ISSN: 2639-7455

Research Article

Earth & Environmental Science Research & Reviews

Building Tools for further Investigating Acid Mining Production: Intercomparison of Four Hydrological Model Versions through a Scoring Technique on the Niger River Basin, in west Africa

Salif Koné [a*], Gil Mahé[b], Fatogoma Bamba[a], Jean Emmanuel Paturel[b], Alain Dezetter[b] and Eric Servat[c]

[a]*Department of Geology and Mines, National School of Engineers - Abderhamane Baba Touré (ENI-ABT), 410, Av. Van Vollenhoven, BP: 242 – Phone: (+223) 76 39 60 09 / (+223) 20 22 27 10 36, Bamako, Mali. http://orcid.org/0000-0001-7001-7447*

[b]*HydroSciences Montpellier (HSM), Faculty of Pharmaceutical and Biological Sciences, University of Montpellier, Hydropolis Building, 15, avenue Charles Flahault, Montpellier CS 13 14491, 34093 Montpellier cedex 05, France.*

[c]*UNESCO ICIREWARD Center, University of Montpellier - CC 57, 163 rue Auguste Broussonnet, 34090 Montpellier, France.*

***Corresponding Author**
KONE Salif, Ph.D, Abderhamane Baba Toure National School of Engineers Bamako, MALI.

Submitted: 2023 July 02; **Accepted**: 2023 July 30; **Published**: 2023 Aug 18

Citation: Koné, S., Mahé, G., Bamba, F., Paturel, J. E., Dezetter, A., et al. (2023). Building Tools for further Investigating Acid Mining Production: Intercomparison of Four Hydrological Model Versions through a Scoring Technique on the Niger River Basin, in west Africa. *Envi Scie Res & Rev, 6*(4), 566-594.

Abstract

We modify a framework to model time series of discharges, on watersheds of the Niger River and its tributaries, over 1901-2020. This framework is built from a French hydrological model – Genie Rural with 2 parameters at Monthly time step – GR2M. The first parameter, X1, concerns the input variables. This parameter is applied to the rainfall and the evapotranspiration data; its reciprocal, 1/X1, is applied to the Soil's Water Holding Capacity data (assimilated to reservoir height). The second parameter, X2, is applied to the time series of discharges, which is the output variable in modelling. Modifications lead to a new model version called SimulHyd, which stands for Simulation of Hydrological systems. A weighting rule on grids is inserted into the framework to create a gridded variant of semi-distributed modelling from a lumped model. Following a World Meteorological Organization's goal, we develop a scoring technique for intercomparisons of model versions on both one single and several catchments. Over the 16 watersheds, a three-model intercomparison shows the preponderance of SimulHyd semi distributed (with 117.83 out of 192 total scoring-points, equivalent to 61.37%) over both GR2M non-distributed (with 61.33 scoring-points, or 31.94%) and GR2M semi-distributed (with 12.83 scoring-points, or 6.68%).

Keywords: The Niger River; Hydrological Modelling; Conceptual framework GR2M and SimulHyd; intercomparison; Scoring Technique; Accuracy.

1. Introduction

Hydrological modelling aims at producing diverse gridded simulation variables during each time step and on different spatial scales. It folds in two major approaches: physical-based modelling and conceptual modelling with reservoirs [1-3]. Between both major approaches, a range of other mixed approaches exists: analogic modelling, mathematical modelling, and statistical or stochastic modelling from probability theories. In addition, hydrological models are distributed on two major pole structures: some are lumped [4] and some others are distributed [5-6]. Hybrid structures like semi-distributed and semi-lumped are common in literature; for instance, [7] applied a likely semi-distributed modelling in the Free State of Saxony (USA) in using the model "BROOK90 (R version)". However, [8] made another classification. They discriminated mathematical models and systematic models. Performing intercomparison investigations and developing efficiency criteria are ways to make a methodical selection of models or model versions for a selected project [9].

1.1. A succinct timeline on hydrological modelling

A World Meteorological Organization's report [10], in 1975, indicated the need to develop objective criteria of models' performances that can be applied in comparing one or several models both on a same catchment and on several catchments at once [11]. In 1983, [12] made propositions to improve rainfall-runoff model in Australia. Five years afterward, Pilgrim et al. [13] assembled problems of rainfall-runoff modelling.

In 1994, [14] presented "CELMOD5 - a semi-distributed cell model for conversion of rainfall into runoff in semi-arid watersheds". Three years later, [15] presented a "distributed Rainfall / Runoff" hydrological modelling in the watershed of Guadiana (in South of Portugal). In 2000, Hernandez et al. [16] was interested in "land cover and rainfall spatial variability" impact on runoff response. Two year later, [17] published works on the Nakambe River in establishing the climatic and anthropogenic impacts on the flow regime, and [18] exposed "An alternative IUH" (instantaneous unit hydrograph) "for the hydrological lumped models". The following year, [19] paved the way to investigate, in West Africa, the outcome of hydrological modelling receiving some distributed input data. In 2004, [20] theoretically compared a lumped and two semi-distributed model versions using Nash criterion [21] to assess efficiency. A year later, [22] presented works on surface water estimates through hydrological modelling using a French hydrologic model – Genie Rural with 2 parameters at Monthly time step – GR2M. In 2006, [23] published works on assessment of climatic scenarios through hydrological modelling.

Pursuing research undertook by [19], we developed an intercomparison through a scoring technique on model versions that stem from each other (Kone [24]). Further, [25] worked on how embedding available versions of a model to produce a single optimal model version.

Relatively to [26] succeeded in presenting the "fuse: an R package for ensemble Hydrological Modelling" meanwhile [27-28] and [29] presented an "assemble of GR model" under the packages airGR and airGRteaching. Further, [30] published their results obtained through using airGR packages, while [31] explored machine learning to assess hydrologic model performances in 2022.

1.2. From perspectives to problem statement

While [20] theoretically compared a lumped model and two semi-distributed model versions, Yao et al. [32] compared distributed and lumped hydrological models similarly to Santos et al. [33]. Theses authors used criteria other than a scoring technique. [34] and Askew [35] previously performed hydrological model intercomparison studies following recommendations of the World Meteorological Organization [10]; however, neither they scarcely explored a scoring technique specifically on model versions that stem from each other nor had a same gridded semi-distribution structure as presented in this article.

Comparative analyses on hydrological models are widely performed since the WMO's 1975 statement on model performances [10]. In 2022, [36] undertook comparative studies between lumped and semi-distributed models through conceptual hydrological modelling. Three years earlier, [37] performed a comparison of hydrological simulations using synthetic rainfall data. [38] includes an approach called "event scale", which has inspired [39] in constructing a comparison method of multiple models through criteria that involved a Monte Carlo simulation.

We previously used a French hydrological model – Genie Rural with 2 parameters at Monthly time step – GR2M to develop a gridded semi-distributed version of a lumped model, which leads to model version intercomparison issues [23-24]. Using the GR2M model, four model variants were generated through two approaches. First, we modify GR2M to produce a threeinput variable model – instead of two – (called SimulHyd, which stands for Simulation of Hydrological systems). Second, we change a lumped model version to produce a gridded semidistributed model version. GR2M is a hydrological model developed by Edijatno [2] and further by Kabouya & Michel [22] from French laboratories, while SimulHyd model is our variant of GR2M. Hence, the assessment of the four model versions must be performed on the studied catchments. These model versions, which derive from each other, are GR2M non-distributed (lumped form), SimulHyd non-distributed, GR2M semi-distributed (lumped form), and SimulHyd semi-distributed.

Appropriating the previously mentioned World Meteorological Organization's target [10] leads to establish an intercomparison method suitable to approaches that consists of varying a single model to obtain four or more different model versions [38]. Moreover, this established intercomparison method allows to manage cases where an equifinality occurred between models when calculating conventional criteria, such as Nash criterion [21]. Diverse aspects of hydrological modelling have been subject to comparison or intercomparison studies [40]. Compared modelling through statistical regressions for forecasting baseflow, while [41] performed comparison in post-process stages working on streamflow. In Gosling et al. [42], comparison was about climate change scenarios through varying temperature, whereas in [43] the intercomparison concerned model robustness relative to streamflow forecasting. An intercomparison between "a lumped model and a distributed model" was performed by Liu et al. [44] while regional scale hydrological models were under intercomparison a year later by [45] – with a target to assess climate change impacts. Authors like Sittner et al. [46] and [47] furthermore performed hydrological models intercomparison studies according to WMO recommendations [10], as previously stated. [48] applied a multi-criterion validation to a semidistributed conceptual model while [49] undertook comparative studies of two hydrological models in the northwestern part of Algeria.

*Table 1 : Watersheds and Simulation periods according to runoff data choices. The first period end is preferably 1970 if data allowed that and the beginning of the second period is 1971. * Number of grids concerned by a basin. ** Codification used by the French research institute IRD (Institute of Research for Development). *** A third period running from 1996 to 2018 is associated to the hydrometric station of Koulikoro; on this watershed, observed discharge time series will be used from 1907 to 2018.*

Basins****	Characteristics			Discharge time series						Precipitation (mm)		
				First period			Second period					
	Area [Km^2]	Grid * 0.5x 0.5	Station code IRD **	Start year	End year	Gap [%]	Start year	End year	Gap [%]	1901 to 1970	1971 to 2020	Relative variation
Banankoro	73458	49	1271500110	1971	1980	46	1981	1999	23	162	146	-10
Baranama	6608	8	1171503506	1973	1980	55	1981	1995	9	192	177	-8
Baro	13108	15	1171501805	1950	1970	8	1971	1995	51	181	149	-17
Dialakoro	70591	49	1171500110	1957	1970	39	1971	1980	57	164	147	-10
Faranah	3178	4	1171500115	1969	1980	27	1981	1995	8	202	182	-10
Gouala	33075	26	1271502005	1961	1970	53	1971	1978	74	154	142	-8
Iradougou	1824	5	1091504003	1965	1970	0	1971	1995	4	146	136	-7
Kankan	10080	14	1171501705	1950	1970	3	1971	1995	4	184	169	-8
Kerouane	1423	4	1171501707	1973	1980	53	1981	1995	59	213	196	-8
Kissidougou	1400	5	1171501810	1960	1970	17	1971	1995	29	199	180	-9
Koulikoro***	120603	73	1271500142	1950	1970	0	1971	1995	1	154	141	-8
Kouroussa	17201	15	1171500120	1950	1970	22	1971	1995	42	177	158	-10
Mandiana	21952	16	1171501201	1957	1970	23	1971	1995	33	168	156	-8
Ouaran	19777	21	1171502505	1957	1970	25	1971	1985	18	141	125	-12
Siguiry	71064	50	1171500130	1954	1970	16	1971	1995	36	164	148	-10
Tinkisso	6569	10	1171502510	1958	1970	43	1971	1995	48	154	137	-11

Our investigation is about how comparing the performance of hydrological model versions on one single or multiple catchments at once; these versions specifically derive from each other.

Results from these model versions are so close that intercomparison methods from literature, which use commonly known criteria, detect scarcely no difference between them. We hence develop a scoring technique in intercomparing model versions that stem from each other.

2. Materials and Methods

2.1.Watersheds and Data

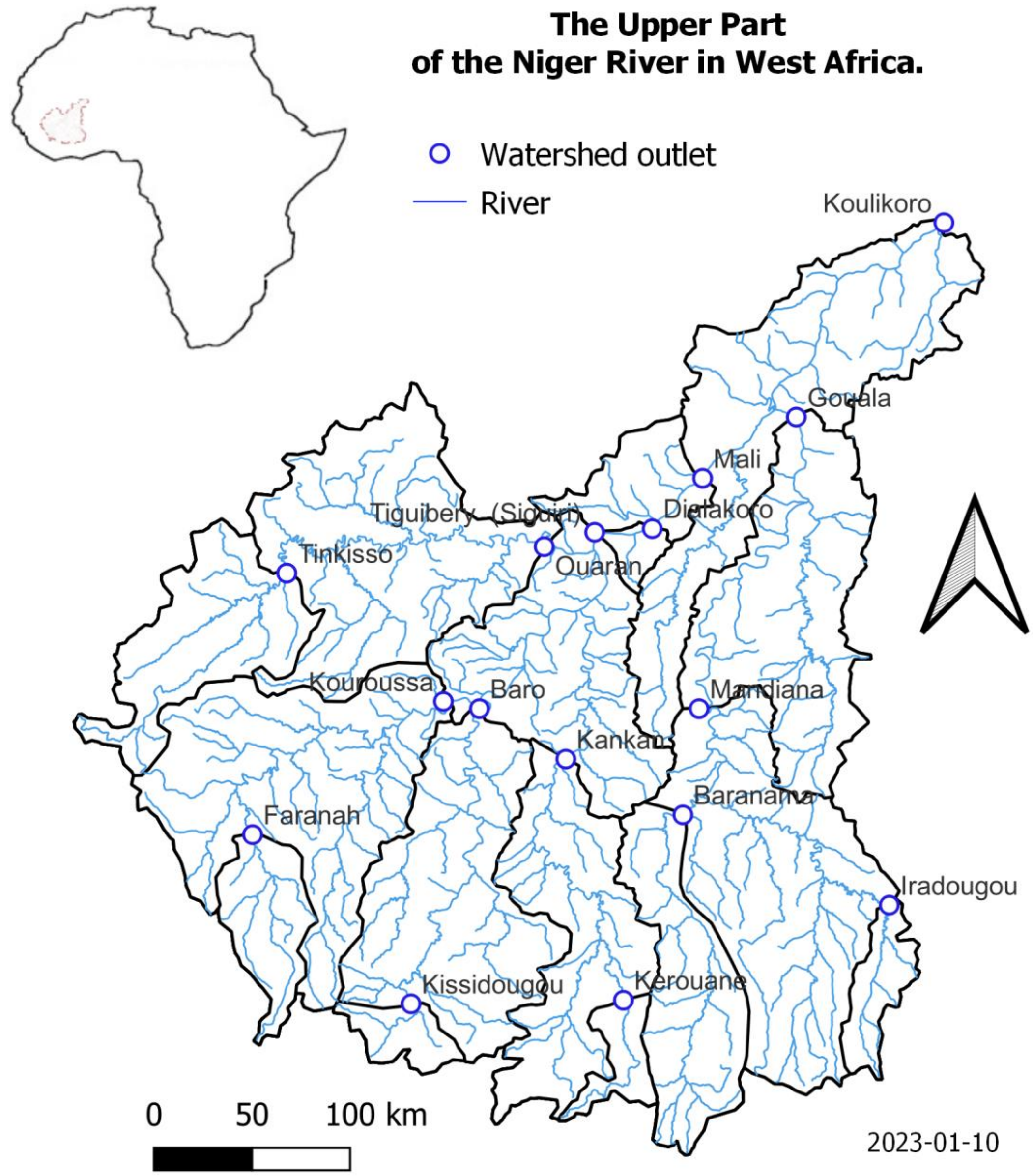


Figure 1: Maps of 16 watersheds on the Niger River, in West Africa (Boyer et al. [50]; Koné [24,38]). The rivers order is decided in ESRI software, the main river being the first. Only the first three order of rivers are drawn on the map.

This study focuses on the Niger River basin starting from Guinea downward to the hydrometric station of Koulikoro in Mali, covering some 15 other hydrometric stations (Figure 1 and Table 1); its drainage area at this station is 120 603 km². This basin spreads geographically from the Fouta Djallon forested Highlands in Guinea, and the hilly Northwestern Côte d'Ivoire, to the forested savannah plains of the Southwest of Mali. This basin connects the population from these countries both economically and culturally. The Climatic Research Units (CRU) published variable indexes in matrix 360.latx720.long under the designation cru_ts3.20 – the current version is cru_ts4.06 (see online at the British Atmospheric Data Center - BADC). Using the R programming language, we extracted the rainfall data on the grids 0.5x0.5 degree and built a precipitation database called PluieCRU. Similarly, we built an evapotranspiration database called EtpCRU. The SIEREM database also supplies similar Precipitation data called PluieIRD. As reported by Boyer et al. [50], SIEREM stands for *Système d'Informations Environnementales sur les Ressources en Eaux et leur Modélisation.*

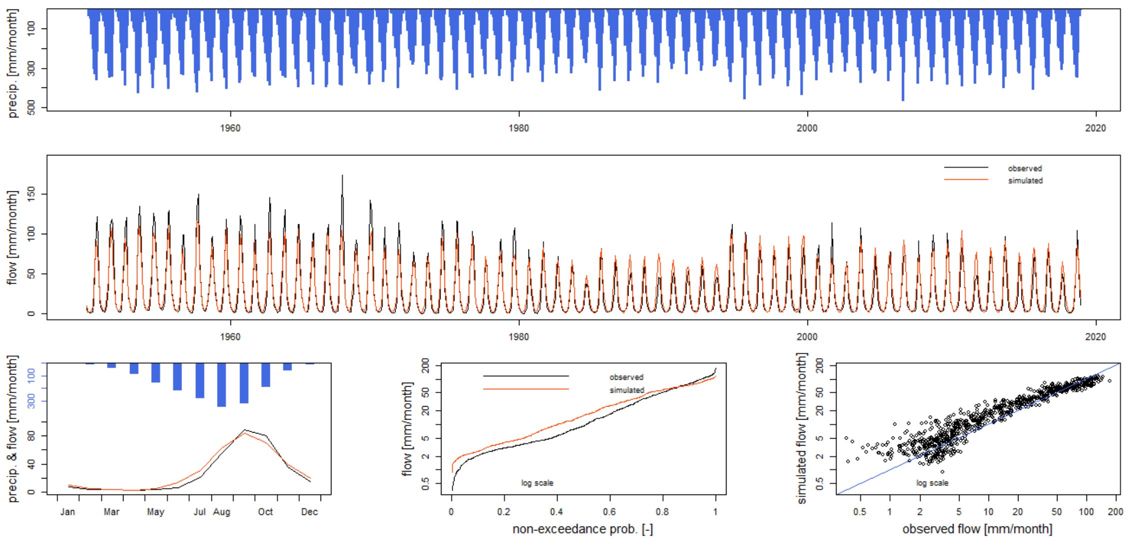


Figure 2: Modelling data at the hydrometric station of Koulikoro in Mali with a drainage area 120 603 km2 over ears 1950 to 2018. In the first row, the monthly average rain (in millimeter). In the second row, the monthly average time series of discharges (in millimeter): observation data (in black) and simulation through SimulHyd semi-distributed model (in red). In the third row and from left to right: the inter-annual monthly average rainfall (blue histogram) and runoff (curves); runoff rate non-exceedance probability; diagram observed runoff data versus simulated runoff data. Blue color pertained to rainfall, black to observed runoff, and red to simulated runoff.

We extracted the soil's Water Holding Capacity (WHC) from SIEREM database [50] in using raster's attribute tables; nevertheless, [51] handles soils data differently. The IRD's codification is used to identify watershed outlets as illustrated on the Table 1. This table concerns the set of 16 time series of discharges selected on the Niger River and its tributaries. In 1986, Brunet et al. [52] previously published a part of these pre-cited time series on the Niger River.

A visualization capability, included in airGR packages [27-28], permits to draw the Figure 2 that shows hydro-meteorological and hydrometric time series at the hydrometric station of Koulikoro. Analyses based on the Table 1 and Figure 2 corroborate a previous remark from literature: the year 1970 is a turning-point [23] in West Africa as dropping occurred in precipitation since this year up to around 1993.

2.2. Models

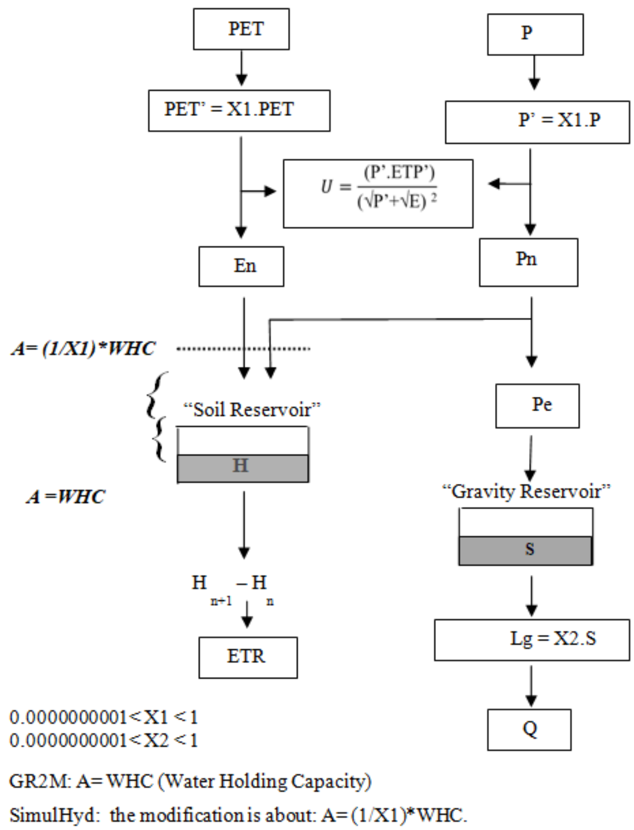


Figure 3: A French hydrologic model – Genie Rural with 2 parameters at Monthly time step – GR2M –the version of Kabouya [22]. Its modification leads to the SimulHyd model, which stands for Simulation of Hydrological systems [24,38]. X1 and X2 are the model parameters.
P = Precipitation. PET = Potential Evapotranspiration. Modulated entry data are P' = X1.P and E' = X1.E. The water intercepted by canopy and other obstacles is U. The net rainfall and the net evapotranspiration are respectively: Pn = X1.P – U and En = X1.E – U. The levels of the soil reservoir are the H. The complement of Pn is Pe = Pn – (H1 – H). The estimate of the real evapotranspiration is RET. The water levels of the gravity reservoir are the S. Water delivered (during the month) by the gravity reservoir is Lg. The time series of discharge is Q. See appendix 1.

GR2M model went through diverse modification propositions leading to different formulations [22, 53] without changing its anagram that stands for Genie Rural with 2 parameters running at monthly time step. It was developed inside INRA – Institut National de Recherche Appliqué in French language, which was previously known as CEMAGREF - Centre d'Étude du Machinisme Agricole et du Génie Rural des Eaux et Forêts in French language. Edijatno formulation of GR2M [2], implemented by the laboratory HydroSciences Montpellier is used in this work along with a new version called SimulHyd (Figure 3 and appendix 1).

GR2M is a conceptual framework with two reservoirs (called Soil reservoir and Gravity reservoir) and its functioning is like a funnel that receives water in its superior part and loses it in outflow through its bottom part. Previously, the initial filling up of the first reservoir (Soil reservoir) was fixed to a constant value depending on the world region (e.g., A=200 in France) and expressed in millimeters. In the version of Edijatno [2], GR2M is a model with two parameters (X1, X2), with a constant A, and with two entry data, the rainfall, and the evapotranspiration. Further, works undertook at the laboratory HydroSciences Montpellier drove to the proposition of a physical quantity in replacement of the constant A. This quantity noted WHC in further works [23-25], characterizes the capacity of soil to holding water. WHC is thus considered as a third entry data of the GR2M model. Enlarging the influence of the first parameter to the three input data (rainfall, evapotranspiration, Water Holding Capacity), leads to meaningful improvements in hydrological modelling with GR2M. We built a new model version in applying the reciprocal of the first parameter to the Soil reservoir. The constant A is hence replaced by the term (1/X1), WHC, which leads to the proposition of a new version of GR2M named SimulHyd (Simulation of Hydrological Systems). We explain a weighting rule of different grid influences supporting a semi-distributed modelling structure (Formula 1).

2.3. Semi-globality - a gridded variant of semi-distributed modelling - and weighting factor

This paper adopts the term (gridded) semi-distributed modelling to designate a modified spatial process introduced in lumped modelling – This distribution on grids were earlier called as semiglobality at HydroSciences Montpellier – a French laboratory. Therefore, our semi-distributed models produce gridded intermediary output data on grids: it turns lumped models to semidistributed models with gridded intermediary output. [20] studied three model structures that are “lumped Approach”, “[Semi distributed Approach]” and “[Semi lumped Approach]”; the lasts produce simulation output data on meshes in using distributed input. We, therefore, modify the last concept of these authors to produce intermediary output data on grids – mimicking a distributed modelling.

In its gridded semi-distributed structure, GR2M is a model with two reservoirs considered m times; m is the number of entry grids covering a watershed [23-24]. An analogy of its functioning would be a "puppet" funnels having several superior openings sharing one lower opening. Each superior opening, which represents an individual input grid, receives the corresponding modelling entry data (Rain, PET, and WHC). Using the entry data of every individual grid to estimate separately the time series of discharge at the outlet of the whole watershed constitutes the philosophical concept behind these semi-distributed models.

Hence, we run a lumped model on the global watershed using the data from one grid, and that is repeating several times according to the number of grids. The global model runs m times on the watershed to produce m time series of discharges. It is thus necessary to define a weighting rule (Formula 1) for averaging these m time series of discharge [24]. The result represents the simulation output from the semi-distributed modelling. A fundamental problem is the quantification of grid influences in these semi-distributed modelling. Moreover, both the choice of the weighting rule (Formula 1) and the criteria for reaching an optimal weighting factor are also crucial. We maintain a criterion that permits to a semi-distributed model to produce a higher value of the objective function and concomitantly to reach the same set of optimal model parameters (X1, X2) as obtained with the lumped model (non-distributed). Kone [24] demonstrated that weighting rule through WHC values is far more impacting than through both precipitation and potential evapotranspiration.

$$\beta_i^t = \frac{\left(WHC_i^t - \overline{WHC^t}\right)^2}{\sum_{i=1}^{N}\left(WHC_i^t - \overline{WHC^t}\right)^2} \qquad (1)$$

β_i^t *is a weighting rule built from Soil’s Water Holding Capacity data (WHC)*

Index number “i” concerns spatial grids.

N is the total number of grids covering a river basin.

Index number “t” pertains to time steps (in cases the soil surface evolution is considered)

2.4. Intercomparison Method – A Proposed Scoring Technique

2.4.1. Justification

Model comparison studies most often conclude on which percentage of basins a model A is more performing than a model B, and not which percentage a model A is better than a model B either on a single watershed or on several watersheds at once [11,53]. These methods usually compare models based on a single known characteristic such as Nash criterion [21], sensitivity to samples, or robustness [56]) but it is difficult using them to discriminate which model is most efficient than others based on differences showed through their first or second decimal. Therefore, we propose an intercomparison method of models through a scoring technique, which combines existing criteria and a proposed new protocol. This technique works even in cases where models do not show straightforward differences in terms of efficiency using conventional comparison methods. It aims to compare two or several models at once, on one or several catchments, and therefore addresses a 1975’s WMO proposed challenge on hydrologic model intercomparison issues [10-11].

Moreover, a range of other model structures exists between both poles presented in introduction (lumped and distributed). For illustration, [20] studied three model structures that are “Lumped Approach”, “[Semi-distributed Approach]” and “[Semi-lumped Approach]”. We adopt the term “structure” instead of the terms “approach” and develops a semi-lumped modelling with output on grids, instead of on meshes (see sub-section 2.3). Therefore, the manuscript proposes a gridded semi-lumped modelling adopted as a semi-distributed modelling on grids. Consequently, a suitable intercomparison method to assess a lumped model turned to suck a gridded semi-distributed model is needed as we develop it.

2.4.2. Methodology and components of the intercomparison method

It is possible to summarize the modelling process to the bias-variance trade-offs [54-55]. Models with less bias (or a greater number of parameters) show most often a great flexibility to simulate discharge time series during their calibration phases, on the other hand they show a big dependence to the calibration samples; their parameters vary more when changing the calibration period. This bias-variance trade-offs leads to the essential points in model comparison studies. First, the suppleness measure of a model (or performance in calibration) is estimated through efficiency criteria and, the measure of the actual performance (in validation) using samples that did not served during the calibration phase. Second, the dependence of model parameters to samples are estimated; if parameters change much in value from a first calibration period to a second, it means the model has a high variance: both calibrated forms of the model would drove to two distinct results.

Our intercomparison method integrates information from the Nash criterion compared results [21], from the robustness criteria through a double samples technique [56], from a reformulation of the Nash-Garrick criteria in our work [11], and eventually from information based on a compared evaluation of model parameter sensitivity to samples through a proposed double calibration protocol (DCP). Therefore, we present two protocols: the double calibration protocol, which serves as a component of the intercomparison protocol, which is the second. Finally, the intercomparison method succeeds in the calculation of scoring-points distributed between the models. In this paper, the sum-up of the intercomparison scores is 12 on a single watershed. Further consideration is based on the following relation: the Root Means Square Error equals the bias component (power 2) plus the variance component (RMSE = Bias2 + Variance) [54-55].

2.4.3. The Double Calibrations' Protocol (DCP)- a proposition

Models could differ from each other through their variance, which is the sensitiveness of a model to samples. The Double Calibration Protocol (DCP) is the variance part of the biasvariance trade-offs in modelling; 3 intercomparison scoring-points (out of 12) are thus allocated to this component according to the preceding RMSE's relation.

Our protocol intercompares model parameter sensitivities to samples. It consists in defining two distinct calibration periods, the first calibration period (CP1) is the usual calibration period, and the second calibration period (CP2) coincides with the usual validation period. During calibration, for each model in competition, we consider the difference between the two values of the parameter X1: (a) that obtained on the first calibration period (CP1), (b) and the other obtained on the second calibration period (CP2).

The model with the least value of this difference gains one scoring-point, and each of the other model(s) will have zero point, so on with parameter X2. At this step, these operations of the Double Calibration Protocol (DCP) take two scoring-points. In validation, the difference between the values of Nash criterion, obtained during the 'cross-validation' of a model, between the first and second calibration periods (CP1 and CP2), is considered.
We perform the model validation on the second period using the set of model parameters obtained from the calibration process on the first period, vice-versa. The model with the least

value, of the difference of both precedent processes, in terms of Nash criterion, gains one scoring point, and each of the other model(s) are given zero point. Eventually, the double calibrations' protocol puts three scoring points in competition. In a similar case of a three-model competition, where more models could gain a scoring-point leading to an equality issue: a set of two-model competitions could be performed to discriminate which model is the less sensitive to samples.

2.4.4. A reformulation of the Nash-Garrick criteria [11]- Nash-Garrick modif

$$R_m^2 = 1 - \frac{\sum_{t=1}^{T}\left(Q^{m,t} - Q^{m,t}\right)^2}{\sum_{t=1}^{T}\left(Q^{m,t} - \overline{Q^m}\right)^2} \qquad (2)$$

R_m^2 is the Nash-Garrick efficiency criterion [11].
Index number "t" concerns time steps in years.
Index number m concerns months (e.g., m = January, or February… or December).
T is relative to the total number of years in the modelling period.
$Q^{m,t}$ is the observed time series of discharges during the months m, and at the year t.
$Q^{m,t}$ is the simulated time series of discharges during the months m, and at the year t.
$\overline{Q^m}$ is the average discharge during the months m; we use its twelvefold in this paper.

Garrick et al. [11] proposed two criteria on model efficiency assessments, which was motivated by a report published in 1975 by the WMO on intercomparison issues of rainfall-runoff models [10]. This report of the WMO indicates the need to develop objective criteria for model performances that are appropriated to intercompare: (i) a set of models on a single watershed, (ii) the efficiency of a single model on a set of watersheds, (iii) and the performances of a set of models on a set of watersheds. Their work aimed to improve the Nash criterion while proposing in a first modification to replace the reference model by new reference models that carry the information of the intra-temporary variability in observed discharge data between the twelve months (2). Nash-Garrick criteria is accurate; nevertheless, its passive calculation on simulation results obtained using Nash criterion drove to weird results. When calculating the Nash-Garrick criteria, we deliberately underestimated the twelve inter-annual values calculated on discharge time series.

For instance, we divide the sum of the observed runoff during the January months by the number of all months in the modelling period – months with no observation data are excluded (2). Thus, these calculated quantities differ from the twelve inter-annual values by twelvefold approximately (e.g., m = January, or February…or December). Using these quantities, instead of the twelve reference models, leads to range of values like that obtained with either Nash criterion (Formula 4) or robustness calculation (Formula 3). Our modification on the Nash Garrick criteria reduces its twelve reference models in dividing each of them approximately by twelve in ideal cases. In this paper differential values of Nash-Garrick criterion and robustness are used − modifications are no impacts on results, except on figure 4.

2.4.5. Known criteria in literature: Nash criterion by Nash et al. [21] and a robustness calculation (a double samples technique by Klemes [56] and formulated by Mouelhi et al. [57])

$$R^2 = 1 - \frac{\sum_{t=1}^{T}\left(Q^t - Q'^{t}\right)^2}{\sum_{t=1}^{T}\left(Q^t - \overline{Q}\right)^2} \quad (3)$$

R^2 is the Nash efficiency criterion [21].
Index number "t" concerns time steps.
T is the total time.
Q^t is the observed time series of discharges.
Q^t is the simulated time series of discharges during the months m, and at the year t.
$\overline{Q}$ is the average discharge.

$$F = 1 - \frac{\sum_{t=T1+1}^{T2}\left(\sqrt{Q^t} - \sqrt{Q'^{t}}\right)^2 + \sum_{t=T2+1}^{T}\left(\sqrt{Q^t} - \sqrt{Q'^{t}}\right)^2}{\sum_{t=T1+1}^{T}\left(\sqrt{Q^t} - \frac{1}{T-(T1+1)} \times \sum_{t=T1+1}^{T}\sqrt{Q^t}\right)^2} \quad (4)$$

F is the robustness through a double samples' technique by Klemes et al. [56], as presented in 356 Mouelhi et al. [57]
F' overestimates the value of calculated robustness, and its reference model unfortunately 358 includes simulation result; it is not calculated in this paper.
Index number "t" concerns time steps.
T is the total number of time steps.
T_1 is the number of time steps in the running period.
T_2 is the end of the calibration period, and the start of the validation period is T_2 +1.
Q^t is the observed time series of discharges.
Q^t is the simulated time series of discharges.

Nash criterion (3) is based on the principle of least squares [21] and is commonly used in model assessments. The Double Samples Technique (DST) permits to assess model performances through a robustness calculation formulated by [56] and presented in [57]. Its formulation (4) contrasts with Nash criterion (3) by combining observed and simulated results from two different periods, calibration, and validation. Moreover, a variant of robustness calculation (5) overestimates calculated values compared to Formula in (4); in addition, its reference model unfortunately includes simulation results; it is not calculated in this paper.

Table 2: Template of intercomparison method – a scoring technique. An (number)-models competition where n models are compared on one or several catchments at once. [1]Double Samples Technique, [2]Nash criterion differences assessment through a "double calibration protocol". Row one through row nine are described in the reference [38].

<table>
<tr><td rowspan="2">watersheds</td><td rowspan="10">Periods: Calibration & Validation</td><td colspan="2" rowspan="2">Actions \ Models</td><td colspan="5">Number-model Competition</td></tr>
<tr><td>Model 1</td><td>Model 2</td><td>Model 3</td><td>........</td><td>Model n</td></tr>
<tr><td rowspan="5">Bias</td><td colspan="2">Evaluation Nash Garrick(Calibration)</td><td colspan="5">Row 1</td></tr>
<tr><td colspan="2">Evaluation Nash Garrick (Validation)</td><td colspan="5">Row 2</td></tr>
<tr><td colspan="2">Calibration Nash</td><td colspan="5">Row 3</td></tr>
<tr><td colspan="2">Validation Nash</td><td colspan="5">Row 4</td></tr>
<tr><td colspan="2">Robustness Method DST [1]</td><td colspan="5">Row 5</td></tr>
<tr><td rowspan="3">Variance</td><td rowspan="3">Less sensitivity to Samples</td><td>X1</td><td colspan="5">Row 6</td></tr>
<tr><td>X2</td><td colspan="5">Row 7</td></tr>
<tr><td>Nash (val.) [2]</td><td colspan="5">Row 8</td></tr>
<tr><td colspan="4">Scoring-points</td><td colspan="5">Row 9</td></tr>
</table>

2.4.6. Implementation of the intercomparison method: a structured table as a scoring technique

Table 3: Models intercomparison scoring technique table at the hydrometric station of Koulikoro on the Niger River, over both periods 1950 to 1970 and 1971 to 1995 (first round).
[1]Double Samples Technique, [2]a double calibration protocol on Nash criterion differences, [3] Non-Distributed model, and [4] Semi-Distributed model.

Koulikoro		Models / Actions		Four-model Competition				Three-model Competition			Two-model Competition	
				GR2M ND [3]	SimulHyd ND	GR2M SD [4]	SimulHyd SD	GR2M ND	GR2M SD	SimulHyd SD	GR2M ND	GR2M SD
Bias	Periods: 1950 to 1970 & 1971 to 1995	Evaluation Nash Garrick (Calibration)		0.25	2	0	0.75	0.5	0	2.5	1.75	1.25
		Evaluation Nash Garrick (Validation)		0	2	0.25	0.75	0.5	0.25	2.25	1.75	1.25
		Calibration Nash		0	1	0	0	0	0	1	1	0
		Validation Nash		0	1	0	0	0	0	1	0	1
		Robustness Method DST [1]		0	1	0	0	0	0	1	1	0
Variance		Less sensitivity to Samples	X1	0	1	0	0	0	0	1	1	0
			X2	0	0	1	0	0	1	0	0	1
			Nash (val.) [2]	0	0	0	1	0	0	1	0	1
Scoring-points				0.25	8	1.25	2.5	1	1.25	9.75	6.5	5.5

The table approach (see the section 3.2.1) allows to summarize the remarks made on the compared model versions in competition. The intercomparison template structure on (Table 2) and the simulation protocol are thoroughly exposed in in thesis [38] (p.61-63n, in French). 12 scoring-points are available for competing model versions on a single watershed, and this number is multiplied according to the number of the watersheds on which intercomparison is undertaken.

3. Results and discussion

3.1.Classical presentation of models' comparison results

The Figure 4 presents the averaged value of both the Nash criterion and the robustness on the 16 watersheds; on it, cross-validation are performed between two periods, before 1970 and after 1970 (Table 1). It shows results respectively from these four models: GR2M Non-Distributed, SimulHyd Non-Distributed, GR2M Semi-Distributed and SimulHyd Semi-Distributed. The set of the four simulation results from these models are very similar both in robustness calcultions [56] than in Nash criterion calculations [21] and in calibration than in validation. It is thus difficult to discriminate which model is more efficient than the others when using these two pre-cited criteria, robusteness (4) and Nash criterion (3).

3.2. Intercomparison method results

In the previous section, we demonstrate the difficulty to compare the efficiency of a set of models deriving one from the others. This kind of derivation is performed either in extending the influence of the parameter that control model input to the entire entry variable, or in constructing a semi-distributed modelling concept from a non-distributed one. To succeed in the classification of these models, we use intercomparison method through a scoring technique.

3.2.1. Table approach in intercomparison results'presentation – a proposition

On the Table 3, we note the equality of three models in term of sensitivity to samples through the variance part in the four-model competition. Each of these three models wins one scoringpoint leading to an apparent equality between them; three scoring-points are assigned to the variance component of the intercomparison (out of the total 12 scoring points available for a watershed). It is when a model won more than 50% of the available scoring-points that it is qualified as the best from a (number)-model competition step. Therefore, subsequent (number1)-model competition steps necessarily include non-qualified models from the precedent (number)-model competition. Table containing the (number)-model competition is designated as the first round and the one that includes the immediate subsequent (number-1)-model competition is the second round table.

Table 3 is thus relative to the table approach (at the hydrometric station of Koulikoro on the Niger River). SimulHyd Non-Distributed wins 8 scoring-points during the four-model competition (in left), which is more than 50% of the 12 available scoring-points on a single watershed (Table 3). This intercomparison method accurately detects which is relatively the best model that win 50% of the available scoring-points during a competition. Therefore, other subsequent intercomparisons are needed to have a complete ranking of the remaining models. On the Table 3, there are a three-model intercomparison (in middle) and a two-model intercomparison (in right). In middle, SimulHyd Semi-Distributed wins more than 50% of the 12 available scoring-points facing both GR2M Non-Distributed and Semi-Distributed. In right, GR2M Non-Distributed wins over GR2M Semi-Distributed.

Table 4 : *Models intercomparison scoring technique table at the hydrometric station of Koulikoro on the Niger River, over both periods 1950 to 1970 and 1971 to 1995 (second round). From left to right: three three-model (A, B, and C) and a two-model intercomparison results. 1 Double Samples Technique, 2 a double calibration protocol on Nash criterion differences, 3 Non-Distributed model, and 4 Semi-Distributed model.*

Koulikoro	Periods: 1950 to 1970 & 1971 to 1995	Models / Actions		Set of three-model Competitions								
				A			B			C		
				GR2M ND [3]	SimulHyd ND	SimulHyd SD [4]	GR2M ND	SimulHyd ND	GR2M SD	GR2M ND	GR2M SD	SimulHyd SD
Bias		Evaluation Nash Garrick (Calibration)		0.25	2	0.75	0.25	2.5	0.25	0.5	0	2.5
Bias		Evaluation Nash Garrick (Validation)		0	2	1	0	2.5	0.5	0.5	0.25	2.25
Bias		Calibration Nash		0	1	0	0	1	0	0	0	1
Bias		Validation Nash		0	1	0	0	1	0	0	0	1
Bias		Robustness Method DST [1]		0	1	0	0	1	0	0	0	1
Variance		Less sensitivity to Samples	X1	0	1	0	0	1	0	0	0	1
Variance		Less sensitivity to Samples	X2	0	0	1	0	0	1	0	1	0
Variance		Less sensitivity to Samples	Nash (val.) [2]	0	0	1	0	1	0	0	0	1
Scoring-points				0.25	8	3.75	0.25	10	1.75	1	1.25	9.75

In decreasing order (from the best), we classify models in terms of efficiency as following: SimulHyd Non-Distributed (in first position, as the first qualified through a four-model intercomparison), SimulHyd Semi-Distributed (second position, as the second qualified through a three-model intercomparison), GR2M Non-Distributed (third position, as the third qualified through a two-model intercomparison) and finally GR2M Semi-Distributed (the last position). Thus, a set of (number-1)-model competitions allows to accurately build a ranking of models in terms of efficiency; "number" is the number of models involved in the competition. In case of an apparent equality, the method leads scarcely to a significant conclusion without doing a set of (number-1)-model competitions; eventually, followed by a set of (number-2)model competitions, and so on. Apparent equality appeared sometimes in the sensitiveness to samples intercomparison processes.

A set of two-model intercomparisons would suffice to classify the four models in competition without doing either a four-model intercomparison or a set of three-model intercomparisons. However, these last inform whether it is necessary to perform many two-model competitions.

Table 5: Model intercomparison through the scoring technique at 16 hydrometric stations on the Niger River and on its tributaries, over both periods: before 1970 and after 1970, (first round). From left to right: a four-model, a three-model, and a two-model intercomparison results on sixteen watersheds of the Niger River and its tributaries. 1Double Samples Technique, a double calibration protocol on Nash criterion differences, 3Non-Distributed model, and 4Semi-Distributed model.

16 watersheds	Periods: before 1970 and after 1970	Actions \ Models		Four-model Competition				Three-model Competition			Two-model Competition	
				GR2M ND [3]	SimulHyd ND	GR2M SD [4]	SimulHyd SD	GR2M ND	GR2M SD	SimulHyd SD	GR2M ND	GR2M SD
Bias		Evaluation Nash Garrick (Calibration)		10.13	25.13	3.88	8.88	15.42	3.92	28.67	29.75	18.25
		Evaluation Nash Garrick (Validation)		10.13	25.38	2.63	9.88	15.92	2.92	29.17	31	17
		Calibration Nash		1	12	0	3	5	0	11	10	6
		Validation Nash		0	11	0	5	4	0	12	9	7
		Robustness Method DST [1]		2	13	0	1	3	0	13	15	1
Variance		Less sensitivity to samples	X1	1	9	0	6	1	0	15	11	5
			X2	5	3	5	3	8	5	3	9	7
			Nash (val.) [2]	5	6	1	4	9	1	6	11	5
Scoring-points				34.25	104.5	12.5	40.75	61.33	12.83	117.83	125.75	66.25

Table 4 is relative to the table approach (at the hydrometric station of Koulikoro on the Niger River). In decreasing order (from the best), we classify models in terms of the less sensitive to samples as (Table 4): SimulHyd Semi-Distributed (as the first qualified through a set of two three-model intercomparisons in A and in C); SimulHyd Non-Distributed (as the second qualified through a three-model intercomparison in B). Finally, GR2M Semi-Distributed (the third qualified through a two-model intercomparison on Table 3) and GR2M Non-Distributed (the fourth).

Table 6 : Model intercomparison through the scoring technique at 16 hydrometric stations on the Niger River and on its tributaries, over both periods: before 1970 and after 1970, (second round). Left to right: three three-model competitions (A, B, C) and a two-model competition. [1] Double Samples Technique, [2] Nash criterion differences, [3] Non-Distributed model, and [4] SemiDistributed model.

16 watersheds		Models / Actions		Set of three-model competitions								
				A			B			C		
				GR2M ND [3]	SimulHyd ND	SimulHyd SD [4]	GR2M ND	SimulHyd ND	GR2M SD	SimulHyd ND	GR2M SD	SimulHyd SD
Bias	Period length 18 t0 46 years	Evaluation Nash Garrick (Calibration)		12.42	25.67	9.92	10.17	32.42	5.42	29.17	9.92	8.92
		Evaluation Nash Garrick (Validation)		10.67	25.42	11.92	10.67	32.92	4.42	29.67	7.92	10.42
		Calibration Nash		1	12	3	1	15	0	12	1	3
		Validation Nash		0	11	5	0	16	0	11	0	5
		Robustness Method DST [1]		2	13	1	2	14	0	15	0	1
Variance		Less sensitivity to samples	X1	1	9	6	1	14	0	10	0	6
			X2	8	3	5	5	6	5	6	7	3
			Nash (val.) [2]	5	6	5	5	10	1	10	2	4
Scoring-points				40.08	105.1	46.83	34.83	140.3	15.83	122.8	27.83	41.33

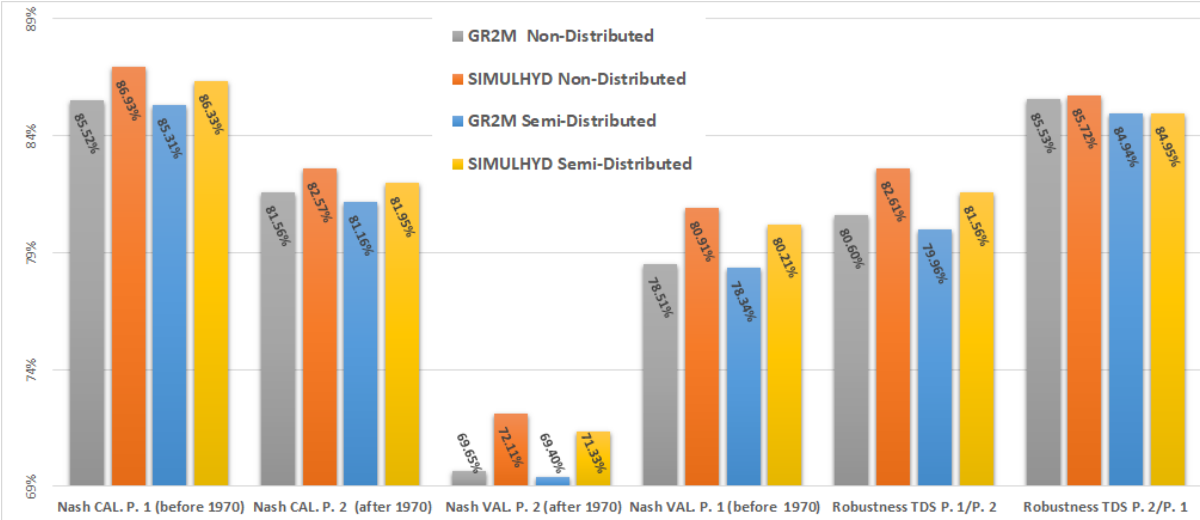


Figure 4: Classical presentation of models' comparison results. Average Nash and Robustness results on 16 watersheds of the Niger River and tributaries over 18 years to 46 years, using four different models. CAL.: Calibration; VAL.: Validation; Robustness TDS (Technique of Double Samples) P. 1/P. 2: robustness calculation on simulation results from both a calibration period ending in 1970 (period, P. 1) and a validation period beginning after 1970 (period, P.2); vice-versa.

Tables 5 and 6 are relative to the table approach (at the sixteen hydrometric stations). The equality issue on table 3 disappeared on Tables 5; it was relative to the less sensitiveness to samples involving three models in a four-model competition. The Table 3 shows intercomparison results on a single watershed whereas results on the Table 5 concerned sixteen watersheds at once (Table 1). So, in decreasing order (from the best), we classify models in terms of the less sensitive to samples. SimulHyd Non-Distributed is the first qualified through a set of three three-model intercomparisons (Tables 6 : A, B and C). Afterward, SimulHyd Semi-Distributed is the second qualified through a three-model intercomparison (Tables 5); it wins 50% of the available 48 scoring-points; and the two-model competition shows that neither GR2M Non-Distributed nor GR2M Semi-Distributed could individually has the remaining 50% of the remaining scoring-points. Finally, GR2M Non-Distributed is the third qualified through a two-model intercomparison (Tables 5) and GR2M Semi-Distributed is the fourth classified in the intercomparison process.

3.2.2. Sum-up results on the 16 watersheds (a Totalizing approach)

The totalizing approach consists in performing the intercomparison of the four model versions on the studied 16 watersheds at once. Results from this approach are summarized on the Tables 5 and 6, and on the Figure 5.

Tables 5 and 6, and figure 5 are relative to the sum-up approach (at the 16 hydrometric stations). Globally, intercomparison results through our scoring technique leads to, both in terms of high efficiency and in terms of low sensitivity (to samples), the following ordering from the best to the worse model version, on the studied 16 watersheds on the Niger River and some of its tributaries (Tables 5 and 6, and Figure 5). SimulHyd Non-Distributed ranks first, SimulHyd Semi-Distributed second, GR2M Non-Distributed third and GR2M Semi-Distributed last.

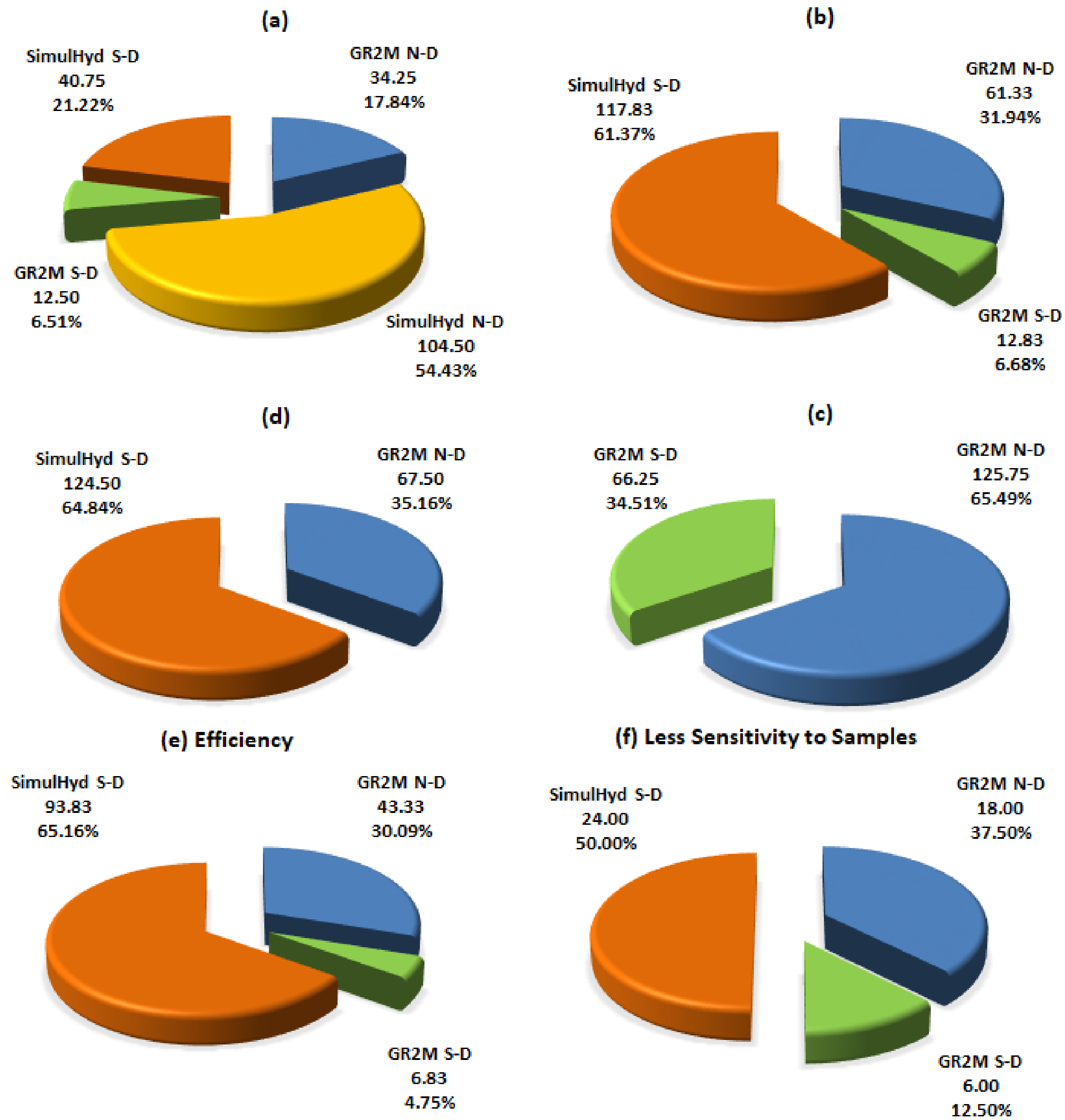


Figure 5: Intercomparison results (totalizing approach) on 16 watersheds of the Niger River.
*Clockwise, the first graph (top left, a) permits to discriminate SimulHyd Non-Distributed as the most efficient model against the three others (it obtains more than 50% out of the 16*12 = 192 points). The second graph (top right, b) shows SimulHyd Semi-Distributed in the wining position over the two other models (it also wins more than 50% out of the 16*12 = 192 points of the three-model intercomparison); eventually the first 2-model intercomparison (middle right, c) exhibits the predominance of GR2M Non-Distributed over GR2M Semi-Distributed (125.75 points against 66.25 points). The second 2-model intercomparison (middle left, d) shows which percent SimulHyd Semi-Distributed excels over GR2M Non-Distributed, it gains 64.84% out of the 192 points (against 35.16%). The last row dissects the 3-model intercomparison (on top right) in terms of both efficiency (e) and less sensitivity to samples (f).*

When intercomparing models on the single Koulikoro watershed (the largest one), the precedent ordering is respected in terms of model efficiency but not in term of les sensitivity to samples. Hence, both the ranking of SimulHyd Non-Distributed versus SimulHyd Semi-Distributed and the ranking of GR2M Non-Distributed versus GR2M Semi-Distributed are respectively interchanged (Tables 3 and 4). We hence enhance the accuracy of the conceptual framework GR2M in proposing a set of two model versions SimulHyd Non-Distributed and SimulHyd Semi-Distributed.

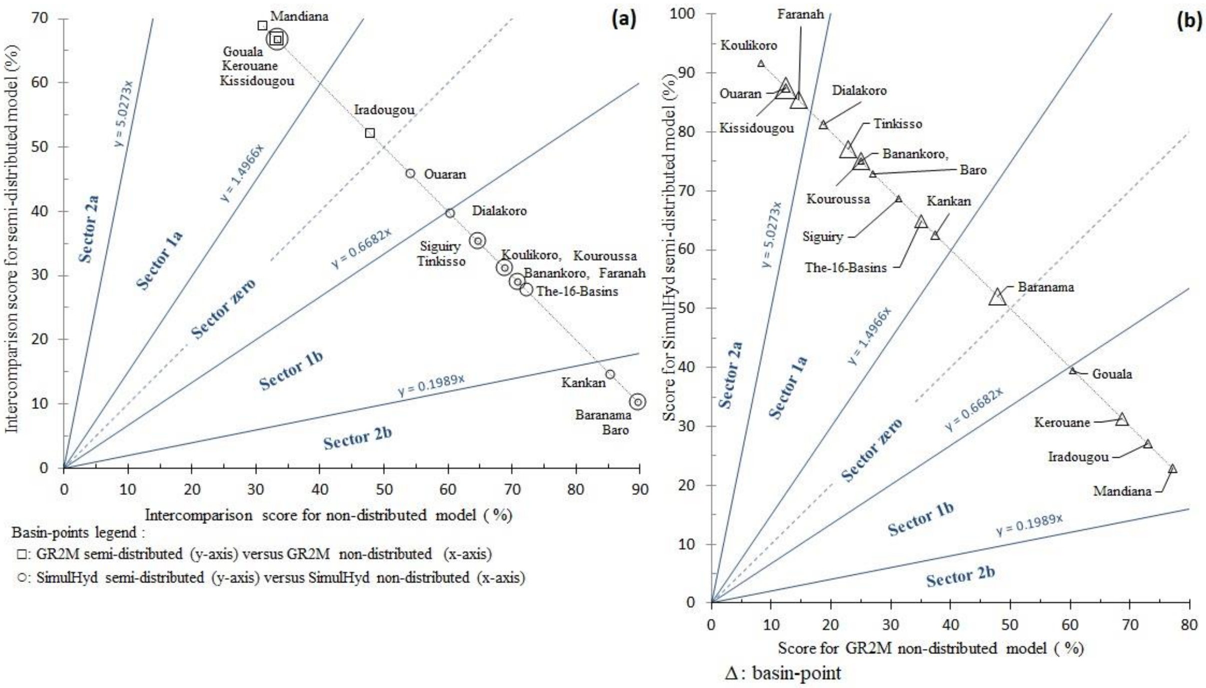


*Figure 6: intercomparison results in graphic (Basin approach 1). In case of superposition of basin-points where a same symbol varied in size, the smallest symbol represents the first written basin-point (either horizontally or vertically). On the x-axis the sector 2a is in the interval [from x-axis = 0 to x-axis = 100/ (1+ tan (7*PI/16))] whereas on the y-axis it is in the interval [from y-axis = 100*(1-1/ (1+ tan (7*PI/16))) to y-axis = 100], so on for the other sectors. Hence, the boundaries of these sectors are for sector 2a, [from x-axis = 0 to x-axis = 16.59, and from yaxis = 83.41 to y-axis = 100]. For sector 1a, they are [from x-axis = 16.59 to x-axis = 40.05, and from y-axis = 59.95 to y-axis = 83.41]. For sector zero, interval boundaries are [from xaxis = 40.05 to x-axis =59.95, and from y-axis =40.05 to y-axis = 59.95]; for sector 1b: [from x-axis = 59.95 to x-axis = 83.41, and from y-axis = 16.59 to y-axis = 40.05]; and for sector 2b: [from x-axis = 83.41 to x-axis = 100, and from y-axis = 0 to y-axis = 16.59]. A diagonal line is drawn. Finally, a line following the opposite diagonal (with Formula y-axis = -x-axis + 100) passes through all the basin-points. Left graphic (a) is for both semi-distributed models (SimulHyd andGR2M) against their respective non-distributed versions. Right graphic (b) shows the intercomparison results of SimulHyd semi-distributed against GR2M nondistributed.*

3.2.3. Basin approach in intercomparison results presentation

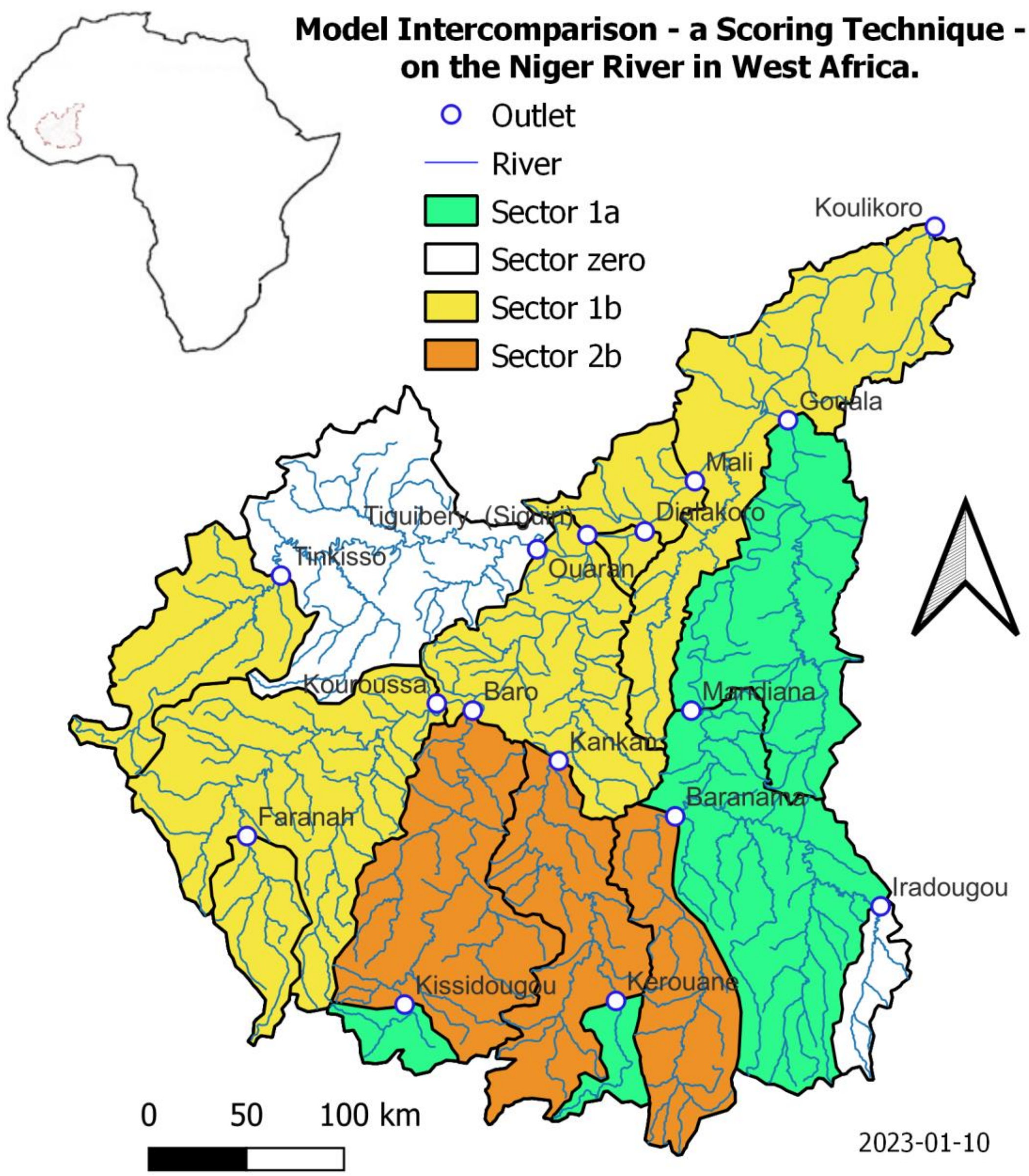


Figure 7: intercomparison results on map (basin approach 2 based on the mapping of the sectors on the graphic in basin approach 1).

On the Figure 6a, the sector 2b runs from x-axis= 83.41% to x-axis= 100% (and from y-axis= 0% to y-axis= 16.59%). It is where the qualification of "very highly more efficient" is attributed to the model SimulHyd Non-Distributed relatively to the model SimulHyd Semi-Distributed on the Kankan (on the Milo River), Baranama (on the Dion River) and Baro (on the Niandan River) watersheds. The sector zero runs from x-axis= 40.05% to x-axis= 59.95%, and from y-axis= 40.05% to y-axis= 59.95%. On its below part (from x-axis= 50% to x-axis= 59.95 %, and from y-axis= 40.05% to y-axis= 50%) the qualification of "equally to or more efficient" is attributed to the model SimulHyd Non-Distributed relatively to the model SimulHyd Semi-Distributed on the Ouaran watershed (on the Tinkisso River).

This twofold approach summarizes graphically, and on a map, the intercomparison results of several models on diverse watersheds at once. Figure 6a shows results from a set of two-model intercomparisons. Its sectors are delimited in trigonometric angle from x-axis (PI/16, 3* PI/16, 5*PI/16 and 7* PI/16). Respectively, sectors 2a and 1a are favorable to the y-axis component of the two-model intercomparison; sectors 2b and 1b are for the model that occupies the x-axis; and the sector zero would be in some case considered not in favor of either models. Sectors, mapped on the Figure 7, permit to understand spatially how strong a model is more efficient than the other is on a basin. These twofold intercomparison presentations are complementary and more informative if analyzed together: graphically through the basin approach 1 (Figure 6a) and cartographically through the basin approach 2 (Figure 7).

The combined analyses of the three-model competitions (in A and in C) and of the three-model competition in B allows to display on Table 4 that SimulHyd semi-distributed showed less sensitivity to samples over the three other models; followed by SimulHyd non-distributed which wins over the two other models. Eventually the two-model competition allows to show GR2M semi-distributed as less sensitive to samples than GR2M non-distributed.

Tables 5 and 6, and figures 6a and 7 are relative to the basin approach (at the 16 hydrometric stations). A combined analysis of both Figures (6a and 7) leads to some remarks: SimulHyd non-distributed is the most efficient of the four models, at the studied hydrometric stations located on the main stream of the Niger River (Banankoro, Dialakoro, Faranah, Koulikoro, Kouroussa and Siguiry); followed by SimulHyd semi-distributed. These pre-cited hydrometric stations belong to the sector 1b (Figure 6a). This sector (1b) runs from x-axis= 59.95% to xaxis= 83.41% (and from y-axis= 40.05% to y-axis= 16.591%) when lecturing the intersection between the anti-diagonal line and the sector limiting lines. These percentages are about twelve scoring points in competition on each basin-point located along the intercomparison line or space. The qualification of "highly more efficient" is attributed to the model SimulHyd NonDistributed relatively to the model SimulHyd Semi–distributed, on these pre-cited watersheds on the Niger River, and on the Tinkisso watershed (on the Tinkisso River).

These remarks are also valid when adapted to the sectors 2a, 1a, and the upper part of the sector zero. In the present study, Sector 2a remains empty. The sector 1a is from x-axis= 16.59% to x-axis= 40.05%, and from y-axis= 59.95% to y-axis= 83.41%. In this sector the model SimulHyd Semi-Distributed is highly more efficient than the model SimulHyd Non-Distributed on the Kissidougou watershed (on the Niandan River), and the model GR2M Semi-Distributed is highly more efficient than the model GR2M Non-Distributed on the Gouala (on the Sankarani River), Kerouane (on the Niandan River), and Mandiana (on the Sankarani River) watersheds. The upper part of the sector zero extends from x-axis= 40.05% to x-axis= 50 %, and from yaxis= 50% to y-axis= 59.95%. Here, the model GR2M Semi-Distributed is more efficient than the model GR2M Non-Distributed on the Iradougou watershed (on the Kouroukele River).

Figure 6b (right) shows the intercomparison results of SimulHyd semi-distributed against GR2M non-distributed whereas Figure 6a (left) is for both semi-distributed models against their respective non-distributed versions. On Figure 6b, basin-points become available in sector 2a and more in sector 1a (comparatively to Figure 6a): it shows the preponderance of SimulHyd semi-distributed over GR2M non-distributed. However, a set of basin-points (Mandiana, Kerouane and Gouala, in Guinea) shows that GR2M semi-distributed is more performing than SimulHyd semi-distributed (Figure 6a and 6b).

The Figure 7 is a way to exhibit the preceding remarks on an intercomparison map that displays both the non-empty sectors and the studied watersheds.

3.3.Developed Scoring technique in relation to methods in literature.

Table 7: Parameterized SimulHyd semi-distributed on 16 watersheds. The parameterization uses the calibrated parameters from the SimulHyd lumped version calibrated on a reference period (preferably before 1970 in West Africa). The A column shows the variation of simulated runoff between both periods (before and after 1970). The B column shows the variation of simulated runoff at a hydrometric station compared with the one at Koulikoro station on the reference period (ideally before 1970) while the column C concerns the period ideally after 1970 up to 2020.

Hydrometric Stations	Area in Km^2	Runoff simulations through SimulHyd semi-distributed using parameters (X1, X2) of its lumped version (calibrated on a first period)						Statistics on simulated Runoff: variation (A) and compared with Koulikoro (B and C)		
		Parameters and Modulus On a reference period ideally before 1970				Modulus on a period after 1970 and up to 2020				
		Number of years in reference period	X1	X2	First simulated Runoff [mm]	Number of years in the selected period	Second simulated Runoff [mm]	A [%]	B [%]	C [%}
Banankoro	73458	10	0.601174	0.603419	30	40	26	-13	-11	-3
Baranama	6608	8	0.478479	0.570661	40	40	37	-7	18	37
Baro	13108	21	0.634033	0.631035	55	50	45	-18	62	67
Dialakoro	70591	14	0.626513	0.567465	40	50	31	-23	18	14
Faranah	3178	12	0.586313	0.585309	55	40	50	-8	60	87
Gouala	33075	10	0.776532	0.4234319	59	50	51	-14	74	89
Iradougou	1824	6	0.286000	0.484884	42	50	39	-7	24	44
Kankan	10080	21	0.602021	0.670842	59	50	48	-19	74	78
Kerouane	1423	8	0.663168	0.755662	75	40	79	5	121	193
Kissidougou	1400	11	0.795414	0.5812902	90	50	76	-16	165	181
Koulikoro***	120603	21	0.594934	0.521148	34	50	27	-21	0	0
Kouroussa	17201	21	0.519973	0.565106	39	50	29	-26	15	7
Mandiana	21952	14	0.523756	0.561184	38	50	31	-18	12	15
Ouaran	19777	14	0.520571	0.503660	29	50	22	-24	-15	-19
Siguiry	71064	17	0.632367	0.569243	42	50	33	-21	24	22
Tinkisso	6569	13	0.535108	0.574188	35	50	26	-26	3	-4

The present scoring technique performs intercomparison in assigning scoring-points to models based on differences of calculated efficiency criteria, such as Nash criterion . The technique has a bias component for assessing efficiency and a variance component for assessing sensitivity to samples. Models are commonly assessed based on these two components; however, Baroni et al. [58] compared distributed models in just analyzing their sensitivity to "evapotranspiration and soil moisture at different soil depths" when observing their "models agree in the simulated river discharge".

The bias component of our scoring technique is based on differences of criteria calculated using runoff simulation whereas its variance component is based on the difference of both model parameters and Nash criterion obtained between two calibration periods. However, Garavaglia et al. [59] suggested analyzing more than one output data in comparing model structures.

The literature scarcely elucidates both concepts investigated in this paper: the intercomparison between gridded semi-distributed modelling, as we have developed it, and the interchangeability of set of parameters between model versions. Therefore, critics are in the next sub-section.

3.4. Critique about the developed scoring technique

The developed scoring technique tends to give equal weight to large and small differences between models in terms of calculation of existing performance criteria, such as Nash criterion and robustness calculation. Therefore, the proposed method produces scoring-points that bring semi-quantitative information relative to a set of models involved in an intercomparison process. Moreover, scoring-points won by a model could change when varying the number of models in the competition: it is thus a relative method as the results depend on the selected set of models to be intercompared.

However, regional projects that imply hydrological modelling could require the developed intercomparison method to choose the adequate models or model versions for reaching optimum results in a study area.

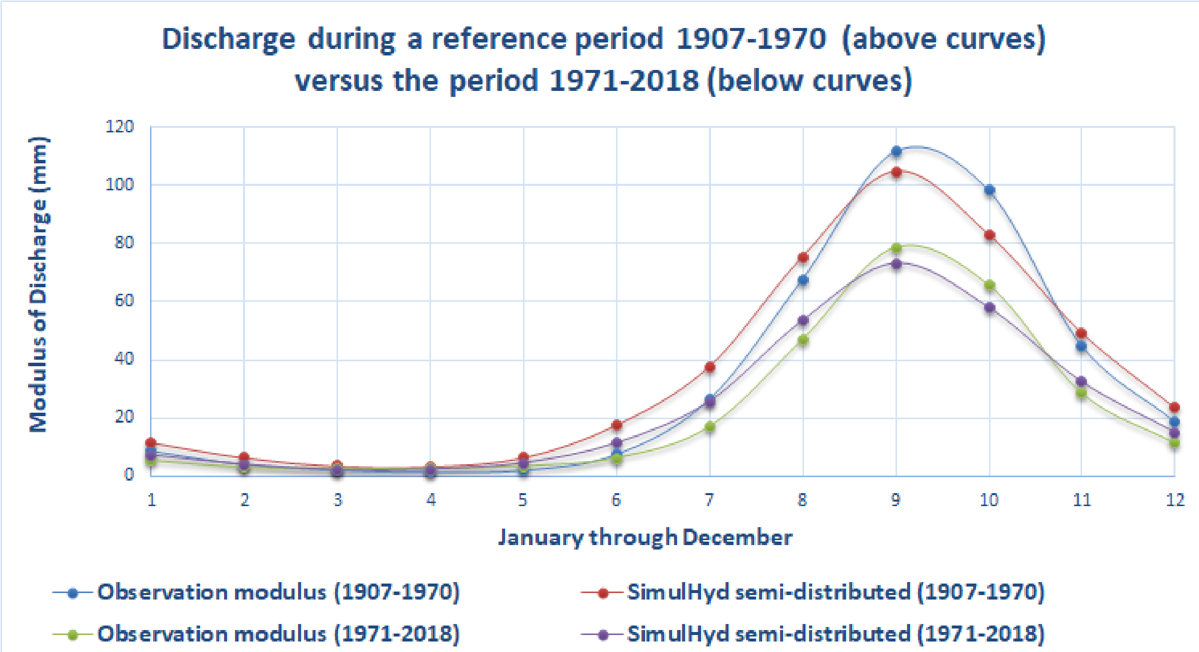


Figure 8: Discharge data observed on 1971-2018 compared to a reference period (1907-1970) at the hydrometric station of Koulikoro.

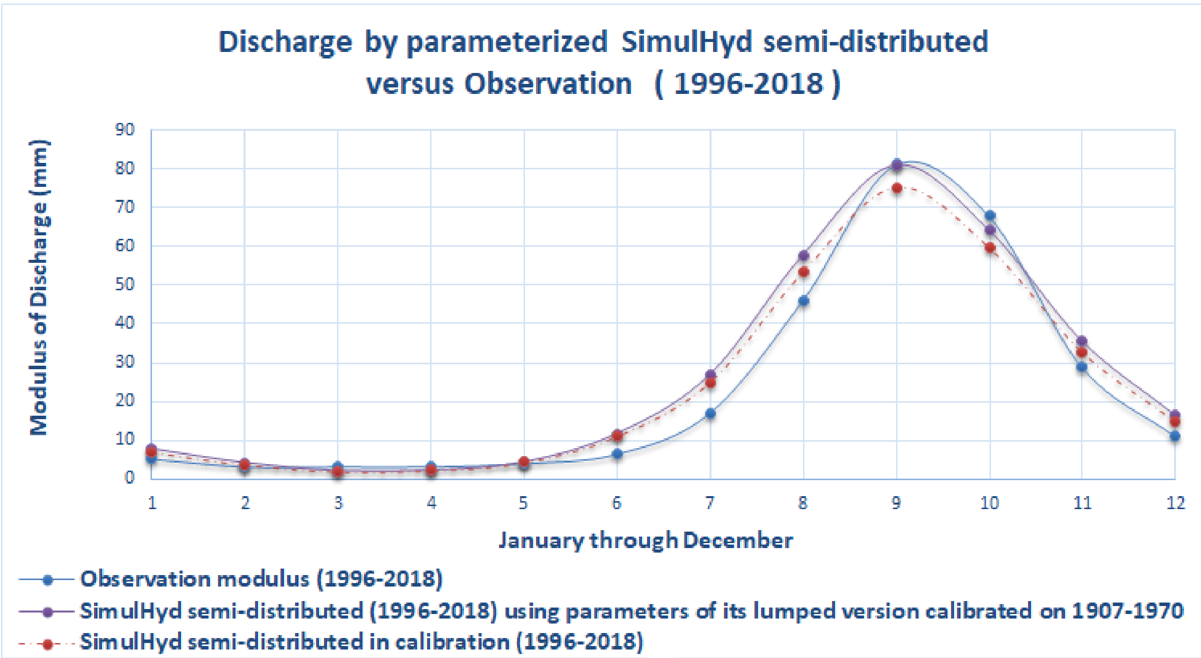


Figure 9: Simulation through SimulHyd semi-distributed using parameters from its lumped version at the hydrometric station of Koulikoro.

3.5. Parameterized SimulHyd semi-distributed on the Niger River Basin

In West Africa, a rainfall drop occurred around 1970 that is considered as a turning point [2324]. Kone [38] considered a time span before this date as a reference period in the hydro-system functioning as illustrated in the Figure 8. The period 1907-1970 is thus a reference period on which calibrated parameters are given, as shown on the Figure 9. At the Koulikoro hydrometric station on the Niger River, we obtained X1 = 0.575025 and X2 = 0.533399 as the calibrated parameters of the SimulHyd non-distributed model over 1907-1970.

The set of parameters (X1, X2) is injected in SimulHyd semi-distributed model to illustrate both the interchangeability of parameters between model versions and the use of this parameterized model on a third period (1996-2018). Simulated runoff on the table 7 are obtained using the parameterized SimulHyd semi-distributed. The parameterization uses the calibrated parameters from its lumped version on a reference period.

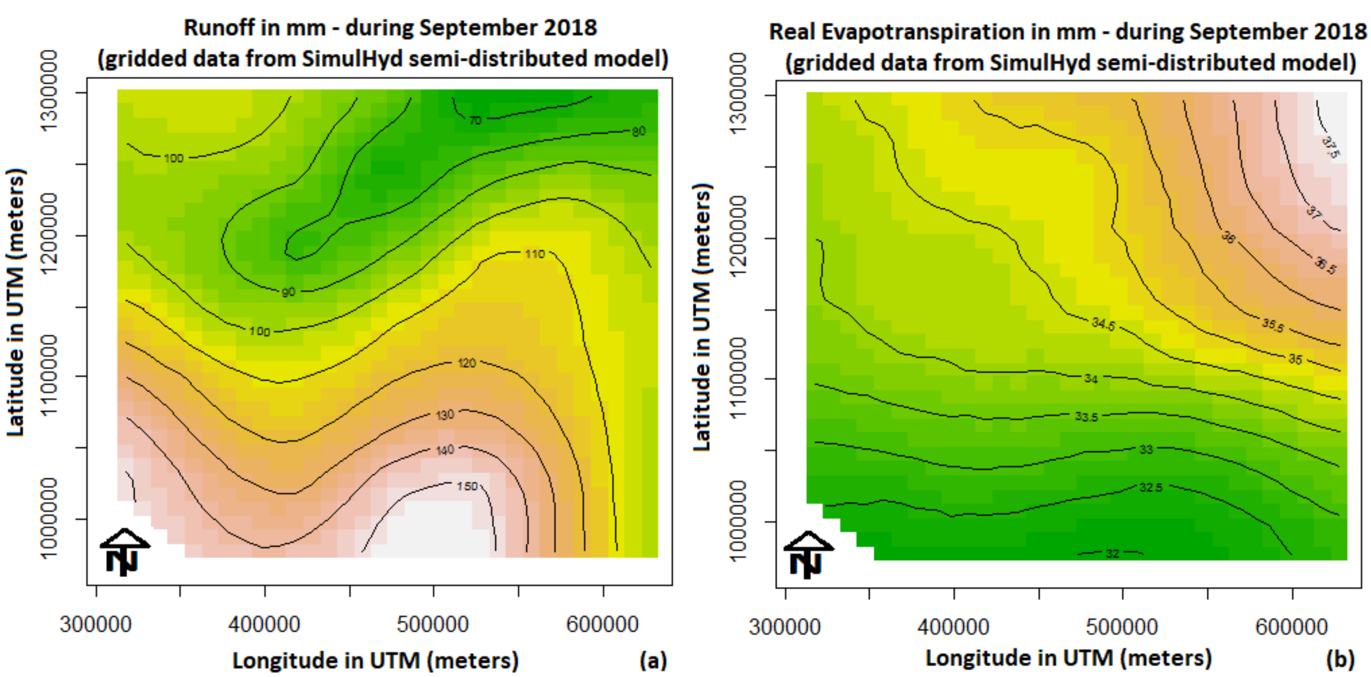
As a way forward in gridded semi-distribution modelling, in West Africa, we build a parameterized SimulHyd semi-distributed model in replacing its parameters by the set of parameters from its lumped version as illustrated on the figure 9 at the hydrometric station of Koulikoro, on the Niger River in West Africa.

Simulation results from the parameterized SimulHyd semi-distributed lead to runoff data in Table 7. Moreover, the parameterization uses the calibrated parameters from the SimulHyd lumped version calibrated on a reference period (preferably before 1970 in West Africa).

Table 1 and 7 demonstrate that discharge time series (runoff) were affected by the dropping occurred in precipitation (rainfall) around 1970, in West Africa: precipitation dropping varies between 7% and 12 % (Table 1, last column) while runoff dropping stresses between 8% and 26% % (Table 7, column A). These remarks are consistent with previous results from literature.

Mahé et al. [60] demonstrated the non-linearity of rainfall-runoff relation by analyzing the groundwater depletion curves. Their results led to conclude that the diminution of groundwater resources affects the baseflow contribution to discharge at hydrometric stations.

3.6. Gridded data generated from SimulHyd semi-distributed – their potential implication in extractive industries in West Africa



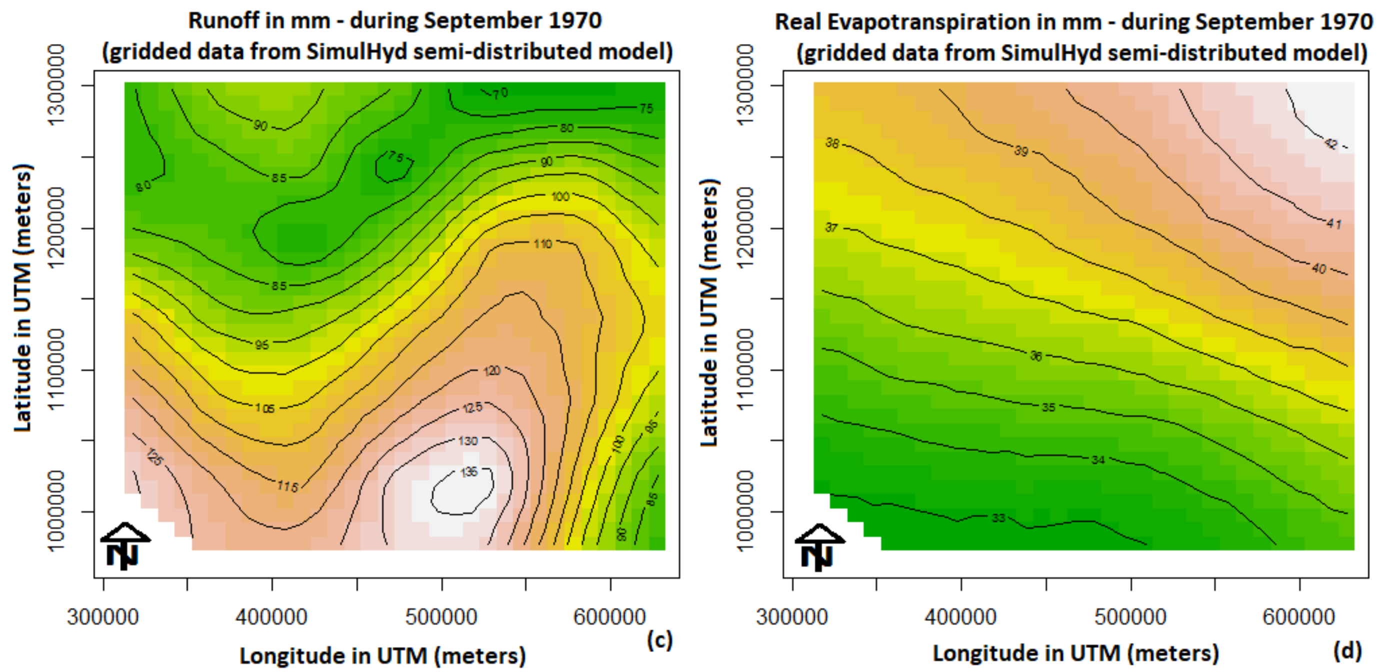


Figure 10: Gridded modelling results of both runoff (left) and real evapotranspiration (right) on the Niger River Basin at the hydrometric station over 1907-2018 using SimulHyd semidistributed in calibration phase – extracted maps concerned September 2018 (first row) and September 1970 (second row, c and d). 2018 compared to 1970, surface runoff increased while precipitation decrease as stated in Mahé et al. [60].

Economic mining ores are becoming deeper and are being confronted on water issues from surface runoff water as well as from groundwater flow around the world. We introduce this problematic in investigating its surface runoff aspect and its other aspects would need further investigation, in West Africa. As presented in Figure 10, gridded maps are intermediary produced when running SimulHyd model in its semi-distributed version at the hydrometric station of Koulikoro (in calibration phase), over 1907-2018, on the Niger River. Modelling input are mainly hydroclimate data (precipitation and evapotranspiration) from climate research unit and are partially explained in Harris et al. [61].

On Figure 10, runoff and real evapotranspiration are thus modelling output both during September 2018 (first rows, a and b) and during September 1970 (second rows, c and d). On it, our mapping uses 48 simulated runoff values (on 0.5 degree x 0.5 degree grid) to generate maps through interpolation processes between 1046 spatial nodes. Table 8 synthetizes the statistics on the 48 values and presents variograms that underlying the Figure 10. Furthermore, suck information on runoff permits informing about mining acid evolution and environmental pollutions.

Figure 10 (a to d) corroborate that surface runoff increased on a major part of the study area when considering September 2018 relatively to September 1970. The 48 grids (0.5-degree x 0.5 degree) in consideration on Figure 10 belong partially to other basins. Therefore, a protocol would be further needed to average runoff information on a specific grid when different simulations at different hydrometric stations concerned the same grid.

We present on the Figure 11 both the relative variation of runoff (a) and the relative variation of real evapotranspiration (b) during September when comparing year 2018 to year 1970 (as a reference). Both relative variations are mapped around 957 spatial nodes. The Figure 11 allows

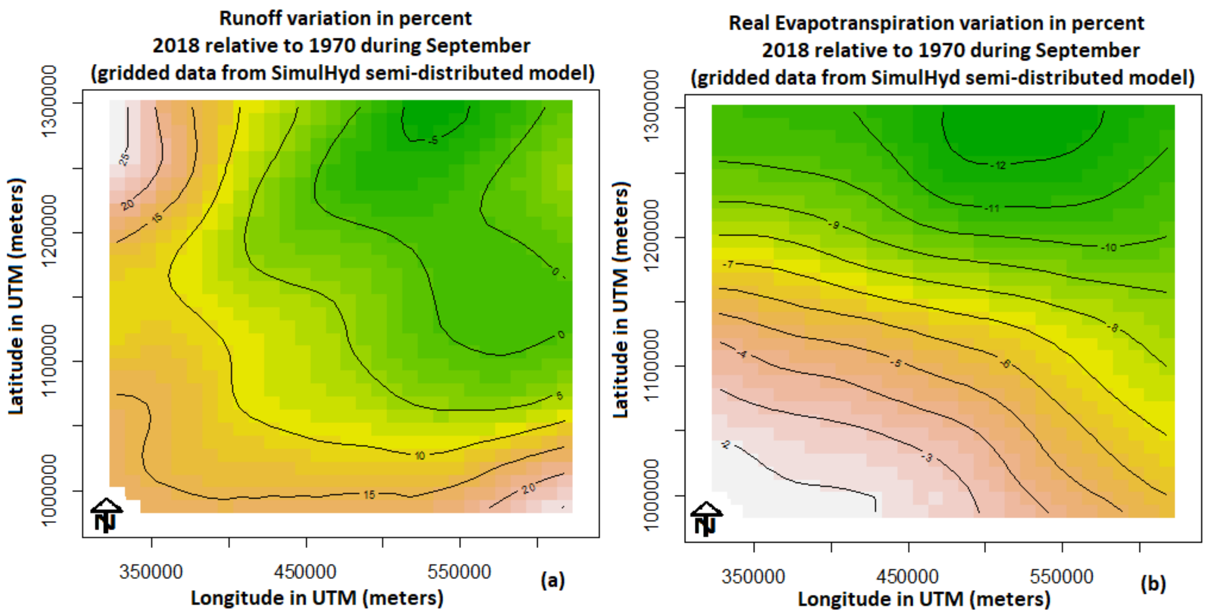


Figure 11: Relative variation of runoff (a) and the relative reduction of real evapotranspiration during September when comparing 2018 to 1970.

assessing climate change between these two periods separated by 48 years.

Table 8 : Statistics and variograms built from 48 spatial gridded data, which are mapped on figure 10: runoff and real evapotranspiration for September 2018 (a) and for September 1970 (b) – extracted from a 111- years simulated series at the hydrometric station of Koulikoro – using SimulHyd semi-distributed in calibration phase during 1907-2018. Variograms equation are on row c. The distance in meters is h.

September 2018 (a)	Data summary (in millimeters)					
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Runoff	49.792	95.284	105.979	109.810	122.393	164.781
real evapotranspiration	31.346	32.945	34.266	34.242	35.520	39.495

September 1970 (b)	Data summary (in millimeters)					
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Runoff	53.373	85.272	99.879	99.993	117.326	148.526
real evapotranspiration	31.653	34.632	36.787	36.961	39.198	43.959

(c)	Variogram Gamma (h in meters)	
	September 2018	September 1970
Runoff	124 + 217*(1.5*h/ 128063-0.5*(h/ 128063) ^3)	135 + 223*(1.5*h/ 141043-0.5*(h/ 141043) ^3)
real evapotranspiration	6.796*10 ^-1 + h*3.009*10 ^-6	5.539*10 ^-1 + h* 5.539*10 ^-6

5. Conclusion

This paper confirms a new efficient model version (SimulHyd – Simulation of Hydrological systems) on the Niger River and its tributaries. It proposes and validates an intercomparison method in hydrological modelling as targeted by the World Meteorological Organization since 1975. Our method is a scoring technique that integrates harmoniously previous known model performance criteria and provides a unique score of performance, for a specific model version in relation to others. It is an intercomparison method for comparing models on both a single watershed and on several watersheds at once.

Two stratified groups of components characterize this scoring technique that sum up to 12 scoring-points on each watershed: an accentuated bias component with nine scores and a variance component with three scores. On 16 watersheds, results show the preponderance of SimulHyd semi-distributed (with 117.83 out of 192 score-points, equivalent to 61.37%) over both GR2M lumped non-distributed (with 61.33 out of 192 score-points, equivalent to 31.94%) and GR2M semi-distributed (with 12.83 out of 192 score-points, equivalent to 6.68%). Specifically, the watershed at Koulikoro hydrometric station (on the Niger River) should better be investigated and accordingly to the study objectives, in using SimulHyd non-distributed for producing and completing runoff chronological data gaps and in using SimulHyd SemiDistributed for environmental and hydroclimate variability assessments. The study leads further to model parameters interchangeability possibilities using the above-mentioned four model versions.

The parameterized SimulHyd semi-distributed is recommended for investigating future scenario through using output from climate simulation models. Its simulation results permit to corroborate the non-linearity in rainfall-runoff relation in West Africa as precipitation dropping is around 10% while runoff decreasing is around 17%.

This work enriches the literature in methodologies that are reproducible and adaptable on other models and in other geographical contexts.

Author Contributions: Authors participate through years to the formulation of research hypotheses that underlying findings included in the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: hydrometric data belong to national hydraulics services from countries in study area.

Acknowledgments: The Malian Government through the doctoral formation conventions Number 002/CFD/2011, renewed two times, sponsored this work.

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

Abbreviations

The following abbreviations are used in this manuscript:
BADC: British Atmospheric Data Center, http://badc.nerc.ac.uk/
CAG24: 24th Colloquium of African Geology
CAL.: Calibration
CP: Calibration Period
CRU: Climatic Research Unit, www.cru.uea.ac.uk/
GR2M: Genie rural at 2 parameters and monthly time step (a hydrologic model)
SimulHyd: Simulation of hydrological systems (a hydrologic model)
HSM: HydroSciences Montpellier: a French laboratory
INRA: Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Env.
IRD : Institut de Recherche pour le Développement
GR2M ND: GR2M Non-Distributed model or lumped version
GR2M SD: GR2M Semi-Distributed model version
SimulHyd ND: SimulHyd Non-Distributed model or lumped version
GR2M ND: SimulHyd Semi-Distributed model version
N-MODELS: competition where N number of models are involved.
RMSE : Root Mean Squared Error
SIEREM: Système d'Informations Environnementales sur les Ressources en Eaux et leur Modélisation
VAL.: Validation

WHC: Water Holding Capacity
WMO: World Meteorological Organization

References

1. Beven, K. (1989). Changing ideas in hydrology—the case of physically-based models. Journal of hydrology, 105(1-2), 157-172.
2. Edijatno, & Michel, C. (1989). Un modèle pluie-débit journalier à trois paramètres. La Houille Blanche, 75(2), 113-122.
3. Hughes, D., Birkinshaw, S., & Parkin, G. (2021). A method to include reservoir operations in catchment hydrological models using SHETRAN. Environmental Modelling & Software, 138, 104980.
4. Martina, M. L. V., Todini, E., & Liu, Z. (2011). Preserving the dominant physical processes in a lumped hydrological model. Journal of Hydrology, 399(1-2), 121-131.
5. Zhu, D., Wang, G., Ren, Q., & Ilyas, A. M. (2020). Hydrological evaluation of hourly merged satellite–station precipitation product in the mountainous basin of China using a distributed hydrological model. Meteorological Applications, 27(2), e1909.
6. Jadidoleslam, N., Mantilla, R., & Krajewski, W. F. (2021). Data assimilation of satellite-based soil moisture into a distributed hydrological model for streamflow predictions. Hydrology, 8(1), 52.
7. Luong, T. T., Pöschmann, J., Vorobevskii, I., Wiemann, S., Kronenberg, R., & Bernhofer, C. (2020). Pseudo-spatially-distributed modeling of water balance components in the Free State of Saxony. Hydrology, 7(4), 84. Available from: https://www.mdpi.com/2306-5338/7/4/84.
8. Pandi, D., Kothandaraman, S., & Kuppusamy, M. (2021). Hydrological models: a review. International Journal of Hydrology Science and Technology, 12(3), 223-242.
9. Clarke, R. T. (2008). Issues of experimental design for comparing the performance of hydrologic models. Water resources research, 44(1).
10. World Meteorological Organization. (1975). Intercomparison of conceptual models used in operational hydrological forecasting. The Organization.
11. Garrick, M., Cunnane, C., & Nash, J. E. (1978). A criterion of efficiency for rainfall-runoff models. Journal of Hydrology, 36(3-4), 375-381.
12. Sukvanachaikul, Y., & Laurenson, E. M. (1983, January). An improved rainfall-runoff model for semi-arid regions. In Hydrology and Water Resources Symposium 1983: Preprints of Papers (pp. 108-112). Barton, ACT: Institution of Engineers, Australia.
13. Pilgrim, D. H., Chapman, T. G., & Doran, D. G. (1988). Problems of rainfall-runoff modelling in arid and semiarid regions. Hydrological Sciences Journal, 33(4), 379-400.
14. Karnieli, A. M., Diskin, M. H., & Lane, L. J. (1994). CELMOD5—a semi-distributed cell model for conversion of rainfall into runoff in semi-arid watersheds. Journal of Hydrology, 157(1-4), 61-85.
15. Gomes, F., Lacerda, M., Duarte, M. L., & Almeida, J. A. (1997). Hydrological Modelling of Distributed Rainfall/Runoff in the Watershed of Guadiana. In geoENV I—Geostatistics for Environmental Applications: Proceedings of the Geostatistics for Environmental Applications Workshop, Lisbon, Portugal, 18–19 November 1996 (pp. 213-221). Springer Netherlands.
16. Hernandez, M., Miller, S. N., Goodrich, D. C., Goff, B. F., Kepner, W. G., Edmonds, C. M., & Bruce Jones, K. (2000). Modeling runoff response to land cover and rainfall spatial variability in semi-arid watersheds. In Monitoring Ecological Condition in the Western United States: Proceedings of the Fourth Symposium on the Environmental Monitoring and Assessment Program (EMAP), San Francisco, CA, April 6–8, 1999 (pp. 285-298). Springer Netherlands.
17. Mahe, G., Dray, A., Paturel, J. E., Cres, A., Koné, F., Manga, M., ... & Servat, E. (2002). Climatic and anthropogenic impacts on the flow regime of the Nakambe River in Burkina Faso. International Association of Hydrological Sciences, Publication, (274), 69-76.
18. Szymkiewicz, R. (2002). An alternative IUH for the hydrological lumped models. Journal of Hydrology, 259(1-4), 246-253.
19. Paturel, J. E., Ouedraogo, M., Mahé, G., Servat, E., Dezetter, A., & Ardoin, S. (2003). The influence of distributed input data on the hydrological modelling of monthly river flow regimes in West Africa. Hydrological sciences journal, 48(6), 881-890.
20. Andréassian, V., Oddos, A., Michel, C., Anctil, F., Perrin, C., & Loumagne, C. (2004). Impact of spatial aggregation of inputs and parameters on the efficiency of rainfall-runoff models: A theoretical study using chimera watersheds. Water Resources Research, 40(5).
21. Mandeville, A. N., O'connell, P. E., Sutcliffe, J. V., & Nash, J. E. (1970). River flow forecasting through conceptual models part III-The Ray catchment at Grendon Underwood. Journal of Hydrology, 11(2), 109-128.
22. Kabouya, M., & Michel, C. (1991). Estimation of surface water resources at monthly and annual time steps, application to a semi-arid country. Review of Water Sciences , 4 (4), 569-587.
23. Ardoin-Bardin, S., Dezetter, A., Servat, E., Dieulin, C., Casenave, L., Niel, H., ... & Mahe, G. (2006). Application of climatic scenarios in hydrological modelling: using GCMs outputs. Climate variability and change: hydrological impacts, 436-441.
24. Koné, S. (2007). Master' s thesis (DEA) : "Comparaison entre une modélisation semiglobale et une modélisation globale". Université Montpellier II, (2006-2007).
25. Dezetter, A., Girard, S., Paturel, J. E., Mahé, G., Ardoin-Bardin, S., & Servat, E. (2008). Simulation of runoff in West Africa: Is there a single data-model combination that produces the best simulation results?. Journal of Hydrology, 354(1-4), 203-212.
26. Vitolo, C., Wells, P., Dobias, M., & Buytaert, W. (2016). fuse: An R package for ensemble Hydrological Modelling. Journal

of Open Source Software, 1(8), 52.
27. Coron, L., Thirel, G., Delaigue, O., Perrin, C., & Andréassian, V. (2017). The suite of lumped GR hydrological models in an R package. Environmental modelling & software, 94, 166-171.
28. Givati, A., Thirel, G., Rosenfeld, D., & Paz, D. (2019). Climate change impacts on streamflow at the upper Jordan River based on an ensemble of regional climate models. Journal of Hydrology: Regional Studies, 21, 92-109.
29. Delaigue, O., Thirel, G., Coron, L., & Brigode, P. (2018, July). airGR and airGRteaching: two open-source tools for rainfall-runoff modeling and teaching hydrology. In 13th International Conference on Hydroinformatics (HIC 2018) (Vol. 3, pp. 541-548). Goffredo La Loggia, Gabriele Freni, Valeria Puleo and Mauro De Marchis.
30. Oueldkaddour, F. Z. E., Wariaghli, F., Brirhet, H., & Yahyaoui, A. (2021). Hydrological modeling of rainfall-runoff of the semi-arid Aguibat Ezziar watershed through the GR4J model. Limnological Review, 21(3), 119-126.
31. Rozos, E., Dimitriadis, P., & Bellos, V. (2021). Machine learning in assessing the performance of hydrological models. Hydrology, 9(1), 5.
32. YAO, H., HASHINO, M., TERAKAWA, A., & SUZUKI, T. (1998). Comparison of distributed and lumped hydrological models. Proceedings of Hydraulic Engineering, 42, 163-168.
33. dos Santos FM, de Oliveira RP, Mauad FF. Lumped versus Distributed Hydrological Modeling of the Jacaré-Guaçu Basin, Brazil. J Environ Eng. 2018;144(8).
34. Refsgaard, J. C., & Knudsen, J. (1996). Operational validation and intercomparison of different types of hydrological models. Water resources research, 32(7), 2189-2202.
35. Askew, A. J. (1989). Real-time intercomparison of hydrological models. In New Directions for Surface Water Modeling. Proceedings of a Symposium held in Baltimore, Maryland, May 1989. IAHS Publication (No. 181).
36. Okiria, E., Okazawa, H., Noda, K., Kobayashi, Y., Suzuki, S., & Yamazaki, Y. (2022). A Comparative Evaluation of Lumped and Semi-Distributed Conceptual Hydrological Models: Does Model Complexity Enhance Hydrograph Prediction?. Hydrology, 9(5), 89.
37. Fang, Z. N., Shultz, M. J., Wienhold, K. J., Zhang, J., & Gao, S. (2019). Case study: Comparative analysis of hydrologic simulations with areal-averaging of moving rainfall. Hydrology, 6(1), 12.
38. Koné, S. (2015). Apport de la Géostatistique à la Modélisation Hydrologique: Etude de Bassins Hydrologiques en Zone Soudano-Sahélienne (Doctoral dissertation, Institut Supérieur de Formation et de Recherche Appliquée, ISFRA-USJPB, Bamako-Mali).
39. Haruna, A., Garambois, P. A., Roux, H., Javelle, P., & Jay-Allemand, M. (2022). Does flash flood model performance increase with complexity? Signature and sensitivity-based comparison of conceptual and process-oriented models on french Mediterranean cases. Hydrology, 9(8), 141.
40. Zhang, J., Zhang, Y., Song, J., Cheng, L., Gan, R., Shi, X., ... & Zhao, P. (2017). Comparing hydrological modelling, linear and multilevel regression approaches for predicting baseflow index for 596 catchments across Australia. Hydrology and Earth System Sciences Discussions, 1-40.
41. Biondi, D., & Todini, E. (2018). Comparing hydrological postprocessors including ensemble predictions into full predictive probability distribution of streamflow. Water Resources Research, 54(12), 9860-9882.
42. Gosling, S. N., Zaherpour, J., Mount, N. J., Hattermann, F. F., Dankers, R., Arheimer, B., ... & Zhang, X. (2017). A comparison of changes in river runoff from multiple global and catchment-scale hydrological models under global warming scenarios of 1 C, 2 C and 3 C. Climatic Change, 141, 577-595.
43. Loganathan, P., & Mahindrakar, A. B. (2021). Intercomparing the robustness of machine learning models in simulation and forecasting of streamflow. Journal of Water and Climate Change, 12(5), 1824-1837.
44. Liu, G., He, Z., Luan, Z., & Qi, S. (2018). Intercomparison of a lumped model and a distributed model for streamflow simulation in the Naoli River Watershed, Northeast China. Water, 10(8), 1004.
45. Krysanova, V., Vetter, T., Eisner, S., Huang, S., Pechlivanidis, I., Strauch, M., ... & Hattermann, F. F. (2017). Intercomparison of regional-scale hydrological models and climate change impacts projected for 12 large river basins worldwide—a synthesis. Environmental Research Letters, 12(10), 105002.
46. Sittner, W. T. (1976). WMO project on intercomparison of conceptual models used in hydrological forecasting. Hydrological Sciences Journal, 21(1), 203-213.
47. Cavadias G, Morin G. The combination of simulated discharge of hydrological models: application to WMO intercomparison of conceptual models of snowmelt runoff. Nord Hydrol. 1986;17(1).
48. Ambroise, B., Perrin, J. L., & Reutenauer, D. (1995). Multicriterion validation of a semidistributed conceptual model of the water cycle in the Fecht Catchment (Vosges Massif, France). Water Resources Research, 31(6), 1467-1481.
49. ZENNAKI A, BABA-HAMED K, BOUANANI A. Étude comparative des modèles hydrologiques conceptuels globaux GR et Gardénia appliqués au bassin versant de l'oued Boukiou (Nord-Ouest algérien). Tech Sci Méthodes. 2021;(12).
50. Boyer, J. F., Dieulin, C., Rouche, N., Cres, A., Servat, E., Paturel, J. E., & Mahe, G. (2006). SIEREM: an environmental information system for water resources. IAHS publication, 308, 19.
51. Feki, M., Ravazzani, G., Barontini, S., Ceppi, A., & Mancini, M. (2020). A comparative assessment of the estimates of the saturated hydraulic conductivity of two anthropogenic soils and their impact on hydrological model simulations. Soil and Water Research, 15(3), 135-147.
52. Brunet-Moret, Y., Chaperon, P., Lamagat, J. P., & Molinier, M. (1986). Monographie hydrologique du fleuve Niger. I: Niger supérieur. Monographies hydrologiques ORSTOM, 1(8).

53. Perrin, C. (2002). Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. La Houille Blanche-Revue internationale de l'eau, (6-7), 84-91.
54. Brandt S. Data analysis: Statistical and computational methods for scientists and engineers, fourth edition. Data Analysis: Statistical and Computational Methods for Scientists and Engineers, Fourth Edition. 2014.
55. López-Bermeo C, González-Palacio M, Sepúlveda-Cano L, Montoya-Ramírez R, Hidalgo-Montoya C. Comparison of machine learning parametric and non-parametric techniques for determining soil moisture: Case study at las palmas andean basin. Adv Sci Technol Eng Syst. 2021;6(1).
56. Klemeš, V. (1986). Operational testing of hydrological simulation models. Hydrological sciences journal, 31(1), 13-24.
57. Mouelhi, S. (2003). Vers une chaîne cohérente de modèles pluie-débit aux pas de temps pluriannue, annuel, mensuel et journalier (Doctoral dissertation, ENGREF (AgroParisTech)).
58. Baroni, G., Schalge, B., Rakovec, O., Kumar, R., Schüler, L., Samaniego, L., ... & Attinger, S. (2019). A comprehensive distributed hydrological modeling intercomparison to support process representation and data collection strategies. Water Resources Research, 55(2), 990-1010.
59. Garavaglia, F., Le Lay, M., Gottardi, F., Garçon, R., Gailhard, J., Paquet, E., & Mathevet, T. (2017). Impact of model structure on flow simulation and hydrological realism: from a lumped to a semi-distributed approach. Hydrology and Earth System Sciences, 21(8), 3937-3952.
60. Mahe, G., Lienou, G., Descroix, L., Bamba, F., Paturel, J. E., Laraque, A., ... & Khomsi, K. (2013). The rivers of Africa: witness of climate change and human impact on the environment. Hydrological Processes, 27(15), 2105-2114.
61. Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Scientific data, 7(1), 109.

Appendix 1: ***Mathematic definitions of GR2M and SimulHyd models relative to Figure 3***

GR2M is a French hydrologic model standing for Genie Rural with 2 parameters at Monthly time step – the version of Kabouya and Michel [22].

Modulation of the entry data

X1 parameter is applied to the hydro-climatic entry data: precipitation (P) and potential Evapotranspiration (E):

P' = X1.P and E' = X1.E; modulated precipitation is (P') and modulated potential Evapotranspiration is (E').

For the SimulHyd model alone, the reciprocal of the X1 parameter is applied to the Water Holding Capacity data:

A = (1/X1). WHC; the constant A is replaced by a physical quantity called Water Holding Capacity relative to soil.

Neutralization of a part of the hydro-climatic entry data

A same quantity U is subtracted from the modulated hydro-climatic entry data; it represents the water intercepted by canopy and other obstacles: **U = { P'.E'} / { [(P') (1/ɣ) + (E') (1/ɣ)] ɣ }**

The value ɣ=2 was adopted in studies on watersheds in West Africa.

After the neutralization the net rainfall and the net evapotranspiration are respectively: **Pn = X1.P – U and En = X1.E – U**

Impact of Pn on the initial water level H of the soil reservoir

The contribution of the net rainfall (Pn) makes the water level of the soil reservoir increased from H to H1:

H1 = {H + A.V} / {1 + H. (V/A)} avec V = Tanh (Pn/A)

The complement of Pn is then: **Pe = Pn – (H1 – H)**

Impact of En on the new water level H1 (of the soil reservoir)

The new water level H1 of the soil reservoir diminishes under the evapotranspiration effect and becomes:

H2 = {H1. (1 – W)} / {1 + W. [1 - (H1/A)]}, with W = Tanh (En/A)

H2 corresponds to the water level of the soil reservoir for the following time step. The estimate of the real evapotranspiration (RET) can be deducted by difference of the levels of the soil reservoir at the beginning and at the end of every step of time: **RET = Hn+1 – Hn**

Partition of water between direct out-flow and out-flow by gravity

The soil reservoir of the GR2M model receives the piece of water that didn't contributes to the direct outflow. The water level S of the gravity reservoir at the beginning of the considered month increases and becomes S1 with: **S1 = S + (1 – α).Pe**

The coefficient α being for the direct out-flow: **Sr = α.Pe**

Water delivered (during the month) by the gravity reservoir is: **Lg = X2.S1**

The water level S2 of the gravity reservoir at the end of the time step is then: **S2 = S1 - Lg S2** corresponds to the water level of the gravity reservoir for the following time step.

The total out-flow (during the month) is: **L = Lg + Sr**