

Daily Rainfall-Runoff Modeling at Watershed Scale: A Comparison Between Physically-Based and Data-Driven Models

Federico Vilaseca(✉), Alberto Castro, Christian Chreties, and Angela Gorgoglione

Universidad de la República, 11300 Montevideo, Uruguay
fvilaseca@fing.edu.uy

Abstract. In the last decades, data-driven (DD) machine-learning models have been rapidly developed and widely applied to solve hydrologic problems. To explore DD approaches' capability in rainfall-runoff modeling compared to knowledge-driven models, we conducted a thorough comparison between Soil & Water Assessment Tool (SWAT) and Random Forest (RF) models. They were implemented to simulate the daily surface runoff at Santa Lucía Chico watershed in Uruguay. Aiming at making a fair comparison, the same input time series for RF and SWAT models were considered. Both approaches are able to represent the daily surface runoff adequately. The RF model shows a higher accuracy for calibration/training, while the SWAT model yields better results for validation/testing, indicating that the latter has a better generalization capacity. Furthermore, RF outperforms SWAT in terms of computational time needed for a proper calibration/training. Strategies to improve RF performance and interpretability should include feature selection, feature engineering and a more sophisticated sensitivity analysis technique.

Keywords: Hydrology · SWAT · Random Forest · Machine learning

1 Introduction

Due to heavy rainfall events, floods are among the most catastrophic natural hazards in terms of losses of lives and economic damages [1]. Therefore, to reduce natural disasters, it is of paramount importance to have accurate and reliable predictions of daily runoff discharges [2]. With the aim of dealing with pragmatic needs, such as warning alerts, a substantial amount of research effort has been invested in developing and implementing reliable hydrologic models as tools for better understanding and characterizing the complex and non-linear rainfall-runoff process at watershed scale [3–5]. Its multifaceted nature is due to the high temporal and spatial variability and its dependence on different factors: climatic conditions, land use/land cover, soil type, and physical watershed characteristics [6].

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12955, pp. 18–33, 2021.
https://doi.org/10.1007/978-3-030-87007-2_2

Numerous methods have been adopted to simulate and/or predict the rainfall-runoff process. Overall, such methods can be divided into physically-based (PB) and data-driven (DD) models. The former approach, also called knowledge-driven models, simplifies the system under study and represents its inner processes by solving mathematical equations assessed from physics. In river hydraulics, these are the 1D- and 2D-hydrodynamic models, and in hydrology, these are the conceptual and distributed models [7]. The DD approach, or empirical models, involves mathematical equations evaluated not from the physical process in the catchment but from the analysis of simultaneous input and output time series [8].

During the last decades, among the PB models, the ones based in geographical information systems (GIS) couple the descriptions of hydrologic features on a spatial scale with the simulation and predictive power of models [7]. The Soil & Water Assessment Tool (SWAT) is one of these models, widely used worldwide [9–11]. It is a process-based model that embeds most hydrologic processes using the water-balance principle [9]. Its versatility ensures its multiple uses in hydrologic matters. In recent years, it has been used to assess the impact and variation of different runoff-induced processes such as erosion [12], suspended-sediment transport [13], and nutrient contribution to water bodies from non-point sources [14] at a watershed scale. Its spatially distributed scheme gives the possibility to simulate different land-use and water-management scenarios and assess their impact on water quantity and quality [15–17].

However, several researchers have stated that the PB modeling approach deals with the need of numerous input data and a higher computational time to reach a reasonable accuracy and reliability [1, 18, 19]. For this reason, recently, DD models have received significant attention from environmental engineers and researchers. Among several techniques, Random Forest (RF) is one of the most adopted since it has the advantage of being a white box method and provides feature importance based on out of the bag sampling, which allows a better understanding of the learning process [20].

Many successful applications of RF to hydrologic rainfall-runoff modeling have been reported in recent years [20–23]. However, there still exist challenges about improving the modeling performance with daily (and sub-daily) time-step, particularly in complex watersheds with highly non-linear behavior [24]. The proper selection of input variables (*feature engineering*), and their influence on model accuracy, is another point currently under discussion [25, 26].

It is clear that a rigorous analysis and comprehensive discussion are needed to understand which are the capabilities and drawbacks of both modeling approaches when applied to the same case study. In particular, it is important to extend the body of knowledge with crucial insights into the application of machine-learning techniques for modeling physical-water-related processes at watershed scale.

Based on these considerations, the main objective of this study is to provide an in-depth comparison between a PB distributed model, SWAT, and a DD model, RF, in simulating the daily rainfall-runoff process at watershed scale. In particular, the comparison will tackle the following aspects: *(i)* accuracy in simulating the physical rainfall-runoff process, *(ii)* computational cost in terms of hardware resource requirements and computational time, and *(iii)* model sensitivity (most influencing parameters).

2 Materials

2.1 Study Area

The study area selected for this work is Santa Lucía Chico watershed, with the outlet located at the closure point of Paso Severino dam (Fig. 1). It is positioned in the south-central region of Uruguay, covering an area equal to 2478 km^2. Its main urban center is Florida city (Fig. 1), and the major land covers are grassland (82.4% of total area) and agriculture (9.4%). The climate is temperate with a total annual precipitation that ranges between 1000 mm and 1500 mm and an air temperature that varies between 3 °C and 30 °C [27].

It is a watershed of national importance due to the intense agricultural activity that takes place in its territory and because it is one of the primary sources of drinking water for the country. Paso Severino dam, in activity since the '80s, was created as a reservoir to support the water treatment plant of Aguas Corrientes during dry periods.

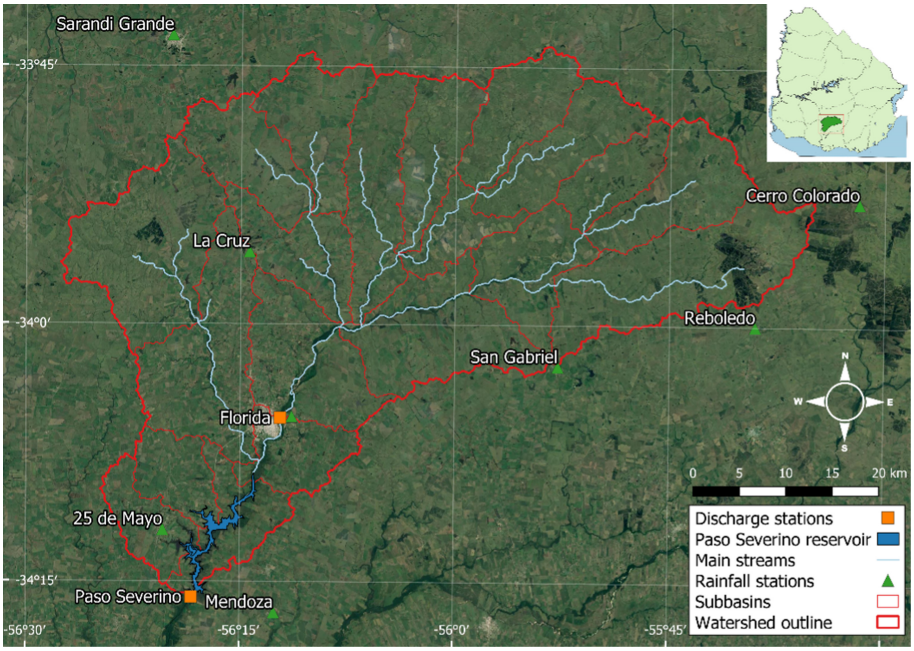

Fig. 1. Santa Lucía Chico watershed and location of the discharge stations (orange squares) and rainfall stations (green triangles). The box in the top right corner shows the location of the watershed (dark green) in Uruguay. Grid's coordinates are Latitude and Longitude in degrees. (Color figure online)

2.2 Data Description

Time Series: Mean daily streamflow was recorded at the station located in Florida city, which is owned and managed by the Uruguay National Water Board (DINAGUA). Data were available from 1/1/1980 to 6/30/2020. In addition, the mean daily discharge of Paso Severino dam was available, covering the period from 1/1/1990 to 12/31/2015. The data was provided by the State Sanitary Infrastructures (OSE), which is the state company in charge of the distribution of drinking water and dam management.

Accumulated daily rainfall records were obtained for eight pluviometric stations located around the watershed at Cerro Colorado, Florida, La Cruz, Mendoza, Reboledo, San Gabriel, Sarandí Grande, and 25 de Mayo (Fig. 1). Most of them have records from 1/1/1980 to 6/30/2020, except for Florida, the only meteorological station, where records start from 5/1/1989. The information was provided by the National Institute of Meteorology (INUMET), the Uruguayan state agency in charge of the meteorological forecasting.

Other climatic variables were also considered to compute evapotranspiration: relative humidity (RH), solar radiation (SR), maximum and minimum daily temperatures (T_{max} and T_{min}), and wind speed (WS). The time series of such variables, covering the period from 1/1/1980 to 6/30/2020 were obtained from the database of the National Institute of Agricultural Research (INIA). Records from Las Brujas meteorological station were used (Lat: −34.67, Lon: −56.34), as it was the closest to the study area.

Spatial Information: Spatial information used for this study includes land use/land cover, soil types, and topography. The three digital maps were obtained in raster format, with cell size 30 m × 30 m from the Ministry of Livestock, Agriculture, and Fisheries (MGAP) (Fig. 2).

The current national land use/land cover map corresponds to the year 2018 and has a scale of 1:50,000 [28]. The main land use, covering 82.4% of the total area, is grassland, which is mostly used for extensive livestock farming. The second important land use in terms of surface is agriculture, which occupies 9.4% of the area. Crops include oats, soy, wheat, and corn. The remaining is occupied by forestry (4.9%), dairy farming (1.2%), urban areas (0.9%), water bodies (0.7%) and wetlands (0.5%) (Fig. 2(a)).

The soil map adopted for this study uses soil classification according to a national productivity index (CONEAT) [29]. Nineteen different soil types are present, mainly loamy soils belonging to hydrological groups B and C (Fig. 2(b)).

The digital terrain model comes from a national topographical map with scale 1:50000 [30]. Watershed slopes are low, with less than 5% for 77% of the area and more than 10% for only 2% (Fig. 2(c)).

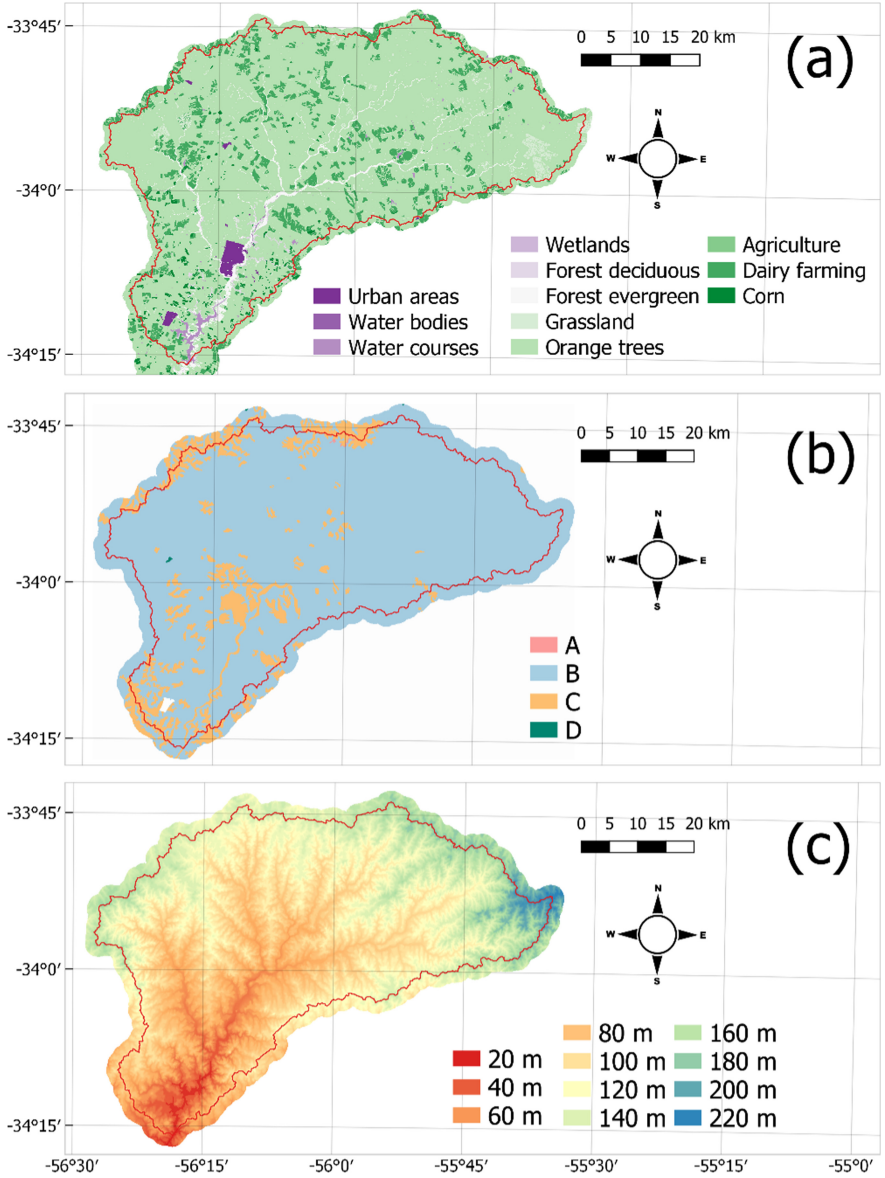

Fig. 2. (a) Land use/land cover map; (b) map of hydrological group of soils; (c) digital terrain model of Santa Lucía Chico watershed. Grid's coordinates are Latitude and Longitude in degrees. (Color figure online)

2.3 Software and Computational Equipment

To build the PB SWAT model, the QGIS based tool QSWAT was used. It consists in a graphical interphase which allows to easily generate the model input files from spatial data shapefiles. The calibration and validation process were carried out using the SWAT-CUP software [31], which is designed specifically for SWAT calibration and uncertainty analysis, enabling the use of several optimization functions. SWAT computations ran in a 3 GHz Intel i5 processor with 8 GB of memory.

The RF algorithm [32] was coded in Python 3.8 and ran on a 2 GHz Intel i3 PC with 8 GB of memory. The library *scikit-learn* [33] was used for this implementation.

3 Methods

3.1 Physically-Based Model Setup

The SWAT hydrologic model was chosen for the PB implementation. It is widely used for rainfall-runoff simulation at watershed scale and has been proven to yield accurate results for both monthly and daily time steps. Its capabilities exceed rainfall-runoff modelling, as it incorporates models for sediment and nutrient generation and transport [34].

Calculations are made in a spatially distributed scheme, considering the physical characteristics of the watershed. Its surface is divided into two scales: *i)* subbasins, which are determined based on the distribution of the channel network, and *ii)* hydrological response units (HRUs), which are a sub-division of the subbasin. HRUs are determined by the intersection of land use, soil type, and slope maps. Runoff calculations are made for each HRU, and the result is then aggregated for each subbasin and for the entire watershed. For this implementation 20 subbasins (Fig. 1) and 317 HRUs were delineated. HRUs were determined by setting a filter threshold of 5% for land use, soil type, and elevation maps.

SWAT includes different methods to perform the calculations of the hydrological processes. In this case, daily surface-runoff computation was based in the Soil Conservation Service (SCS) Curve Number (CN) method. For evapotranspiration, the Penman-Monteith option was chosen, and the Variable Storage method was selected to simulate transport from the subbasins to the watershed's outlet.

Information on land-use management was obtained from a previous SWAT model setup in the same basin [35]. Climatic input for the model includes rainfall at the eight stations with available data (Fig. 1), and relative humidity, solar radiation, maximum and minimum daily temperatures, and wind speed from INIA Las Brujas station. Paso Severino reservoir, located downstream, close to the watershed outlet, was included in the model as an uncontrolled reservoir with an average annual release rate. In this way, its output is not fixed, allowing to calibrate the model including reservoir parameters in the iteration.

The initial model setup, previous to the calibration and validation process, was carried out covering the period between 1/1/1990 and 12/31/2015 (26 years), where all the required information is available, leaving the first three years for model warmup.

3.2 Data-Driven Model Setup

The chosen algorithm for the DD model was RF. It consists in an ensemble of several classification and regression trees (CART) [32]. Each CART is trained from a random subset of features and variables taken with replacement from the full training dataset (bootstrap). After the RF is trained, the value of the output variable corresponding to a new observation is calculated for all the trees. Then the final output of the RF is the average of the output of each individual tree (RF regressor). It is proven that this approach yields models able to represent non-linear behaviors without overfitting to the training data [24].

It is regarded as a "white-box" machine-learning algorithm since information about the training process can be easily extracted. It also has the advantage of being able to provide a sensitivity score for each variable, based on the error calculated from out-of-the-bag samples during the training process [22].

To provide a fair comparison between PB and DD models, for the latter, we considered the same input variables adopted by SWAT: rainfall, relative humidity, solar radiation, maximum and minimum daily temperatures, and wind speed. The streamflow time series of Florida was also included among the inputs, and discharge at Paso Severino was defined as the model output. To provide the model with time awareness, past values of all the selected input variables for the three previous days (lagged variables) were also included. Finally, date information was incorporated through categorical input variables (year, month, week, day, day of the week, and boolean variables indicating if it is the start or the end day of the month), and elapsed date (continuous variable). Table 1 shows a summary of the input variables for the RF model.

Table 1. Summary of input variables for the RF model.

Variable	Stations	Type	Lag
Runoff	Florida	Continuous	$t-1, t-2, t-3$
Rainfall	Cerro Colorado, Florida, La Cruz, Mendoza, Reboledo, San Gabriel, Sarandí Grande y Villa 25 de Mayo	Continuous	$t-1, t-2, t-3$
Relative humidity	Las Brujas	Continuous	$t-1, t-2, t-3$
Solar radiation	Las Brujas	Continuous	$t-1, t-2, t-3$
Max. temperature	Las Brujas	Continuous	$t-1, t-2, t-3$
Min. temperature	Las Brujas	Continuous	$t-1, t-2, t-3$
Wind speed	Las Brujas	Continuous	$t-1, t-2, t-3$
Year		Categorical	
Month		Categorical	
Week		Categorical	
Day		Categorical	

(*continued*)

Table 1. (*continued*)

Variable	Stations	Type	Lag
Day of week		Categorical	
Is month end		Categorical	
Is month start		Categorical	
Elapsed date		Continuous	

3.3 Calibration, Validation and Sensitivity Analysis of the Physically-Based Model

For SWAT calibration/validation, SWAT-CUP [31] was adopted, by exploiting the SUFI-2 algorithm [36] for model optimization. A set of 17 parameters was selected based on literature recommendations [34, 36, 37] and consideration of the main physical processes that influence watershed hydrology. The optimization considered the output at both Paso Severino and Florida, using Nash-Sutcliffe efficiency (NSE) as the objective function. The process involved over seven sets of 100 model-runs each. The calibration period considered was from 1/1/1993 to 12/31/2006. After calibration, the model was validated by making one run with the best parameter ranges resulting from calibration, for the period 1/1/2010 to 12/31/2015. The period between 1/1/2007 and 12/31/2009 was not used for calibration or validation due to the high amount of missing data in the Florida runoff time series.

Model performance was evaluated using coefficient of determination (R^2), Kling-Gupta Efficiency (KGE), and root mean squared error (RMSE).

Calibration in SWAT-CUP with SUFI-2 algorithm yields a sensitivity analysis for calibration parameters. This tool was used for the first 100 iterations to identify the parameters and physical processes that have a major influence on watershed hydrology. Sensitivity is characterized by p-value and t-stat resulting from a t-test over model iterations [34]. Lower values of p-value and higher absolute values of t-stat indicate a higher sensitivity.

3.4 Training, Testing and Sensitivity Analysis of the Data-Driven Model

For the DD model, the training/validation process consisted of fitting several RFs to the data varying the hyperparameters of the algorithm to find the ones that determine the most accurate model. To this end, several instances of 5-fold cross-validation were run, combining the hyperparameter values shown in Table 2. The best model was chosen based on the resulting NSE.

Table 2. Value of hyperparameters tested during 5-fold cross-validation.

Hyperparameter	Values
min_samples_leaf	1, 5, 10
max_features	$n^{0.5}$, n
max_samples	0.5 × m, 0.75 × m, m
n_estimators	1×10^3, 1×10^4

For a better comparison, the data period considered is the same as in the calibration of the SWAT model (from 1/1/1993 to 12/31/2006). Once the best hyperparameters for the model were determined, it was fit to the data of the calibration period and then used to predict the output for a testing period, which matched the one chosen to validate the PB model (from 1/1/2010 to 12/31/2015). The same metrics (R^2, KGE and RMSE) were used for performance evaluation.

The RF out-of-the-bag feature-importance score was used to assess the model's output sensitivity to input variables. It takes values between 0 and 1, with bigger values indicating higher sensitivity. In the Discussion section, results were interpreted from a physical point of view and compared with the resulting parameter sensitivity from the PB SWAT model.

4 Results

4.1 Simulation Accuracy

Table 3 shows comparative model performance at Paso Severino, based on the selected goodness-of-fit indicators, while Fig. 3 displays comparative time series plots of both models results and measured streamflow at Paso Severino. Two-year periods with varying hydrological conditions were chosen for the calibration and validation comparison.

Table 3. Comparison of performance measures for both models.

Model	Instance	NSE	R^2	KGE	RMSE
SWAT	Calibration	0.61	0.62	0.66	60.8
SWAT	Validation	0.47	0.56	0.67	64.0
Random Forest	Training	0.90	0.91	0.75	31.5
Random Forest	Testing	0.43	0.47	0.57	67.5

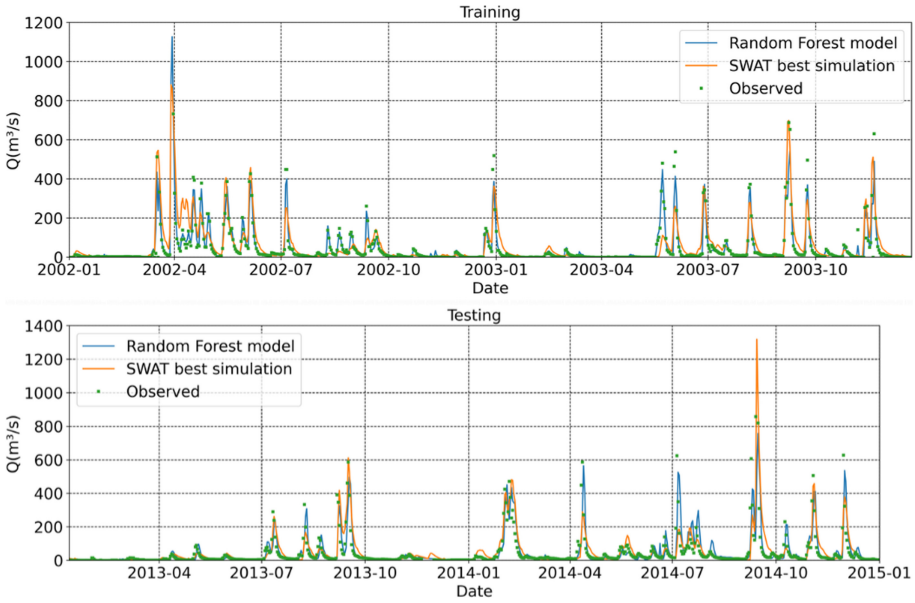

Fig. 3. Comparison of observed flow rate (green dots) with the simulations of RF (blue line) and SWAT model (orange line) at Paso Severino. Selected two-year windows for training/calibration (2002–2004) and testing/validation (2013–2015). (Color figure online)

4.2 Computational Cost

In terms of computational time, running 17 years of the SWAT model takes 70 s with the used equipment. Taking into account that over 700 runs were required for a full calibration process, the overall time is around 817 min (nearly 14 h) to obtain a working version of the model. The training process of the RF model needed 88 s to run. The hyperparameter search through cross-validation required 36 training instances correspondent to the different hyperparameter value combinations, so, in total it took 53 min to obtain the model.

4.3 Sensitivity Analysis

Parameter sensitivity of SWAT model, and feature importance from RF model are shown respectively in Table 4 and Table 5.

Table 4. Results from SWAT-CUP parameter sensitivity. Bigger absolute values of t-stat and smaller values of p-value indicate higher sensitivity.

Parameter	Definition	P-value	T-stat
ALPHA_BF.gw	Recession constant of base flow	0.00	13.95
CH_K2.rte	Hydraulic conductivity of channels (outside subbasins)	0.00	−4.77
CN2.mgt	Curve number associated to soil cover	0.00	4.75
CH_N1.sub	Manning roughness coefficient (inside subbasins)	0.00	−4.59
GWQMN.gw	Threshold water content in underground reservoir for return flow to occur	0.01	2.75
ESCO.hru	Compensation factor for soil water evapotranspiration	0.01	−2.71
GW_REVAP.gw	Return coefficient for underground water to soil saturated zone	0.01	2.60
CH_K1.sub	Hydraulic conductivity of channels (inside subbasins)	0.01	−2.52
EVRCH.bsn	Evapotranspiration adjustment factor	0.02	−2.47
CH_N2	Manning roughness coefficient (outside subbasins)	0.05	−2.01

Table 5. Feature importance resulting from Random Forest out of the bag scoring.

Variable	Location	Lag	Feature importance
Runoff	Florida		0.593
Runoff	Florida	t − 1	0.124
Runoff	Florida	t − 2	0.018
Elapsed date			0.015
Runoff	Florida	t − 3	0.014
Rainfall	La Cruz		0.009
Rainfall	Cerro Colorado		0.008
Rainfall	Sarandí Grande		0.007
Rainfall	25 de Mayo	t − 1	0.007
Wind speed	Las Brujas		0.007
Rainfall	Cerro Colorado	t − 1	0.007

5 Discussion

The DD model has a better performance for the training/calibration set than the PB, while for the testing/validation set the opposite is observed. This indicates a tendency to overfit by RF model, which is a common issue in DD modeling, while the SWAT model shows a better generalization capacity. This result illustrates the ongoing discussion between

both types of models. While DD models usually show better accuracy, they lack the level of generalization usually achieved by PB models, which reduces their reliability to simulate fictitious scenarios. In this case, the RF model's accuracy is below expected for the testing set. This could be due to the complexity of the watershed, which has an irregular shape and a large reservoir near its output. Another reason could be the lack of a proper analysis for input variable selection. In this case, they were chosen with a comparative criterion, but the inclusion of excess information could have led to the introduction of noise in the training process.

According to the guidelines set by Moriasi et al. [38], a threshold value of 0.5 for NS should be considered to rate a hydrological model as satisfactory. With this reference, both presented models achieve satisfactory performance for calibration/training (NS = 0.61 for PB model and NS = 0.90 for DD model) but are unsatisfactory for validation/testing (NS = 0.47 for PB model and NS = 0.43 for DD model).

A review of the performance of other recently developed watershed-scale rainfall runoff models, taken from the cited works of this paper, shows that NS values of SWAT models range between 0.36 and 0.99, being 0.78 the average value, while R^2 varies between 0.59 and 0.99, with an average value of 0.82 [9–15]. For DD models, NS is in the range of 0.36 to 0.99, with an average of 0.79, and R^2 varies between 0.50 and 0.93, averaging 0.72 [22, 23, 25, 26]. In comparison with these results, the PB model falls below average regarding performance metrics, while the DD model has an above-average performance for calibration, but not for validation, indicating, as stated before, that there might be overfitting issues.

Another important reference to be taken into account for comparison is the work of Narbondo et al. [3], who implemented a PB lumped model in the same watershed, using the GR4J model, and obtained calibration values of 0.68 and 0.83 for NS and R^2 respectively. This implementation shows a better performance for calibration than the SWAT model, while being also below average in comparison with the other results for PB models taken from the state-of-the-art. That may indicate that the performance of PB hydrological models is limited for the case study, due to the physical characteristics of the watershed, the high temporal and spatial variability that characterizes the hydro-meteorological variables of the country, and the quality issues in the input available data. However, there is still room for improvement of the performance of both models.

The time series plots show that both models are able to adequately represent the main physical processes such as base flow and hydrogram increment and descent. This is expected from the PB model, which has such processes considered in its parametrization, but not from the DD model in which they are learned from the data themselves. In both cases, an underestimation of peak flows is observed for major floods. This behavior is a common issue in SWAT but is not reported as such for RF models.

Sensitivity analysis for the SWAT model yields that ALPHA_BF.gw is the most sensitive parameter. It takes part in the process of the contribution of groundwater to base flow, indicating the relevance of this component in the total outflow of the watershed. CH_K2.rte and CH_N1.sub are also among the most sensitive parameters, and they intervene in the concentrated transport process. Then, CN2.mgt is the SCS runoff curve number, which has relevance in initial abstraction and flow concentration speed during rainfall events.

RF feature importance shows that runoff in Florida, for the actual and previous days, has major relevance in the output at Paso Severino. This is to be expected, due to the high correlation between both variables, and highlights the importance of including past values of relevant variables as inputs, to provide the algorithm with insight on the temporality of hydrological processes. The following variables in importance are rainfall measurements at Cerro Colorado, Sarandí Grande, and 25 de Mayo. It is interesting that they surpass in importance those that are located closer to the watershed's geometrical center. Wind speed, representative of the evapotranspiration process, shows similar importance to rainfall. The extracted feature importance does not show a high potential for physical interpretation, as it resumes correlation in a single value. In that sense, maybe other more complex machine-learning sensitivity analysis techniques, such as SHAP value analysis [39] could be implemented to enrich model interpretation.

In terms of computational time, the RF model clearly outperformed the SWAT model during the training/calibration stage, even when it ran in less powerful processing hardware. This highlights a strong point of DD modeling. However, it should be noted that the overall working time to develop an adequate model is more even, from a qualitative perspective. While the calibration/training process is faster for DD models, it is compensated with the selection of input features and feature engineering, which is often case-specific and necessary to develop the best possible model. This stage is not needed for PB modelling, where input variables are determined by the selected model. In this case, the RF model was developed considerably faster than the SWAT model, because input variables were chosen to be the same in both cases. It is also worth mentioning that GIS-based distributed models such as SWAT take a considerable amount of time for model setup, due to the complexity of their parametrization. This is not the case for DD models or simpler (lumped) PB models.

6 Conclusions

A PB model, SWAT, and a DD model, RF, were successfully implemented and compared for the Santa Lucía Chico watershed in Uruguay. The comparison was carried out in terms of simulation accuracy, computational cost, and model sensitivity.

RF model showed a better performance for calibration/training, while the SWAT model outperformed RF in validation/testing, considering the selected goodness-of-fit indicators. This indicated a tendency to overfit by the RF model and a better generalization capacity for the SWAT model.

Sensitivity analysis showed that the most relevant parameters for SWAT are linked to the processes of groundwater that contribute to base flow, concentrated transport, and flow concentration during rainfall events. For the RF model, feature importance shows the relevance of present and past values of runoff at Florida, located upstream, for prediction of discharge at the watershed's outlet.

The RF model outperformed the SWAT model in terms of computational calculation time. However, this difference in required modeling time is reduced if input feature selection and engineering are taken into account.

The comparison with other similar watershed-scale rainfall-runoff models, showed that both SWAT and RF models have room for improvement. Performance and interpretability of the RF model could be improved through feature engineering on input variables and the incorporation of more complex sensitivity analysis techniques respectively. These aspects will be taken into account in our future works.

References

1. Pappenberger, F., et al.: Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). Hydrol. Earth Syst. Sci. **9**(4), 381–393 (2005)
2. Young, C.C., Liu, W.-C., Wu, M.-C.: A physically based and machine learning hybrid approach for accurate rainfall-runoff modeling during extreme typhoon events. Appl. Soft Comput. **53**, 205–216 (2017)
3. Narbondo, S., Gorgoglione, A., Crisci, M., Chreties, C.: Enhancing physical similarity approach to predict runoff in ungauged watersheds in sub-tropical regions. Water **12**, 528 (2020)
4. Gorgoglione, A., Crisci, M., Kayser, R.H., Chreties, C., Collischonn, W.: A new scenario-based framework for conflict resolution in water allocation in transboundary watersheds. Water **11**, 1174 (2019)
5. Chang, W., Chen, X.: Monthly rainfall-runoff modeling at watershed scale: a comparative study of data-driven and theory-driven approaches. Water **10**, 1116 (2018)
6. Wang, W., Ding, J.: Wavelet network model and its application to the prediction of hydrology. Nat. Sci. **1**(1), 67–71 (2003)
7. Solomatine, D., Ostfeld, A.: Data-driven modelling: some past experiences and new approaches. J. Hydroinform. **10**(1), 3–22 (2008). https://doi.org/10.2166/hydro.2008.015
8. Wu, C.L., Chau, K.L.: Rainfall-runoff modelling using artificial neural network coupled with singular spectrum analysis. J. Hydrol. **399**(3–4), 394–409 (2003)
9. Apostel, A., et al.: Simulating internal watershed processes using multiple SWAT models. Sci. Total Environ. **759**, 143920 (2021). https://doi.org/10.1016/j.scitotenv.2020.143920
10. Chen, H., Luo, Y., Potter, C., Moran, P.J., Grieneisen, M., Zhang, M.: Modeling pesticide diuron loading from the San Joaquin watershed into the Sacramento-San Joaquin Delta using SWAT. Water Res. **121**, 374–385 (2017)
11. Pang, S., Wang, X., Melching, C., Feger, Karl-Heinz.: Development and testing of a modified SWAT model based on slope condition and precipitation intensity. J. Hydrol. **588**, 125098 (2020). https://doi.org/10.1016/j.jhydrol.2020.125098
12. Zhu, X., Wang, X., Zeng, S., Zhou, W.: Simulation of runoff and sediment yield in Zhuxi watershed in Changting county of China based on SWAT model. In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, Texas (2017)
13. Hallouz, F., Meddi, M., Mahé, G., Alirahmani, S., Keddar, A.: Modeling of discharge and sediment transport through the SWAT model in the basin of Harazza (Northwest of Algeria). Water Sci. **32**(1), 79–88 (2018)
14. Cheng, J., et al.: Modeling the sources and retention of phosphorus nutrient in a coastal river system in China using SWAT. J. Environ. Manag. **278**, 111556 (2021). https://doi.org/10.1016/j.jenvman.2020.111556
15. Busico, G., Colombani, N., Fronzi, D., Pellegrini, M., Tazioli, A., Mastrocicco, M.: Evaluating SWAT model performance, considering different soils data input, to quantify actual and future runoff susceptibility in a highly urbanized basin. J. Environ. Manag. **266**, 110625 (2020). https://doi.org/10.1016/j.jenvman.2020.110625

16. Marin, M., et al.: Assessing the vulnerability of water resources in the context of climate changes in a small forested watershed using SWAT: a review. Environ. Res. **184**, 109330 (2020). https://doi.org/10.1016/j.envres.2020.109330
17. Zhang, H., Wang, B., Liu, De Li., Zhang, M., Leslie, L., Qiang, Y.: Using an improved SWAT model to simulate hydrological responses to land use change: a case study of a catchment in tropical Australia. J. Hydrol. **585**, 124822 (2020). https://doi.org/10.1016/j.jhydrol.2020.124822
18. Rezaeianzadeh, M., et al.: Assessment of a conceptual hydrological model and artificial neural networks for daily outflows forecasting. Int. J. Environ. Sci. Technol. **10**(6), 1181–1192 (2013). https://doi.org/10.1007/s13762-013-0209-0
19. Ahmadi, M., Moeini, A., Ahmadi, H., Motamedvaziri, B., Zehtabijan, G.R.: Comparison of the performance of SWAT, IHACRES and artificial neural networks models in rainfall-runoff simulation (case study: Kan watershed, Iran). Phys. Chem. Earth A/B/C **111**, 65–77 (2019)
20. Li, M., Zhang, Y., Wallace, J., Campbell, E.: Estimating annual runoff in response to forest change: a statistical method based on random forest. J. Hydrol. **589**, 125168 (2020). https://doi.org/10.1016/j.jhydrol.2020.125168
21. Desai, S., Ouarda, T.: Regional hydrological frequency analysis at ungauged sites with random forest regression. J. Hydrol. **594**, 125861 (2021). https://doi.org/10.1016/j.jhydrol.2020.125861
22. Muñoz, P., Orellana-Alvear, J., Willems, P., Célleri, R.: Flash-flood forecasting in an Andean Mountain catchment—development of a step-wise methodology based on the random forest algorithm. Water **10**(11), 1519 (2018). https://doi.org/10.3390/w10111519
23. Pini, M., Scalvini, A., Liaqat, M.U., Ranzi, R., Serina, I., Mehmood, T.: Evaluation of machine learning techniques for inflow prediction in Lake Como, Italy. Procedia Comput. Sci. **176**, 918–927 (2020)
24. Schoppa, L., Disse, M., Bachmair, S.: Evaluating the performance of random forest for large-scale flood discharge simulation. J. Hydrol. **590**, 125531 (2020). https://doi.org/10.1016/j.jhydrol.2020.125531
25. Abbasi, M., Farokhnia, A., Bahreinimotlagh, M., Roozbahani, R.: A hybrid of random forest and deep auto-encoder with support vector regression methods for accuracy improvement and uncertainty reduction of long-term streamflow prediction. J. Hydrol. (2020). https://doi.org/10.1016/j.jhydrol.2020.125717
26. Adnan, R.M., Petroselli, A., Heddam, S., Guimaräes, C.A., Kisi, O.: Comparison of different methodologies for rainfall-runoff modelling: machine learning vs conceptual approach. Nat. Hazards **105**, 2987–3011 (2021)
27. Gorgoglione, A., Gregorio, J., Ríos, A., Alonso, J., Chreties, C., Fossati, M.: Influence of land use/land cover on surface-water quality of Santa Lucía River, Uruguay. Sustainability **12**(11), 4692 (2020)
28. MGAP: Uruguayan Integrated Land Use/Land Cover Map. Available online: https://www.gub.uy/ministerio-ganaderia-agricultura-pesca/comunicacion/publicaciones/mapa-integrado-coberturauso-del-suelo-del-uruguay-ano-2018 (2018). Accessed 22 March 2021
29. MGAP: Soil classification with CONEAT productivity index for Uruguayan soils. Available: https://www.gub.uy/ministerio-ganaderia-agricultura-pesca/tramites-y-servicios/servicios/consulta-coneat. Accessed 24 March 2021
30. MGAP: Digital terrain map of Uruguay. Available: https://www.gub.uy/ministerio-ganaderia-agricultura-pesca/tramites-y-servicios/servicios/modelo-digital-terreno. Accessed 24 March 2021
31. Abbaspour, K., et al.: Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. J. Hydrol. **333**(2–4), 413–430 (2007). https://doi.org/10.1016/j.jhydrol.2006.09.014

32. Brieman, L.: Random forests. Mach. Learn. **45**, 5–32 (2001)
33. Pedregosa, F., et al.: Scikit-learn: machine learning in Python. J. Mach. Learn. Res. **12**(85), 2825–2830 (2011)
34. Abbaspour, K., Vaghefi, S., Srinivasan, R.: A guideline for successful calibration and uncertainty analysis for soil and water assessment: a review of papers from the 2016 International SWAT Conference. Water **10**(1), 6 (2017). https://doi.org/10.3390/w10010006
35. Mer, F., et al.: SWAT Subcuenca Santa Lucia. https://doi.org/10.17605/OSF.IO/UQB5J (2020, October 11)
36. Abbaspour, K.C., Johnson, C.A., van Genuchten, M.: Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. Vadose Zone J. **3**(4), 1340–1352 (2004). https://doi.org/10.2136/vzj2004.1340
37. Abbaspour, K.C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H., Kløve, B.: A continental-scale hydrology and water quality model for Europe: calibration and uncertainty of a high-resolution large-scale SWAT model. J. Hydrol. **524**, 733–752 (2015). https://doi.org/10.1016/j.jhydrol.2015.03.027
38. Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L.: Model evaluation guidelines for systematic quantification and accuracy in watershed simulations. Trans. ASABE **50**(3), 885–900 (2007)
39. Lundberg, S.M., Lee, S.: A unified approach to interpreting model predictions. In: 31st Conference on Neural Information Processing Systems (NIPS), Long Beach, California, USA. arXiv:1705.07874 (2017)