water

MDPI

Article

Optimizing Spatial Discretization According to Input Data in the Soil and Water Assessment Tool: A Case Study in a Coastal Mediterranean Watershed

Mathilde Puche [1,2,*], Magali Troin [1,2], Dennis Fox [1] and Paul Royer-Gaspard [1,2]

[1] UMR CNRS ESPACE, Université Côte d'Azur, Campus Carlone, 06204 Nice, France
[2] Hydroclimat, 365 Chemin du Camp Sarlier, 13400 Aubagne, France
* Correspondence: mathilde.puche@etu.univ-cotedazur.fr

Abstract: Spatial discretization in hydrological models has a strong impact on computation times. This study investigates its effect on the performance of the Soil and Water Assessment Tool (SWAT) applied to a French Mediterranean watershed. It quantifies how spatial discretization (the number of sub-basins and hydrological response units (HRUs)) affects the SWAT model's performance in simulating daily streamflow and whether this effect depends on the choice of soil and land use input datasets. Sixty-eight SWAT model configurations were created using various soil and land use datasets and 17 discretization setups, evaluated from 2001 to 2021 with the Kling–Gupta efficiency (KGE) metric. The key findings include (1) while the number of sub-basins does not impact model performance, increasing HRUs significantly degrades it (KGE loss of 0.13 to 0.26) regardless of the number of sub-basins or input datasets. (2) SWAT is found to be more sensitive to variations in soil datasets than in land use datasets, but the observed performance decline with more HRUs is attributed to the calibration process and the increased heterogeneity in soil types rather than input dataset spatial resolution. (3) Minimizing the number of HRUs may improve both the accuracy of streamflow simulations and the computational efficiency of the SWAT model.

Keywords: watershed discretization; computational efficiency; soil type; land use; soil and water assessment tool; hydrological optimization; hydrological modeling

check for updates

Academic Editor: Athanasios Loukas

Received: 26 November 2024
Revised: 10 January 2025
Accepted: 11 January 2025
Published: 16 January 2025

Citation: Puche, M.; Troin, M.; Fox, D.; Royer-Gaspard, P. Optimizing Spatial Discretization According to Input Data in the Soil and Water Assessment Tool: A Case Study in a Coastal Mediterranean Watershed. *Water* **2025**, *17*, 239. https://doi.org/10.3390/w17020239

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Hydrological models support water-related risk decisions such as flood prediction [1,2], drought monitoring [3,4], and the hydrological impacts of climate change [5,6]. Based on the spatial representation of the watershed, hydrological models can be classified as lumped, semi-distributed, or fully distributed models [7]. Lumped models define the watershed as a single unit with spatially averaged parameters and meteorological inputs. In contrast, distributed models take into account the spatial variability in the hydrological processes, input data, and characteristics of the watershed by discretizing it into a regular or irregular grid. The advantages and disadvantages of lumped and distributed approaches are extensively discussed in the literature (e.g., [8–11]). The consensus is that lumped models fail to capture the spatial variability in hydrological processes, while fully distributed models are intensively data-demanding and time-consuming. Semi-distributed models are recognized as a good compromise between lumped and fully distributed approaches [12]. In such models, the watershed is divided into sub-basins or smaller sub-units called hydrological response units (HRUs). Each HRU is considered as a spatially homogeneous

hydrological entity. Therefore, the delimitation of these spatial units determines both the model complexity and the level of parameter aggregation within the watershed.

The Soil and Water Assessment Tool (SWAT) is a small watershed-to-river basin scale model widely used to assess water resource and quality issues [13,14]. It is a semi-distributed model with a two-level discretization scheme. The first spatial discretization of the watershed creates sub-basins based on topographic watershed limits, and this is followed by further discretization into HRUs using land use, soil characteristics, and slope inclination. Assumed to respond similarly to climate inputs [15], HRUs are the basic computational units in SWAT. Consequently, the accuracy and resolution of the input datasets influence the two levels of discretization of SWAT and, subsequently, the spatial aggregation of the hydrological processes [16].

The official SWAT website [17] provides global soil and land use datasets for constructing SWAT models. The Digital Soil Map of the World (DSMW) from the Food and Agriculture Organization of the United Nations [18] is the suggested soil input dataset. DSMW describes the physical and chemical soil properties at a spatial resolution of 5 km in a format compatible with SWAT GIS interfaces (e.g., QSWAT in QGIS [19]). The proposed land use map is derived from the United States Geological Survey (USGS) Global Land Cover Characterization (GLCC) database at a spatial resolution of 400 m [20]. However, these global maps should only be used when more reliable data are unavailable, as more detailed and local datasets may be better suited to enhance model accuracy [17]. Multiple studies have emphasized the benefits of employing finer-resolution soil [21–24] and land use [25–27] datasets in SWAT for simulating streamflow and water quality. On the contrary, other recent studies have demonstrated that high-spatial-resolution soil and land use datasets do not necessarily improve model accuracy for streamflow simulations [28–31]. Therefore, although it is commonly accepted that the model should be configured with a 'fit for purpose' mindset, for streamflow simulations, there is currently no consensus on how finer-scale soil and land use inputs affect SWAT model accuracy.

Nonetheless, enhancing the level of details in land use, DEM, and soil datasets leads to a greater number of HRUs across the watershed, thereby increasing model complexity [32]; this, in turn, increases the computational time and power required [33]. To reduce the computational time for a more efficient model, a subset of HRUs can be selected. This is achieved either by considering threshold values for soil, land use, and slope or by selecting a specific number of dominant HRUs [14]. However, the selection of a subset of HRUs reduces the spatial heterogeneity in the watershed through input dataset aggregation, so considerable information about the watershed characteristics can be lost [34]. Previous studies reported that the number of sub-basins only slightly affects water yield simulations but has a strong impact on sediment or nutrient loads [35–38]. For instance, the study by Xiao et al. [32] indicated that increasing the aggregation thresholds in SWAT led to almost a 30% increase in annual sediment loads. Similarly, Jha et al. [39] observed that runoff showed less sensitivity to the division of sub-basins, whereas sediment load was more affected by the number of sub-basins. In addition, Alsilibe and Bene [40] evaluated the impact of watershed subdivision and weather input datasets on monthly and yearly streamflow simulations. They found that varying the number of sub-basins did not affect streamflow predictions, but differences in climate datasets significantly impacted the results. However, their study did not explore daily simulations or subdivisions beyond sub-basins. More generally, only a few studies have examined the effects of further subdividing sub-basins into HRUs. Among them, Her et al. [34] investigated how soil, land use, and slope thresholds in HRU definition affect watershed representation, streamflow, and water quality simulations. They suggested minimizing spatial discretization when water quality is not a primary focus since streamflow is less sensitive to thresholds compared to nutrients

or sediments. Similar conclusions were reached by Xiao et al. [32], who analyzed the influence of soil spatial aggregation on SWAT simulations. Although these studies provide insights into the effects of generalizing soil, land use, and slope through HRU aggregation on simulated hydrological processes and water quality, no assessment of its impact on model performance in terms of streamflow simulation and its alignment with observed streamflow has been conducted to date.

Furthermore, hydrological processes in SWAT depend inherently on the spatial scale. While finer spatial discretization captures small-scale variability in land use and soil properties, it introduces spatial heterogeneity and input uncertainties that might not necessarily correspond to improvements in output accuracy. Given that computational time increases with the level of discretization, it is essential to assess whether the increase in processing time leads to an improvement in model performance. Additionally, despite the strong dependence of HRU definition on soil, land use, and DEM inputs, there is a lack of investigation into the relationships between the optimal watershed discretization (number of HRUs) and the characteristics of the soil and land use input datasets in terms of class distribution, hydrological properties, and spatial resolution. Thus, questions such as "At what scale does spatial discretization stop to enhance or decrease accuracy?" and "Does this scale depend on the input data used?" remain unanswered.

To fill this gap, the present study assesses how dividing the watershed into smaller units using different soil and land use data impacts the accuracy of the SWAT model in a large Mediterranean catchment. Specifically, this study has two main objectives: (1) to explore how the number of sub-basins and HRUs affects SWAT model performance in simulating daily streamflow and (2) to determine if this impact is altered by the attribute information and spatial resolution of soil and land use datasets. Intuitively, it is hypothesized that increasing the level of discretization improves model accuracy, particularly when high-resolution and detailed soil and land use maps are employed, as finer spatial representations are expected to better capture watershed heterogeneity. To test this hypothesis, four input sets—representing distinct combinations of low and moderate spatial resolution soil and land use datasets—with a varying number of sub-basins and HRUs, are applied to the Argens river basin in the coastal Mediterranean region. The study results will help modelers to optimize discretization and input datasets in order to gain both model accuracy and computation efficiency.

2. Materials and Methods

This section provides an overview of the study area and the characteristics of input datasets, as well as the fundamentals of the SWAT model. The experimental design used to achieve the study objectives is described in three parts, including the comparison of soil and land use input datasets after translation into SWAT (Section 2.4), the generation of the 68 model configurations (Section 2.5), and the performance evaluation metrics (Section 2.6).

2.1. Study Area

The Argens river basin (2505 km^2) is the second-largest coastal Mediterranean watershed on the southeast coast of France (Figure 1). It is subject to flooding, as demonstrated by the extreme floods of June 2010, November 2011, and November 2019, which caused significant human and material damage on the French Mediterranean coast [41].

The Argens river flows into the Mediterranean Sea at Fréjus. The elevation ranges from 0 m to 1165 m, with an average slope of 9%. The dominant land cover is natural vegetation (69%), followed by agriculture (23%, mainly Vineyards), and urbanized areas (8%) [42]. The geology of the watershed is dominated by limestones in the north and

western parts, whereas the south-east part is characterized by metamorphic schists and gneisses and to a lesser extent by sandstones.

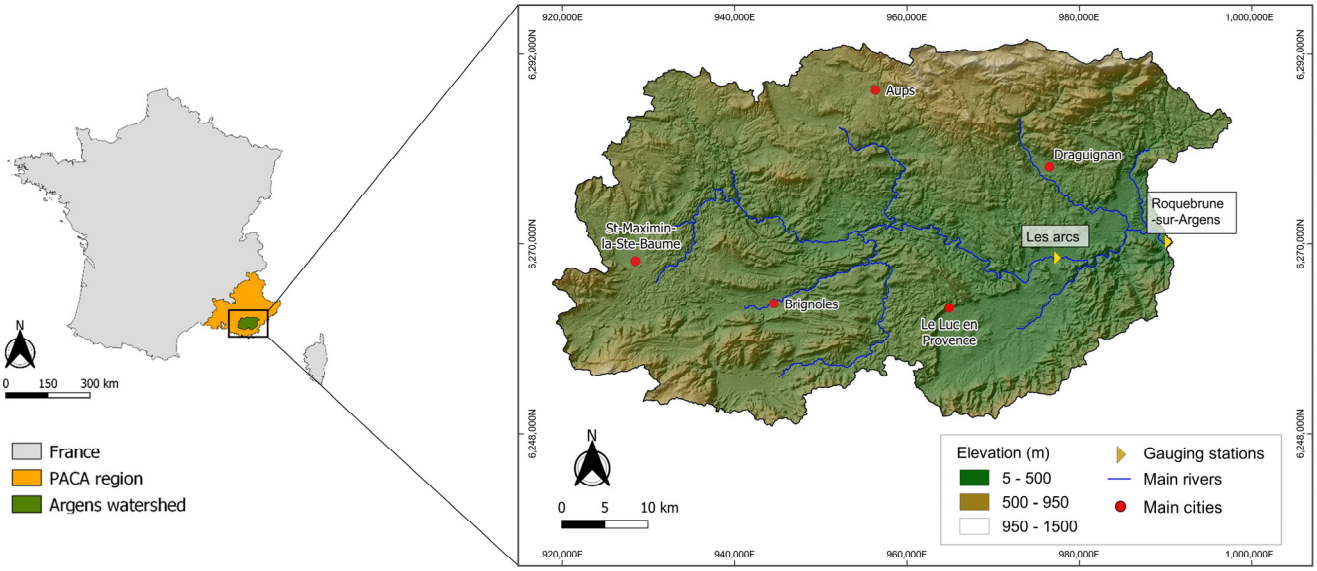


Figure 1. Location of the Argens watershed in France.

The Argens watershed is subject to a typical Mediterranean climate characterized by hot dry summers and cool rainy winters [43]. The annual average daily temperature is about 11 °C in the north and 15 °C on the coast. The annual average precipitation is about 800 mm, falling mainly in autumn and winter, though intense rainfall events can occur in spring [43].

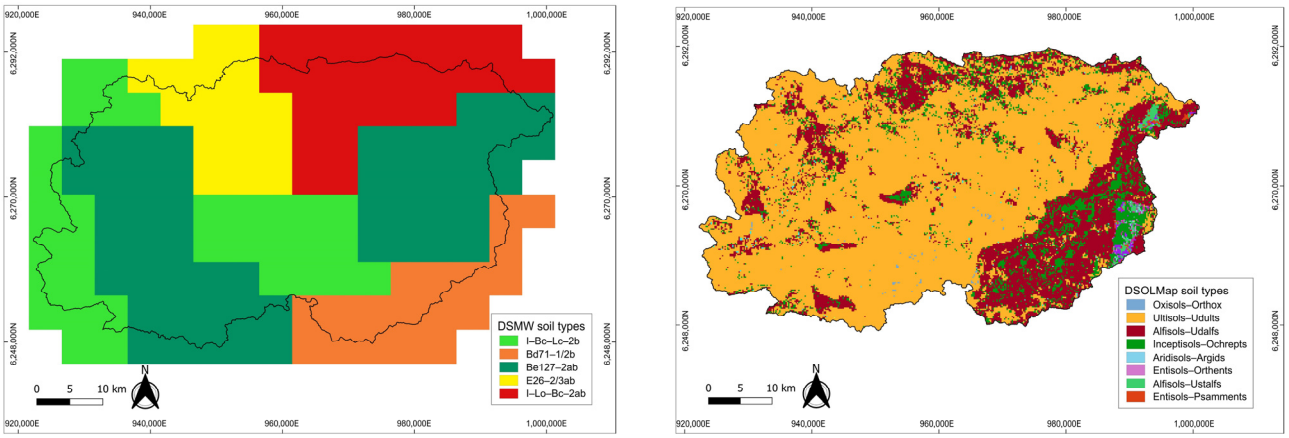
2.2. Input Datasets

Topographic, land use, soil, and meteorological datasets are the four types of data input required to perform SWAT simulations. Calibration and evaluation of the model additionally involves observed streamflow at a gauging station.

The DEM was obtained from the Shuttle Radar Topography Mission (SRTM) [44]. It covers the entire globe with a spatial resolution of 1 arcsecond (approximately 30 m). While other DEMs with higher spatial resolution, such as the "RGE ALTI 5M" from the French National Geographic Institute (IGN) at 5 m, were available for the study region, preliminary tests using the latter did not yield improved model performance in streamflow simulations despite an exponential increase in calculation time. Consequently, the SRTM DEM was retained as the most suitable choice for the current study.

Two soil datasets were used for the purpose of this study (Figure 2). The first is the Digital Soil Map of the World (DSMW) developed by the Food and Agriculture Organization of the United Nations (FAO) with the International Institute for Applied Systems Analysis (IIASA) [18]. It describes the physical and chemical soil properties of two layers (topsoil: 0–30 cm and subsoil: 30–100 cm) for the entire globe with a spatial resolution of 5 km. DSMW is the soil dataset offered to users through the SWAT website. The provided lookup table (matching the soil map to the SWAT soil database) is designed to integrate the DSMW dataset automatically. The second map, the Digital Soil Open Land Map (DSOLMap), was recently developed by López-Ballesteros et al. [22] to address the lack of high-resolution digital soil property maps compatible with SWAT GIS interfaces on a global scale. DSOLMap has a spatial resolution of 250 m with a soil profile divided into six horizons. It is freely available on the website of WateriTech [45] together with a SWAT-compliant database of

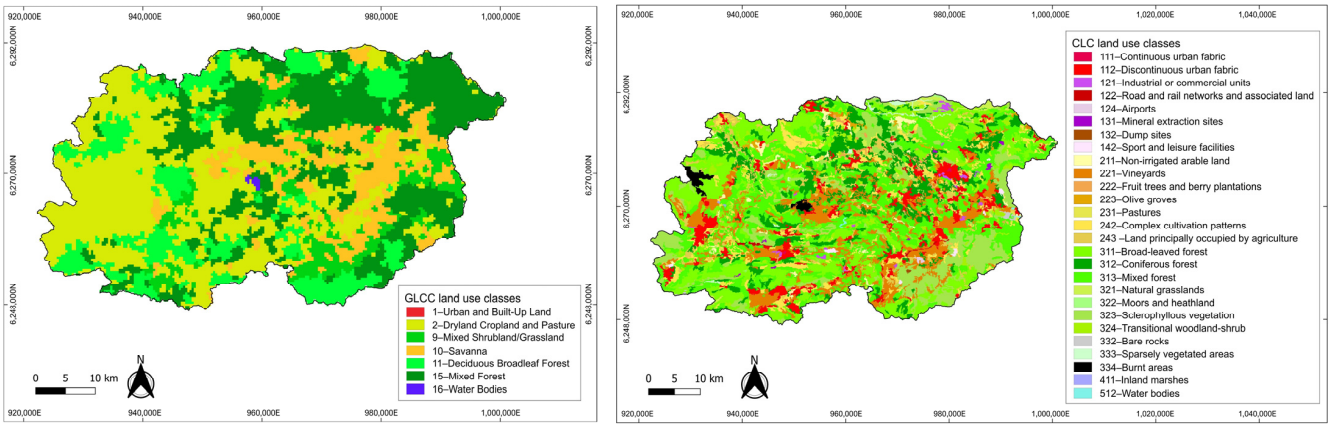
hydrological soil properties and the corresponding lookup table. DSWM is referred to here as the low-resolution soil map and DSOLMap as the moderate-resolution soil map.



(**a**) DSMW (**b**) DSOLMap

Figure 2. Soil datasets: (**a**) low-resolution soil map and (**b**) moderate-resolution soil map.

Two land use datasets were (Figure 3), namely, the USGS Global Land Cover Characterization (GLCC) land use map [20] and the 2018 CLC Land Cover (CLC) map [46]. The GLCC map has a spatial resolution of 360 m at the Argens location and classifies the entire globe according to a scheme of 25 land use/land cover (LULC) classes. It is the land use dataset offered to users through the SWAT website with a ready-to-use lookup table [17]. This makes the GLCC map easy to manipulate, which justifies its widespread use in the SWAT community. The CLC dataset is provided by the Copernicus Land Monitoring Service and offers a European land cover and land use inventory with 44 thematic classes at a spatial resolution of 100 m. This map seems to be better suited to European case studies, but its use inside SWAT is less straightforward since the CLC classification is not directly compatible with SWAT GIS interfaces. For the present study, a lookup table was built to match the CLC categories to their closest SWAT counterpart (Table S1 in the Supplementary Materials). GLCC is referred to here as the low-resolution land use map and CLC as the moderate-resolution land use map.



(**a**) GLCC (**b**) CLC

Figure 3. Land use datasets: (**a**) low-resolution land use map and (**b**) moderate-resolution land use map.

The meteorological data come from the land component of the fifth generation of the European Reanalysis (ERA5-Land) produced by the European Centre for Medium-Range Weather Forecasts [47]. ERA5-Land provides a total of 50 variables describing the water and energy cycles globally and covers the period from January 1950 to the present. Daily precipitation as well as minimum and maximum daily temperatures are available on a regular 10 km grid. SWAT GIS interfaces such as QSWAT treat every cell as a weather station; for each sub-basin, they retain the value of the cell closest to the sub-basin polygon centroid to run the simulations. If the spatial variability in the meteorological conditions over the watershed is significant, the climate data can be aggregated by sub-basin before using them as an input into SWAT. Both approaches (with and without climate data aggregation) were initially tested over the Argens watershed; no clear impact on the streamflow simulations was obtained, and SWAT was subsequently configured without considering an aggregation of the climate data.

Daily observational data of streamflow are available at two gauging stations (Les arcs and Roquebrune-sur-Argens; Figure 1) over the Argens watershed. Values were extracted from HydroPortail [48], the French portal for hydrometric data. The daily time series are available from January 1979 to the present, which is considered long enough for model calibration and validation.

2.3. SWAT Model

SWAT is a semi-distributed process-based hydrological model created in 1998 by the United States Department of Agriculture Agricultural Research Service (USDA ARS) and Texas A&M University [13,49]. It was initially developed to assess the impact of land management practices on water quantity and quality in large complex watersheds with varying soil, land use, and management conditions over long periods [14]. Since then, it has been applied widely around the world for a range of hydrological simulation purposes, including water resource management [50,51], land use and/or climate change impact assessment [52–54], agricultural water pollution modeling [55,56], and flood and drought prediction [57,58].

As a semi-distributed model, SWAT divides a watershed into sub-basins and then sub-divides these into HRUs according to land use, soils, and slope inclination. Based on the user's selection, a subset of HRUs is retained. At least one HRU is defined by sub-basin, and HRUs that are not retained are reapportioned into the other qualified HRUs [34].

SWAT simulates watershed hydrology in two phases [14]. The first phase is the land phase of the hydrologic cycle, which calculates the water balance of each HRU to provide the amount of water available for each sub-basin main channel at a given time step. The second phase is channel routing, which determines the progress of water through the river network towards the basin outlet. The HRU water balance is expressed as:

$$SW_t = SW_0 + \sum_{i=1}^{t} \left(R_{day} - Q_{surf} - E_a - W_{seep} - Q_{gw} \right) \quad (1)$$

where SW_t is the soil water content at time t (mm of water); SW_0 is the initial soil water content (mm of water); R_{day} is the precipitation on day i (mm of water); Q_{surf} is the surface runoff on day i (mm of water); E_a is the evapotranspiration on day i (mm of water); W_{seep} is the percolated water through the soil profile on day i (mm of water); and Q_{gw} is the groundwater flow on day i (mm of water). Although SWAT's water balance is operated on a daily time step, the results can be provided on a daily, monthly, or annual basis.

SWAT calculates surface runoff at the HRU level using the SCS curve number (CN) procedure. Channel routing is estimated using the variable storage routing method, and evapotranspiration is estimated with the Hargreaves method. With such methods, only

daily precipitation and daily minimum and maximum temperature are required as meteorological inputs. For streamflow simulations, 28 parameters describing the hydrological cycle can be calibrated, but this list is usually reduced depending on the study region and its hydrologic characteristics. In the present study, QSWAT was used for preparing input data, while SWAT version 2012 was employed for simulating daily streamflow.

2.4. Soil and Land Use Datasets Translated into SWAT Hydrological Properties

The soil and land use input maps initially contain alphanumeric codes specific to their respective classification systems, making them not directly comparable. To integrate these maps into SWAT and assess their impact on the model, key SWAT properties were extracted and analyzed. These properties include soil texture, hydrologic soil group (HSG), and the moisture condition II curve number (CN2). They were used to compare the datasets before implementing SWAT on the Argens watershed. While the first two properties are derived only from the soil map, CN2 results from the association of both the land use and soil datasets. CN2 is the SCS curve number for the average moisture condition and varies from 0 to 100, where high values correspond to high runoff potential.

Using the user-provided soil lookup table, QSWAT matches soil raster codes in the input map to classes in the soil database. Each class is associated with key properties such as soil texture and HSG. Soil texture (clay, silt, and sand) affects porosity and soil available water content. The percentage values of clay, silt, and sand were extracted from DSMW and DSOLMap. The standard HSGs (A, B, C, and D) indicate runoff potential based on soil permeability, texture, and depth and are fundamental components of the CN method for estimating runoff. To evaluate differences in the input soil maps, the distribution of HSGs across the watershed was compared between the DSMW and DSOLMap datasets. This comparison provided insight into how soil properties influence runoff potential in SWAT.

For land use, QSWAT uses a similar process: from the land use lookup table, it translates the input land use map classes into SWAT-compatible land use categories. This enables SWAT to associate each land use type with hydrological properties, particularly the CN2 values for the four HSGs (CN2A, CN2B, CN2C, and CN2D). To compare predominant land use types in the watershed across the DLCC and CLC datasets, the proportion of each class after translation was calculated.

Since the CN2 value depends on both land use and soil type, the two land use datasets (DLCC and CLC) were combined with the two soil maps (DSMW and DSOLMap). This resulted in four combinations (GLCC-DSMW, GLCC-DSOLMap, CLC-DSMW, and CLC-DSOLMap), which were used to derive the distribution of CN2 values across the watershed. Statistics such as mean, median, and standard deviation were calculated for each CN2 distribution to assess variations in runoff potential based on different input map combinations.

2.5. Model Experiments

In this study, four input sets were created by combining the soil and land use datasets: low–low, moderate–low, low–moderate, and moderate–moderate spatial resolutions of the soil and land use maps (Figure 4). For each input set, 17 configurations were obtained with an evolving number of sub-basins (4, 12, and 18) and HRUs (from 4 to 320). The number of sub-basins was defined based on the area threshold value (ATV): 4 sub-basins with an ATV value of 300 km^2; 12 sub-basins with an ATV value of 200 km^2; and 18 sub-basins with an ATV value of 100 km^2. It is important to note that even if the delineation of HRUs is also dependent on the DEM input, the DEM was kept unchanged for all configurations in this study.

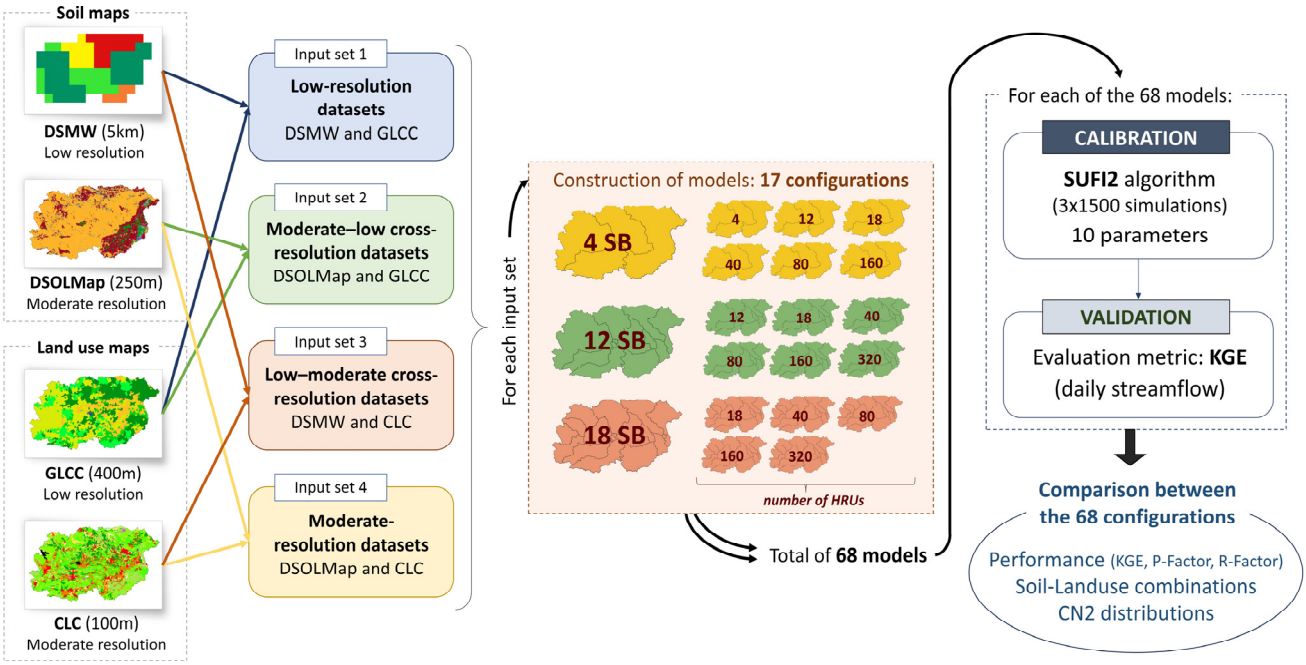


Figure 4. Workflow for model experiments.

The ensemble of 68 models (4 input sets × 17 configurations) was calibrated on the odd years over the 2001–2021 period and validated on the even years of the same period. Simulations were conducted at a daily time step, and a one-year warm-up period was used to dissipate the initialization errors. The calibration process was performed with R-SWAT, an open-source web-based application written in the R language [59]. The SUFI-2 algorithm [60] was used to identify the ten most sensitive parameters for streamflow simulations in SWAT (Table 1) and to adjust their values based on the Kling–Gupta Efficiency (KGE) objective function [61]. Since two gauging stations are considered in the present study (Figure 1), the objective function was computed as the average of the KGE values at each station. Three iterations of 1500 simulations were run to calibrate these parameters.

Table 1. Summary of the ten SWAT parameters used for calibration.

Group	Parameter	Description	Default Value	Range	Change
Infiltration	CN2	SCS curve number	Soil data	[−0.25, 0.25]	Relative
Soil	SOL_AWC	Available water capacity	Soil data	[−0.25, 0.25]	Relative
Evaporation	ESCO	Soil evaporation compensation factor	0.95	[0, 1]	Replace
Groundwater	GWQMN	Threshold water depth in the shallow aquifer for return flow to occur (m)	0	[0, 5000]	Replace
Groundwater	GW_REVAP	Groundwater revap coefficient	0.02	[0.02, 0.2]	Replace
Groundwater	REVAPMN	Threshold water depth in the shallow aquifer for revap or percolation to the deep aquifer to occur (mm)	1	[0, 750]	Replace
Groundwater	ALPHA_BF	Baseflow alpha factor	0.048	[0.01, 1]	Replace
Groundwater	GW_DELAY	Groundwater delay time (days)	31	[0, 500]	Replace
Groundwater	RCHRG_DP	Deep aquifer percolation fraction	0.05	[0.01, 0.99]	Replace
Soil	DEP_IMP	Depth to impervious layer in soil profile	0.5	[−0.25, 0.25]	Relative

2.6. Model Performance Evaluation

The performance of the 68 models was evaluated by comparing daily observed and simulated streamflow using three metrics: KGE, P-factor, and R-factor. These metrics were computed for the two gauging stations on the watershed. KGE decomposes the Nash–Sutcliffe efficiency (NSE) and the mean squared error (MSE) into a three-dimensional criterion (Equation (2)). It has no units and ranges from $-\infty$ to 1, where a value approaching 1 implies a more accurate model.

$$KGE = \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \quad (2)$$

where r is the Pearson correlation coefficient between observed and simulated streamflow; β and γ are, respectively, the bias ($\beta = \mu_s/\mu_o$) and variability ratio ($\gamma = \sigma_s/\sigma_o$) between simulation and observation. μ and γ are the mean and standard deviations of the variable, and s and o are indices meaning simulation and observation, respectively. Following the thresholds established by Moriasi [62] for NSE and the comparison with KGE carried out by Knoben et al. [63], a model simulating daily streamflow is considered as satisfactory for $KGE > 0.3$, good for $KGE > 0.6$, and very good for $KGE > 0.7$.

The P-factor is the fraction of the observed data included in the 95% prediction uncertainty (95PPU) of the model [60]. The 95PPU is calculated based on the cumulative distribution of the output variable at the 2.5% and 97.5% levels, eliminating 5% of the very poor simulations. It varies from 0 to 1, where 1 indicates a perfect model simulation. The R-factor is the ratio of the average width of the 95PPU band and the standard deviation of the observations. It ranges from 0 to $+\infty$, with values closer to 0 corresponding to a thinner band (i.e., a more accurate model).

3. Results

In this section, a comparison of hydrological properties (texture, HSG, and CN2) and land use types derived from the soil and land use input datasets is first presented. The performance of the 68 models is then compared, taking into account their discretization setup (number of sub-basins and HRUs) and the properties resulting from this segmentation (soil–land use pairs and CN2 distribution). Finally, the performance evolution is examined according to the characteristics of the soil and land use maps used in model construction.

3.1. Comparison of Hydrological Properties from Soil and Land Use Datasets

The results related to the distribution of soil textures over the Argens watershed are summarized in Table 2 and reveal that the two soil maps have similar average textural properties. This is particularly true for silt, which differs by only 2%. The difference in the clay content is slightly greater (5%). The greatest difference is found in the sand content, with 7% more sand in the DSOLMap dataset. As sand has a higher permeability, the models based on the DSOLMap dataset are expected to show higher infiltration and lower runoff compared to those built with the DSMW dataset. This emphasizes the importance of selecting soil maps that align with actual hydrological behaviors.

Table 2. Soil texture values in the Argens watershed.

	Mean Clay (%)	Mean Silt (%)	Mean Sand (%)
DSMW	28	34	38
DSOLMap	23	32	45
Difference	5	2	7

As soil texture influences the HSG category, the HSG distributions are shown in Figure 5. While DSOLMap assigns 99% of the watershed to group C, DSMW attributes 56% of the watershed to group D and only 44% to group C. Consequently, a higher CN2 value, and, therefore, a greater runoff potential, is attributed to the initial SWAT configurations based on the DSMW data. These findings suggest that soil datasets can significantly influence hydrological modeling outcomes, particularly in runoff estimations.

On the other hand, Table 3 displays the land use classes assigned by the two land use maps. Overall, mixed forest and deciduous broadleaf forest cover similar proportions of the watershed in both maps, accounting for approximately 25% and 17%, respectively. However, some discrepancies can be seen between the two land use maps. While "Dryland, Cropland, and Pasture" is the dominant land use class in GLCC (35% of the watershed area; Table 3(a)), CLC divides the agricultural class into several sub-classes (e.g., "Vineyard",

"CropLand/Grassland mosaic", "Olives", "Barren/Sparsely vegetated", "Orchard", and "Pasture") representing 25% of the watershed (i.e., 10% less than in GLCC). Finally, the "Savanna" class in GLCC represents 12% of the watershed (Table 3(a)), but this type of land use does not exist in the south of France. It has no equivalent in the CLC map, where it is replaced by two land use types (evergreen forest (11%) and residential areas (7%); Table 3(b)). As the Argens watershed is mainly covered by forest and agricultural areas [42], CLC seems more reliable. These differences underscore the limitations of generalized datasets like GLCC for watershed-scale studies and highlight the improved accuracy of CLC for capturing realistic land use types in the Argens watershed.

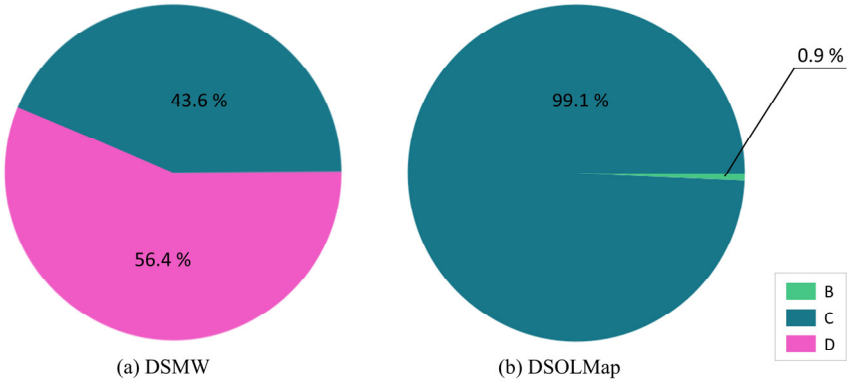


Figure 5. Distribution of the HSGs over the Argens watershed in the (**a**) DSMW and (**b**) DSOLMap datasets. HSGs B, C, and D correspond to moderately low, moderately high, and high runoff potential, respectively.

Table 3. Land use classes and derived CN2 values for HSGs C and D from (a) GLCC and (b) CLC.

(a) From GLCC classes				
Area (%)	**Class code** (*)	**Class Name** (*)	**CN2C**	**CN2D**
35.2	CRDY	Dryland, Cropland, and Pasture	81	85.5
28.0	FOMI	Mixed Forest	73	79
16.7	FODB	Deciduous Broadleaf Forest	77	83
12.2	SAVA	Savanna	76.5	82
7.8	MIGS	Mixed Shrubland/Grassland	76.5	82
Weighted average CN2		———	77	82.5
(b) From CLC classes				
Area (%)	**Class code** (*)	**Class Name** (*)	**CN2C**	**CN2D**
24.4	FRST (FOMI)	Forest—Mixed (Mixed Forest)	73	79
17.2	FODB	Deciduous Broadleaf Forest	77	83
13.9	SHRB	Shrubland	74	80
11.6	GRAP	Vineyard	77	83
11.4	FRSE	Forest—Evergreen	70	77
10.6	CRGR	CropLand/GrassLand mosaic	81	85.5
6.5	URLD	Residential—Low Density	72	79
0.8	OLIV	Olives	77	83
0.8	BARR	Barren	91	94
0.6	GRAS	Grassland	79	84
0.6	UIDU	Industrial	72	79
0.5	BSVG	Baren or Sparsely vegetated	74	80
0.5	ORCD	Orchard	77	83
0.3	UTRN	Transportation	72	79
0.3	PAST	Pasture	79	84
Weighted average CN2		———	75	81

* Codes and names translated from the initial land use map into SWAT classes.

Despite these differences in land use classes, the influence on the CN2C and CN2D values is relatively small, with weighted averages differing by only 2 and 1.5 units, respectively (Table 3). However, the final CN2 value depends on both land use (to obtain the CN2C and CN2D) and soil characteristics (to obtain the hydrological soil group). As DSMW and DSOLMap do not assign the same HSGs to the watershed soils (Figure 5), the attributed

CN2 values differ slightly according to the soil dataset used: the two combinations based on DSMW provide higher mean CN2 values (80 and 79; Figure 6a,c) than those based on DSOLMap (77 and 75; Figure 6b,d). As for land use, the CLC-based combinations show a wider CN2 distribution due to a greater class heterogeneity than in the GLCC dataset (Figure 6c,d and Table 3). While these differences appear small, they reflect little changes in runoff potential. This finding highlights the influence of soil and land use datasets on the definition of hydrological parameters in SWAT, particularly for runoff modeling.

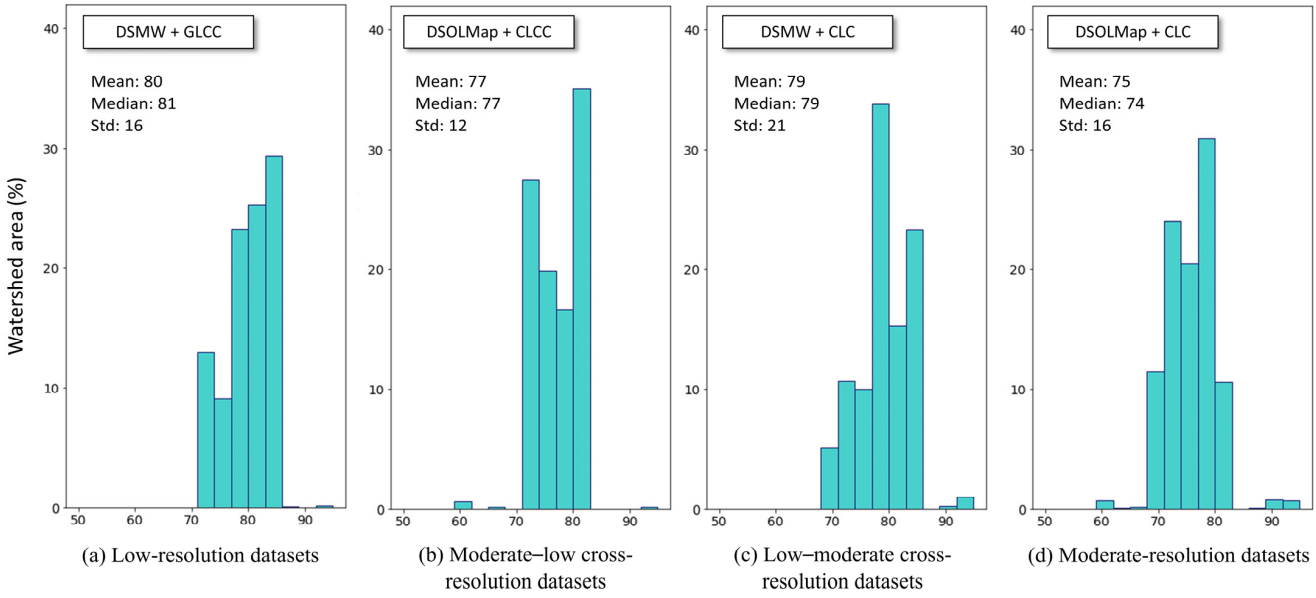


Figure 6. CN2 distribution over the Argens watershed for the four combinations of soil and land use maps: (**a**) low-resolution for both (DSMW and GLCC), (**b**) moderate–low cross-resolution (DSOLMap and GLCC), (**c**) low–moderate cross-resolution (DSMW and CLC), and (**d**) moderate-resolution for both (DSOLMap and CLC).

For the remainder of this study, attention must be paid to the meaning of these distributions (Figure 6), representing the CN2 values resulting from the raw combination of soil and land use maps, without any aggregation or simplification. Subsequently, when the user selects a subset of HRUs for SWAT, both soil and land use data are aggregated, causing CN2 distributions to be modified (Figure S1).

3.2. Model Simulations

This section presents the evolution of model performance as a function of the number of sub-basins and HRUs before focusing on the impact of soil and land use datasets on the optimal spatial discretization.

3.2.1. Performance Evolution with Increasing Discretization

The 68 models perform well during the calibration with KGE values up to 0–0.5 units greater than the validation values. For greater clarity and interpretability, only the validation results are shown in Figure 7. Likewise, as the model's performance was found to be similar at the upstream (Les Arcs station) and downstream (the Roquebrune-sur-Argens station) levels, only the metrics corresponding to the downstream station are presented.

For each input set, at least 4 configurations out of 17 exceed the very good performance threshold during the validation, i.e., KGE $\geq$ 0.7 (4 configurations for the low-resolution and low–moderate cross-resolution input sets, 5 configurations for the moderate–low cross-resolution input set, and 8 configurations for the moderate-resolution input set; Figure 7). The P-factor and R-factor values are similar across the 68 models (not shown). The P-factor

remains in the range of 0.80 to 0.91, regardless of the input dataset and the watershed discretization. Similarly, the R-factor shows values from 0.71 to 0.86, without a clear relationship between its variations and the input set or the watershed configuration. For this reason, this study focuses on the performance evaluation in terms of KGE.

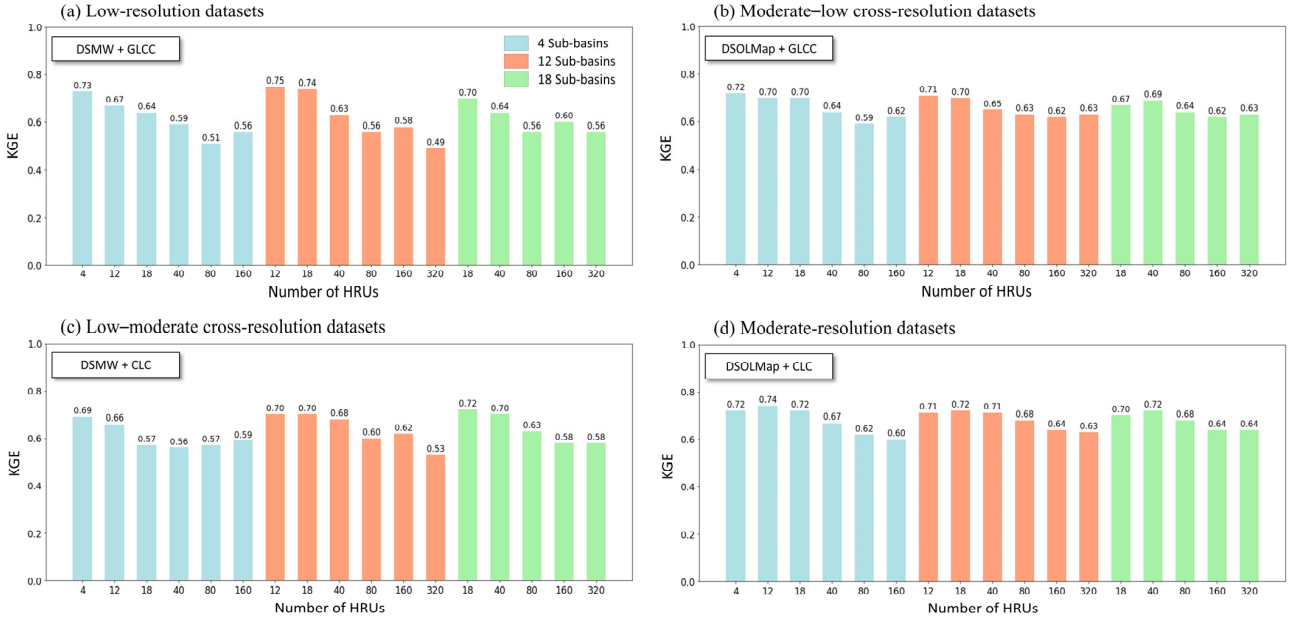


Figure 7. Performance of the model over the validation period according to the number of sub-basins and HRUs: (**a**) low-resolution input set (DSMW and GLCC), (**b**) moderate–low cross-resolution input set (DSOLMap and GLCC), (**c**) low–moderate cross-resolution input set (DSMW and CLC), and (**d**) moderate-resolution input set (DSOLMap and CLC). The KGE value is indicated above the histograms.

The performance of the model follows a similar trend for 4, 12, and 18 sub-basins within each input set (Figure 7). Hence, the number of sub-basins has no impact on model performance. However, the KGE values decrease with an increasing number of HRUs for the four input sets regardless of the number of sub-basins. For instance, KGE decreases from 0.75 to 0.49 for the models using the low-resolution datasets (Figure 7a) and from 0.74 to 0.60 for those based on the moderate-resolution datasets (Figure 7d). As for the cross-resolution input sets, the minimum values of KGE are 0.59 and 0.53 for the moderate–low and low–moderate sets, respectively (Figure 7b,c).

To further understand the performance differences among the configurations, an examination of the associated HRUs was conducted. As the present study focuses on the influence of the soil and land use inputs on SWAT performance, the soil–land use pairs (soil type and land use type) defining each HRU were analyzed without considering slopes. For each of the 68 models, the percentage of watershed area allocated to each soil–land use pair was compared (Figure 8 and Table S2). As the number of HRUs increases, new pairs of soil and land use types are introduced. However, after a certain threshold of HRUs, few new soil–land use pairs are added, and they represent a very small portion of the catchment ($\leq$1%). So, the overall distribution of the soil–land use pairs does not evolve significantly afterward. The HRU threshold depends on the input set and the number of sub-basins. With the low-resolution input set (Figure 8a), the HRU threshold is 80 whatever the number of sub-basins. For the DSOLMap-based models, the HRU threshold ranges between 80 and 160 according to the number of sub-basins (Figure 8b,d). Compared to the other input sets, no HRU threshold is clearly identified for the moderate–low cross-resolution input sets using DSMW and CLC (Figure 8c), probably because this threshold

is above the maximum number of HRUs considered in the present study. As the HRUs are defined by the combination of soil, land use, and slope, the lack of evolution in the soil–land use pairs with an increase in the number of HRUs indicates that only the slope varies. This finding reveals that between 160 and 320 HRUs, the number of units doubles, but no additional information from the soil and land use datasets is added. Nonetheless, this finding alone does not explain the lower performance of the model as the number of HRUs increases.

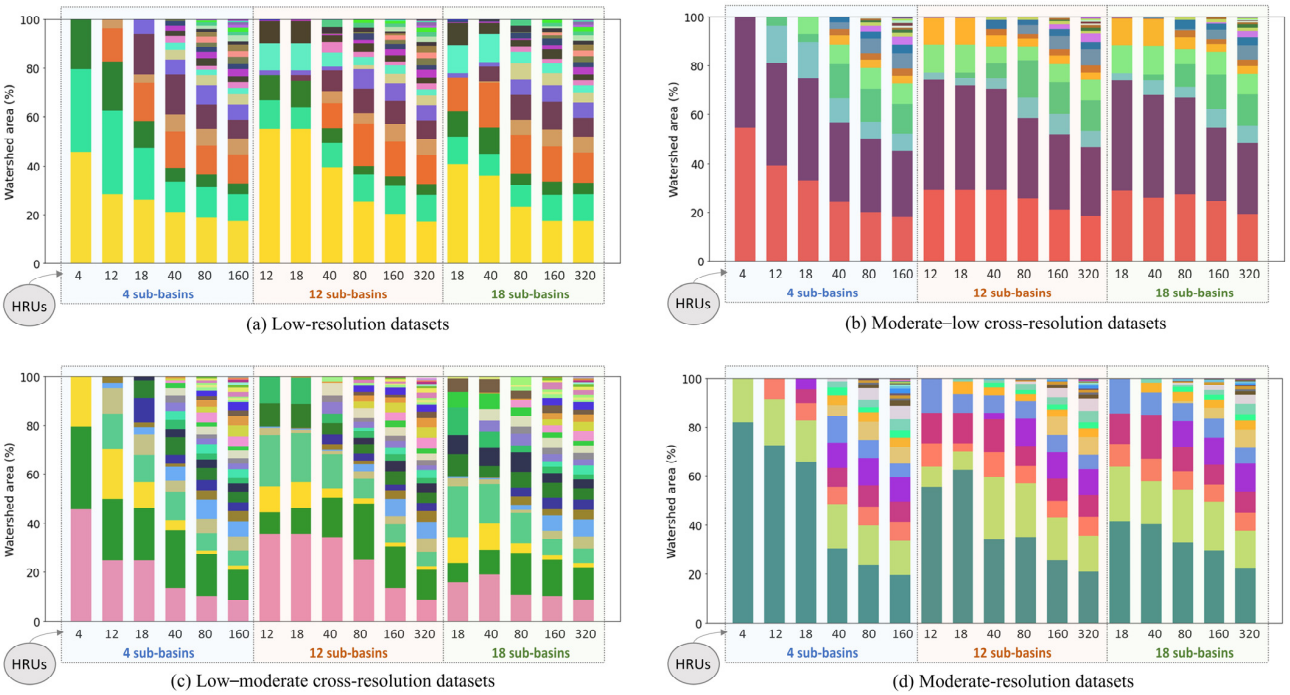


Figure 8. Area of soil–land use pairs depending on the number of sub-basins and HRUs for each input set: (**a**) low-resolution (DSMW and GLCC), (**b**) moderate–low cross-resolution (DSOLMap and GLCC), (**c**) low–moderate cross-resolution (DSMW and CLC), and (**d**) moderate-resolution (DSOLMap and CLC). Each color corresponds to a single soil–land use pair. Since each input set is different between (**a**), (**b**), (**c**), and (**d**), all soil–land use pairs are different from one input set to another.

CN2 is derived from the soil–land use pairs. While land use provides the CN2A, CN2B, CN2C, and CN2D values for each spatial unit, soil type establishes the HSG and thus determines the final CN2 value of each HRU. The different watershed discretizations modify the distributions of soil and land use types, thus altering the spatial aggregation of the CN2 parameter. For the 68 model configurations, the CN2 distribution across the watershed was compared to evaluate the influence of watershed discretization on the CN2 values. To ease the analysis and interpretation, only the CN2 distributions for the minimum and maximum number of HRUs per sub-basin configuration are presented (Figure 9). The full panel of CN2 distributions is available in the supplementary information (Figure S1). As expected, the greater the number of HRUs, the more the CN2 distributions are close to the case where all HRUs are conserved (Figure 6). In addition, the greater the number of HRUs, the broader the distribution of CN2 values for all input sets. This is particularly obvious for the configurations with four sub-basins (Figure 9). This issue will be discussed further in Section 4.1.

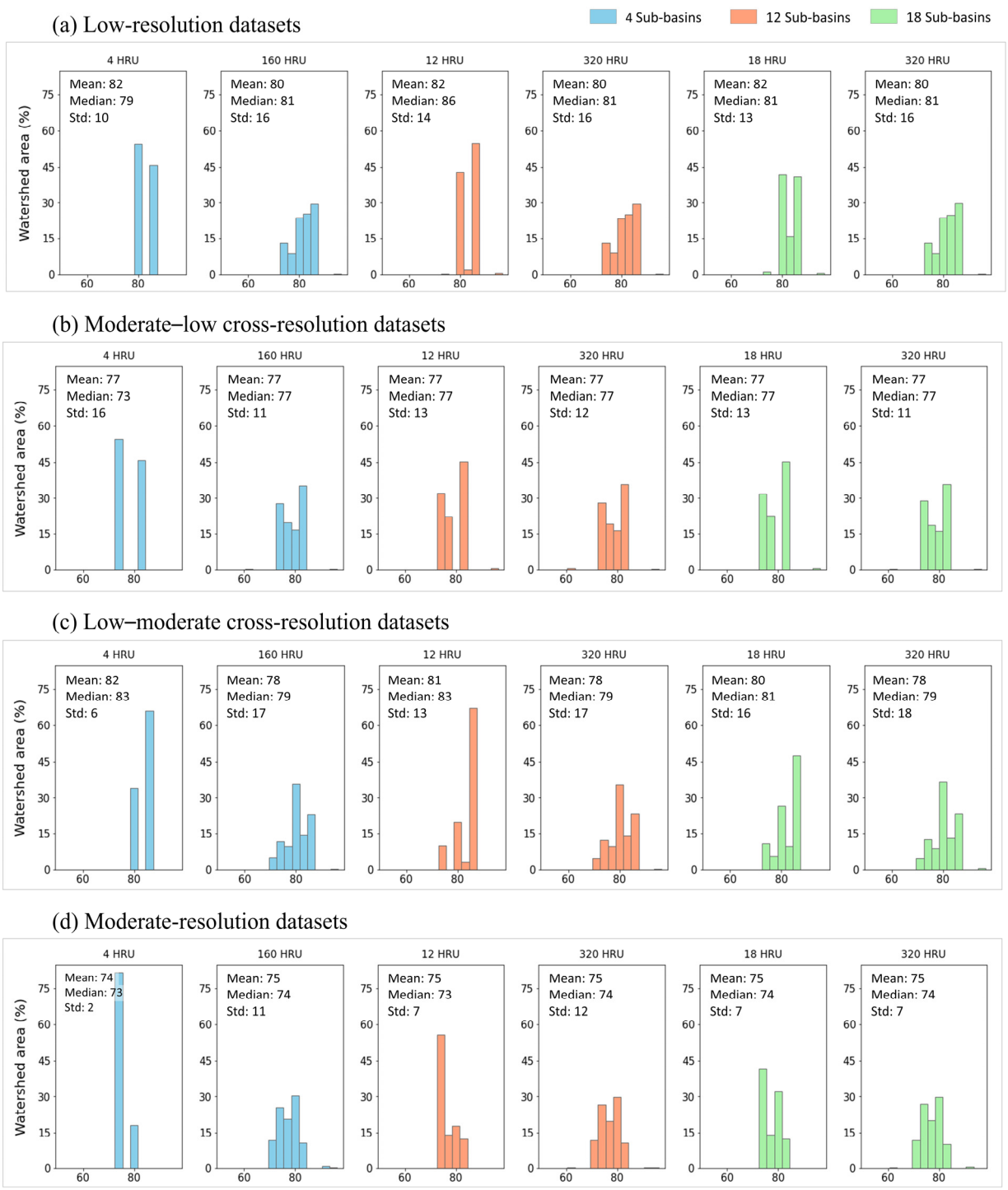


Figure 9. (**a**–**d**) CN2 distribution by input set over the Argens watershed according to the number of sub-basins and HRUs. The full CN2 distribution is presented in the supplemental information (Figure S1). The mean, median, and standard deviation (std) are annotated on each histogram.

3.2.2. Performance Evolution with Different Soil and Land Use Input Datasets

The results described above show a decrease in model performance as the number of HRUs increases. For the configurations with the minimum number of HRUs per sub-basin (i.e., one HRU per sub-basin), the performance is similar between the four input schemes, with a KGE difference of only 0.05 (Figure 7). However, greater discrepancies between the input schemes are found when the discretization is higher than 40 HRUs. These differences tend to increase with the number of HRUs. Consequently, the influence of the soil and land

use datasets on model performance is greater when the watershed is divided into a larger number of spatial units.

The soil map has the greatest impact on the performance evolution of the models. The KGE values for the models based on DSOLMap remain satisfactory whatever the watershed discretization (KGE > 0.59; Figures 7 and 7d). On the contrary, models using DSMW as input exhibit a greater decrease in KGE (Figure 7a,c). Overall, the models built from input sets with DSOLMap perform better than those with DSMW in simulating daily streamflow, especially in the configurations with many HRUs (i.e., more than 160 HRUs in the present study).

The influence of the land use dataset on model performance is more contrasted. For configurations using the moderate-resolution soil map (DSOLMap), the CLC-based models outperform those based on GLCC, achieving KGE differences up to 0.06 for the same discretization scheme (i.e., the number of sub-basins and HRUs, Figure 7b,d). However, the models built from input sets using the low-resolution soil map (DSMW) reveal no clear advantage for CLC over GLCC, and the DSMW-based configurations even perform slightly better for some discretization schemes (Figure 7a,c). These results suggest that the benefits depend on the soil dataset and watershed discretization, rather than on the land use map.

When comparing pairs of soil and land use types, a greater heterogeneity is observed for models with the DSMW-based input sets (Figure 8a,c). Indeed, models with the low-resolution and low–moderate cross-resolution input sets have, respectively, 13 and 19 pairs of soil and land use types that cover more than 4% of the watershed area (Figure S1 in the Supplementary Materials). The models with the DSOLMap-based input sets give only 9 and 10 of such pairs. This is because only two main soil types are attributed to the Argens watershed with DSOLMap, while DSMW assigns five soil types (Figure 2). While this diversity offers greater detail, it may also introduce complexity that undermines model performance, particularly for a high number of HRUs.

4. Discussion

This study provides insights regarding the relationships between watershed spatial discretization and soil–land use input datasets in a Mediterranean watershed. In particular, the aim is to assess the evolution of model performance with increasing watershed discretization (Section 4.1) and to examine how soil and land use input datasets influence this evolution (Section 4.2). In addition, given the impact of discretization on computational time, a third section is dedicated to the implications of opting for a high spatial discretization of the watershed in distributed hydrological modeling (Section 4.3).

4.1. Impact of Increasing Discretization on Model Performance

This study confirmed that discretizing the watershed into a greater number of topographic sub-basins does not improve the performance of the SWAT model. This finding corresponds to the results of Bingner et al. [36], who investigated the effect of the number and size of sub-basins in SWAT for a watershed in northern Mississippi and concluded that runoff volume is not highly affected by the changing discretization. It also agrees with the results of Barbarosa et al. [64], who found no variation in the performance of the semi-distributed HEC-HMS model applied to the Brazilian Uberaba River with different levels of watershed spatial discretization (2 to 32 sub-basins).

On the other hand, and unexpectedly, the present study demonstrated that the SWAT model performance decreases as the number of HRUs increases, contradicting the initial hypothesis proposed in the Introduction. The hypothesis suggested that greater discretization would enhance the model's accuracy, especially with higher-resolution soil and land use maps. These findings, however, extend the findings of Githui and Thayalakumaran [65],

who investigated the impact of aggregating land use and soil type (through the number of HRUs) on both catchment-scale and within-catchment outputs. While they also observed a drop in SWAT model performance correlated to the increase in HRU number for the simulation of monthly streamflow (i.e., catchment-scale output), they inversely noted an improvement in within-catchment water flux results. These results suggest that, depending on the objective of the model, the addition of HRUs may be beneficial (e.g., to study processes within the watershed) or detrimental (e.g., to study watershed output streamflow, with a loss of accuracy, an increase in calculation time and an increase in the complexity of the model). However, the conclusions of Githui and Thayalakumaran [65] are only valid for the monthly time step, and the present study therefore allows these results to be extended to daily catchment-scale outputs. In addition, Vema and Sudheer [33] sought to minimize computational demands by assessing the SWAT model's performance under various discretization and parametrization methods. While maintaining a very good NSE value, they noted a decline in model performance with increasing discretization. Gong et al. [66] evaluated the effect of various delineation schemes on SWAT performance by incrementally adjusting the threshold drainage area. In agreement with the present study, they showed that the simulation under the finest delineation scheme did not yield the highest model accuracy. However, in their approach, they simultaneously increased the number of sub-basins and HRUs, which prevented the assessment of the HRU impact on model simulations independently of the number of sub-basins. Hence, the first added value of the present study is that it separates the impact of HRU discretization from sub-basin discretization. By focusing on HRU effects alone, it appeared that increasing the number of HRUs reduces model performance for daily streamflow simulations, while modifying the number of sub-basins had no significant effect. This clarifies how discretization can impact the accuracy of the SWAT model, but it does not explain the underlying reasons why.

From the analysis of the soil–land use pairs, no clear explanation of the model performance decrease was found (Section 3.2.1). However, the CN2 distributions resulting from these pairs highlighted an interesting point: the higher the number of HRUs, the larger the spread in the CN2 distribution whatever the input set (Figure 9). CN2 is a relevant parameter in SWAT that needs to be calibrated. It is adjusted in a relative way, meaning that the CN2 value of each HRU decreases or increases by a certain percentage from the initial value. However, SWAT calibration is not performed at the HRU level since the adjustment in the CN2 value is applied uniformly to all HRUs. For example, if the relative change in the CN2 value is calibrated to +5%, a 5% increase in CN2 is applied to all HRUs even though the optimum CN2 value for a specific HRU may differ (e.g., −3%). The consequence of this uniform adjustment is that the calibration process seeks a relative CN2 value that would be optimal for all HRUs simultaneously (the best compromise) without achieving the optimum CN2 value of each individual HRU. Hence, as the number of HRUs increases, it becomes unlikely that the optimal CN2 value is allocated to each individual HRU during the calibration process. Zhou et al. [67] assessed the effects of both the watershed partitioning level and the choice of optimization algorithm on the hydrological simulation uncertainty with the semi-distributed TOPMODEL model. They found that the primary factor influencing uncertainty in runoff simulation was the optimization algorithm, especially for the flood periods. Therefore, the observed decline in model performance with an increasing number of HRUs could be attributed to the calibration process rather than to the discretization itself. The calibration process used in this study is the most commonly used for SWAT (e.g., [28,68–70]), and few authors opt for an HRU-level calibration [71,72] because this approach is highly time-consuming and difficult to automate. Testing it here was beyond the scope of this study, but it remains a fundamental research issue for SWAT users and for semi-distributed modeling in general.

4.2. Impact of the Soil and Land Use Datasets on the Evolving Performance

The second significant contribution of this study lies in testing the combined effect of input datasets and discretization on SWAT model performance. Regardless of the chosen input data, KGE performance declines as the number of HRUs increases (Figure 7). However, the rate of decline varies depending on the input datasets used. For the four input sets, the models have similar KGE values in the configurations with one HRU per sub-basin, but the performance of the models differs as the number of HRUs increases. The influence of the soil and land use input maps on model performance is therefore greater when the watershed is divided into a larger number of units (Figure 7). Several studies have shown that the use of different datasets for the same region in SWAT led to contrasted streamflow and water quality simulations [73,74]. However, contrary to the present study, the relationships between the soil–land use input datasets and the spatial discretization have not been investigated in any of the previous studies.

The results show an influence of the soil datasets on the model performance through a greater decrease in KGE for the DSMW-based models compared to the DSOLMap-based models (Figure 7). There are three major differences between the two soil maps: the spatial resolution, the heterogeneity in soil types, and the hydrologic soil groups attributed to the Argens watershed. DSOLMap has a spatial resolution twenty times finer than DSMW (250 m to 5 km, respectively). However, despite its lower spatial resolution, DSMW attributes a greater diversity of soils to the Argens watershed than DSOLMap. It results in a higher heterogeneity in the soil–land use pairs for models using the DSMW map (Figure 8). Yet, the DSOLMap-based models perform better at simulating daily streamflow, especially in the configuration with many HRUs (Section 3.2.2). This observation could be linked to the study by Busico et al. [75], who evaluated the SWAT model performance in simulating monthly streamflow using three different soil maps with varying resolutions and soil information. Their results highlighted that performance worsens as more soil types are included in the calibration process. This suggests that the higher performance of the DSOLMap-based models in the present study is not due to the higher spatial resolution of the soil map but rather to its lower soil type heterogeneity across the watershed. Moreover, the impact of different HSGs was assessed through CN2 distributions (Figure 9), revealing no significant differences between models using either soil dataset. This suggests that HSGs have no influence on the SWAT model parametrization, especially since CN2 is a calibrated parameter. Thus, the variations in model performance between models based on DSOLMap and on DSMW are not attributed to the HSG differences but likely result from the lower heterogeneity in the DSOLMap soil types. The global calibration process, which implies correcting the entire watershed by the same relative or absolute value, may yield better results when the input data are more homogeneous.

On the other hand, a low impact of the land use datasets on the model performance was shown. The land use maps used in this study differ according to their spatial resolution and the types of land use attributed to the Argens watershed. The improvement in the spatial resolution between the two land cover maps is relatively weak (from 400 m for GLCC to 100 m for CLC) compared to that of the soil datasets (5 km to 250 m). But, unlike soil maps, the increase in the spatial resolution of the land cover datasets is accompanied by an increase in the number and types of classes (from 6 to 16 Table 3). Moreover, the analysis of land use maps highlighted some inconsistencies in the GLCC dataset, such as the substantial part of the watershed area attributed to the Savanna class (12%, Table 3a), a land cover type that is not present in the study watershed. However, such differences in land use classes did not significantly affect the CN2C and CN2D values attributed to each spatial unit and, therefore, did not impact the accuracy of the streamflow simulations. Hence, the modification of the provided global land use map in SWAT (i.e., DLCC) into a more local

map at a finer spatial resolution did not lead to an improvement in SWAT performance in simulating daily streamflow. The low sensitivity of the SWAT model to the resolution of land use datasets was also reported in other studies [28,30,76]. In addition, Wu et al. [77] demonstrated that while high-resolution DEM data can reduce model uncertainty, land use resolution has less impact on parameter sensitivity, with only slight changes in the ranking sensitivity of parameters. These findings underscore the small impact of land use resolution on parameter sensitivity and model performance and suggest that efforts to improve SWAT simulations may be better focused on other sources of uncertainty, such as other input data and calibration strategies.

4.3. Optimal Discretization, Model Complexity, and Computation Time

In the SWAT model, spatial discretization and spatial complexity are interconnected, with the number of HRUs being a determining factor of computational efficiency and simulation accuracy. Section 4.1 highlighted that while an increase in the number of HRUs can improve the representation of watershed heterogeneity, it also increases spatial complexity and the number of parameters to calibrate. As explained by Zhou et al. [67], this can lead to over-parameterization, amplified uncertainties in input data, and significant challenges in calibration. This finding is further supported by Section 4.2, where the increase in soil type heterogeneity was found to be a factor of performance decline. In addition, as depicted in Figure 10, increasing spatial discretization greatly extends computational time without improving model performance. Instead, a rise in spatial complexity is correlated with poorer performance. This decline is contrary to initial expectations but might be due to the added noise and redundancy introduced by the larger number of HRUs, which do not contribute to improvements in simulating water processes. According to Al-Khafaji et al. [28], a higher number of HRUs introduces more hydrologic parameters calibrated against a single observed variable—streamflow—thereby amplifying uncertainty. Adding more variables to the calibration process might improve the performance of the SWAT model when the number of HRUs is high.

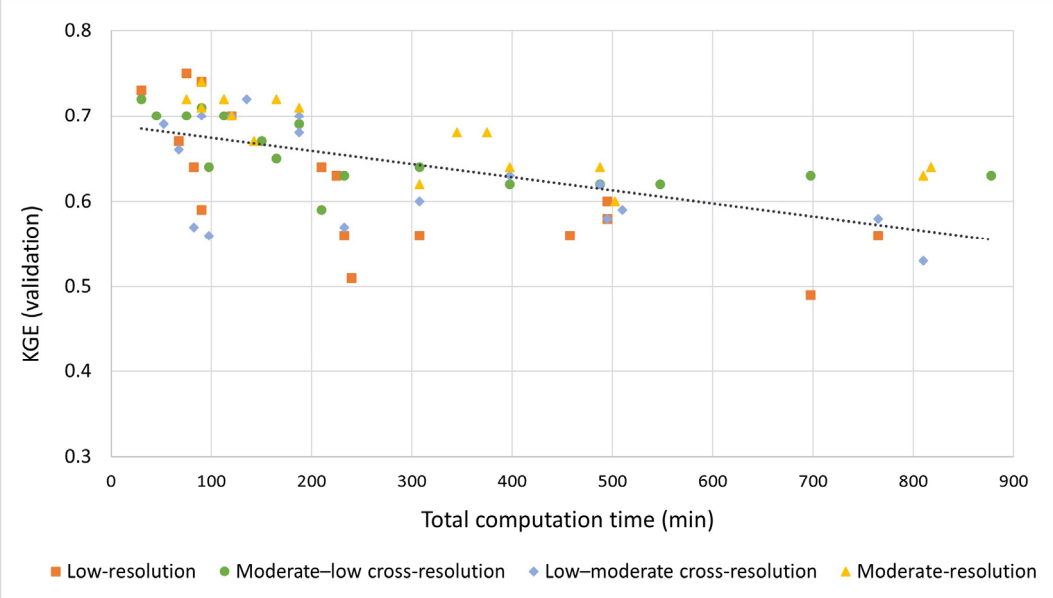


Figure 10. KGE performance as a function of computational time. Each symbol corresponds to an input set, and the dotted line corresponds to the general line trend.

Thus, the model performance decrease likely comes from increased uncertainty and calibration challenges at a higher number of HRUs. In addition to the model discretization,

the choice of the calibration algorithm itself influences the computational times. In particular, SUFI-2 was shown to be less efficient in terms of speed compared to the dynamically dimensioned search (DDS) algorithm [78,79]. Despite its slower performance, SUFI-2 was selected in this study for its greater ability to deal with uncertainty from different sources, including model input, model structure, parameters, and observed data [80]. In particular, Wu et al. [77] demonstrated better applicability and lower uncertainty resulting from the SUFI-2 algorithm using the SWAT model. Further investigations are required to determine if the impact of the number of HRUs on model performance remains the same when the SUFI-2 algorithm is configured to operate at the HRU level.

Finally, while minimizing HRUs improves computational efficiency and the accuracy of daily streamflow simulations, applications requiring spatially detailed outputs, such as conservation planning or analysis of land use change in specific areas, may still benefit from a higher sub-basin number associated with a limited number of HRUs. This approach ensures that spatial complexity is optimally managed, providing necessary detail without exploding computational resources.

4.4. Limitations and Future Work

This work highlighted a decline in SWAT model performance with a greater number of HRUs, a trend that could not be attributed to input data accuracy, as it was observed regardless of the input datasets used. Sections 4.1 and 4.3 suggested the calibration process as a potential cause for this decline, particularly due to calibration being conducted at the basin level rather than the HRU level, but also because SUFI-2 might be of poorer performance than other methods such as DDS. Indeed, calibration methods in SWAT may interact with spatial discretization in ways that influence model performance [81]. Future research should explore the performance of various calibration algorithms under different spatial discretization scenarios to extend the findings of this study.

Furthermore, while this study considered all types of daily streamflow, future research could focus more on extremes (i.e., floods and droughts) and particularly assess how spatial discretization affects the simulation accuracy of extreme hydrological events, such as high-water peaks and low flows. It would also be valuable to investigate whether the spatial heterogeneity in soil and land use plays a more critical role during flood events when runoff processes are dominant.

Finally, this work focused on the Argens, a specific Mediterranean watershed. Expanding this study to include other types of watersheds would help to ensure the observed trends are universally applicable.

5. Conclusions

The implementation of the semi-distributed SWAT hydrological model involves a division of the watershed into sub-basins and HRUs to capture the spatial variability in watershed characteristics, such as topography, soil characteristics, land use, and slope. However, increasing discretization leads to a rise in the number of model parameters and results in greater computation times. Nonetheless, since HRUs rely on soil, land use, and topography inputs, any changes in these datasets modify the characteristics of the HRUs (size and properties), which can impact SWAT performance in simulating daily streamflow.

Focusing on the Argens Mediterranean watershed, the main findings of this study are as follows:

1. The performance of SWAT in simulating streamflow remains unaffected by the variation in the number of sub-basins. However, an increase in the number of HRUs results in a decrease in model performance, whatever the number of sub-basins or the input datasets.

2. A higher sensitivity of the SWAT model to variations in the soil input is identified compared to changes in the land use dataset, with a faster decline in KGE performance with increasing discretization for the models based on the provided DSMW soil map than on the new DSOLMap. This sensitivity is attributed to the heterogeneity in soil types rather than to the spatial resolution of the inputs.
3. Increasing discretization expands computation time and reduces model performance. It is therefore recommended to minimize the number of HRUs during watershed subdivision for optimal model accuracy of catchment-scale outputs.

Future research could be dedicated to exploring different calibration strategies, such as automated and efficient approaches at the HRU level, to further address the observed performance decrease in the SWAT model with an increasing number of HRUs. Incorporating additional variables into the calibration process, such as soil moisture or evapotranspiration, is another area worth exploring.

The results have important implications for an accurate and fast simulation of daily streamflow in Mediterranean watersheds subjected to extreme and rapid floods. Given the critical importance of computation time for fast streamflow simulations, the unexpected finding concerning the low benefit of a high spatial discretization for SWAT streamflow simulation should be considered in future operational model developments in the Mediterranean region. The results, however, are also of broader interest to any practitioner using semi-distributed hydrological models.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/w17020239/s1: Table S1: Lookup table used for the translation of CLC values into SWAT land use classes; Table S2: Area of soil–land use pairs for each input set (for readability, only combinations that cover more than 4% of the watershed area in at least one configuration were kept); Figure S1: CN2 distribution over the watershed depending on the number of sub-basins and HRUs for each input set (full).

Author Contributions: Conceptualization, M.P., M.T., D.F. and P.R.-G.; methodology, M.P., M.T. and D.F; software, M.P.; validation, M.T. and D.F.; formal analysis, M.P.; investigation, M.P.; resources, M.T.; data curation, M.P.; writing—original draft preparation, M.P.; writing—review and editing, M.P., M.T., D.F. and P.R.-G.; visualization, M.P. and P.R.-G.; supervision, M.T. and D.F.; project administration, M.T. and D.F.; funding acquisition, M.T. and D.F. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partly funded by the SUD-PACA region as part of the "Emplois Jeunes Doctorants" program and by Hydroclimat. The authors are also grateful for partial funding from the national "France Relance" recovery plan of the CNRS through its financing of the POP-RISK project.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Rajib, A.; Liu, Z.; Merwade, V.; Tavakoly, A.A.; Follum, M.L. Towards a Large-Scale Locally Relevant Flood Inundation Modeling Framework Using SWAT and LISFLOOD-FP. *J. Hydrol.* **2020**, *581*, 124406. [CrossRef]
2. Kauffeldt, A.; Wetterhall, F.; Pappenberger, F.; Salamon, P.; Thielen, J. Technical Review of Large-Scale Hydrological Models for Implementation in Operational Flood Forecasting Schemes on Continental Level. *Environ. Model. Softw.* **2016**, *75*, 68–76. [CrossRef]
3. Dash, S.S.; Sahoo, B.; Raghuwanshi, N.S. Improved Drought Monitoring in Teleconnection to the Climatic Escalations: A Hydrological Modeling Based Approach. *Sci. Total Environ.* **2023**, *857*, 159545. [CrossRef] [PubMed]
4. Tran, T.-N.-D.; Tapas, M.R.; Do, S.K.; Etheridge, R.; Lakshmi, V. Investigating the Impacts of Climate Change on Hydroclimatic Extremes in the Tar-Pamlico River Basin, North Carolina. *J. Environ. Manag.* **2024**, *363*, 121375. [CrossRef]

5. Hagemann, S.; Chen, C.; Clark, D.B.; Folwell, S.; Gosling, S.N.; Haddeland, I.; Hanasaki, N.; Heinke, J.; Ludwig, F.; Voss, F.; et al. Climate Change Impact on Available Water Resources Obtained Using Multiple Global Climate and Hydrology Models. *Earth Syst. Dyn.* **2013**, *4*, 129–144. [CrossRef]
6. Tran, T.-N.-D.; Nguyen, B.Q.; Grodzka-Łukaszewska, M.; Sinicyn, G.; Lakshmi, V. The Role of Reservoirs under the Impacts of Climate Change on the Srepok River Basin, Central Highlands of Vietnam. *Front. Environ. Sci.* **2023**, *11*. [CrossRef]
7. Jajarmizadeh, M.; Harun, S.; Salarpour, M. A Review on Theoretical Consideration and Types of Models in Hydrology. *J. Environ. Sci. Technol.* **2012**, *5*, 249–261. [CrossRef]
8. Boyle, D.P.; Gupta, H.V.; Sorooshian, S. Toward Improved Calibration of Hydrologic Models: Combining the Strengths of Manual and Automatic Methods. *Water Resour. Res.* **2000**, *36*, 3663–3674. [CrossRef]
9. Carpenter, T.M.; Georgakakos, K.P. Intercomparison of Lumped versus Distributed Hydrologic Model Ensemble Simulations on Operational Forecast Scales. *J. Hydrol.* **2006**, *329*, 174–185. [CrossRef]
10. Refsgaard, J.C.; Knudsen, J. Operational Validation and Intercomparison of Different Types of Hydrological Models. *Water Resour. Res.* **1996**, *32*, 2189–2202. [CrossRef]
11. Vilaseca, F.; Narbondo, S.; Chreties, C.; Castro, A.; Gorgoglione, A. A Comparison between Lumped and Distributed Hydrological Models for Daily Rainfall-Runoff Simulation. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *958*, 012016. [CrossRef]
12. Savvidou, E.; Efstratiadis, A.; Koussis, A.D.; Koukouvinos, A.; Skarlatos, D. The Curve Number Concept as a Driver for Delineating Hydrological Response Units. *Water* **2018**, *10*, 194. [CrossRef]
13. Arnold, J.G.; Srinivasan, R.; Muttiah, R.S.; Williams, J.R. Large Area Hydrologic Modeling and Assessment Part I: Model Development. *J. Am. Water Resour. Assoc.* **1998**, *34*, 73–89. [CrossRef]
14. Neitsch, S.L.; Arnold, J.G.; Kiniry, J.R.; Williams, J.R. *SWAT Theoretical Documentation—Version 2009*; SWAT: College Station, TX, USA, 2011.
15. Li, E.A.; Shanholtz, V.O.; Contractor, D.N.; Carr, J.C. Generating Rainfall Excess Based on Readily Determinable Soil and Landuse Characteristics. *Trans. ASABE.* **1977**, *20*, 1070–1078. [CrossRef]
16. Zhong, K.; Chen, X.; Chen, Y.; Liu, M. Simulation of effects of topography and soil/land use spatial aggregation on sediment yield and runoff using AnnAGNPS. *Trans. CSAE.* **2016**, *32*, 127–135. [CrossRef]
17. TAMU, (Texas A&M University) SWAT | Soil & Water Assessment Tool. 2024. Available online: https://swat.tamu.edu/ (accessed on 4 October 2023).
18. FAO/UNESCO Digital Soil Map of the World. 2024. Available online: https://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/faounesco-soil-map-of-the-world/en/ (accessed on 31 July 2023).
19. Dile, Y.; Srinivasan, R.; George, C. *QGIS Interface for SWAT (QSWAT)*; SWAT: College Station, TX, USA, 2019.
20. EROS Global Land Cover Characterization (GLCC). 2017. Available online: https://doi.org/10.5066/F7GB230D (accessed on 14 January 2025).
21. Geza, M.; McCray, J.E. Effects of Soil Data Resolution on SWAT Model Stream Flow and Water Quality Predictions. *J. Environ. Manage.* **2008**, *88*, 393–406. [CrossRef]
22. López-Ballesteros, A.; Nielsen, A.; Castellanos-Osorio, G.; Trolle, D.; Senent-Aparicio, J. DSOLMap, a Novel High-Resolution Global Digital Soil Property Map for the SWAT + Model: Development and Hydrological Evaluation. *Catena.* **2023**, *231*, 107339. [CrossRef]
23. Rivas-Tabares, D.; de Miguel, Á.; Willaarts, B.; Tarquis, A.M. Self-Organizing Map of Soil Properties in the Context of Hydrological Modeling. *Appl. Math. Model.* **2020**, *88*, 175–189. [CrossRef]
24. Zhao, A. Effect of Different Soil Data on Hydrological Process Modeling in Weihe River Basin of Northwest China. *Arab. J. Geosci.* **2016**, *9*, 664. [CrossRef]
25. Escamilla-Rivera, V.; Cortina-Villar, S.; Vaca, R.A.; Golicher, D.; Arellano-Monterrosas, J.; Honey-Rosés, J. Effects of Finer Scale Soil Survey and Land-Use Classification on SWAT Hydrological Modelling Accuracy in Data-Poor Study Areas. *J. Water Resour. Prot.* **2022**, *14*, 100–125. [CrossRef]
26. Jin, X.; Jin, Y.; Yuan, D.; Mao, X. Effects of Land-Use Data Resolution on Hydrologic Modelling, a Case Study in the Upper Reach of the Heihe River, Northwest China. *Ecol. Model.* **2019**, *404*, 61–68. [CrossRef]
27. Romanowicz, A.A.; Vanclooster, M.; Rounsevell, M.; La Junesse, I. Sensitivity of the SWAT Model to the Soil and Land Use Data Parametrisation: A Case Study in the Thyle Catchment, Belgium. *Ecol. Modell.* **2005**, *187*, 27–39. [CrossRef]
28. Al-Khafaji, M.; Saeed, F.H.; Al-Ansari, N. The Interactive Impact of Land Cover and DEM Resolution on the Accuracy of Computed Streamflow Using the SWAT Model. *Water Air Soil Pollut.* **2020**, *231*, 416. [CrossRef]
29. Bouslihim, Y.; Rochdi, A.; El Amrani Paaza, N.; Liuzzo, L. Understanding the Effects of Soil Data Quality on SWAT Model Performance and Hydrological Processes in Tamedroust Watershed (Morocco). *J. Afr. Earth Sci.* **2019**, *160*, 103616. [CrossRef]
30. Huang, J.; Zhou, P.; Zhou, Z.; Huang, Y. Assessing the Influence of Land Use and Land Cover Datasets with Different Points in Time and Levels of Detail on Watershed Modeling in the North River Watershed, China. *Int. J. Environ. Res. Public Health* **2013**, *10*, 144–157. [CrossRef] [PubMed]

31. Ngeang, L.; Tantanee, S.; Anlauf, R. Comparison of FAO and SOILGRID Data Application on Streamflow and Sus-Pended Sediment Study Using SWAT Model: A Case Study of Upper Yom Basin, Thailand. *GMSARN Int. J.* **2019**, *13*, 104–111.
32. Xiao, J.; Wang, Y.; Sun, J.; Xie, S.; Huang, Y.; Wan, Z.; Meng, L.; Li, X.; Zhong, K. Effect of Soil Spatial Aggregation Caused by the Calculation Unit Division on Runoff and Sediment Load Simulation in the SWAT Model. *J. Hydrol.* **2023**, *626*, 130345. [CrossRef]
33. Vema, V.K.; Sudheer, K.P. Towards Quick Parameter Estimation of Hydrological Models with Large Number of Computational Units. *J. Hydrol.* **2020**, *587*, 124983. [CrossRef]
34. Her, Y.; Frankenberger, J.; Chaubey, I.; Srinivasan, R. Threshold Effects in HRU Definition of the Soil and Water Assessment Tool. *Trans. ASABE.* **2015**, *58*, 367–378. [CrossRef]
35. Arabi, M.; Govindaraju, R.S.; Hantush, M.M.; Engel, B.A. Role of Watershed Subdivision on Modeling the Effectiveness of Best Management Practices with Swat1. *J. Am. Water Resour. Assoc.* **2006**, *42*, 513–528. [CrossRef]
36. Bingner, R.; Garbrecht, J.; Arnold, J.; Srinivasan, R. Effect of Watershed Subdivision on Simulation Runoff and Fine Sediment Yield. *Trans. ASAE.* **1997**, *40*, 1329–1335. [CrossRef]
37. FitzHugh, T.W.; Mackay, D.S. Impacts of Input Parameter Spatial Aggregation on an Agricultural Nonpoint Source Pollution Model. *J. Hydrol.* **2000**, *236*, 35–53. [CrossRef]
38. Kumar, S.; Merwade, V. Impact of Watershed Subdivision and Soil Data Resolution on SWAT Model Calibration and Parameter Uncertainty. *J. Am. Water Resour. Assoc.* **2009**, *45*, 1179–1196. [CrossRef]
39. Jha, M.; Gassman, P.W.; Secchi, S.; Gu, R.; Arnold, J. Effect of Watershed Subdivision on Swat Flow, Sediment, and Nutrient Predictions. *JAWRA J. Am. Water Resour. Assoc.* **2004**, *40*, 811–825. [CrossRef]
40. Alsilibe, F.; Bene, K. Watershed Subdivision and Weather Input Effect on Streamflow Simulation Using SWAT Model. *Pollack Period.* **2021**, *17*, 88–93. [CrossRef]
41. Auda, O.; Bouisson, S.; Bourgeois, N.; Decuq, E.; Grimoin, V.; Jiguet, M.; Lantenois, J.M.; Metge, J.; Pelissier, J.; Thévenot, A.; et al. Livret historique sur les crues du Bassin de l'Argens 2015. Available online: https://var.fr/documents/20142/110788/PAPI+le+livret+sur+les+crues+historiques.pdf/b4c0f5d4-5e52-f6ca-fdc1-b9aabba83482 (accessed on 25 November 2024).
42. SMA Le Territoire | Syndicat Mixte de l'Argens. 2017. Available online: http://syndicatargens.fr/le-sma/le-territoire/ (accessed on 11 December 2023).
43. EauRMC Etude d'évaluation Des Volumes Prélevables Globaux—Bassin Versant de l'Argens (Phases 1, 2 et 3). 2013. Available online: https://www.rhone-mediterranee.eaufrance.fr/l-argens (accessed on 25 November 2024).
44. NASA. *Shuttle Radar Topography Mission (SRTM) Global*; NASA: Washington, DC, USA, 2013.
45. WateriTech WateriTech. Official Home Page. Available online: https://www.wateritech.com/resources (accessed on 22 November 2024).
46. Büttner, G.; Kosztra, B.; Maucha, G.; Pataki, R.; Kleeschulte, S.; Hazeu, G.; Vittek, M.; Schröder, C.; Littkopf, A. *Copernicus Land Monitoring Service Copernicus Land Monitoring Service—CORINE Land Cover User Manual*; European Environment Agency: Copenhagen, Denmark, 2021.
47. Muñoz-Sabater, J.; Dutra, E.; Agustí-Panareda, A.; Albergel, C.; Arduini, G.; Balsamo, G.; Boussetta, S.; Choulga, M.; Harrigan, S.; Hersbach, H.; et al. ERA5-Land: A State-of-the-Art Global Reanalysis Dataset for Land Applications. *Earth Syst. Sci. Data.* **2021**, *13*, 4349–4383. [CrossRef]
48. SCHAPI, HydroPortail. Official Home Page. Available online: https://www.hydro.eaufrance.fr/ (accessed on 22 November 2024).
49. Srinivasan, R.; Ramanarayanan, T.S.; Arnold, J.G.; Bednarz, S.T. Large Area Hydrologic Modeling and Assessment Part II: Model Application. *J. Am. Water Resour. Assoc.* **1998**, *34*, 91–101. [CrossRef]
50. Haro-Monteagudo, D.; Palazón, L.; Beguería, S. Long-Term Sustainability of Large Water Resource Systems under Climate Change: A Cascade Modeling Approach. *J. Hydrol.* **2020**, *582*, 124546. [CrossRef]
51. Krysanova, V.; White, M. Advances in Water Resources Assessment with SWAT—An Overview. *Hydrol. Sci. J.* **2015**, *60*, 771–783. [CrossRef]
52. Moghadam, S.H.; Ashofteh, P.-S.; Singh, V.P. Sensitivity Analysis of Streamflow Parameters with SWAT Calibrated by NCEP CFSR and Future Runoff Assessment with Developed Monte Carlo Model. *Theor Appl Clim.* **2024**, *155*, 8797–8813. [CrossRef]
53. Ougahi, J.H.; Karim, S.; Mahmood, S.A. Application of the SWAT Model to Assess Climate and Land Use/Cover Change Impacts on Water Balance Components of the Kabul River Basin, Afghanistan. *J. Water Clim. Change* **2022**, *13*, 3977–3999. [CrossRef]
54. Moghadam, S.H.; Ashofteh, P.-S.; Loáiciga, H.A. Investigating the Performance of Data Mining, Lumped, and Distributed Models in Runoff Projected under Climate Change. *J. Hydrol.* **2023**, *617*, 128992. [CrossRef]
55. Dwivedi, S.; Kothari, M.; Singh, P.K.; Chhipa, B.G.; Gupta, T. SWAT Model Applications in Hydrology: A Systematic Review. *Int. J. Adv. Biochem. Res.* **2024**, *8*, 603–607. [CrossRef]
56. Darwiche-Criado, N.; Comín, F.A.; Masip, A.; García, M.; Eismann, S.G.; Sorando, R. Effects of Wetland Restoration on Nitrate Removal in an Irrigated Agricultural Area: The Role of in-Stream and off-Stream Wetlands. *Ecol. Eng.* **2017**, *103*, 426–435. [CrossRef]

57. CARD SWAT Literature Database for Peer-Reviewed Journal Articles. 2023. Available online: https://www.card.iastate.edu/swat_articles/ (accessed on 1 October 2023).
58. Tan, M.L.; Gassman, P.W.; Yang, X.; Haywood, J. A Review of SWAT Applications, Performance and Future Needs for Simulation of Hydro-Climatic Extremes. *Adv. Water Resour.* **2020**, *143*, 103662. [CrossRef]
59. Nguyen, T.V.; Dietrich, J.; Dang, T.D.; Tran, D.A.; Van Doan, B.; Sarrazin, F.J.; Abbaspour, K.; Srinivasan, R. An Interactive Graphical Interface Tool for Parameter Calibration, Sensitivity Analysis, Uncertainty Analysis, and Visualization for the Soil and Water Assessment Tool. *Environ. Modell. Software.* **2022**, *156*, 105497. [CrossRef]
60. Abbaspour, K.C.; Johnson, C.A.; van Genuchten, M.T. Estimating Uncertain Flow and Transport Parameters Using a Sequential Uncertainty Fitting Procedure. *J. Hydrol.* **2004**, *3*, 1340–1352. [CrossRef]
61. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling. *J. Hydrol.* **2009**, *377*, 80–91. [CrossRef]
62. Moriasi, D.N. Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria. *Trans. ASABE.* **2015**, *58*, 1763–1785. [CrossRef]
63. Knoben, W.J.M.; Freer, J.E.; Woods, R.A. Technical Note: Inherent Benchmark or Not? Comparing Nash–Sutcliffe and Kling–Gupta Efficiency Scores. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 4323–4331. [CrossRef]
64. Barbarosa, J.; Fernandes, A.; Lima, A.; Assis, L. The Influence of Spatial Discretization on HEC-HMS Modelling: A Case Study. *J. Hydrol.* **2019**, *3*, 442–449. [CrossRef]
65. Githui, F.; Thayalakumaran, T. The Effect of Discretization of Hydrologic Response Units on the Performance of SWAT Model in Simulating Flow and Evapotranspiration. In Proceedings of the 19th International Congress on Modelling and Simulation—Sustaining Our Future: Understanding and Living with Uncertainty, Perth, Australia, 12–16 December 2011; pp. 3412–3418. [CrossRef]
66. Gong, Y.; Shen, Z.; Liu, R.; Wang, X.; Chen, T. Effect of Watershed Subdivision on SWAT Modeling with Consideration of Parameter Uncertainty. *J. Hydrol. Eng.* **2010**, *15*, 1070–1074. [CrossRef]
67. Zhou, S.; Wang, Y.; Guo, A.; Li, Z.; Chang, J.; Meng, D. Impacts of Changes in the Watershed Partitioning Level and Optimization Algorithm on Runoff Simulation: Decomposition of Uncertainties. *Stoch. Environ. Res. Risk Assess.* **2020**, *34*, 1909–1923. [CrossRef]
68. Bhatta, B.; Shrestha, S.; Shrestha, P.K.; Talchabhadel, R. Evaluation and Application of a SWAT Model to Assess the Climate Change Impact on the Hydrology of the Himalayan River Basin. *Catena.* **2019**, *181*, 104082. [CrossRef]
69. Tuo, Y.; Duan, Z.; Disse, M.; Chiogna, G. Evaluation of Precipitation Input for SWAT Modeling in Alpine Catchment: A Case Study in the Adige River Basin (Italy). *Sci. Total Environ.* **2016**, *573*, 66–82. [CrossRef]
70. Veettil, A.V.; Mishra, A.K. Water Security Assessment Using Blue and Green Water Footprint Concepts. *J. Hydrol.* **2016**, *542*, 589–602. [CrossRef]
71. Abbaspour, K.C.; Vaghefi, S.A.; 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.* **2018**, *10*, 6. [CrossRef]
72. Whittaker, G.; Confesor, R., Jr.; Luzio, M.D.; Arnold, J.G. Detection of Overparameterization and Overfitting in an Automatic Calibration of SWAT. *Trans. ASABE.* **2010**, *53*, 1487–1499. [CrossRef]
73. 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.* **2015**, *524*, 733–752. [CrossRef]
74. Kamali, B.; Abbaspour, K.C.; Yang, H. Assessing the Uncertainty of Multiple Input Datasets in the Prediction of Water Resource Components. *Water.* **2017**, *9*, 709. [CrossRef]
75. 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. Manage.* **2020**, *266*, 110625. [CrossRef]
76. El-Sadek, A.; İRvem, A. Evaluating the Impact of Land Use Uncertainty on the Simulated Streamflow and Sediment Yield of the Seyhan River Basin Using the SWAT Model. *Turk. J. Agric. For.* **2014**, *38*, 515–530. [CrossRef]
77. Wu, L.; Xu, Y.; Li, R. Effects of Input Data Accuracy, Catchment Threshold Areas and Calibration Algorithms on Model Uncertainty Reduction. *Eur. J. Soil Sci.* **2024**, *75*, e13519. [CrossRef]
78. Tolson, B.A.; Shoemaker, C.A. Dynamically Dimensioned Search Algorithm for Computationally Efficient Watershed Model Calibration. *Water Resour. Res.* **2007**, *43*. [CrossRef]
79. Zamani, M.; Shrestha, N.K.; Akhtar, T.; Boston, T.; Daggupati, P. Advancing Model Calibration and Uncertainty Analysis of SWAT Models Using Cloud Computing Infrastructure: LCC-SWAT. *J. Hydroinf.* **2020**, *23*, 1–15. [CrossRef]

80. Abbaspour, K.C.; Yang, J.; Maximov, I.; Siber, R.; Bogner, K.; Mieleitner, J.; Zobrist, J.; Srinivasan, R. Modelling Hydrology and Water Quality in the Pre-Alpine/Alpine Thur Watershed Using SWAT. *J. Hydrol.* **2007**, *333*, 413–430. [CrossRef]
81. Tapas, M.R.; Etheridge, R.; Tran, T.-N.-D.; Le, M.-H.; Hinckley, B.; Nguyen, V.T.; Lakshmi, V. Evaluating Combinations of Rainfall Datasets and Optimization Techniques for Improved Hydrological Predictions Using the SWAT+ model. *J. Hydrol. Reg. Stud.* **2025**, *57*, 102134. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.