF are suitable models for simulating fluvial sediment production since they improved the results of the SWAT model, which also requires more time and data to set up and calibrate. However, since SWAT does not require observed streamflow as an input, it remains a useful model, achieving acceptable results in basins with limited streamflow data.

1. Introduction

In river and dam engineering studies, the volume of sediment transported by a river is of particular interest due to its effect on hydraulic structures and water resource management (Kisi, 2004). River sediments are of additional interest in environmental engineering, especially if the sediments transport pollutants. Suspended sediment (SS) inputs are among the main factors contributing to water quality degradation (Zeiger and Hubbart, 2016). The transport of soil particles through surface waters has negative consequences for the fauna and flora of rivers since such sediments can contain organic pollutants and/or absorbed heavy metals, as well as increasing water turbidity (Sirabahenda et al., 2020).

To assess the extent of the soil erosion problem and identify problem areas within a basin, the amount of sediment transported by streams or rivers must first be reliably quantified. However, finding accurate tools to estimate suspended sediment load (SSL) is challenging since fluvial sediment transport is a complex, non-linear process influenced by hydrographic, hydraulic, climatic, and anthropogenic factors in the river basin (Zounemat-Kermani et al., 2020). To simulate sediment transport processes in river basins and water bodies, several empirical, physically based, and conceptual methods are available (Borrelli et al., 2021;

* Corresponding author.
E-mail addresses: pjimeno@ucam.edu (P. Jimeno-Sáez), raquel.m.e@um.es (R. Martínez-España), jcs@unavarra.es (J. Casalí), julio.sanchez@ulpgc.es (J. Pérez-Sánchez), jsenent@ucam.edu (J. Senent-Aparicio).

https://doi.org/10.1016/j.catena.2021.105953
Received 28 July 2021; Received in revised form 24 November 2021; Accepted 13 December 2021
Available online 28 December 2021

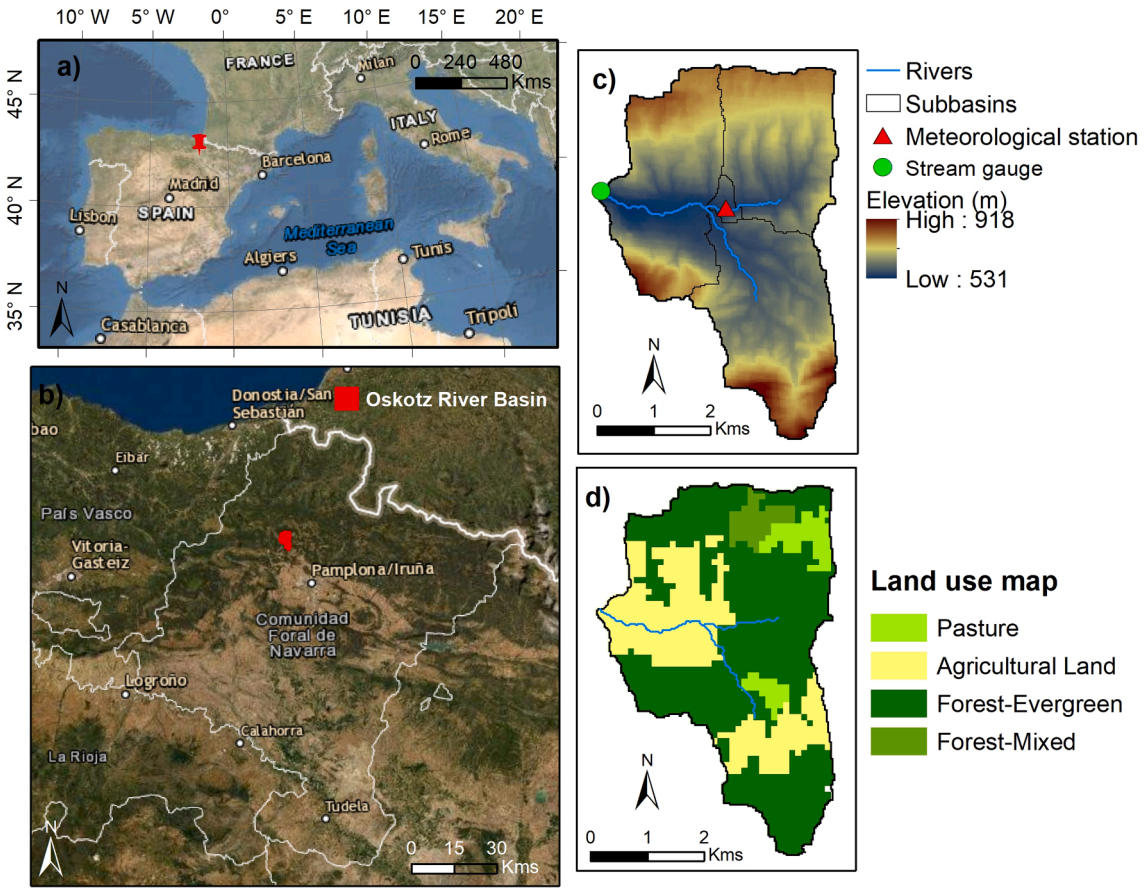

Fig. 1. (a) Location of the Oskotz river basin in Spain and (b) in the province of Navarra, (c) digital elevation model (DEM), and (d) land use map of the Oskotz river basin.

Merritt et al., 2003). Such models are useful for estimating sediment concentrations and/or loads generated under different climatic conditions, land use or management strategies, or in ungauged river basins (Fu et al., 2019).

The Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998) is an efficient tool for many types of water resource and land management applications (Gassman et al., 2014) and the most popular physically based model applied at the river-basin scale (Fu et al., 2019). Numerous studies conducted worldwide have concluded that the SWAT model provides a satisfactory estimate of sediment load (Duru et al., 2018; Dutta and Sen, 2018; López-Ballesteros et al., 2019; Pulighe et al., 2019). Physically based models simulate sediment produced by rainfall events in a river basin based on the laws of conservation of mass and energy. They provide an understanding of river basin processes, which is beneficial for assessing the impacts of soil and water conservation measures (Singh et al., 2014). However, a disadvantage of these models is that they require extensive information for development, calibration, and validation. In addition, calibrating the numerous parameters in such models is highly complex, due to the non-linearity of sediment transport processes, and requires expert knowledge and high computational time compared to data-driven models (Hamaamin et al., 2016; Khosravi et al., 2020). These drawbacks demand the exploration of more efficient methods of sediment computation.

Due to their capacity to model complex non-linear systems, data-driven methods such as machine learning (ML) have emerged as a powerful alternative to physically based models. In water resource management, such models can be used to estimate variables such as streamflow (Jimeno-Sáez et al., 2018; Minns and Hall, 1996; Srivastava et al., 2006) and sediment transport (Chen and Chau, 2016; Gupta et al., 2021; Kumar et al., 2016; Olyaie et al., 2015; Zounemat-Kermani et al., 2020). In contrast to physical models, ML models uses mathematical functions to connect inputs to outputs, ignoring the physical, logical relationship between variables (Ji et al., 2021). Kisi (2005) used a neural network approach and a neuro-fuzzy technique to estimate current suspended sediment values based on previous streamflow and sediment data, finding that the best results were obtained using neuro-fuzzy techniques. Al-Mukhtar (2019) predicted suspended sediments in the Tigris-Baghdad river using random forest (RF), support vector machine, and neural network techniques. The results showed that RF performed best. Ghasempour et al. (2021) used ML techniques to predict sediment, presenting a kernel-based approach based on the Gaussian process and the extreme learning kernel. The former was used for linear processes, while the latter was applied to non-linear processes. Sihag et al. (2021) developed a study to evaluate the best model using M5P and RF regression techniques to estimate sediment. The M5P-based model performed best in the study.

To date, several sediment transport models using SWAT and ML algorithms have been studied individually. However, few studies have compared both approaches. Singh et al. (2014, 2012) compared SWAT with a multilayer perceptron artificial neural network model and a radial basis neural network by simulating monthly sediment yields and obtained better results using the artificial neural network. Kim et al. (2012) observed that using artificial neural networks to predict total suspended solids was a useful alternative to SWAT. Sirabahenda et al. (2020) evaluated the SWAT model and the adapted neuro-fuzzy inference system (ANFIS) for SS prediction concluding that ANFIS obtained higher accuracy than SWAT.

In Navarra, Spain, soil erosion is a major problem on agricultural land (Casalí et al., 1999; De Santisteban et al., 2006). For this reason, the Government of Navarra created a network of experimental basins to

obtain data on water quality and soil erosion and evaluate the impact of agriculture in different areas of the region. In terms of morphology, soils, climate, land use, and management, these experimental river basins are representative of large areas of Navarra and Spain (Casalí et al., 2008). Oskotz is one of the four pilot basins in this network. This study uses data recorded at the Oskotz river basin to study SSL.

Modeling the specific processes that occur in a basin is highly complex, and no model works perfectly in all basins. Identifying models that accurately simulate the complexity of basin processes using available data is a challenge for decision-makers in basin management (Nguyen et al., 2019). This study individually tests two ML methods, M5P and RF, as alternatives to the SWAT model to estimate SSL in the Oskotz river basin. To our knowledge, no previous studies have compared SWAT with these models for SSL estimation. The application of ML versus SWAT in the study basin is also novel. The main objective of this research is to determine efficient models for SSL estimation by comparing the results obtained using ML models with those obtained using the physically based SWAT model. This study follows four steps: (i) SWAT model calibration for streamflow and sediment, (ii) ML model training, (iii) SSL estimation using the calibrated SWAT model and trained ML models, and (iv) comparison of the results and model performance analysis.

2. Materials and methods

2.1. Study area and data source

This research was conducted in the Oskotz river basin, located in Navarra (northern Spain) between 1°47′–1°44′ W longitude and 42°55′–42°58′ N latitude (Fig. 1). It is a small experimental basin covering 16.74 km^2. The climate in the basin is sub-Atlantic with an average annual rainfall of 1200 mm and an average annual temperature of 12 °C (Casalí et al., 2010). The Oskotz river basin has a wet season in autumn and winter and a drier season in summer. Nevertheless, summer rainfall is still important, accounting for approximately 11% of the annual precipitation. Most of the basin is covered by forest (66%) while the remaining area consists of pasture and crops. Pastures are used for grazing animals and are both natural and cultivated. The dominant soil class depends on the landscape type. According to Casalí et al. (2010), accumulation hillslopes are dominated by Typic Ustochrepts, eroded hillslopes by Lythic and Typic Ustochrepts and the valley plain by Fluventic Ustochrepts. Soils are fine with a thickness of more than 1 m, except on the eroded hillslopes where soils are shallow, between 0.5 and 1 m. The elevation in the basin is between 531 and 918 m a.s.l. and most of the slopes are in the range of 8–30%.

The basin is equipped with an automatic meteorological station and a hydrological station that measures streamflow and water quality parameters and takes water samples. The data from these stations are available in the experimental basins portal from the local Government of Navarra (http://cuencasagrarias.navarra.es/) and include precipitation (mm), maximum and minimum temperatures (°C), and observed streamflow and SSL data. Hydrometeorological data are available from 2002 to the present while sediment data are available from 2004. These data, obtained from observation stations in the basin, were used to construct the ML methods and the hydrological model. The SWAT hydrological model also required spatial inputs such as digital elevation model (DEM) data (Fig. 1), land use, and soil data. The 25 × 25 m DEM was obtained from the National Geographic Institute (IGN) in Spain. Land-use map was extracted from Corine Land Cover (2012) with a scale of 1:100,000, and a soil map with a resolution of 1 km was implemented from the Harmonized World Soil Database (HWSD) (Nachtergaele et al., 2010).

2.2. Hydrological model

2.2.1. SWAT model description

SWAT is a physically based, hydrological model developed to simulate water, sediment, and agricultural chemical production in a river basin in a semi-distributed form. The SWAT model divides the basin into sub-basins, each of which is further divided into several Hydrological Response Units (HRUs), areas of land that are homogeneous and have similar responses to meteorological inputs. Each HRU is a combination of a specific soil type, land use, and slope. The hydrological part of the model simulates a catchment's hydrological cycle, based on the water balance equation, and calculates the runoff from each HRU. The curve number and Muskingum methods are employed for runoff computation and channel routine respectively. SWAT calculates the sediment yield for each HRU using the Modified Universal Soil Loss Equation (MUSLE) (Williams, 1975), which predicts erosion as a function of a runoff factor representing the energy used in the detachment and transport of sediment (Neitsch et al., 2009). SWAT computes water and sediment yield for each HRU individually and aggregates them at the sub-basin level.

2.2.2. SWAT model set up, calibration, and validation

The Oskotz river basin was subdivided into four sub-basins and homogenous sections, resulting in 53 HRUs. The Hargreaves-Samani approach was chosen to estimate potential evapotranspiration. The SUFI-2 algorithm was used to automatically calibrate the SWAT model parameters in SWAT-CUP (Abbaspour et al., 2007). Since sediment transport is dependent on runoff, a sequential calibration approach was applied (recommended by Arnold et al., (2015)), in which streamflow generation parameters were calibrated first, followed by sediment parameters. One thousand simulations were performed twice, and the parameters were readjusted after the second iteration. The selected parameters were calibrated using daily time steps and adjusted using the Nash-Sutcliffe efficiency (NSE) as an objective function to ensure that the simulation results were as close as possible to the streamflow and SSL observations. The periods 2002–2012 (11 years) and 2013–2020 (8 years) were used to calibrate and validate the streamflow, respectively. In the case of SSL, 2004–2012 (9 years) was the calibration period and 2013–2020 (8 years) was the validation period. In both calibrations, three years were used as a warm-up period.

2.3. Machine learning algorithms

2.3.1. M5P

The M5P technique (Wang and Witten, 1997) is a remodeling of Quinlan's M5 (Quinlan, 1992) for induction trees in regression models. This technique combines a traditional decision tree with the possibility of performing linear regression functions at the nodes. First, an induction decision tree is constructed, applying the splitting criterion, which minimizes the variance of a subset of class values at each branch, at each node. This process is stopped if the values of each branch vary slightly or there are a minimum number of instances at the node. Secondly, a pruning process is performed, in which a regression function converts the internal nodes into a leaf node. Finally, to avoid discontinuities, a smoothing process is applied that combines the leaf model prediction with each node encountered on the way to the root node.

2.3.2. Random forest (RF)

RF (Breiman, 2001) is defined as an ensemble based on decision trees. Decision trees have the advantage of good interpretability in both the constructed model and inference. However, they have the disadvantage of bias and variance problems. These complications are resolved using the ensemble to merge and combine information from the decision trees. On the one hand, the data variability and amount of stored information are increased. On the other hand, the interpretability of the constructed model is maintained, although in a more complex manner. In general, the RF ensemble has the following characteristics:

- Given a dataset of |N| samples to construct each tree a, |N| cases are randomly selected using replacement as a training dataset. The

Table 1
The correlation coefficients (CC) between the input variables and SSL_t.

Variable	Q_t	Q_{t-1}	Q_{t-2}	Q_{t-3}	Q_{t-4}	Q_{t-5}	P_t	P_{t-1}	P_{t-2}	P_{t-3}	P_{t-4}	P_{t-5}
CC	0.64	0.34	0.18	0.16	0.15	0.11	0.40	0.36	0.17	0.12	0.12	0.11

Table 2
Input scenarios for the ML models.

Scenario	Model inputs
I	Q_t
II	Q_t, P_t
III	Q_t, P_t, P_{t-1}
IV	Q_t, Q_{t-1}
V	Q_t, P_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}, P_{t-1}, P_{t-2}
VI	Q_t, P_t, Q_{t-1}
VII	Q_t, P_t, Q_{t-1}, Q_{t-2}
VIII	Q_t, P_t, Q_{t-1}, Q_{t-2}, P_{t-1}, P_{t-2}
IX	Q_t, P_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, P_{t-1}, P_{t-2}, P_{t-3}

process of sampling with replacement is called bootstrapping. One third of the data is excluded from training and used for testing (Schonlau and Zou, 2020). These data are known as out-of-bag (OOB) samples. Each tree has an OOBa set with which it is tested. The testing result provides a weighting for each tree used in the combination of the information.

- To choose the splitting decision for a node, the M attributes to be studied are chosen at random from the M input variables at each tree node.
- Each tree grows to its maximum possible extent and can be configured so that no pruning is required. New instances are predicted by aggregating the predictions of the A trees (i.e., a majority vote for classification, an average for regression).

2.3.3. ML model inputs

Concerning ML model inputs, streamflow and precipitation were selected to estimate SSL according to previous studies (Cobaner et al., 2009; Kumar et al., 2016; Singh et al., 2014). The precipitation and streamflow variables are correlated but complementary since precipitation transports sediment in the drainage basin and streamflow regulates concentrations and downstream transport (Sirabahenda et al., 2020). Correlation coefficients (CCs) were calculated to analyze the dependence between the observed sediment data and the flow and precipitation data on the same and previous days (Table 1).

The SSL for day t (SSL_t) was strongly correlated with the streamflow on that day (Q_t), with a CC of 0.64. The CC for SSL_t and the precipitation for day t (P_t) had a value of 0.40. From time t-1, the correlations between the Q and P variables and SSL (t) decreased considerably, with values below 0.2. Therefore, to estimate daily SSL, several input combinations were constructed and tested (Table 2), including the daily streamflow and precipitation of the current day t and previous days.

Table 3
Model performance metrics.

Measure	Equation	Range	Optimal value
R^2	$\left[\frac{\sum_{t=1}^{n}(O_t-\overline{O})\cdot(S_t-\overline{S})}{\sqrt{\sum_{t=1}^{n}(O_t-\overline{O})^2}\cdot\sqrt{\sum_{t=1}^{n}(S_t-\overline{S})^2}}\right]^2$	[0, 1]	1
MAE	$\frac{\sum_{t=1}^{n}\lvert O_t-S_t\rvert}{n}$	[0, ∞]	0
RSR	$\sqrt{\frac{\sum_{t=1}^{n}(O_t-S_t)^2}{\sum_{t=1}^{n}(O_t-\overline{S})^2}}$	[0, ∞]	0
NSE	$1-\frac{\sum_{t=1}^{n}(O_t-S_t)^2}{\sum_{t=1}^{n}(O_t-\overline{O})^2}$	[−∞, 1]	1
PBIAS	$\frac{\sum_{t=1}^{n}(O_t-S_t)\cdot 100}{\sum_{t=1}^{n}(O_t)}$	[−∞, ∞]	0

The first scenario considered the streamflow for day t (i.e. the variable most strongly correlated with SSL_t) as the only predictor variable. Scenario II considered the two variables most correlated with sediment (Q_t and P_t). Scenario III was similar to scenario two but with the addition of the third most correlated variable (P_{t-1}). Scenario IV considered only the hydrological variables on the same day and the previous day (Q_t and Q_{t-1}). In addition, scenario V included all variables with a CC greater than 0.15. The remaining scenarios were other combinations of these variables. The advantage of employing only streamflow and precipitation as inputs is that the collection and availability of these data are often easier than other data in many basins. Other studies (Kumar et al., 2016; Sihag et al., 2021; Zounemat-Kermani et al., 2020) have also used sediment from previous days as inputs with the disadvantage that this type of data is measured with less frequency and is, therefore, more

Table 4
The SWAT model parameters used to calibrate streamflow and sediment.

Parameter	Description	Initial range used in calibration	Calibrated value
Parameters used to calibrate streamflow			
r_CN2.mgt	SCS runoff curve number	–0.2 to 0.2	−0.16
v_ESCO.bsn	Soil evaporation compensation factor	0.1 to 1	0.89
v_EPCO.bsn	Plant uptake compensation factor	0.1 to 1	0.56
v_SURLAG.bsn	Surface runoff lag time (days)	0.05 to 24	15.40
v_LAT_TTIME.hru	Lateral flow travel time (days)	0 to 30	2.20
v_ALPHA_BF.gw	Baseflow alpha factor (days-1)	0 to 1	0.53
v_GW_DELAY.gw	Groundwater delay (days)	0 to 100	13
v_GW_REVAP.gw	Groundwater revap coefficient	0.02 to 0.20	0.09
v_GWQMN.gw	Threshold depth of water in the shallow aquifer for return flow to occur (mm)	0 to 5000	1863.12
v_REVAPMN.gw	Threshold depth of water in the shallow aquifer for revap to occur (m)	0 to 1000	252.13
v_RCHRG_DP.gw	Deep aquifer percolation fraction	0 to 1	0.07
r_SOL_AWC.sol	Available water capacity of the soil layer (mm H_2O/mm soil)	–0.2 to 0.2	0.12
r_SOL_BD.sol	Moist bulk density	–0.2 to 0.2	0.11
r_SOL_K.sol	Saturated hydraulic conductivity (mm h-1)	–0.2 to 0.2	0.12
Parameters used to calibrate SSL			
v_CH_COV1.rte	Channel erodibility factor	0 to 1	0.35
v_CH_COV2.rte	Channel cover factor	0 to 1	0.42
v_SPCON.bsn	Linear parameter for calculating the maximum amount of sediment that can be re-entrained during channel sediment routing	0.0001 to 0.01	0.0001
v_SPEXP.bsn	Exponent parameter for calculating sediment re-entrained in channel sediment routing	1 to 1.5	1.45
v_USLE_P.mgt	USLE equation support practice factor	0 to 1	0.85
r_USLE_K.sol	USLE equation soil erodibility (K) factor	−0.2 to 0.2	0.04

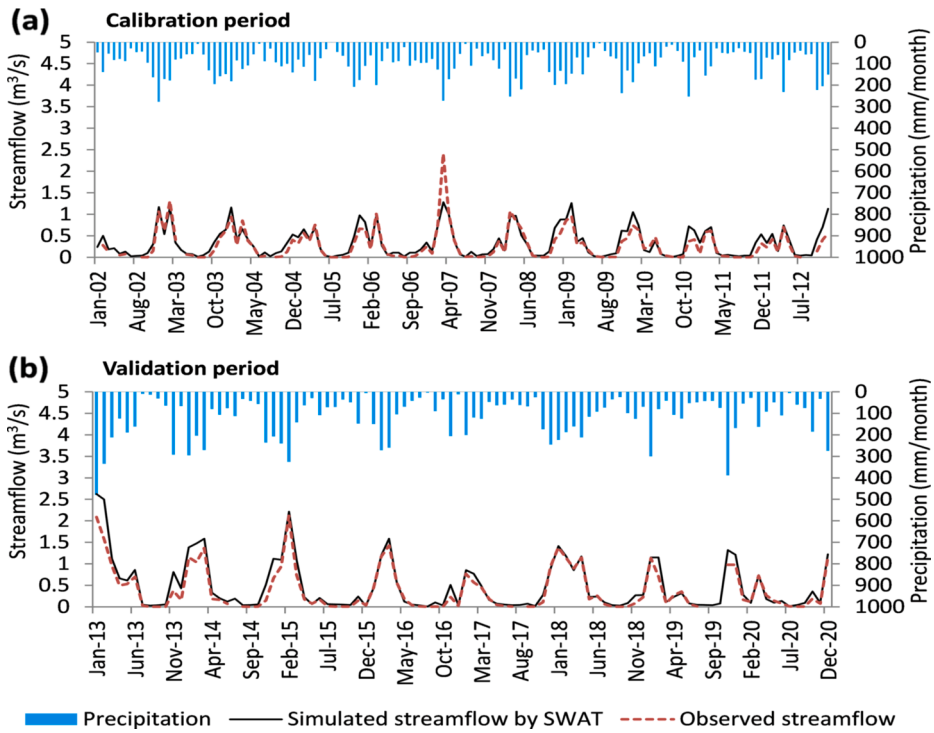

Fig. 2. Monthly precipitation, observed streamflow, and streamflow simulated by SWAT at a hydrological station for (a) the calibration period and (b) the validation period.

difficult to collect. Therefore, nine scenarios were tested for each ML model. The training period was from September 2004 to December 2012, and the test period was from January 2013 to December 2020.

2.4. Model performance metrics and evaluation criteria

Five statistical criteria were used to evaluate the performance of the models on two different time scales (daily and monthly). The coefficient of determination (R^2), the mean absolute error (MAE), the root mean square error observations standard deviation ratio (RSR), the NSE, and the percent bias (PBIAS) were calculated using the equations listed in Table 3, where O_t is the observed data at time t, $\overline{O}$ is the mean of the observed data, S_t is the simulated data at time t, $\overline{S}$ is the mean of the simulated data, and n is the total number of observed data. The R^2 denotes the degree of collinearity between the simulated and observed data, as well as the fraction of variance in the observed data explained by the model. The MAE measures the overall deviation of the models. The RSR uses the standard deviation of the observations to standardize the mean square error. The variance between the observed and simulated data is quantified using NSE. The PBIAS indicates the average tendency of the simulated data to be higher (negative values) or lower (positive values) than the observed data.

3. Results and discussion

3.1. SWAT calibration

Fourteen commonly used streamflow calibration parameters and their ranges were selected based on available studies close to the study area (Epelde et al., 2015; Meaurio et al., 2015) and our previous experience. Following hydrological calibration, sediment calibration was performed using six specific parameters. The calibration process substantially reduced the disparity between the observed and simulated streamflow and SSL. The specific parameters used to calibrate streamflow and SSL are listed in Table 4, which describes each parameter, its range, and its final calibrated value.

The calibrated parameter analysis presented in Table 4 shows that CN2 decreased by 16% compared to the default value, thus increasing infiltration and decreasing runoff. The adjusted ESCO value was 0.89, a high value typical of climates in which evapotranspiration is not highly relevant (Jimeno-Sáez et al., 2018). The LAT_TTIME value was low, indicating that water pathways through the soil profile are short. The RCHRG_DP value was very low at 0.07 and characteristic of land with no relevant groundwater storage (Senent-Aparicio et al., 2019). The fitted values for ESCO, LAT_TTIME, RCHRG_DP, SPCON, and SPEXP in the Oskotz river basin were similar to those used in nearby basins (Epelde et al., 2015; Meaurio et al., 2015).

3.2. ML model parameters and computational time

The following M5P parameters were used in the experiments: a minimum number of four examples for the leaf nodes and no pruning for the trees. Regarding the parameters used in the RF experiments, there were 100 trees (A), each having a maximum depth expansion, and for each node, the number of attributes selected to divide de node was $\log_2(M) + 1$, where M was the maximum number of input attributes. The computational times were very low for both the M5P and RF techniques. In the scenario with the most variables, the model building times for M5P and RF were 0.33 and 0.58 s respectively. The inference was 0.08 and 0.03.

3.3. SWAT streamflow estimation

The calibrated SWAT model simulated daily streamflow using R^2 values of 0.71 and 0.79 for the calibration and validation periods, respectively. The RSR, NSE, and PBIAS were respectively 0.54, 0.70, and −20.95% for the calibration period and 0.48, 0.76, and −24.38% for the validation period. At the monthly scale, the performance statistics were more satisfactory, displaying R^2 values of 0.84 for calibration and 0.94 for validation. The monthly RSR, NSE, and PBIAS values were 0.46, 0.81, and −20.55% for the calibration period and 0.36, 0.86, and −24.85% for the validation period. Therefore, the model was evaluated

Table 5
The daily performance of the models in SSL estimation.

Input scenario	Model	Performance metrics Calibration (Validation)				
		R^2	MAE (ton/day)	RSR	NSE	PBIAS (%)
–	SWAT	0.27 (0.28)	5.64 (11.17)	0.86 (0.86)	0.28 (0.26)	−7.60 (−11.98)
I	M5P	0.41 (0.49)	4.63 (8.83)	0.77 (0.75)	0.41 (0.44)	−5.25 (−5.79)
	RF	0.88 (0.31)	2.15 (10.14)	0.37 (0.83)	0.86 (0.32)	1.58 (−7.55)
II	M5P	0.41 (0.48)	4.62 (8.86)	0.77 (0.75)	0.41 (0.44)	−6.12 (−6.84)
	RF	0.90 (0.41)	1.97 (8.95)	0.36 (0.78)	0.87 (0.39)	0.12 (2.11)
III	M5P	0.50 (0.51)	4.12 (8.02)	0.71 (0.71)	0.50 (0.50)	−0.60 (0.50)
	RF	0.91 (0.45)	1.94 (8.58)	0.35 (0.77)	0.88 (0.41)	−0.47 (4.79)
IV	M5P	0.42 (0.50)	4.55 (8.73)	0.76 (0.75)	0.42 (0.44)	−3.11 (−1.23)
	RF	0.89 (0.37)	1.99 (9.52)	0.37 (0.80)	0.86 (0.35)	0.93 (−2.78)
V	M5P	0.54 (0.44)	4.03 (8.93)	0.68 (0.75)	0.54 (0.44)	−0.21 (−17.87)
	RF	0.92 (0.41)	1.91 (8.94)	0.34 (0.79)	0.88 (0.38)	−0.73 (0.40)
VI	M5P	0.42 (0.50)	4.55 (8.73)	0.76 (0.75)	0.42 (0.44)	−3.11 (−1.23)
	RF	0.92 (0.34)	1.92 (9.23)	0.35 (0.82)	0.88 (0.33)	1.2 (0.99)
VII	M5P	0.42 (0.50)	4.55 (8.74)	0.76 (0.75)	0.42 (0.44)	−3.06 (−1.25)
	RF	0.91 (0.41)	1.91 (8.85)	0.35 (0.79)	0.88 (0.38)	−0.4 (2.06)
VIII	M5P	0.43 (0.49)	4.70 (8.87)	0.75 (0.75)	0.43 (0.44)	−7.13 (−3.71)
	RF	0.90 (0.43)	1.95 (8.80)	0.34 (0.78)	0.88 (0.40)	−1.91 (1.60)
IX	M5P	0.68 (0.51)	3.55 (8.10)	0.58 (0.71)	0.66 (0.50)	0.68 (−6.80)
	RF	0.92 (0.43)	1.96 (8.77)	0.34 (0.78)	0.88 (0.39)	−1.4 (−2.52)

Table 6
The monthly performance of the models in SSL estimation.

Input scenario	Model	Performance metrics Calibration (Validation)				
		R^2	MAE (ton/day)	RSR	NSE	PBIAS (%)
–	SWAT	0.75 (0.57)	116.20 (181.24)	0.52 (0.65)	0.72 (0.57)	1.25 (−1.27)
I	M5P	0.62 (0.62)	118.25 (153.12)	0.62 (0.62)	0.62 (0.62)	8.55 (4.27)
	RF	0.90 (0.58)	59.97 (161.97)	0.40 (0.60)	0.84 (0.58)	12.13 (1.08)
II	M5P	0.64 (0.64)	116.72 (153.09)	0.60 (0.61)	0.64 (0.63)	7.75 (3.30)
	RF	0.92 (0.59)	59.81 (162.63)	0.37 (0.65)	0.86 (0.57)	9.79 (12.25)
III	M5P	0.75 (0.72)	101.08 (136.08)	0.50 (0.55)	0.74 (0.70)	8.83 (6.04)
	RF	0.94 (0.65)	58.33 (157.04)	0.35 (0.63)	0.88 (0.60)	8.62 (14.24)
IV	M5P	0.62 (0.62)	118.70 (157.55)	0.62 (0.63)	0.61 (0.60)	11.21 (7.76)
	RF	0.92 (0.61)	58.17 (158.25)	0.38 (0.63)	0.86 (0.60)	11.19 (4.39)
V	M5P	0.78 (0.69)	92.90 (142.02)	0.46 (0.56)	0.78 (0.69)	4.86 (−4.48)
	RF	0.93 (0.69)	60.56 (146.24)	0.36 (0.58)	0.87 (0.66)	9.36 (8.99)
VI	M5P	0.62 (0.62)	118.70 (157.55)	0.62 (0.63)	0.61 (0.60)	11.21 (7.76)
	RF	0.93 (0.64)	59.58 (157.31)	0.39 (0.61)	0.85 (0.62)	11.66 (10.41)
VII	M5P	0.62 (0.62)	118.68 (157.60)	0.62 (0.63)	0.61 (0.60)	11.25 (7.74)
	RF	0.92 (0.67)	59.35 (151.35)	0.38 (0.56)	0.85 (0.63)	10.31 (10.63)
VIII	M5P	0.66 (0.65)	119.16 (159.31)	0.60 (0.57)	0.64 (0.62)	8.14 (5.06)
	RF	0.92 (0.72)	60.45 (145.51)	0.38 (0.53)	0.86 (0.67)	8.83 (10.10)
IX	M5P	0.83 (0.73)	89.13 (135.04)	0.43 (0.54)	0.81 (0.71)	8.90 (5.19)
	RF	0.94 (0.72)	57.95 (143.39)	0.35 (0.57)	0.88 (0.67)	8.56 (11.60)

as very good regarding monthly streamflow simulation during both calibration and validation, according to the RSR and NSE criteria described by Moriasi et al. (2007). The good agreement between simulated and observed monthly streamflow is presented graphically in Fig. 2, which displays both the calibration and validation periods. However, the model was classified as satisfactory according to PBIAS (Moriasi et al., 2007) since calibration and validation both obtained negative values, indicating that the model overestimated the streamflow (see Fig. 2). However, the simulated streamflow fitted the observed flows very satisfactorily, matching the low flows and most of the peaks well.

3.4. SWAT and ML sediment estimation

Nine combinations of inputs were used in the ML models. The daily calibration and validation results for each model are presented in Table 5.

In all scenarios, the M5P and RF models obtained better daily performance metrics than the SWAT model, during both calibration and validation. In the training (calibration) phase, RF performed particularly well compared to the other models. However, RF was less reliable in the test phase (validation) due to model overfitting during training. This problem can occur if the model is not tested using cross-validation (Cai et al., 2020). The M5P models were more stable, demonstrating similar results during both phases. Compared to the RF models, the M5P models improved the test-phase statistics in many scenarios, indicating that the model did not overfit during the training phase. Regarding the different scenarios, at the daily level, scenario III, which included the three variables most correlated with sediment, presented the best results during M5P and RF model validation. Scenario V obtained one of the best results for both models, albeit during calibration only. Scenario IX was the most successful of the nine scenarios for both ML models in both phases (training and validation). The SWAT model did not reach the NSE value of 0.3. However, the ML models obtained an NSE value greater than 0.3 in all cases, reaching the satisfactory classification according to the daily criteria established by Kalin et al. (2010). The two ML models and SWAT achieved PBIAS values of less than 25%, the criterion for very good daily models set by Kalin et al. (2010).

The statistics in Table 6 indicate that the SWAT and ML approaches performed well in SSL estimation at the monthly scale. According to criteria set by Moriasi et al. (2007), SWAT simulated monthly SSL well during calibration and satisfactorily during validation. Conversely, PBIAS had very good values in both cases. The monthly SWAT results were significantly better than the daily results. Similar findings have been observed in other studies (Choukri et al., 2020; Nunes et al., 2018), in which problems in SWAT's sediment transport module caused inaccuracies in daily sediment estimates, which were averaged out and smoothed on a monthly scale. The M5P model was superior to SWAT during calibration in scenarios III, V, and IX but better in all scenarios during validation. In all cases, RF outperformed SWAT. As expected, the best results for both models were obtained in scenarios III, V, and IX, which is similar to the daily scale results. Again, both ML models performed best in scenario IX, with performance statistics classified as very good during calibration and good during validation, according to the

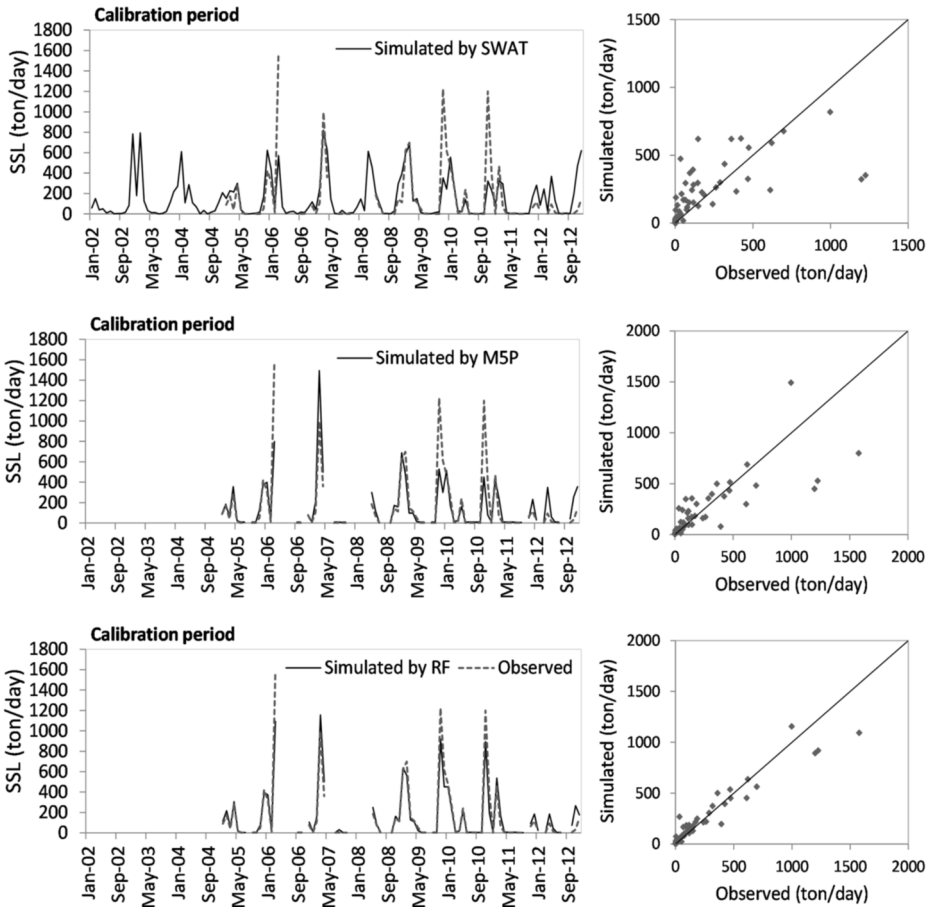

Fig. 3. A comparison of observed SSL and monthly SSL simulated using SWAT, M5P, and RF during the calibration period.

monthly criteria established by Moriasi et al. (2007).

The monthly results of the SWAT and ML models for the best-case scenario (scenario IX) are presented graphically in Fig. 3 and Fig. 4. The temporal variations in monthly SSL at the Oskotz river basin are given for the calibration period (Fig. 3) and the validation period (Fig. 4).

In all cases, the PBIAS values were positive, indicating that the models simulated less sediment load than observed (shown in the figures on the right). The exception was the SWAT model validation (PBIAS = -1.27%). Significant R^2 relationships between the observed and simulated SSLs are seen on the right of the figures in the scatter plots. The satisfactory monthly NSE (NSE $>$ 0.5) values suggest that the monthly SSLs simulated by all the models were close to the observed quantities. Scatter plots compare the estimated and observed SSLs, providing another tool for evaluating the model. The more dispersed the data points, the worse the model performance. The higher the NSE value, the better the data fit the 1:1 line. Graphically, it is evident that the RF calibration provided the best fit. The scatter plot of the RF model calibration, with an NSE of 0.88, shows that the data points are very close to the diagonal line. During validation, however, M5P gave the best fit, although the graphical results are less evident. The scatterplots show that the ML models provided the best estimates of low SSL values. For high SSL events, both ML models predicted SSL more accurately than the SWAT model, although all the models underestimated most of the peaks, as found by Sirabahenda et al. (2020). Benaman and Shoemaker (2005) analyzed thirty-five high-flow events using SWAT. They found that SWAT generally underestimated high sediment load events, as is the case in this study. Conversely, a clear advantage of SWAT over the ML models is reflected graphically; SWAT can simulate the sediment load over the entire calibration period without the need for data for the present or previous days.

The effect of precipitation variability on SSL production can be observed in Fig. 5, which presents the mean monthly values for precipitation, streamflow, and SSL in the basin for the calibration and validation periods.

The sediment loads simulated by all the models followed a seasonal distribution similar to the observed data. In both the calibration (Fig. 5a) and validation (Fig. 5b) periods, all the models estimated the highest sediment loads during the months with the greatest rainfall and, therefore, the highest flows. In the Oskotz, sediment loads displayed significant inter-annual variability, which is logical since they are largely regulated by precipitation. The production of SSL is directly related to precipitation and streamflow. During months with higher precipitation (autumn and winter), more streamflow is generated and, therefore, more SSL. During the summer months, however, monthly flows are very low, even when there is precipitation, and SSL is therefore insignificant. Despite accounting for 11% of the annual rainfall and having the most erosive precipitation, discharge during the dry season barely represents 1% of the total annual flow, since seasonal vegetation cover provides canopy interception and evapotranspiration, resulting in less runoff (Casalí et al., 2010). Most of the Oskotz Basin is covered by forest (66%). Therefore, according to Gallart and Llorens (2003), the negligible streamflow generation is also explained by high infiltration rates caused by the macro-porosity of the forest soil and drier soil conditions before any rainfall event. In the Oskotz basin, most sediment is generated during the winter season, as observed by Casalí et al. (2010), since during winter, basin soils are more saturated, resulting in increased runoff. Furthermore, this study concludes that the sediment load in the Oskotz could be explained by vegetation and soil conditions rather than the erosivity of precipitation episodes. Owens et al. (1997)

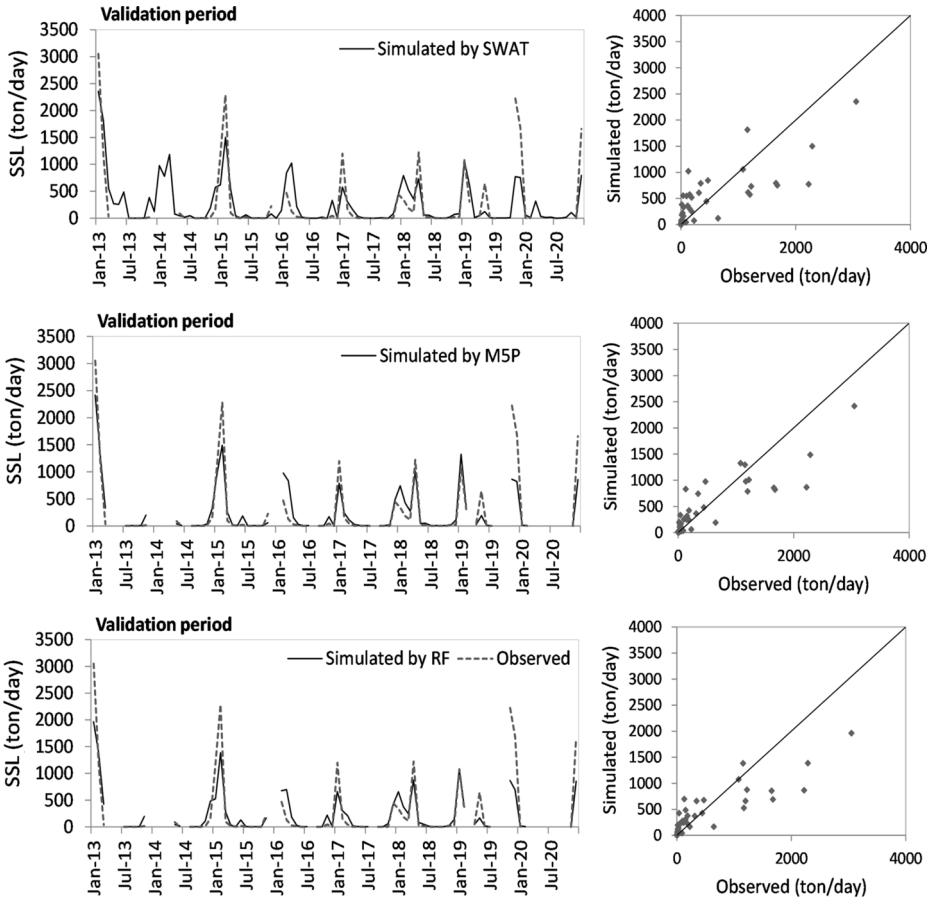

Fig. 4. A comparison of observed SSL and monthly SSL simulated using SWAT, M5P, and RF during the validation period.

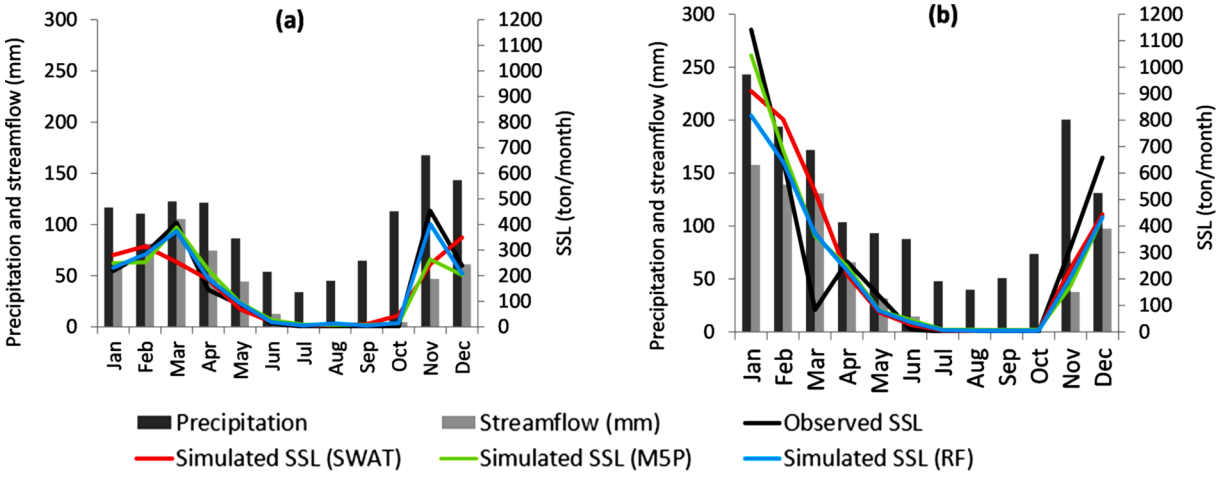

Fig. 5. The seasonal distribution of observed precipitation, streamflow, and SSL and SSL simulated using SWAT and ML models for (a) calibration and (b) validation.

demonstrated that winter grazing can produce a large increase in sediment in small basins with pasture because livestock damage the soil structure when the soil is wet, compacting it and reducing its infiltration capacity and vegetation cover. During calibration, the models estimated that between 48% and 51% of the sediment was produced in winter compared to 49% of the observed sediment. In validation, the values were higher, between 65% and 69%, compared to 57% of the observed sediment. Specifically, winter sediment loads estimated by SWAT, M5P, and RF were 849.92, 891.77, and 888.42 tons respectively during the calibration period (2005–2012), which was drier than the later years used for validation (2013–2020). During winter, the observed mean SSL was 911.35 tons. Therefore, the two ML models produced better estimates for this season. The same was true for the validation period, as the average winter sediment loads for SWAT, M5P, and RF were 2,242.67, 2,094.04, and 1,831.08 tons, compared to an observed SSL of 1,878.16 tons.

The successful performance of the ML models is supported in previous literature. In their study estimating bed-load transportation rates,

Khosravi et al. (2020) found that M5P models could replace process-based models since they effectively reproduce the highly stochastic behavior of sediment transport and have low build and running costs. Al-Mukhtar (2019) and Sihag et al. (2021) found that RF had a superior performance among the other algorithms used to model sediment load. Part of the reason why ML models improve on SWAT may be different input data. The SWAT model simulates streamflow data and then uses the values as input in the sediment model, whereas M5P and RF used observed streamflow data to achieve the same outcome.

Regarding the advantages and disadvantages of each model, the SWAT model's utility is limited compared to the other techniques because this model requires a large amount of data (Pandey et al., 2021) and considerable calibration time. However, SWAT has the advantage of simulating sediment on days where no streamflow data exist, which allows filling in historical data or generating future sediments just by using climate data from the future. Conversely, the ML models require fewer data less and computational time. In addition, the proposed ML models share the advantage of easy result interpretation. It should be noted, however, that RF, being an ensemble, is more difficult to interpret than M5P. Furthermore, although RF is fast to run, it takes several seconds longer to display the results compared to M5P. An advantage of M5P is the potential to display the model using a set of rules or a decision tree, providing flexibility in interpreting the results as well as the variables used in each scenario.

The main limitation to the development of sediment estimation models is the availability of climate and streamflow data, which can be very limited, incomplete and sometimes non-existent (Mapes and Pricope, 2020). ML models constructed using small data sets are fitted to the local conditions of a study area rather than universally applicable laws of physics. Therefore, such ML models are not transferable outside the training region (Shen et al., 2021). ML models are obviously more advanced and powerful than traditional forecasting models, but they are less interpretable (Sarkar and Pal, 2021). In addition to data availability, the SWAT model has several other limitations. Data accuracy is crucial for the proper functioning of SWAT, but so is understanding the parameters. The SWAT model requires the user to have a high level of hydrological knowledge. Moreover, SWAT simulates sediment based on the MUSLE equation, which tends to underestimate large sediment events and overestimate smaller ones (Ma et al., 2021). However, SWAT has the potential to be applied in different assessment scenarios, such as climate change impacts, land-use changes and land management practices (Li et al., 2017).

4. Conclusions

Accurately estimating SSL is important for understanding the hydrodynamics of rivers. Due to the non-linear behavior of sediment transport, ML algorithms demonstrate considerable potential for accurately estimating sediment loads. This study has explored the performance of two ML models (M5P and RF) and compared them to the SWAT hydrological model for SSL simulation, using data collected in the Oskotz river basin, an experimental basin in Navarra, northern Spain. This study presents the following conclusions:

1. The sediment loads estimated by the three models generally provide satisfactory approximations.
2. Regarding the ML models, nine input scenarios were tested exclusively using daily precipitation and streamflow as input data and considering scenarios with and without past data. The best scenario included streamflow and precipitation data from day t to day t-3 as inputs. Moreover, both ML models were equally valid with similar statistics.
3. At the daily scale, the ML models outperformed the SWAT hydrological model in all scenarios. At the monthly scale, both ML models achieved better validation results than SWAT.
4. The ML models significantly reduced the time and computational effort required compared to SWAT, making sediment estimation easier. Since ML models are data-driven, they can be trained to describe complex processes without spatial data. In contrast, SWAT requires multiple spatial data and a complete description of the physical processes that govern the hydrological behavior of a river basin.

Overall, the results indicate that ML techniques provide a more accurate prediction of SSL than the SWAT model. Therefore, for water basin managers and stakeholders, ML can be a useful method for simulating sediment production, analyzing soil degradation, and designing appropriate measures for soil and water conservation. Such models could be particularly useful in basins where limited spatial data are available or knowledge of the processes taking place in the basin is limited or unknown. The novelty of this study is that it is the first to compare the performance of SWAT with M5P and RF models to estimate SSL. Moreover, no previous studies using these models have been carried out in the study area. Furthermore, SSL estimation can be extended to other ML models or the estimation of other water quality variables. In addition, future research could improve the models by exploring the influence of other inputs related to uptake and sedimentation processes.

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.

Acknowledgements

The authors acknowledge the Department of Rural Development and Environment of the Government of Navarra for the collection and transfer of data from its network of experimental agricultural basins. Also, thanks to Scribbr editing services for proofreading the text.

Funding

This work was supported by the Spanish Ministry of Science and Innovation, under grants RTC-2017-6389-5, CGL2015-64284-C2-2-R, and the European Unionś Horizon 2020 research and innovation programme within the framework of the project SMARTLAGOON, grant agreement number 101017861.

References

Abbaspour, K.C., Vejdani, M., Haghighat, S., 2007. SWAT-CUP calibration and uncertainty programs for SWAT. In: Proceedings of the Modsim 2007: International Congress on Modelling and Simulation, Christchurch, New Zealand, 3–8 December 2007, pp. 1603–1609.

Al-Mukhtar, M., 2019. Random forest, support vector machine, and neural networks to modelling suspended sediment in Tigris River-Baghdad. Environ. Monit. Assess 191 (11). https://doi.org/10.1007/s10661-019-7821-5.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment part I: model development. J. Am. Water Resour. Assoc. 34 (1), 73–89. https://doi.org/10.1111/j.1752-1688.1998.tb05961.x.

Arnold, J.G., Youssef, M.A., Yen, H., White, M.J., Sheshukov, A.Y., Sadeghi, A.M., Moriasi, D.N., Steiner, J.L., Amatya, D.M., Skaggs, R.W., Haney, E.B., Jeong, J., Arabi, M., Gowda, P.H., 2015. Hydrological Processes and Model Representation: Impact of Soft Data on Calibration. Trans. ASABE 58, 1637–1660. https://doi.org/10.13031/trans.58.10726.

Benaman, J., Shoemaker, C.A., 2005. An analysis of high-flow sediment event data for evaluating model performance. Hydrol. Process. 19 (3), 605–620. https://doi.org/10.1002/hyp.5608.

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. https://doi.org/10.1016/j.scitotenv.2021.146494.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Cai, J., Xu, K., Zhu, Y., Hu, F., Li, L., 2020. Prediction and analysis of net ecosystem carbon exchange based on gradient boosting regression and random forest. Appl. Energy 262, 114566. https://doi.org/10.1016/j.apenergy.2020.114566.

Casalí, J., Gastesi, R., Álvarez-Mozos, J., De Santisteban, L.M., Lersundi, J.D.V.d., Giménez, R., Larrañaga, A., Goñi, M., Agirre, U., Campo, M.A., López, J.J., Donézar, M., 2008. Runoff, erosion, and water quality of agricultural watersheds in central Navarre (Spain). Agric. Water Manage. 95 (10), 1111–1128. https://doi.org/10.1016/j.agwat.2008.06.013.

Casalí, J., Giménez, R., Díez, J., Álvarez-Mozos, J., Valle, D., de Lersundi, J., Goñi, M., Campo, M.A., Chahor, Y., Gastesi, R., López, J., 2010. Sediment production and water quality of watersheds with contrasting land use in Navarre (Spain). Agric. Water Manage. 97, 1683–1694. https://doi.org/10.1016/j.agwat.2010.05.024.

Casalí, J., López, J.J., Giráldez, J.V., 1999. Ephemeral gully erosion in southern Navarra (Spain). CATENA 36 (1-2), 65–84. https://doi.org/10.1016/S0341-8162(99)00013-2.

Chen, X.Y., Chau, K.W., 2016. A Hybrid Double Feedforward Neural Network for Suspended Sediment Load Estimation. Water Resour. Manag. 30 (7), 2179–2194. https://doi.org/10.1007/s11269-016-1281-2.

Choukri, F., Raclot, D., Naimi, M., Chikhaoui, M., Nunes, J.P., Huard, F., Hérivaux, C., Sabir, M., Pépin, Y., 2020. Distinct and combined impacts of climate and land use scenarios on water availability and sediment loads for a water supply reservoir in northern Morocco. Int. Soil Water Conserv. Res. 8 (2), 141–153. https://doi.org/10.1016/j.iswcr.2020.03.003.

Cobaner, M., Unal, B., Kisi, O., 2009. Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydro-meteorological data. J. Hydrol. 367 (1-2), 52–61. https://doi.org/10.1016/j.jhydrol.2008.12.024.

De Santisteban, L.M., Casalí, J., López, J.J., 2006. Assessing soil erosion rates in cultivated areas of Navarre (Spain). Earth Surf. Process. Landforms 31 (4), 487–506. https://doi.org/10.1002/esp.1281.

Duru, U., Arabi, M., Wohl, E.E., 2018. Modeling stream flow and sediment yield using the SWAT model: a case study of Ankara River basin. Turkey. Phys. Geogr. 39 (3), 264–289. https://doi.org/10.1080/02723646.2017.1342199.

Dutta, S., Sen, D., 2018. Application of SWAT model for predicting soil erosion and sediment yield. Sustain. Water Resour. Manag. 4 (3), 447–468. https://doi.org/10.1007/s40899-017-0127-2.

Epelde, A.M., Cerro, I., Sánchez-Pérez, J.M., Sauvage, S., Srinivasan, R., Antigüedad, I., 2015. Application of the SWAT model to assess the impact of changes in agricultural management practices on water quality. Hydrol. Sci. J. 1–19. https://doi.org/10.1080/02626667.2014.967692.

Fu, B., Merritt, W.S., Croke, B.F.W., Weber, T.R., Jakeman, A.J., 2019. A review of catchment-scale water quality and erosion models and a synthesis of future prospects. Environ. Modell. Softw. 114, 75–97. https://doi.org/10.1016/j.envsoft.2018.12.008.

Gallart, F., Llorens, P., 2003. Catchment Management under Environmental Change: Impact of Land Cover Change on Water Resources. Water Int. 28 (3), 334–340. https://doi.org/10.1080/02508060308691707.

Gassman, P.W., Sadeghi, A.M., Srinivasan, R., 2014. Applications of the SWAT Model Special Section: Overview and Insights. J. Environ. Qual. 43 (1), 1–8. https://doi.org/10.2134/jeq2013.11.0466.

Ghasempour, R., Roushangar, K., Sihag, P., 2021. Suspended sediment load prediction in consecutive stations of river based on ensemble pre-post-processing kernel based approaches. Water Supply 21 (7), 3370–3386.

Gupta, D., Hazarika, B.B., Berlin, M., Sharma, U.M., Mishra, K., 2021. Artificial intelligence for suspended sediment load prediction: a review. Environ. Earth Sci. 80, 346. https://doi.org/10.1007/s12665-021-09625-3.

Hamaamin, Y., Nejadhashemi, A., Zhang, Z., Giri, S., Woznicki, S., 2016. Bayesian Regression and Neuro-Fuzzy Methods Reliability Assessment for Estimating Streamflow. Water 8, 287. https://doi.org/10.3390/w8070287.

Ji, H., Chen, Y., Fang, G., Li, Z., Duan, W., Zhang, Q., 2021. Adaptability of machine learning methods and hydrological models to discharge simulations in data-sparse glaciated watersheds. J. Arid Land. 13 (6), 549–567. https://doi.org/10.1007/s40333-021-0066-5.

Jimeno-Sáez, P., Senent-Aparicio, J., Pérez-Sánchez, J., Pulido-Velazquez, D., 2018. A Comparison of SWAT and ANN Models for Daily Runoff Simulation in Different Climatic Zones of Peninsular Spain. Water 10, 192. https://doi.org/10.3390/w10020192.

Kalin, L., Isik, S., Schoonover, J.E., Lockaby, B.G., 2010. Predicting Water Quality in Unmonitored Watersheds Using Artificial Neural Networks. J. Environ. Qual. 39 (4), 1429–1440.

Khosravi, K., Cooper, J.R., Daggupati, P., Thai Pham, B., Tien Bui, D., 2020. Bedload transport rate prediction: Application of novel hybrid data mining techniques. J. Hydrol. 585, 124774. https://doi.org/10.1016/j.jhydrol.2020.124774.

Kim, R.J., Loucks, D.P., Stedinger, J.R., 2012. Artificial Neural Network Models of Watershed Nutrient Loading. Water Resour. Manage. 26 (10), 2781–2797. https://doi.org/10.1007/s11269-012-0045-x.

Kisi, O., 2005. Suspended sediment estimation using neuro-fuzzy and neural network approaches/Estimation des matière. Hydrol. Sci. J. 50, 683–696. https://doi.org/10.1623/hysj.2005.50.4.683.

Kisi, O., 2004. Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation / Prévision et estimation de la concentration en matières en suspension avec des perceptrons multi-couches et l'algorithme d'apprentissage de Levenberg-Marquardt. Hydrol. Sci. J. 49, 3. https://doi.org/10.1623/hysj.49.6.1025.55720.

Kumar, D., Pandey, A., Sharma, N., Flügel, W.-A., 2016. Daily suspended sediment simulation using machine learning approach. CATENA 138, 77–90. https://doi.org/10.1016/j.catena.2015.11.013.

Li, P., Mu, X., Holden, J., Wu, Y., Irvine, B., Wang, F., Gao, P., Zhao, G., Sun, W., 2017. Comparison of soil erosion models used to study the Chinese Loess Plateau. Earth Sci. Rev. 170, 17–30. https://doi.org/10.1016/j.earscirev.2017.05.005.

López-Ballesteros, A., Senent-Aparicio, J., Srinivasan, R., Pérez-Sánchez, J., 2019. Assessing the Impact of Best Management Practices in a Highly Anthropogenic and Ungauged Watershed Using the SWAT Model: A Case Study in the El Beal Watershed (Southeast Spain). Agronomy 9, 576. https://doi.org/10.3390/agronomy9100576.

Ma, D., Qian, B., Gu, H., Sun, Z., Xu, Y., 2021. Assessing climate change impacts on streamflow and sediment load in the upstream of the Mekong River basin. Int J Climatol 41, 3391–3410. https://doi.org/10.1002/joc.7025.

Mapes, K.L., Pricope, N.G., 2020. Evaluating SWAT Model Performance for Runoff, Percolation, and Sediment Loss Estimation in Low-Gradient Watersheds of the Atlantic Coastal Plain. Hydrology 7, 21. https://doi.org/10.3390/hydrology7020021.

Meaurio, M., Zabaleta, A., Uriarte, J.A., Srinivasan, R., Antigüedad, I., 2015. Evaluation of SWAT models performance to simulate streamflow spatial origin. The case of a small forested watershed. J. Hydrol. 525, 326–334. https://doi.org/10.1016/j.jhydrol.2015.03.050.

Merritt, W.S., Letcher, R.A., Jakeman, A.J., 2003. A review of erosion and sediment transport models. Environ. Modell. Softw. 18 (8-9), 761–799. https://doi.org/10.1016/S1364-8152(03)00078-1.

Minns, A.W., Hall, M.J., 1996. Artificial neural networks as rainfall-runoff models. Hydrol. Sci. J. 41 (3), 399–417. https://doi.org/10.1080/02626669609491511.

Moriasi, D.N., Arnold, J.G., Liew, M.W.V., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Trans. ASABE 50, 885–900. https://doi.org/10.13031/2013.23153.

Nachtergaele, F., van Velthuizen, H., Batjes, N., Dijkshoorn, K., van, V., Fischer, G., Jones, A., Montanarella, L., Petri, M., Prieler, S., Teixeira, E., Wiberg, D., 2010. The harmonized world soil database 4.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2009. Soil and Water Assessment Tool Theoretical Documentation Version 2009. Texas Water Resources Institute Technical Report No. 406.

Nguyen, P., Shearer, E.J., Tran, H., Ombadi, M., Hayatbini, N., Palacios, T., Huynh, P., Braithwaite, D., Updegraff, G., Hsu, K., Kuligowski, B., Logan, W.S., Sorooshian, S., 2019. The CHRS Data Portal, an easily accessible public repository for PERSIANN global satellite precipitation data. Sci. Data 6, 180296. https://doi.org/10.1038/sdata.2018.296.

Nunes, J.P., Naranjo Quintanilla, P., Santos, J.M., Serpa, D., Carvalho-Santos, C., Rocha, J., Keizer, J.J., Keesstra, S.D., 2018. Afforestation, Subsequent Forest Fires and Provision of Hydrological Services: A Model-Based Analysis for a Mediterranean Mountainous Catchment: Mediterranean Afforestation, Forest Fires and Hydrological Services. Land Degrad. Develop. 29 (3), 776–788. https://doi.org/10.1002/ldr.2776.

Olyaie, E., Banejad, H., Chau, K.-W., Melesse, A.M., 2015. A comparison of various artificial intelligence approaches performance for estimating suspended sediment load of river systems: a case study in United States. Environ. Monit. Assess. 187, 189. https://doi.org/10.1007/s10661-015-4381-1.

Owens, L.B., Edwards, W.M., Van Keuren, R.W., 1997. Runoff and sediment losses resulting from winter feeding on pastures. J. Soil Water Conserv. 52 (3), 194–197.

Pandey, S., Kumar, P., Zlatic, M., Nautiyal, R., Panwar, V.P., 2021. Recent advances in assessment of soil erosion vulnerability in a watershed. Int. Soil Water Conserv. Res. 9 (3), 305–318. https://doi.org/10.1016/j.iswcr.2021.03.001.

Pulighe, G., Bonati, G., Colangeli, M., Traverso, L., Lupia, F., Altobelli, F., Dalla Marta, A., Napoli, M., 2019. Predicting Streamflow and Nutrient Loadings in a Semi-Arid Mediterranean Watershed with Ephemeral Streams Using the SWAT Model. Agronomy 10, 2. https://doi.org/10.3390/agronomy10010002.

Quinlan, J.R., 1992. Learning with Continuous Classes. In: Presented at the Proceedings of Australian Joint Conference on Artificial Intelligence, Hobart, pp. 343–348.

Sarkar, T., Tapas, P., 2021. Revisiting the methodological development in soil erosion research. Ensm. 2, 145–165. https://doi.org/10.37948/ensemble-2020-0202-a016.

Schonlau, M., Zou, R.Y., 2020. The random forest algorithm for statistical learning. Stata J. 20 (1), 3–29. https://doi.org/10.1177/1536867X20909688.

Senent-Aparicio, J., Jimeno-Sáez, P., Bueno-Crespo, A., Pérez-Sánchez, J., Pulido-Velázquez, D., 2019. Coupling machine-learning techniques with SWAT model for instantaneous peak flow prediction. Biosyst. Eng. 177, 67–77. https://doi.org/10.1016/j.biosystemseng.2018.04.022.

Shen, C., Chen, X., Laloy, E., 2021. Editorial: Broadening the Use of Machine Learning in Hydrology. Front. Water 3, 681023. https://doi.org/10.3389/frwa.2021.681023.

Sihag, P., Sadikhani, M.R., Vambol, V., Vambol, S., Prabhakar, A.K., Sharma, N., 2021. Comparative study for deriving stage- discharge–sediment concentration relationships using soft computing techniques. J. Achiev. Mater. Manuf. Eng. 2 (104), 57–76.

Singh, A., Imtiyaz, M., Isaac, R.K., Denis, D.M., 2012. Comparison of soil and water assessment tool (SWAT) and multilayer perceptron (MLP) artificial neural network for predicting sediment yield in the Nagwa agricultural watershed in Jharkhand. India. Agric. Water Manage. 104, 113–120. https://doi.org/10.1016/j.agwat.2011.12.005.

Singh, A., Imtiyaz, M., Isaac, R.K., Denis, D.M., 2014. Assessing the performance and uncertainty analysis of the SWAT and RBNN models for simulation of sediment yield in the Nagwa watershed. India. Hydrol. Sci. J. 59 (2), 351–364. https://doi.org/10.1080/02626667.2013.872787.

Sirabahenda, Z., St-Hilaire, A., Courtenay, S.C., van den Heuvel, M.R., 2020. Assessment of the effective width of riparian buffer strips to reduce suspended sediment in an agricultural landscape using ANFIS and SWAT models. CATENA 195, 104762. https://doi.org/10.1016/j.catena.2020.104762.

Srivastava, P., McNair, J.N., Johnson, T.E., 2006. Comparison of process-based and artificial neural network approaches for streamflow modeling in an agricultural watershed. J. Am. Water Resour. Assoc. 42 (3), 545–563. https://doi.org/10.1111/j.1752-1688.2006.tb04475.x.

Wang, Y., Witten, I.H., 1997. Induction of Model Trees for Predicting Continuous Classes. In: Presented at the 9th Eur Conf on Machine Learning, Prague (Czech Republic).

Williams, J.R., 1975. Sediment routing for agricultural watersheds. Water Resour. Bull. 11 (5), 965–974.

Zeiger, S.J., Hubbart, J.A., 2016. A SWAT model validation of nested-scale contemporaneous stream flow, suspended sediment and nutrients from a multiple-land-use watershed of the central USA. Sci. Total Environ. 572, 232–243. https://doi.org/10.1016/j.scitotenv.2016.07.178.

Zounemat-Kermani, M., Mahdavi-Meymand, A., Alizamir, M., Adarsh, S., Yaseen, Z.M., 2020. On the complexities of sediment load modeling using integrative machine learning: Application of the great river of Loíza in Puerto Rico. J. Hydrol. 585, 124759. https://doi.org/10.1016/j.jhydrol.2020.124759.