

---

Article

# Effect of using multi-year LULC and monthly LAI inputs on the calibration of a distributed hydrologic model

Ibrahim Olayode Busari<sup>1,2,3,\*</sup> Mehmet Cüneyd Demirel<sup>4</sup> and Alice Newton<sup>1,5</sup>

<sup>1</sup> Centre of Marine and Environmental Research (CIMA), Gambelas Campus, University of Algarve, 8005-139 Faro, Portugal; [ibrahim.busari@studio.unibo.it](mailto:ibrahim.busari@studio.unibo.it) (I.O.B), [anewton@ualg.pt](mailto:anewton@ualg.pt) (A.N).

<sup>2</sup> Faculty of Marine and Environmental Sciences, University of Cadiz, 11510 Puerto Real (Cádiz), Spain.

<sup>3</sup> Department of Biological, Geological and Environmental Sciences, University of Bologna, 48123 Ravenna, Italy, [ibrahim.busari@studio.unibo.it](mailto:ibrahim.busari@studio.unibo.it) (I.O.B)

<sup>4</sup> Department of Civil Engineering, Istanbul Technical University, Maslak, 34469 Istanbul, Turkey, [demirelmc@itu.edu.tr](mailto:demirelmc@itu.edu.tr) (M.C.D)

<sup>5</sup> Department of Earth, Environmental and Marine Sciences (DCTMA), Gambelas Campus University of Algarve, 8005-139 Faro, Portugal. [anewton@ualg.pt](mailto:anewton@ualg.pt) (A.N).

\* Correspondence: [busaribrahim112@gmail.com](mailto:busaribrahim112@gmail.com), [ibrahim.busari@studio.unibo.it](mailto:ibrahim.busari@studio.unibo.it) (I.O.B).

**Abstract:** This study explores the use of satellite-based LULC (Land Use / Land Cover) data while simultaneously correcting potential evapotranspiration (PET) input with Leaf Area Index (LAI) to increase the performance of a physically distributed hydrologic model. The mesoscale hydrologic model (mHM) was selected for this purpose due to its unique features. Since LAI input informs the model about vegetation dynamics, we incorporated the LAI based PET correction option together with multi-year LULC data. The Globcover land cover data was selected for the single land cover cases, and hybrid of CORINE (coordination of information on the environment) and MODIS (Moderate Resolution Imaging Spectroradiometer) land cover datasets were chosen for the cases with multiple land cover datasets. These two datasets complement each other since MODIS has no separate forest class but more frequent (yearly) observations than CORINE. Calibration period spans from 1990 to 2006 and corresponding NSE (Nash-Sutcliffe Efficiency) values varies between 0.23 and 0.42, while the validation period spans from 2007 to 2010 and corresponding NSE values are between 0.13 and 0.39. The results revealed that the best performance is obtained when multiple land cover datasets are provided to the model and LAI data is used to correct PET, instead of default aspect-based PET correction in mHM. This study suggests that to minimize errors due to parameter uncertainties in physically distributed hydrologic models, adequate information can be supplied to the model with care taken to avoid over-parameterizing the model.

**Keywords:** LULC, mHM, CORINE, MODIS, Hydrologic Modeling, Model Calibration

---

## 1. Introduction

Decision-making about the sustainability of water resources requires supporting tools to manage the resources effectively. Hydrological models are useful tools to capture spatiotemporal variations in hydrological fluxes in many basins. A model consists of various parameters that define the characteristics of the basin. Hydrologic models have become progressive tools for studying the effects of anthropogenic activities and environments on hydrology and ecology [1]. Good modeling should be characterized by a high degree of confidence in the simulated outputs and minimization of uncertainty in model results. These uncertainties may be due to data input integrity, measured output data, non-optimal parameters, and model bias.

The conventional knowledge is that the best simulation is associated with less uncertainty in its output. The best model is the one that gives results close to reality with the use of the least parameters and model complexity, i.e. a parsimonious model. Hydrologists are faced with improving the quality of model outputs by using various procedures to reduce these uncertainties. However, errors due to model bias are difficult to address and are usually mentioned by the modeler when reports are made. To address uncertain-

ties relating to non-optimum parameter values, model calibration is employed. This procedure is targeted at obtaining optimum parameter values for the model setup in the specific basin, such that the error due to non-optimal parameters is insignificant compared with errors due to data integrity [2].

Various strategies have been adopted to calibrate hydrological models. Scientists have introduced series of mathematical functions that guide parameter optimization during the calibration of the model. Hydrologists are also engaged in exploring model setups and data inputs to obtain the best model performances. Various resolutions of Digital Elevation Models (DEM), Land Use Land Cover (LULC) source and resolutions, soil map resolution, and data length have been studied to obtain their effects on model performance. These responses vary among models due to model types and structures [3]. Physically distributed hydrological models are capable of simulating multi-layered hydrological processes. Rather than utilizing extensive hydrological and meteorological data during calibration, physically distributed hydrologic models require the evaluation of many parameters that describes the physical characteristics of the catchment [4]. These models provide output upon which decisions about effective water resources management are based especially with spatiotemporal changes in LULC. Therefore, the accuracy of these outputs is of high importance to hydrologist, hence the need for efficient calibration procedure to design simulations representing the interested basin.

Bouslihim et al. [5] studied the effect of input resolution on the Soil and Water Assessment Tool (SWAT) model's performance. Soil data from different databases with different resolutions were used to calibrate the model, showing that the data's quality and resolution affect the estimation of hydrological fluxes, especially soil water storage and water yield. Ficchi et al. [6] also studied the impact of temporal resolution of input data on model performance using different precipitation time-steps to simulate short duration hydrologic events like flash floods. Storm duration, flood hydrograph peaks, rainfall-runoff lag time and temporal precipitation variability, are the most influential variables for model performance at different time steps. The effect of calibration data and length on model performance was observed by Ilampooranan et al. [7], using crop yield data to calibrate a SWAT model rather than the usual streamflow data to identify a deficiency in model structure. The research highlighted the importance of additional data sources to improve the robustness and predictive ability of distributed models. Li et al. [8] further studied the effect of calibration length on lumped hydrologic model performance. They showed that longer calibration periods do not necessarily result in better model performance and optimum parameter values can be attained with fewer calibrations in humid catchments.

The use of a multi-objective function to calibrate the model to obtain improved model performance was studied by Sahraei et al. [9]. They found that although the performance of the models was improved when calibrated with multi-objective functions, they were unable to identify all aspects of simulation errors relating to the model uncertainty. This led to the use of segmentation-based model-wrapper to group the models together and significantly increased the performance compared to individual models. The influence of farming activities was also studied by Gao et al. [10], where the impacts of dominant crop rotation techniques on SWAT hydrologic model performance were observed. It was concluded that proper sets of crop rotations in the model could improve model performance within the rotation period. However, applying the same set of crop rotations for longer periods in the model decreases the model performance, and the crop rotation pattern often leads to model over-parameterization and equifinality.

Orth et al. [11] studied the influence of model complexities on performance. The study concluded that complexity does not necessarily cause increased model performance. Model performance varies differently with the hydrological variable under consideration. The transfer of information from gauged catchments to ungauged catchments for the prediction of runoff is common in hydrologic modeling, which justifies the need to observe the response of models to varying methods of achieving this. Yang et al. [12] observed the influence of regionalization methods on the performance of hydrological

models in different climatic regions and showed that model parameters vary among regionalization methods and climatic regions with more variations associated with the change of climatic region, which affects the performance of the models.

Land cover depicts the physical state of the land surface, in terms of the spatial distribution of soil, water, vegetation and anthropogenic activities on the earth surface (Parveen et al. [13]; Prasad & Ramesh [14]; Risal et al. [15]). Leaf Area Index (LAI) is an essential parameter of plant community that describes the total area of leaves per unit ground area and is directly related to the amount of light available for interception by a plant. It defines the total one-sided area of leaf tissue per unit ground surface area [16] and influences transpiration, interception of rainfall and sunlight, gas exchange and photosynthetic processes [17], through the canopy structural characteristics [18].

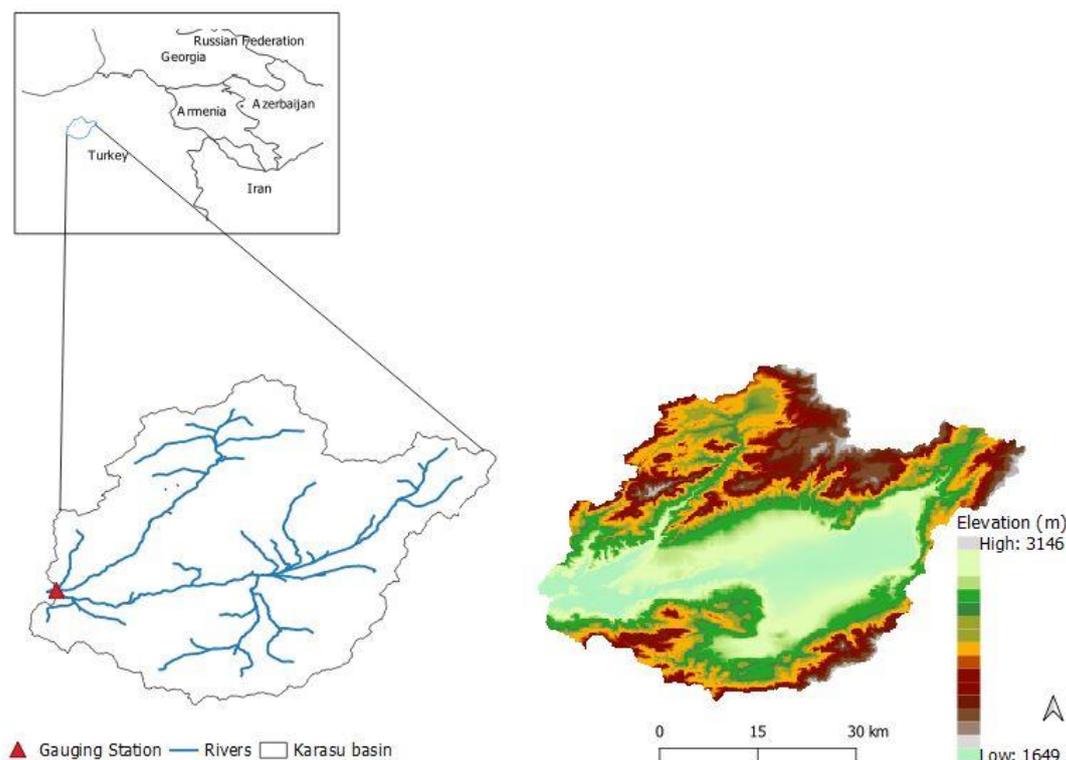
The sensitivity of vegetation parameters to hydrological processes was assessed by Das et al. (2018) using Land Use Land Cover (LULC) changes in the eastern India basin between 1995 and 2005 to set up VIC model. The LAI emerged as the most sensitive parameter that influences hydrological fluxes. LULC changes through deforestation, urbanization, expansion of croplands led to reduction in the extent of canopy cover for interception and transpiration, which affects hydrologic performance. This necessitated the study of the impact of multiple LULC datasets and resolutions on hydrologic performance observed by Jin et al. [1] and suggested using a high-resolution LULC dataset for modeling when a year LULC dataset is available. Further, the use of 10-year LULC dataset to calibrate the SWAT model in the study improved model performance by 2.2% - 13.9%.

The use of yearly datasets identifies the LULC changes better by smoothing deviations over the input data of the semi-distributed model. However, little information exists about the effect of multiple LULC datasets on the performance of fully distributed physically based hydrologic models like the mesoscale Hydrologic Model (mHM). Hence, this study aims at 1) exploring the performance of a distributed hydrologic model by calibrating the hydrologic model using a series of land use land cover datasets and 2) analyzing the influence of leaf area index, which is a significant characteristic of forest areas on the performance of the model. This research involves different model configuration cases with varying LULC, potential evapotranspiration (PET) setup, and comparing the performance and hydrograph of each simulated output.

## **2. Materials and Methods**

### **2.1 Study Area**

The study area is the Karasu basin located at the Euphrates River's headwaters in Turkey (Figure 1). The basin is mostly dominated by pasture, grass and bare land. The total area is about 10,275km<sup>2</sup>. This area is hugely mountainous with an elevation range of 1125–3500m and mean basin slope of 20%. Karasu Basin boundaries are within the longitudes 38° 58' E to 41° 39' E and latitudes 39° 23' N to 40° 25' N. Studies showed that 60-70% of the total annual runoff volume comes during snowmelt season. The annual mean precipitation of the basin fluctuates from 400 to 450mm/yr [20]. Streamflow data for the Karasu basin are collected by General Directorate of Turkish State Hydraulic Works (DSI) at the Karasu Karagöze and Aşağı Kağdıriç gauging station [21]. The basin downstream is characterized by huge dams that make accurate runoff estimation crucial to efficient water resources allocation for flood control, hydropower generation, irrigation, and water supply [22].



**Figure 1:** Location of Karasu Basin and Elevation

## 2.2 Model Description

The mHM is a fully distributed hydrologic model based on numerical approximations of hydrologic processes such as canopy interception, snow accumulation and melting, soil moisture dynamics, infiltration and surface runoff, evaporation, subsurface storage and discharge generation, deep percolation and base flow, and discharge attenuation and flood routing [23]. The model discretization in this setup is based on grid cells of  $0.0625^\circ$  spatial resolution, with each grid generating runoff that is routed using adaptive timestep, constant celerity routing method. The mHM utilizes a Multiscale Parameter Regionalization technique to account for sub-grid variability of catchment morphology (Demirel et al. [24], Dembélé et al. [25], Höllering et al. [26]) which enables the swapping between different scales while calculating all fluxes and routing streamflow on a preferred model scale. The model contains 62 global parameters that can be adjusted during the calibration process and pedo-transfer functions for soil parameterization [24]. Samaniego et al. [27] can be consulted for explicit description of model formulation and parameter description. The mHM was preferred as the interest model due to its three unique features 1) multiscale parameter regionalization technique, which links basin's physical characteristics to parameter value via pedo-transfer functions 2) ability to process multiple LULC data 3) use of leaf area index (LAI) data not only for calculating interception loss but also for correcting PET.

## 2.3 Data Preparation

The basic data for setting up the mHM can be classified into meteorological data, morphological data, land cover data and gauging station data. The essential meteorological variables are the precipitation and average air temperature at hourly to daily temporal resolution. However, based on a strategy of estimating potential evapotranspiration, additional data like net radiation, wind speed, the absolute vapor pressure of air, potential evapotranspiration, maximum and minimum air temperature. The morphological variables are the digital elevation model, soil maps with textural properties, geological maps containing specific yield, permeability and aquifer thickness. For this study, the E-OBS dataset was utilized with the precipitation, average air temperature data and potential

evapotranspiration obtained from Rakovec et al. [28] with a resolution of  $0.25^\circ$  and daily time steps, input directly into the model, hence no need for its estimation in the model (Table 1). All morphological data are provided at a resolution of  $0.001953^\circ$  and the meteorological data at  $0.25^\circ$ . The model is run at a daily time-step with and outputs are internally resampled to a spatial resolution of  $0.0625^\circ$  by the model.

**Table 1:** Overview of input datasets

Variable	Description	Spatial Resolution	Source
Q (daily)	Streamflow	Point	Karasu Karagöze gauging station. Aşağı Kağdarıç gauging station.
P (daily)	Precipitation	$0.25^\circ$	E-OBS 20.0e
$ET_{ref}$ (daily)	Reference evapotranspiration	$0.25^\circ$	E-OBS 20.0e
$T_{avg}$ (daily)	Average air temperature	$0.25^\circ$	E-OBS 20.0e, MODIS
Land cover	Pervious, impervious and forest	$0.001953^\circ$	CORINE, MODIS, GlobcoverV2
DEM data	Slope, aspect, flow accumulation and direction	$0.001953^\circ$	SRTM
Geology class	Two main geological formations	$0.001953^\circ$	EUROPEAN SOIL DATABASE
Soil class	Soil texture data	$0.001953^\circ$	HARMONIZED WORLD SOIL DATABASE

## 2.4 Land Cover Data

The Coordination of Information on the Environment (CORINE), MODIS and GLOBCOVER land cover data were used. The Coordination of Information on the Environment (CORINE) Land Cover (CLC) is the first land cover map in Europe based on photographic interpretation of satellite images for the years 1990, 2000, 2006, 2012 and 2018, producing the CLC 1990, CLC 2000, CLC 2006, CLC 2012 and CLC 2018 products respectively. The spatial resolutions of the products are 100m and 250m, [29] with the land cover inventoried into 44 classes [30]. The main advantage of the CORINE Land Cover (CLC) inventory is its frequent update by European countries, although the level of detail of the source data is a limitation [30]. This limitation suggests that the CLC inventory is a useful resource for land cover change studies on the regional scale. The MODIS Land Cover collection is a global land cover data set designed to support research related to current, seasonal to the decadal status of global land properties, with 17 land cover classes at a 500m spatial resolution and yearly temporal coverage with a 71.6% overall accuracy [29]. The GLOBCOVER dataset is a global project initiated by European Space

Agency (ESA) in conjunction with the Joint Research Centre (JRC), the European Environment Agency (EEA), the Global Observation of Forests Cover and Land Dynamic (GOFC-GOLD), Food and Agriculture Organization (FAO), the United Nations Environment Program (UNEP) and the International Geosphere-Biosphere Program (IGBP) [31]. The project entailed the generation of world land cover map using 300m Medium Resolution Imaging Spectrometer (MERIS) instrument aboard ENVISAT [32]. The mHM identifies only three land cover classification. This classification includes forest (coniferous and deciduous), impervious (urban and built-up areas, water bodies, and consolidated soils), and pervious (grasslands, croplands, and bare soil) [23]. Hence, the land cover data were reclassified to suit the data requirements.

## 2.5 Sensitivity Analysis

Hydrologic models are associated with uncertainties, which have always prompted researchers better to understand the hydrologic process and better representation in models. These models are associated with some level of complexities that vary among models to represent necessary variables correctly. Sensitivity Analysis (SA) is a method employed by modelers to deal with these complexities during calibration by analyzing the contribution of each model parameter to the uncertainty in model prediction. The basis for SA is an effect called the "Sparsity of Factors" principle. This principle, which originates from the field of Statistical Design of Experiments, states that when a process is influenced by several variables, its behavior is likely driven by a small subset of these variables [33]. These subsets of variables characterize sources of uncertainties in model prediction and upon which research towards model performance improvement should be focused [34]. SA can summarily be applied for assessment of model similarities [35], obtaining regions of sensitivity [36], factor and model reduction [37] and uncertainty apportionment [38]. In this study, we used PEST Tool [39] to assess most important parameters based on Jacobian matrix to include in the model calibration.

## 2.6 Model Calibration and Validation

In this study, four different cases (Table 2) based on input data were created to achieve the set objectives. These cases contain different configurations and representations of hydrologic processes. The Case 1 (C1) model was configured using one-year GLOBCOVER land cover data and no correction of PET data using LAI information in the region. The PET data was only subjected to correction with the aspect of the grid cells. The Case 2 (C2) model was also configured with the same land cover data but with LAI correction. This strategy allows the PET data to be corrected based on the LAI data supplied to the model. The Case 3 (C3) model was configured with multiple land cover data from a different database. It contained 1990 and 2000 CORINE land cover data and consecutive MODIS land cover data from 2001 to 2008 for the calibration period. The PET data, in this case, is also subjected to only the aspect correction and ignores the LAI correction. The Case 4 (C4) model contains the same land cover configuration as C3; however, the PET data were subjected to LAI correction. The model was calibrated with the discharge data from the gauging station from 1990 to 2006 while 2007 to 2010 was used for the validation.

**Table 2:** Description of the model cases

Case	Land cover Input	PET correction with LAI
1	1 year Globcover dataset	No
2	1 year Globcover dataset	Yes
3	MODIS and CORINE datasets	No
4	MODIS and CORINE datasets	Yes

## 2.7 Dynamically Dimensioned Search Algorithm (DDS)

The DDS was developed to obtain best global solutions within a model evaluation limit based on heuristic global search algorithm [40]. The DDS algorithm begins its search globally and becomes localized as the iteration numbers approaches the limit of function evaluations. This is achieved by decreasing the number of dimensions in the neighborhood dynamically and probabilistically, and further distribute the reducing probability equally to each decision variable [41]. The need to modify parameters simultaneously during calibration to preserve the current gain in calibration result was the motivation behind the algorithm. The critical feature of DDS was triggered by the poor result associated with manual calibration of watersheds especially with progress in the calibration process. It becomes necessary to calibrate one or few parameters simultaneously as calibration result improves in order to prevent loss of previously gained result [40].

## 2.8 Model Performance

The Nash-Sutcliffe efficiency (NSE) was used to evaluate the integrity of the model simulations for each scenario. This objective function quantifies the accuracy of prediction by comparing the differences between predicted and observed discharge values. NSE is a scaled metric that compares the expected variable with the observed variable's mean as the naïve forecast. An efficiency of zero, indicates that using the observed data as forecast would give the same error as the predicted data, a negative efficiency means that the mean of the observed data is a better predictor than the model prediction [42]. The inverse ( $NSE_{inv}$ ) and logarithm ( $NSE_{ln}$ ) of this metric was also adopted to capture the flow dynamics. Pushpalatha et al. [43] reported that  $NSE_{ln}$  and  $NSE_{inv}$  are useful metrics for the evaluation of low flows, although the latter shows no sensitivity to high flows unlike the  $NSE_{ln}$ , which shows little sensitivity, and NSE, which is highly sensitivity to high flows. The Kling-Gupta Efficiency (KGE) was also adopted to evaluate the annual flows [44]. KGE consists of three components targeted at addressing the flaws observed in NSE. This evaluation criterion is based on NSE decomposition into the linear correlation between observed and simulated flows, the flow variability term and biased term [45]. The optimum value is 1 and unlike NSE, KGE does not have an inherent benchmark against which flows are compared and the value to be compared is subjective to the modeler [46]. Other evaluation criteria used are PBIAS and Mean Absolute Error (MAE), which measured the simulated values' average tendency to be larger or smaller than their observed ones and the absolute difference between observed and simulated flows respectively. The optimum value of MAE is 0 and low magnitude values indicates accurate model prediction. Positive values of PBIAS indicate some overestimation of predicted data while negative values indicates model underestimation of predicted values..

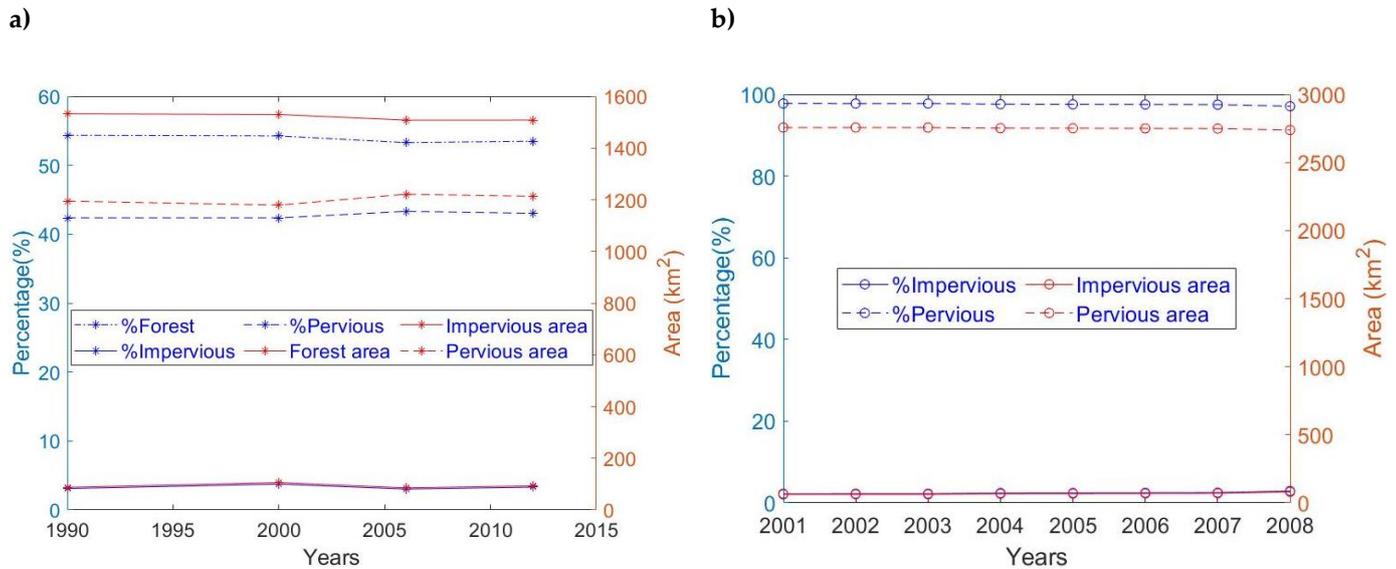
Stagl and Hattermann [47] suggested the use of some hydrological indices for the evaluation of model simulation. For this study, the long-term annual average discharge, flow in dry season, flow in wet season and base flow were selected as indices to evaluate the modeled cases. The wet season was taken from May to October while the rest of the months were considered as dry season. The base flow was estimated as 30% of the average

daily flow and the performance was evaluated using the procedure alighted in Chirachawala et al. [48].

### 3 Results

#### 3.1 Land Use Land Cover Variations

LULC maps in physically distributed hydrological models are essential inputs for efficient modeling of spatial variations in hydrologic fluxes, excess rainfall and infiltration in particular. This study adopted the CORINE and MODIS LULC dataset and the changes in the land cover characteristics shown in Figure 2 using the mHM recognized land cover classes.



**Figure 2:** Land cover evolution in Karasu basin (a) CORINE (b) MODIS

The variation in land cover classes between the upper and lowest land cover map adopted is also shown in Figure A1. Using the conventional CORINE land cover nomenclature to characterize the land cover in the basin, the most dominant land cover obtained in the study period is the natural grasslands, which increased from 718.63km<sup>2</sup> in 1990 to 929.63km<sup>2</sup> in 2012, although a slight decline was noticed in 2000 before the increase in subsequent years. A significant increase in the area covered by water bodies was also observed. In 1990, 0.04% of the watershed, which is 1.1km<sup>2</sup>, is covered by water. This increased to 11.17km<sup>2</sup> in 2000 and varied annually until it reached 9.42km<sup>2</sup> in 2012. Road networks in the basin occupied the least land cover in the studied maps covering 0.013% of the total basin area in 1990 and expanded to 0.054% in 2012. Artificial surfaces like buildings, roads, industry and land cover comprising of urban fabrics increased by 7.35% between 1990 and 2012. This may be responsible for the decrease in the agricultural areas from 1194.53km<sup>2</sup> to 1139km<sup>2</sup>, indicating about 4.7% decline. In the same vein, forest areas declined by 6.3% between 1990 and 2012. Glacier and perpetual snow in the watershed were observed to cover more area in 2012. In 1990, the snow area was around 434.62km<sup>2</sup> and steadily increased annually in the observed maps to 473.98km<sup>2</sup> in 2012. However, the hydrologic model of interest was designed to recognize three land cover classifications; forest, impervious and pervious. MODIS has no separate forest class in the study domain after reclassification but more frequent (yearly) observations than CORINE, hence these two land cover products complement each other.

#### 3.2 Sensitivity Analysis

Sensitivity analysis in this study was used to identify model parameters sensitive to discharge prediction in the basin using the Parameter Estimation and Uncertainty Analysis (PEST) tool [39]. PEST performs all calibration tasks, including sensitivity and uncertainty analysis, upon the user needs. The tool allowed the identification of the sensitive parameters to discharge prediction in the model set up and these were normalized using the maximum sensitivity value. Of the 62 global parameters in the model, 20 parameters were selected as most sensitive as shown in Table A1 and they were afterwards used to calibrate the hydrologic model.

The mHM parameters are divided into groups based on the processes they influence. The soil moisture parameters were observed to be very sensitive to the model output, although this sensitivity varies among the parameters. Eleven of these parameters out of about twenty soil moisture parameters were sensitive in the setup. This actually depends on the method used for modeling soil water movement, the Feddes equation was used in this study. The sensitivity of PET parameters differs with choice of PET estimation method. For this study, PET correction using LAI and aspect maps were eventually adopted. The PET parameters related to correction using LAI parameters were found to be sensitive. The process of interflow was also found to be sensitive to streamflow predictions in the model set up. This was shown in the sensitivity analysis characterized with 3 out of the 5 interflow parameters showing huge sensitivity. Lastly, the percolation process, through the recharge coefficient parameter was found to be sensitive to the streamflow dynamics in the study area. These sensitive parameters were selected to calibrate the model to obtain optimum parameter values with 3000 iterations.

### 3.3 Model Calibration

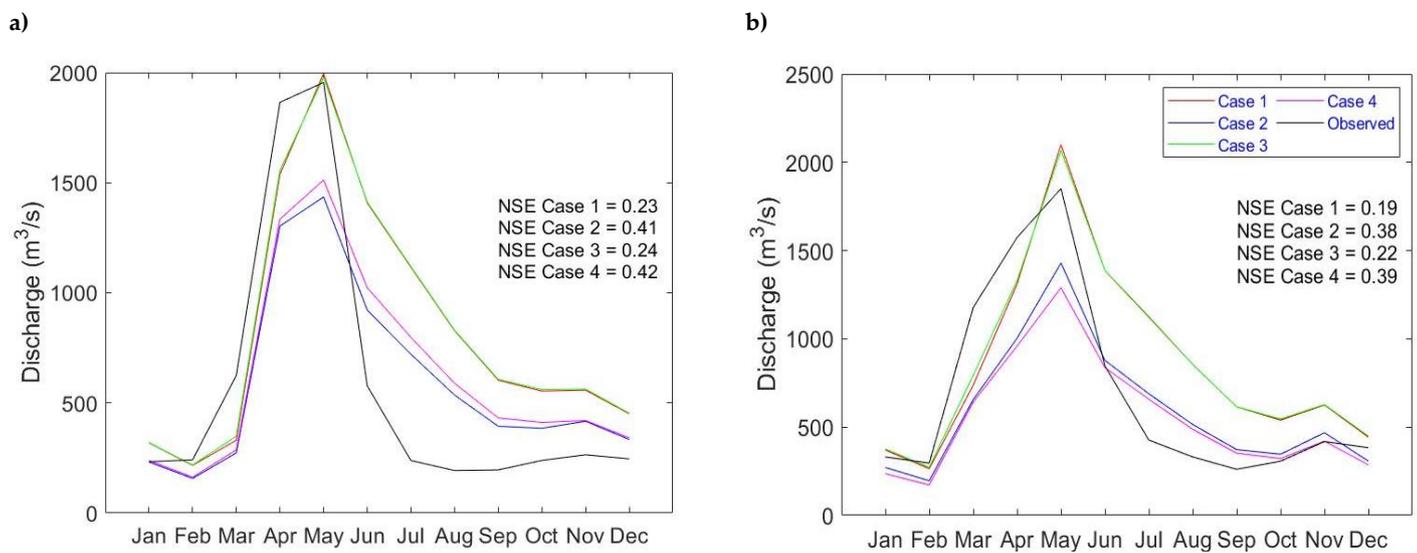
All four cases were calibrated with the sensitive parameters obtained from the sensitivity analysis. These cases are used to consider the simultaneous effect of PET correction technique and multiple LULC maps. The estimated NSE value ranges between 0.226 and 0.422, with the highest in Case 4 and the lowest at Case 1 respectively. PBIAS and MAE value for Case 1 and Case 3 were the highest, indicating a better performance, with Case 2 and Case 4 the better cases. KGE values for the modeled cases during calibration ranges between 0.419 and 0.544 for the calibration, with Case 4 having the highest KGE value indicating the best performance with this metric.  $NSE_{in}$  and  $NSE_{inv}$  were adopted to filter out the high flow sensitivity of conventional NSE during estimation. The range of  $NSE_{in}$  was -0.36 and 0.05 while that of  $NSE_{inv}$  was -0.65 and -0.61. Considering the individual effect of PET correction method on the model performance, Case 1 and Case 2 were configured with the same input data but different PET correction method. Case 3 and Case 4 were also configured with the same input data but different PET correction method. In this vein, the NSE value of Case 2 revealed higher value than that of Case 1 while that of Case 4 also revealed a higher NSE value than that of Case 3. To isolate the effect of multiple land cover data on the model, Case 1 and Case 3 were configured with the same input dataset but different land cover configuration. Case 2 and Case 4 were also configured with the same input data but different land cover configuration. In this comparison, Case 3 and Case 4 proved better with higher NSE values. KGE values for these cases also follows the same pattern with the Case 4 having the best performance among the cases as shown in Table 3.

**Table 3:** Performance assessment of the different cases

	NSE (-)		$NSE_{in}$		$NSE_{inv}$		PBIAS (%)		KGE		MAE (m <sup>3</sup> /s)	
	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val	Cal	Val
<b>Case 1</b>	0.23	0.19	-0.36	-0.08	-0.63	-0.38	43.5	29.2	0.42	0.51	15.9	15.1
<b>Case 2</b>	0.41	0.38	0.05	0.26	-0.61	-0.48	2.8	-10.8	0.54	0.52	12	10.8
<b>Case 3</b>	0.24	0.22	-0.36	-0.07	-0.63	-0.36	44.2	30.2	0.42	0.51	15.9	15
<b>Case 4</b>	0.42	0.39	-0.02	0.23	-0.65	-0.89	8.8	-16.7	0.54	0.48	12.3	10.6

### 3.4 Model Validation

For the validation phase, the NSE value obtained were slightly lower than the calibration phase ranging from 0.19 and 0.39, with the Case 4 also showing the best performance. The  $NSE_{in}$  metric revealed its highest value as Case 2 and the lowest with Case 3, recording efficiency of 0.26 and -0.07 respectively. KGE also recorded its highest value in Case 2 and lowest in Case 4. However, Case 4 performed better under MAE, with the lowest value among the observed Cases as shown in Table 3. The hydrograph of the observed and simulated average monthly discharges for all the cases is shown in Figure 3 and the cumulative distribution frequency of daily discharge data is shown in Figure A2. It should be noted that all cases underestimated the frequency of the discharge until it reached about  $1000\text{m}^3/\text{s}$ , when Case 2 and Case 4 overestimated the frequency, as did the other cases when the discharge accumulated to about  $1400\text{m}^3/\text{s}$ .



**Figure 3:** Observed and simulated average monthly discharge of designed 4 cases. (a) Calibration (b) Validation

### 3.5 Hydrological Indices

The annual discharge, wet season flows, dry season flows and base flow of the modeled cases were used to evaluate the model's performance by comparing them with the observed data from the gauging station, as shown in Table 4. For annual discharge, wet season flows and dry season flows, Case 2 performed better than the others, although Case 4 also performs optimally. For annual discharge, the percentage difference of Case 2 is 2.89% while that of the worst performance is Case 3, with a percentage difference of 44.37%. The dry season flow performance showed a different trend, with Case 3 performing better than the other cases, with a percentage difference of 1.62% against the lowest performing case which was Case 2 with a difference of 22.7%.

To show the effect of misclassification of land cover on discharge, monthly discharge in the basin was predicted by providing the model with CORINE 2006 and MODIS 2006 sequentially. The CORINE land cover product characterized with forest, pervious and impervious land cover predicted higher discharges between May and October, while the MODIS land cover product with just pervious and impervious classes simulated higher discharges for the rest of the months.

**Table 4:** Performance of mHM in simulating hydrological indices

Hydrological Indices	Case	Difference (%)
Annual Discharge	Case 1	43.63
	Case 2	<b>2.89</b>
	Case 3	44.37
	Case 4	8.96
Wet flows	Case 1	91.97
	Case 2	<b>29.50</b>
	Case 3	92.18
	Case 4	40.21
Dry flows	Case 1	2.88
	Case 2	22.7
	Case 3	<b>1.62</b>
	Case 4	21.1
Base flow	Case 1	43.45
	Case 2	<b>2.78</b>
	Case 3	44.19
	Case 4	8.83

**\*\*Bold indicates lowest (best) cases**

#### 4. Discussion

The need to estimate the impacts of climate and land use change on discharge regime has made hydrologic models popular worldwide, especially with their ability to predict flows at gauges and ungauged catchments. However, these predictions are subjected to uncertainty due to model bias, input data errors, and errors in model parameter values [49]. The goal of hydrologic modelers is to reduce uncertainties associated with the model outputs upon which decision about hydrologic fluxes are based. There have been numerous studies focusing on different ways to improve model performance by exploring various input data characteristics that affects the predictions in conceptual and physically distributed models [7]–[11], [50]. Models like mHM are spatially distributed since they consist of equations involving one or more space coordinates, which are used for simulation of spatial variation in hydrological variables within a catchment as well as simple outflows and bulk storage volumes. Such models make considerable demands in terms of computational time and data requirements and are costly to develop and operate [2]. The results presented in this study provide the spatial model's response to varying characteristics of the input data based on the objective under study. Recall that the scenarios labeled as Case 1 to Case 4 and their varying input characteristics. Case 1 and Case 2 were configured with single LULC data, with LAI influence on the Case 2 only. The same applies to Case 3 and Case 4 but with multiple land cover classes, which makes it possible to observe LAI's effect on the model performance. LAI is a key biophysical vegetation property for interception that describes biome-specific canopy structure and an essential variable in evaluating the interrelation of the components of the water-soil-plant-atmosphere system, which influences environmental studies [51]. The vegetation of an area due to its LAI is a significant component of land surface models that influence evapotranspiration simulation and, consequently, affects base flow, recharge, infiltration excess, saturation excess, subsurface storm flow and catchment wetness [52]. As hypothesized, the cases LAI data performed better than their counterparts during calibration and validation process. The implication is that predictions are more accurate when leaf area index is included in the model for correction on evapotranspiration. Precise evaluation of evapotranspiration is a vital aspect of water balance estimation. About 60% of terrestrial precipitation returns to the atmosphere by plant transpiration (40%), or through direct soil evaporation (20%) [53]. The correct simulations of LAI seasonal dynamics and stomatal aperture in an eco-hydrological model are prerequisite of good simulation of canopy radiation exchanges and transpiration fluxes which are important components of the accurate water balance estimation [54].

Furthermore, the impact of single and multiple land cover data was observed. The result associated better model performance with the cases with multiple land cover data. This corroborates Jin et al. [1], where the use of multiple land-use datasets improved model performance although with some complexities in simulation complexity due to greater number of LULC patches. LULC maps are essential input of physically distributed models since they spatially delineate the basin's different morphological characteristics that decide water's fate on the earth's surface. A detailed characterization of land cover is essential for urban areas, as coarse spatial description might result in biased estimates by hiding urban heterogeneous evapotranspiration [55]. Petrucci & Bonhomme [56] also found a similar result to our study where increasing geographical information clearly improves model performances with the land use classification providing the highest benefit. Finally, considering the combined effect of both morphological characteristics, the case with multiple land cover data and LAI to correct evapotranspiration performed better than the other cases and the case with single land cover data, without LAI correction performed worse. It is noteworthy that the discharge data were used in the model's calibration, which leaves the possibility of variation in model performance if different calibration method or data was used. The data-intensive nature of physically distributed models often leads to over-parameterization while effectively representing each hydrologic process that constitutes such a model. This opens up another phase of research

in hydrologic modeling to solve over-parameterization and identify strategies to improve model performance without overwhelming the model with information.

## 5. Conclusions

Hydrologic modeling remains an important decision support tool for efficiently allocating water resources and solving global water security issues, especially with the climate's non-stationarity and changing land use. The model outputs are characterized with some degree of uncertainty by underestimating or overestimating the simulated values. The best models are thus characterized with reduced uncertainties, and different methods are being adopted to achieve this. This study explored the impact of LULC maps and LAI on the performance of a fully distributed hydrologic model. Various model cases were set up to evaluate the independent effects of LAI and LULC map on model performance. The result showed that using LAI for accurate estimation of evapotranspiration values adopted for the simulation of hydrologic processes increased the model performance when the model was calibrated using discharge data. In addition to this, providing mHM with multiple LULC datasets has better performance than those with single LULC data. This study reiterated the observation of Jin et al. [1] about the impact of multiple land cover datasets on model performance. We further suggest using optimum land cover dataset when available to enable the model to capture recent and relevant LULC patches that will be utilized in the simulation of the hydrologic processes. Especially in rapidly developing landscapes, this will be necessary to capture the basin dynamics. Furthermore, in the absence of multiple or yearly LULC maps, other strategies that will reduce model uncertainty like using LAI data for accurate estimation of evapotranspiration can be adopted.

**Author Contributions:** I.O.B and M.C.D provided the morphological and meteorological data, M.C.D performed the sensitivity analysis and I.O.B performed the calibration and validation of the model. I.O.B and M.C.D performed the visualization of the result. The concept of the study and the draft manuscript were written and reviewed by I.O.B, M.C.D and A.N.

**Data Availability Statement:** All data sources are listed in Table 1. Karasu basin model set-up data were kindly provided by Oldrich Rakovec from UFZ in Leipzig, Germany. Discharge data is publicly available at the repository of State Hydraulic Work (DSI) <https://www.dsi.gov.tr/Sayfa/De-tay/744> and daily data from the gauging station was compiled and then provided by Dr Muhammet Yilmaz. Due to data license restrictions, the observed geographical, hydrological and meteorological data cannot be redistributed from the authors, but can be obtained from Rakovec et al. [28], respective organizations and/or UFZ in Leipzig, Germany upon request. The model calibration results can be retrieved from the corresponding author upon request. The source code of the mHM is publicly available at <https://doi.org/10.5281/zenodo.4575390>.

**Acknowledgments:** We acknowledge the financial support for the SPACE project by the Villum Foundation (<http://villumfonden.dk/>) through their Young Investigator Programme (grant VKR023443). The first author is supported by the European Union through the Erasmus Mundus Scholarship program and second author (MCD) is supported by the National Center for High Performance Computing of Turkey (UHeM) under grant number 1007292019.

**Conflicts of Interest:** "The authors declare no conflict of interest."

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

## References

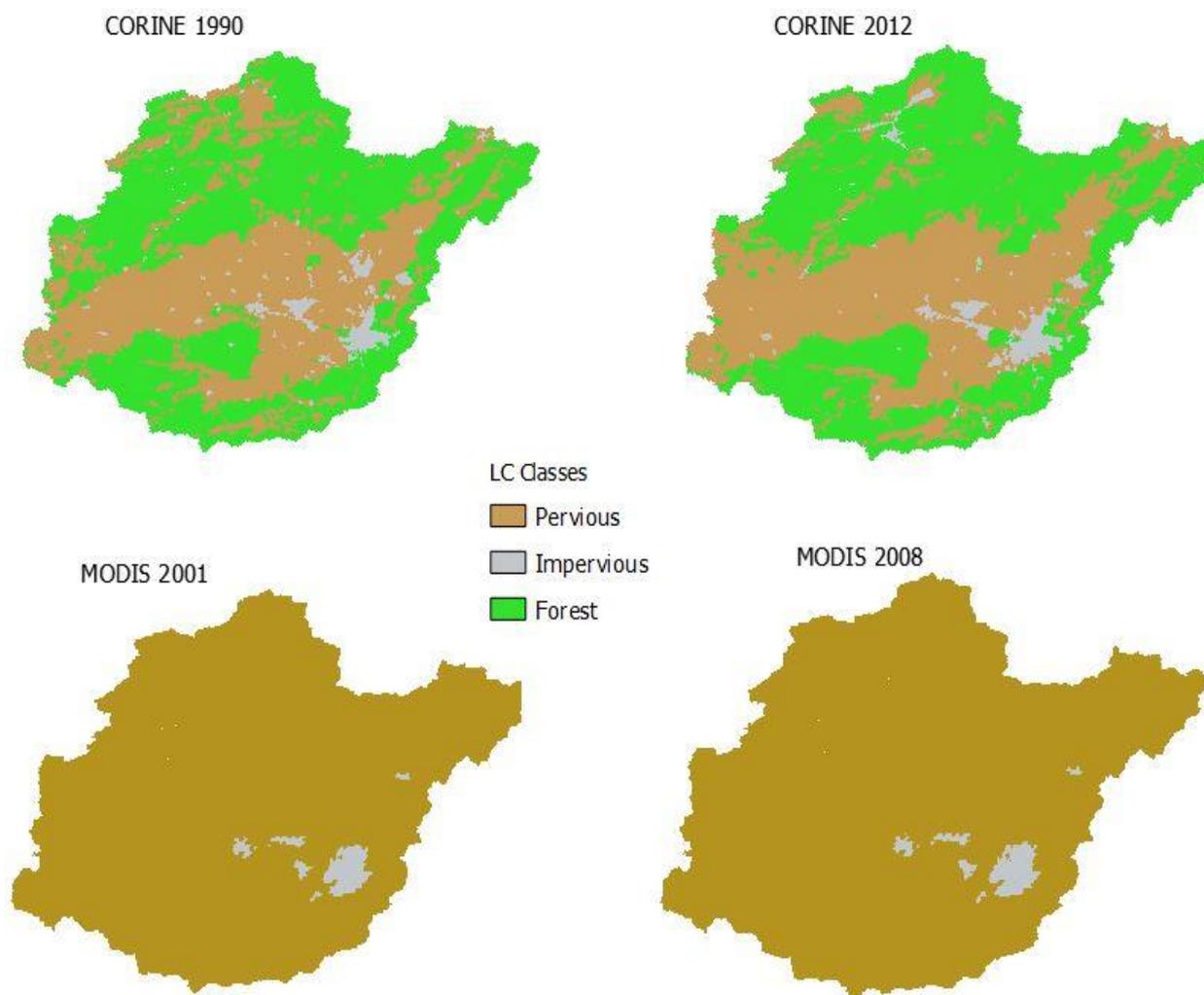
- [1] X. Jin, Y. Jin, D. Yuan, and X. Mao, "Effects of land-use data resolution on hydrologic modelling, a case study in the upper reach of the Heihe River, Northwest China," *Ecol. Modell.*, vol. 404, no. May 2018, pp. 61–68, 2019, doi: 10.1016/j.ecolmodel.2019.02.011.
- [2] A. K. Lohani, "Rainfall-Runoff Analysis and Modelling," National Institute of Hydrology, Roorkee, 2018.
- [3] C. Camargos, S. Julich, T. Houska, M. Bach, and L. Breuer, "Effects of input data content on the uncertainty of simulating water resources," *Water (Switzerland)*, vol. 10, no. 5, pp. 1–18, 2018, doi: 10.3390/w10050621.
- [4] G. K. Devia, B. P. Ganasri, and G. S. Dwarakish, "A Review on Hydrological Models," *Aquat. Procedia*, vol. 4, no. Icwrcoe, pp. 1001–1007, 2015, doi: 10.1016/j.aqpro.2015.02.126.
- [5] Y. Bouslihim, A. Rochdi, N. El Amrani Paaza, and L. Liuzzo, "Understanding the effects of soil data quality on SWAT model performance and hydrological processes in Tamedroust watershed (Morocco)," *J. African Earth Sci.*, vol. 160, no. November 2018, p. 103616, 2019, doi: 10.1016/j.jafrearsci.2019.103616.
- [6] A. Ficchi, C. Perrin, and V. Andréassian, "Impact of temporal resolution of inputs on hydrological model performance: An analysis based on 2400 flood events," *J. Hydrol.*, vol. 538, pp. 454–470, 2016, doi: 10.1016/j.jhydrol.2016.04.016.
- [7] I. Ilampooranan, J. L. Schnoor, and N. B. Basu, "Crops as sensors: Using crop yield data to increase the robustness of hydrologic and biogeochemical models," *J. Hydrol.*, no. May, p. 125599, 2020, doi: 10.1016/j.jhydrol.2020.125599.
- [8] C. Z. Li, H. Wang, J. Liu, D. H. Yan, F. L. Yu, and L. Zhang, "Effect of calibration data series length on performance and optimal parameters of hydrological model," *Water Sci. Eng.*, vol. 3, no. 4, pp. 378–393, 2010, doi: 10.3882/j.issn.1674-2370.2010.04.002.
- [9] S. Sahraei, M. Asadzadeh, and F. Unduche, "Signature-based multi-modelling and multi-objective calibration of hydrologic models: Application in flood forecasting for Canadian Prairies," *J. Hydrol.*, vol. 588, no. May, p. 125095, 2020, doi: 10.1016/j.jhydrol.2020.125095.
- [10] J. Gao, A. Y. Sheshukov, H. Yen, J. H. Kastens, and D. L. Peterson, "Impacts of incorporating dominant crop rotation patterns as primary land use change on hydrologic model performance," *Agric. Ecosyst. Environ.*, vol. 247, no. June, pp. 33–42, 2017, doi: 10.1016/j.agee.2017.06.019.
- [11] R. Orth, M. Staudinger, S. I. Seneviratne, J. Seibert, and M. Zappa, "Does model performance improve with complexity? A case study with three hydrological models," *J. Hydrol.*, vol. 523, pp. 147–159, 2015, doi: 10.1016/j.jhydrol.2015.01.044.
- [12] X. Yang, J. Magnusson, S. Huang, S. Beldring, and C. Y. Xu, "Dependence of regionalization methods on the complexity of hydrological models in multiple climatic regions," *J. Hydrol.*, vol. 582, no. November 2019, p. 124357, 2020, doi: 10.1016/j.jhydrol.2019.124357.
- [13] S. Parveen, J. Basheer, and B. Praveen, "a Literature Review on Land Use Land Cover Changes.," *Int. J. Adv. Res.*, vol. 6, no. 7, pp. 1–6, 2018, doi: 10.21474/ijar01/7327.
- [14] G. Prasad and M. V. Ramesh, "Spatio-Temporal Analysis of Land Use/Land Cover Changes in an Ecologically Fragile Area—Alappuzha District, Southern Kerala, India," *Nat. Resour. Res.*, vol. 28, no. s1, pp. 31–42, 2019, doi: 10.1007/s11053-018-9419-y.
- [15] A. Risal, P. B. Parajuli, P. Dash, Y. Ouyang, and A. Linhoss, "Sensitivity of hydrology and water quality to variation in land use and land cover data," *Agric. Water Manag.*, vol. 241, no. March, p. 106366, 2020, doi: 10.1016/j.agwat.2020.106366.
- [16] N. J. J. Bréda, "Ground-based measurements of leaf area index: A review of methods, instruments and current controversies," *J. Exp. Bot.*, vol. 54, no. 392, pp. 2403–2417, 2003, doi: 10.1093/jxb/erg263.
- [17] C. L. de Almeida, T. R. A. de Carvalho, and J. C. de Araújo, "Leaf area index of Caatinga biome and its relationship with hydrological and spectral variables," *Agric. For. Meteorol.*, vol. 279, no. March, p. 107705, 2019, doi: 10.1016/j.agrformet.2019.107705.
- [18] Z. Bian *et al.*, "Simulation of evapotranspiration based on leaf area index, precipitation and pan evaporation: A case study of

- Poyang Lake watershed, China," *Ecohydrol. Hydrobiol.*, vol. 19, no. 1, pp. 83–92, 2019, doi: 10.1016/j.ecohyd.2018.03.005.
- [19] P. Das *et al.*, "Impact of LULC change on the runoff, base flow and evapotranspiration dynamics in eastern Indian river basins during 1985–2005 using variable infiltration capacity approach," *J. Earth Syst. Sci.*, vol. 127, no. 2, pp. 1–19, 2018, doi: 10.1007/s12040-018-0921-8.
- [20] G. Uysal, A. A. Şorman, and A. Şensoy, "Streamflow Forecasting Using Different Neural Network Models with Satellite Data for a Snow Dominated Region in Turkey," *Procedia Eng.*, vol. 154, pp. 1185–1192, 2016, doi: 10.1016/j.proeng.2016.07.526.
- [21] C. Ertas, B. Akkol, C. Coskun, G. Uysal, A. A. Sorman, and A. Sensoy, "Evaluation of Probabilistic Streamflow Forecasts Based on EPS for a Mountainous Basin in Turkey," *Procedia Eng.*, vol. 154, pp. 490–497, 2016, doi: 10.1016/j.proeng.2016.07.543.
- [22] A. E. Tekeli, Z. Akyürek, A. A. Şorman, A. Şensoy, and A. Ü. Şorman, "Using MODIS snow cover maps in modeling snowmelt runoff process in the eastern part of Turkey," *Remote Sens. Environ.*, vol. 97, no. 2, pp. 216–230, 2005, doi: 10.1016/j.rse.2005.03.013.
- [23] M. Dembélé, M. Hrachowitz, H. H. G. Savenije, G. Mariéthoz, and B. Schaepli, "Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets," *Water Resour. Res.*, vol. 56, no. 1, Jan. 2020, doi: 10.1029/2019WR026085.
- [24] M. C. Demirel, J. Koch, G. Mendiguren, and S. Stisen, "Spatial Pattern Oriented Multi-Criteria Sensitivity Analysis of a Distributed Hydrologic Model," *Water*, vol. 10, no. 9, p. 1188, Aug. 2018, doi: 10.3390/w10091188.
- [25] M. Dembélé, N. Ceperley, S. J. Zwart, E. Salvadore, G. Mariéthoz, and B. Schaepli, "Potential of satellite and reanalysis evaporation datasets for hydrological modelling under various model calibration strategies," *Adv. Water Resour.*, vol. 143, p. 103667, Sep. 2020, doi: 10.1016/j.advwatres.2020.103667.
- [26] S. Höllering, J. Wienhöfer, J. Ihringer, L. Samaniego, and E. Zehe, "Regional analysis of parameter sensitivity for simulation of streamflow and hydrological fingerprints," *Hydrol. Earth Syst. Sci.*, vol. 22, no. 1, pp. 203–220, Jan. 2018, doi: 10.5194/hess-22-203-2018.
- [27] L. Samaniego, R. Kumar, and S. Attinger, "Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale," *Water Resour. Res.*, vol. 46, no. 5, p. W05523, May 2010, doi: 10.1029/2008WR007327.
- [28] O. Rakovec *et al.*, "Multiscale and Multivariate Evaluation of Water Fluxes and States over European River Basins," *J. Hydrometeorol.*, vol. 17, no. 1, pp. 287–307, Jan. 2016, doi: 10.1175/JHM-D-15-0054.1.
- [29] G. Grekousis, G. Mountrakis, and M. Kavouras, "An overview of 21 global and 43 regional land-cover mapping products," *Int. J. Remote Sens.*, vol. 36, no. 21, pp. 5309–5335, 2015, doi: 10.1080/01431161.2015.1093195.
- [30] I. Cieślak, A. Biłozor, A. Żróbek-Sokolnik, and M. Zagroba, "The use of geographic databases for analyzing changes in land cover - A case study of the region of warmia and mazury in Poland," *ISPRS Int. J. Geo-Information*, vol. 9, no. 6, 2020, doi: 10.3390/ijgi9060358.
- [31] Y. An, W. Zhao, and Y. Zhang, "Accuracy assessments of the GLOBCOVER dataset using global statistical inventories and FLUXNET site data," *Acta Ecol. Sin.*, vol. 32, no. 6, pp. 314–320, 2012, doi: 10.1016/j.chnaes.2012.09.001.
- [32] A. T. K. Tchuente, J. L. Roujean, and S. M. de Jong, "Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 13, no. 2, pp. 207–219, 2011, doi: 10.1016/j.jag.2010.11.005.
- [33] G. E. P. Box and R. D. Meyer, "An Analysis for Unreplicated Fractional Factorials," *Technometrics*, vol. 28, no. 1, pp. 11–18, Feb. 1986, doi: 10.1080/00401706.1986.10488093.
- [34] S. Razavi, R. Sheikholeslami, H. V. Gupta, and A. Haghnegahdar, "VARS-TOOL: A toolbox for comprehensive, efficient, and robust sensitivity and uncertainty analysis," *Environ. Model. Softw.*, vol. 112, no. May 2018, pp. 95–107, 2019, doi: 10.1016/j.envsoft.2018.10.005.
- [35] M. P. Clark, D. Kavetski, and F. Fenicia, "Pursuing the method of multiple working hypotheses for hydrological modeling," *Water Resour. Res.*, vol. 47, no. 9, pp. 1–16, 2011, doi: 10.1029/2010WR009827.

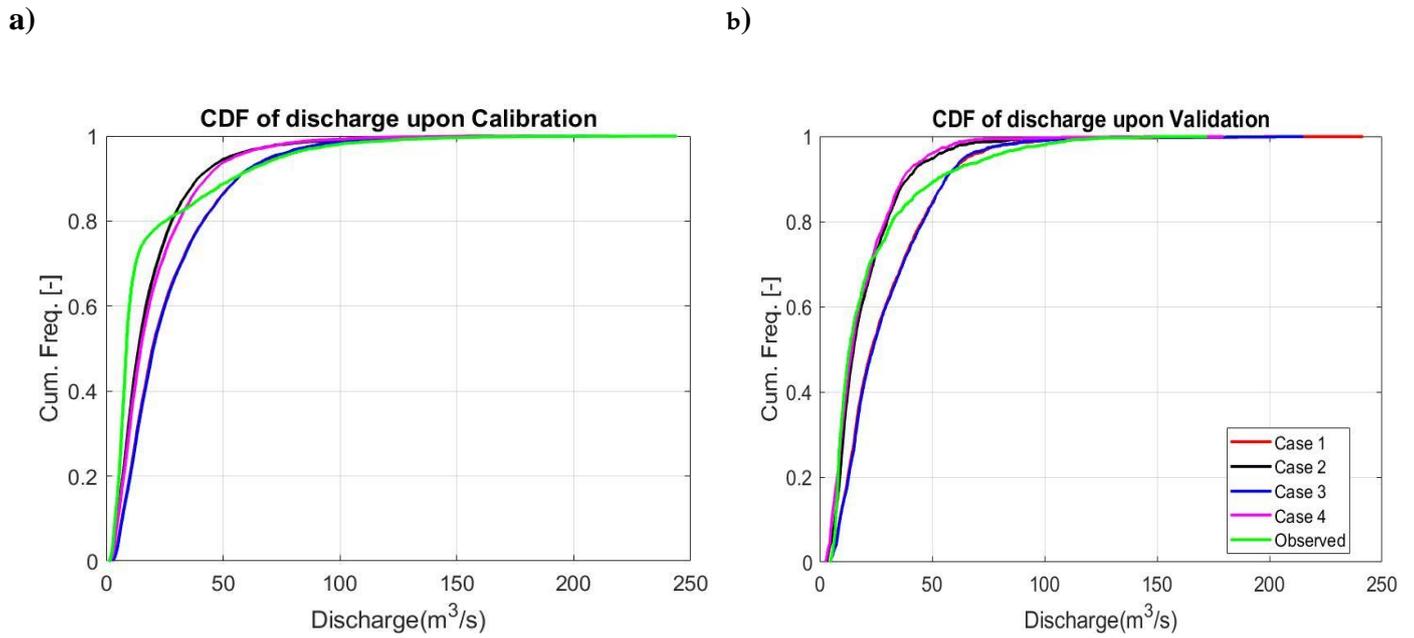
- [36] O. Rakovec, M. C. Hill, M. P. Clark, A. H. Weerts, A. J. Teuling, and R. Uijlenhoet, "Distributed Evaluation of Local Sensitivity Analysis (DELSA), with application to hydrologic models," *Water Resour. Res.*, vol. 50, no. 1, pp. 409–426, Jan. 2014, doi: 10.1002/2013WR014063.
- [37] J. Nossent, P. Elsen, and W. Bauwens, "Sobol' sensitivity analysis of a complex environmental model," *Environ. Model. Softw.*, vol. 26, no. 12, pp. 1515–1525, 2011, doi: 10.1016/j.envsoft.2011.08.010.
- [38] M. L. Chu-Agor, R. Muñoz-Carpena, G. Kiker, A. Emanuelsson, and I. Linkov, "Exploring vulnerability of coastal habitats to sea level rise through global sensitivity and uncertainty analyses," *Environ. Model. Softw.*, vol. 26, no. 5, pp. 593–604, 2011, doi: 10.1016/j.envsoft.2010.12.003.
- [39] J. Doherty, "PEST: Model-Independent Parameter Estimation, User Manual: 5th Edition," 2010.
- [40] B. A. Tolson and C. A. Shoemaker, "Dynamically dimensioned search algorithm for computationally efficient watershed model calibration," *Water Resour. Res.*, vol. 43, no. 1, Jan. 2007, doi: 10.1029/2005WR004723.
- [41] J. Chu, Y. Peng, W. Ding, and Y. Li, "A Heuristic dynamically dimensioned search with sensitivity information (HDDS-S) and application to river basin management," *Water (Switzerland)*, vol. 7, no. 5, pp. 2214–2238, 2015, doi: 10.3390/w7052214.
- [42] E. K. Jackson, W. Roberts, B. Nelsen, G. P. Williams, E. J. Nelson, and D. P. Ames, "Introductory overview: Error metrics for hydrologic modelling – A review of common practices and an open source library to facilitate use and adoption," *Environ. Model. Softw.*, vol. 119, no. May, pp. 32–48, 2019, doi: 10.1016/j.envsoft.2019.05.001.
- [43] R. Pushpalatha, C. Perrin, N. Le Moine, and V. Andréassian, "A review of efficiency criteria suitable for evaluating low-flow simulations," *J. Hydrol.*, vol. 420–421, pp. 171–182, 2012, doi: 10.1016/j.jhydrol.2011.11.055.
- [44] H. Kling and H. Gupta, "On the development of regionalization relationships for lumped watershed models: The impact of ignoring sub-basin scale variability," *J. Hydrol.*, vol. 373, no. 3–4, pp. 337–351, Jul. 2009, doi: 10.1016/j.jhydrol.2009.04.031.
- [45] H. V. Gupta, H. Kling, K. K. Yilmaz, and G. F. Martinez, "Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling," *J. Hydrol.*, vol. 377, no. 1–2, pp. 80–91, Oct. 2009, doi: 10.1016/j.jhydrol.2009.08.003.
- [46] W. J. M. Knoben, J. E. Freer, and R. A. Woods, "Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores," *Hydrol. Earth Syst. Sci. Discuss.*, no. July, pp. 1–7, 2019, doi: 10.5194/hess-2019-327.
- [47] J. C. Stagl and F. F. Hattermann, "Impacts of Climate Change on Riverine Ecosystems: Alterations of Ecologically Relevant Flow Dynamics in the Danube River and Its Major Tributaries," *Water (Switzerland)*, vol. 8, no. 12, 2016, doi: 10.3390/w8120566.
- [48] C. Chirachawala, S. Shrestha, M. S. Babel, S. G. P. Virdis, and S. Wichakul, "Evaluation of global land use/land cover products for hydrologic simulation in the Upper Yom River Basin, Thailand," *Sci. Total Environ.*, vol. 708, p. 135148, 2020, doi: 10.1016/j.scitotenv.2019.135148.
- [49] D. Deckers, "Predicting discharge at ungauged catchments: Parameter estimation through the method of regionalisation," University of Twente, 2006.
- [50] P. Bai, X. Liu, Y. Zhang, and C. Liu, "Incorporating vegetation dynamics noticeably improved performance of hydrological model under vegetation greening," *Sci. Total Environ.*, vol. 643, pp. 610–622, 2018, doi: 10.1016/j.scitotenv.2018.06.233.
- [51] S. R. Hardwick, R. Toumi, M. Pfeifer, E. C. Turner, R. Nilus, and R. M. Ewers, "The relationship between leaf area index and microclimate in tropical forest and oil palm plantation: Forest disturbance drives changes in microclimate," *Agric. For. Meteorol.*, vol. 201, pp. 187–195, 2015, doi: 10.1016/j.agrformet.2014.11.010.
- [52] Z. K. Tesemma, Y. Wei, M. C. Peel, and A. W. Western, "The effect of year-to-year variability of leaf area index on Variable Infiltration Capacity model performance and simulation of runoff," *Adv. Water Resour.*, vol. 83, pp. 310–322, 2015, doi: 10.1016/j.advwatres.2015.07.002.
- [53] N. Di *et al.*, "Modeling soil evaporation and the response of the crop coefficient to leaf area index in mature *Populus tomentosa* plantations growing under different soil water availabilities," *Agric. For. Meteorol.*, vol. 264, no. October 2018, pp. 125–137, 2019, doi: 10.1016/j.agrformet.2018.10.004.

- [54] S. Fatichi *et al.*, "An overview of current applications, challenges, and future trends in distributed process-based models in hydrology," *J. Hydrol.*, vol. 537, pp. 45–60, 2016, doi: 10.1016/j.jhydrol.2016.03.026.
- [55] S. Elga, B. Jan, and B. Okke, "Hydrological modelling of urbanized catchments : A review and future directions," *J. Hydrol.*, vol. 529, pp. 62–81, 2015, doi: 10.1016/j.jhydrol.2015.06.028.
- [56] G. Petrucci and C. Bonhomme, "The dilemma of spatial representation for urban hydrology semi-distributed modelling: Trade-offs among complexity, calibration and geographical data," *J. Hydrol.*, vol. 517, pp. 997–1007, 2014, doi: 10.1016/j.jhydrol.2014.06.019.

## Appendix A



**Figure A1:** Selected CORINE and MODIS Land cover maps of the study area



**Figure A2:** Cumulative distribution function of observed and simulated daily discharge (a) calibration (b) validation

**Table A1:** Normalized sensitivity parameters

<b>Parameters</b>	<b>Range</b>	<b>Normalized Sensitivity</b>
Rotfrcoffore	0.9 - 0.999	1.00000
pet_aforest	0.2999 - 1.3	0.48665
pet_apervi	0.2999 - 1.3	0.14471
Ptflowconst	0.6462 - 0.9506	0.14227
Infshapef	1.0 - 4.0	0.09453
ptflowdb	-0.3726 - -0.1870999	0.07084
pet_aimpervi	0.29999 - 1.3000	0.01222
ptflowclay	0.0001 - 0.00289	0.01091
rotfrcofimp	0.9 - 0.9499	0.00945
ptfkssand	0.006 - 0.02599	0.00339
ptfkconst	-1.2 - -0.28499	0.00841
ptfhigconst	0.5358 - 1.1232	0.00793
rotfrcofclay	0.9 - 0.999	0.00339
expslwintflw	0.05 - 0.3	0.00095
rechargcoef	0 - 50.0	0.00044
intrecesslp	0 - 10.0	0.00040
slwintrecks	1.0 - 30.0	0.00040
pet_bb	0 - 1.5	0.00039
pet_cc	-2.0 - 0	0.00031
Rotfrcofsand	0.001 - 0.09	0.00002