

1 **Streamflow prediction in “geopolitically ungauged” basins using satellite**
2 **observations and regionalization at subcontinental scale**

3
4 Tien L.T. Du^{a,b}, Hyongki Lee^a, Duong D. Bui^{c,*}, Berit Arheimer^d, Hong-Yi Li^a, Jonas Olsson^d,
5 Stephen E. Darby^e, Justin Sheffield^e, Donghwan Kim^a, Euiho Hwang^f
6

7 ^aDepartment of Civil and Environmental Engineering, University of Houston, Houston, TX, USA

8 ^bDa Nang Institute for Socio-Economic Development, Da Nang, Vietnam

9 ^cNational Center for Water Resources Planning and Investigation (NAWAPI), Ministry of Natural
10 Resources and Environment (MONRE), Hanoi, Vietnam

11 ^dSwedish Meteorological and Hydrological Institute (SMHI), Norrköping, Sweden

12 ^eSchool of Geography and Environmental Science, University of Southampton, Southampton, UK

13 ^fWater Resources Research Center, K-Water Institute, K-Water, Daejeon, South Korea
14
15
16

17 *Corresponding author at: Duong D. Bui, NAWAPI, MONRE, Hanoi, Vietnam.

18 Email address: duongdubui@gmail.com (Duong D. Bui)

19 **Highlights**

- 20
- Hydrology of the entire Greater Mekong region was modeled with HYPE.
- 21
- Regionalization was obtained by similarity in physiography and climate.
- 22
- Regionalization reached 80% of performance of local calibration at ungauged basins.
- 23
- Water level based flow correlation helped evaluate model at ungauged basins.
- 24
- Flow correlation framework outperformed the current metrics of modeled water level.
- 25

26 **Abstract**

27 A novel approach of combining regionalization and satellite observations of various
28 hydrological variables were employed to significantly improve prediction of streamflow signatures
29 at “geopolitically ungauged” basins. Using the proposed step-wise physiography and climate-
30 based regionalization approach, the model performance at ungauged basins reached 80% of
31 performance of locally calibrated parameters and significantly outperformed the global
32 regionalization parameters. The proposed water level based flow correlation was found to help
33 diagnose models and outperform the existing performance metrics of simulated water levels at
34 ungauged basins. The study also set up the first multi-national, multi-catchment hydrological
35 model in the Greater Mekong region, the top global biodiversity and major disaster risk hotspot in
36 the world through sequential and iterative refinement of the existing global hydrological model.
37 New model setup or existing models in the poorly-gauged and ungauged basins could be benefited
38 from the proposed approach to predict and evaluate model at ungauged basins.

39 *Keywords:* catchment model, regionalization, flow correlation, satellite observations, altimetry,
40 Mekong

41 **1. Introduction**

42 Adequate and reliable information about streamflows are imperative for effective
43 management of water resources. Streamflow data are required for practical applications such as
44 the design of drainage or water supply infrastructure, as well as planning short-term and long-term
45 water use with respect to changes of land use and climate. However, only a small fraction of
46 catchments in any part of the world, possess a stream gauge (Bloschl et al. 2013). Additionally,
47 the number of actively gauged stations has in recent years declined significantly due to reducing
48 government funds for monitoring networks (Ad Hoc Group et al., 2001; Shiklomanov et al., 2002).
49 Given the scarcity of operational gauging stations, the availability of streamflow data is becoming
50 increasingly limited.

51 In addition to the global trend of declining streamflow gauges, accessing existing data is
52 often more difficult in transboundary river catchments. Unfollowing the human-defined political
53 or administrative boundaries, transboundary river basins account for roughly one-half of the
54 earth's land surface, generate about 60% of the global freshwater flow and are home to nearly 40%
55 of the world's population (UNEP-DHI & UNEP, 2016). At least one transboundary water body
56 exists in almost every non-island state in the world. Even if international agreements enabling data
57 and information sharing among states exist in principle, in practice data sharing is often complex
58 in transboundary waters (Gerlak et al., 2010). For example, in the Okavango River basin, although
59 there is agreement between Namibia and Botswana on sharing river flow data, it is of question
60 how to validate the accuracy of shared data (Turton et al., 2003). Or in the case of the Jordan river
61 basin, where there are asymmetric power relations, intentionally ambiguous mechanisms were
62 designed by stronger states to allow no actual data exchange while diffusing domestic opposition
63 (Fishhendler, 2008). These transboundary river basins are thus considered as “geopolitically

64 ungauged” where data observation networks may exist but data are unavailable for use due to
65 geopolitical constraints (Kibler et al., 2014).

66 Since streamflow observations are not available and accessible for all locations,
67 hydrological models often rely on regionalization approaches to transfer information from gauged
68 to ungauged catchments (see Beck et al., 2016; Razavi and Coulibaly, 2013; Bloschl et al., 2013;
69 Parajka et al., 2013; Hrachowitz et al., 2013; He et al., 2011 for reviews). There are different
70 regionalization approaches with their respective advantages and limitations. In general, approaches
71 that transfer calibrated parameter sets with respect to their climatic and physiographic similarity
72 and/or simultaneously calibrate multiple catchments with those similar characteristics performed
73 better than other approaches (Arheimer et al., 2019; Beck et al., 2016; Donnelly et al., 2016;
74 Garambois et al., 2015; Sellami et al., 2014; Kim and Kaluarachchi 2008; Parajka et al., 2007).
75 Nonetheless, it is of question if this approach would work in the case of physically-based
76 distributed hydrological models, which inevitably have a large number of parameters.

77 Physically-based distributed hydrological models, which have parameters linked to
78 physiography and/or climate in the context of multi-catchment modeling approach (including both
79 gauged and ungauged basins), is a type of regionalization (Donnelly et al., 2016; Abbaspour et al.,
80 2015; Arheimer et al., 2019; [Hossain et al., 2017](#); [Mohammed et al., 2018](#)). Most of these studies
81 used physiography-linked parameter sets, except the most recent study by Arheimer et al. (2019)
82 that included similarity in climate characteristics. Among them, the study by Donnelly et al. (2016)
83 explicitly evaluated physiography-linked parameter sets and concluded that they were useful for
84 prediction at ungauged basins. The most recent study by Arheimer et al. (2019) included climate-
85 linked parameter sets by assigning different potential evapotranspiration algorithms for catchments
86 characterized according to Koppen climate classification. However, it is questionable if the choice

87 of climate regions based on Koppen classification was optimal due to no explicit quantification of
88 improvement in simulating streamflow.

89 The growing availability of spatially distributed remotely sensed data and open global data
90 sources, together with better computational capacity and advanced methods to assure better data
91 quality, has brought the possibility of macroscale hydrological modeling at the continental scale
92 (e.g. Pechlivanidis and Arheimer, 2015; Abbaspour et al., 2015; Donnelly et al., 2016) and the
93 global scale (Arheimer et al, 2019; Doll et al., 2003; Beck et al., 2016). However, it is known that
94 most of the global scale hydrological models do not always have satisfactory performance over all
95 stations within their expansive domains, constraining their application for management purposes.
96 Furthermore, the evaluation of model performance has been undertaken at gauged or pseudo-
97 gauged stations. Accordingly, it is of question how to discern at which station model can
98 satisfactorily capture the observed hydrological regimes. An innovative approach to evaluate
99 model at ungauged basins without using streamflow data is thus required.

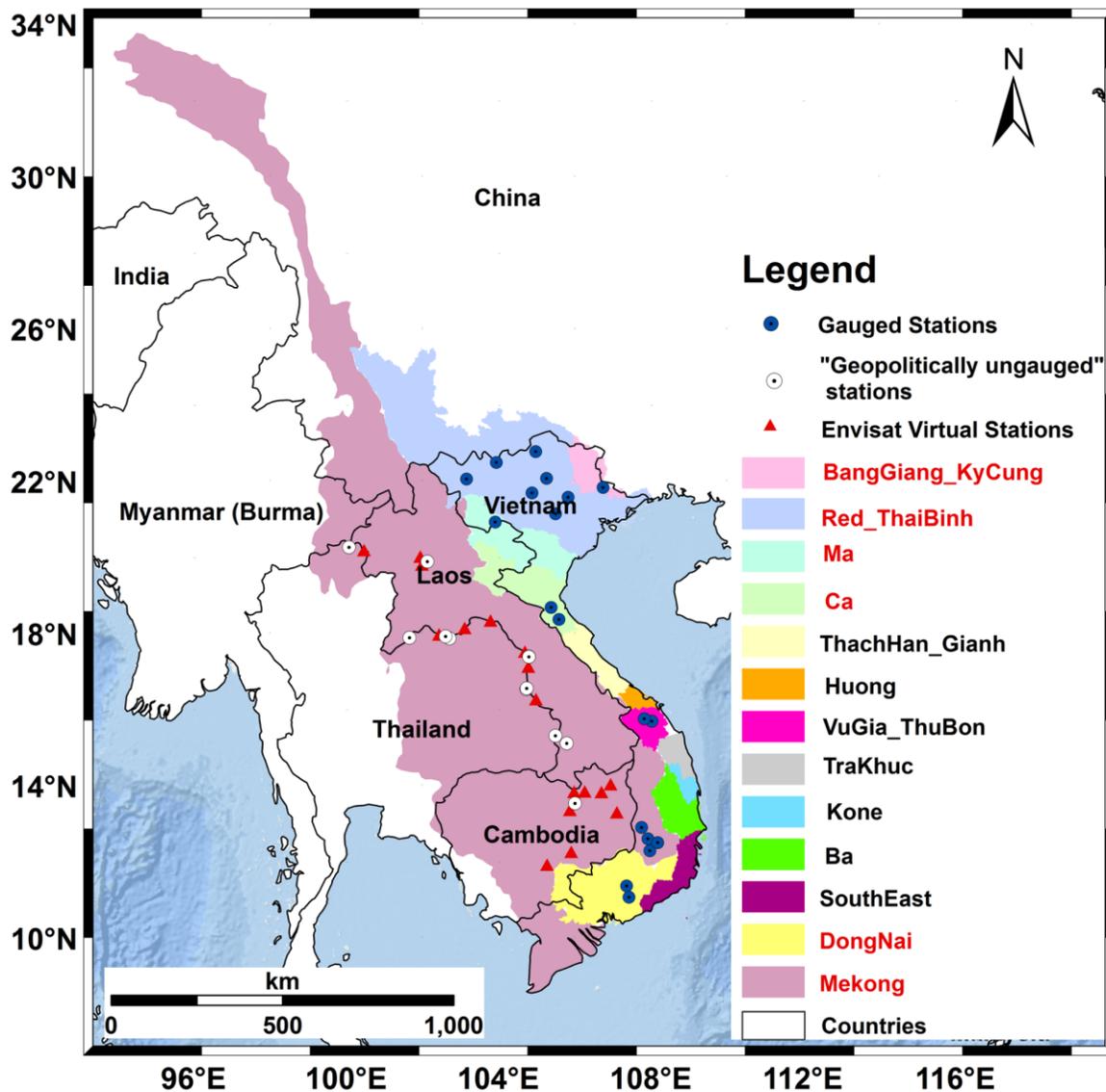
100 While streamflow data are less available and accessible, water level or stream stage data
101 are more widely obtainable because there is less investment of people and equipment to measure
102 them and increasingly more access and coverage of stage data derived from satellite altimetry
103 become available (Okeowo et al., 2017; Lee et al., 2009). Water level observations either from in
104 situ observations or derived from satellite altimetry have been increasingly used to calibrate
105 hydrological models towards replacing streamflow information for poorly and ungauged basins
106 (Getirana, 2010; Sun et al., 2012; Lindstrom, 2016; Jian et al., 2017) and used innovatively to
107 estimate important hydrological information for Mekong river basin (Kim et al., 2019a; Chang et
108 al., 2019). However, the evaluation of hydrological models using the existing performance metrics
109 based on water level only can yield inaccurate results due to inherent numerical problems

110 (Lindstrom, 2016; Jian et al., 2017). Therefore, it is vital to develop methods that more effectively
111 utilize water level data, where available, for evaluation of model at ungauged basins as a surrogate
112 for streamflow.

113 Similar flow dynamics (mean discharge, relative flow variability and catchment response
114 rates) have been found between catchments having high spatial correlation of daily streamflow (p
115 > 0.9), rather than catchments having spatial proximity (Archfield and Vogel, 2010; Betterle et al.,
116 2017; Betterle et al., 2019). Instead of using streamflow for ungauged catchments (receptor), water
117 level observations can be used to find the most highly correlated gauged catchments (donor). This
118 approach is named as water level based flow correlation in this study. If a model can simulate
119 similar correlation to the observed correlation patterns between gauged (using streamflow) and
120 ungauged (using water level), it is hypothesized that performance of simulated ungauged
121 catchments is as similar as gauged catchments. It is thus worth exploring whether this hypothesis
122 is valid.

123 Accordingly, the overarching goal of this work is to develop and test a new method of
124 using satellite observations and regionalization to improve the prediction of streamflow at
125 “geopolitical ungauged” basins using Hydrological Predictions for the Environment (HYPE) semi-
126 distributed hydrological model (Lindstrom et al., 2010). A first subcontinent-scale hydrological
127 model would be setup for the Greater Mekong region, which is a global biodiversity and major
128 disaster risk hotspot but poorly simulated in the existing global hydrological models, constraining
129 their use for pressing management purposes (Tordoff et al., 2012; Dilley et al., 2005; Du et al.,
130 2018). The region covers 13 river basins, of which six international river basins make up 90% of
131 total area, passing the entire territory of Vietnam, Laos, Cambodia and parts of China, Thailand
132 and Myanmar (Figure 1). The specific objectives of this work are to examine: (i) how far a multi-

133 catchment HYPE model using global open data sources including satellite observations can predict
 134 flow signatures for gauged catchments in the region; (ii) identify whether physiography and
 135 climate based regionalized parameters could improve prediction of streamflow signatures at
 136 ungauged catchments; (iii) determine whether water level based flow correlation could help to
 137 evaluate model performance at ungauged catchments.



138 **Figure 1.** The study domain of Greater Mekong, covering thirteen river basins, six of which are international
139 transboundary river basins (red colored legends) (including the entire territory of Vietnam, Laos, Cambodia and part
140 of China, Thailand and Myanmar) whereas the remaining river basins are located inside Vietnam

141 **2. Data**

142 **2.1. Input dataset for HYPE model**

143 The study used HYPE semi-distributed hydrological model, which has been examined in
144 extensive catchment types worldwide (Arheimer et al., 2019). In this study, HYPE for Greater
145 Mekong region is named as Greater Mekong HYPE (GMH), which has been developed
146 incrementally, and the current final version 1.3 (GHMv1.3) was based on the first version
147 (GMHv1). GMHv1 was the result of multiple collaboration works over multiple years between
148 Swedish Meteorological Hydrological Institute (SMHI) and National Center for Water Resources
149 Planning and Investigation (NAWAPI). To be comparable with the Worldwide HYPE model
150 version 1.3 (WWHv1.3, Arheimer et al., 2019), the catchment model HYPE for the Greater
151 Mekong region used the same topography and hydrological databases (Table 1). Additionally,
152 supplementary forcing and gauging data were used. In addition to Hydrological Global Forcing
153 Data (HydroGFD) precipitation and HydroGFD temperature (Berg et al., 2018), Multi-Source
154 Weighted-Ensemble Precipitation (MSWEP) precipitation, Tropical Rainfall Measuring Mission
155 (TRMM 3B42) precipitation and National Centers for Environmental Prediction Climate Forecast
156 System version 2 (NCEP CFSv2) temperature, which have been examined to perform well in the
157 region, were added (Tang et al., 2019; Mohammed et al., 2018). Since different forcing datasets
158 have different spatial resolutions (Table 1), the nearest grid approach was used to assign the
159 characteristics to each sub-catchment. Considering Vietnam to be the country that needs to monitor

160 water resources outside of the country (given 60% of water is generated outside of the country;
 161 World Bank, 2019; Du et al., 2016), any streamflow observations inside Vietnam are named as
 162 gauged catchments (used for calibration) whereas the observations outside of Vietnam, where
 163 available, are named as “geopolitically ungauged” catchments (gauged but not used for
 164 calibration). Sources of the additional ground observations of streamflow, water level and
 165 precipitation were supplemented by project partners to calibrate and validate the model
 166 performance in the region (see Figure 1 for their locations. Details of stations’ names, locations
 167 and basic information are provided in Table Supplementary 1).

168 **Table 1**
 169 Data description and sources used in the Greater Mekong HYPE project

Data type	Source and resolution	Reference
Topography (Flow accumulation, flow direction, digital elevation, river width)	SRTM (3 arcsec)	USGS
	HYDRO1k (30 arcsec)	UGGS
	GWD-LR (3 arcsec)	Yamazaki et al., 2014
Floodplains and Lake	Global Lake and Wetland Database (GLWD)	Lehner and Doll, 2004
Reservoirs and dams	Global Reservoir and Dam database v1.1 (GRanD)	Lehner et al., 2011
Land Cover characteristics	ESA CCI Landcover v1.6.1 epoch 2010 (300m)	ESA Climate Change Initiative – Land Cover project
Precipitation	MSWEP (0.25° grid, 1979 – 2014)	Beck et al. 2017
	TRMM 3B42 (0.25° grid, 2001 – 2015)	Huffman et al., 2006
	HydroGFD (0.5° grid, 1961 – 2015)	Berg et al., 2018
	In-situ precipitation stations in Vietnam (176 stations, 1975 – 2006)	BIG DREAM project (VINIF.2019.DA17)
Temperature	HydroGFD (0.5° grid, 1961 – 2015)	Berg et al., 2018
	NCEP CFSv2 (0.25° grid, 1979 – 2014)	Saha et al., 2011
Potential Evapotranspiration	MOD16A2 (8-day 1 km, 2001 – 2010)	Mu et al., 2016
Streamflow observations in Vietnam (Gauged) (used for calibration)	19 Stations (daily, 1980 - 2010)	BIG DREAM project (VINIF.2019.DA17)
Observations of streamflow and water level in Mekong (“geopolitically ungauged”) (used for independent evaluation)	12 Stations (daily, 1980 – 2007)	Mekong-SERVIR project (ADPC)
Envisat-derived Water Level (Envisat-“ungauged”*) (used for independent evaluation)	17 Virtual Stations (daily every 35 days, 2002-2009)	Okeowo et al., 2017; Lee et al., 2009; Chang et al., 2019; Kim et al., 2019; CTOH_ENVISAT_2014_01

170 * Envisat-“ungauged” catchments are catchments that have virtual stations of Envisat-derived water level but mostly have no
 171 observations of streamflow (except 3 catchments that are located in “geopolitically ungauged” catchments) (More explanations
 172 are provided in Table 2)

173 2.2. Radar Altimetry data

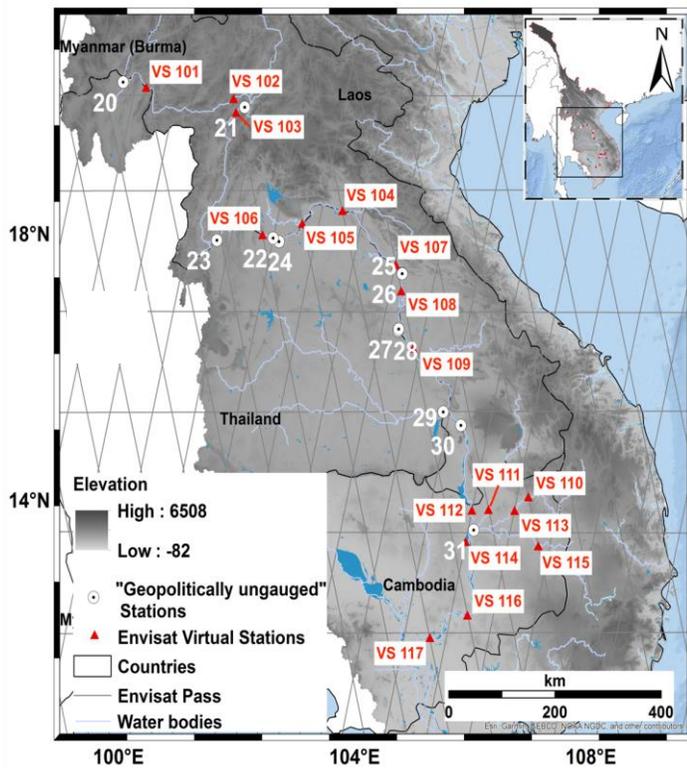
174 The heights of the earth surface every 35 days can be determined using the two-way travel
 175 time of radar pulses by Envisat Radar Altimeter 2 (RA 2) during period from August 2002 to
 176 October 2010 (see Figure 2 and Table 2 for their locations). Altimetric along-track data v2.1 of the
 177 Envisat mission (CTOH_ENVISAT_2014_01) corrected by CTOH (Centre de Topographie des
 178 Océans et de l'Hydrosphère, LEGOS, France) were extracted and time series were generated using
 179 the automation algorithm developed in Okeowo et al. (2017). This algorithm was based on K-
 180 means clustering for the automatic detection of outliers. Their method was found to be
 181 computationally effective compared to other methods, such as Kalman filter approach by Schwatke
 182 et al. (2015) and applicable in the Mekong river basin (Kim et al., 2019b).

183 **Table 2**
 184 List of seventeen virtual stations (VSs) with Envisat pass numbers and their location

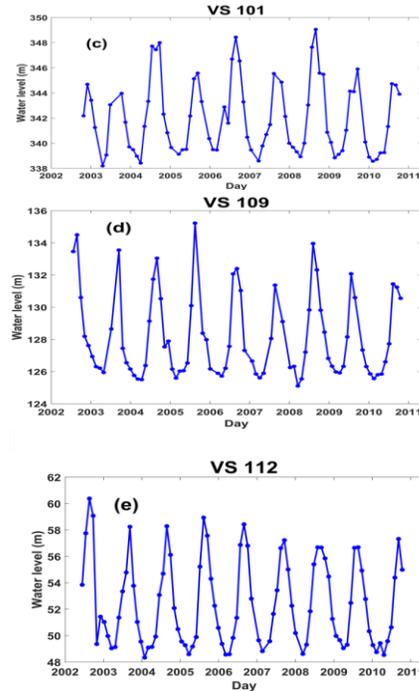
No.	VS	Pass number	Location (Lat/Lon)	Located in “Geopolitically ungauged” catchments?
1.	VS 101	737	20.195°N/100.472°E	No – Envisat-“ungauged” catchment (EU101)
2.	VS 102	651	20.025°N /101.950°E	No – Envisat-“ungauged” catchment (EU102)
3.	VS 103	651	19.817°N /101.994°E	No – Envisat-“ungauged” catchment (EU103)
4.	VS 104	565	18.345°N /103.796°E	No – Envisat-“ungauged” catchment (EU104)
5.	VS 105	107	18.151°N /103.115°E	No – Envisat-“ungauged” catchment (EU105)
6.	VS 106	651	17.980°N /102.443°E	No – Envisat-“ungauged” catchment (EU106)
7.	VS 107	21	17.531°N /104.699°E	Yes – (GU25) Nakhom Phanom Station (EU107)
8.	VS 108	21	17.137°N /104.789°E	No – Envisat-“ungauged” catchment (EU108)
9.	VS 109	21	16.279°N /104.990°E	Yes – (GU27) Savannakhek Station (EU109)
10.	VS 110	937	14.044°N /106.944°E	No – Envisat-“ungauged” catchment (EU110)
11.	VS 111	479	13.856°N /106.269°E	No – Envisat-“ungauged” catchment (EU111)
12.	VS 112	866	13.845°N /105.986°E	Yes – (GU31) Stung Treng Stations (EU112)
13.	VS 113	322	13.842°N /106.709°E	No – Envisat-“ungauged” catchment (EU113)
14.	VS 114	866	13.372°N /105.881°E	No – Envisat-“ungauged” catchment (EU114)
15.	VS 115	939	13.310°N /107.111°E	No – Envisat-“ungauged” catchment (EU115)
16.	VS 116	21	12.270°N /105.911°E	No – Envisat-“ungauged” catchment (EU116)
17.	VS 117	565	11.933°N /105.276°E	No – Envisat-“ungauged” catchment (EU117)

185 Note: “Geopolitically ungauged” catchments are catchments that actually have historical observations of daily streamflow and
 186 water level for only cross-validating the proposed method (not used at all for calibration) (Table 1). Envisat-“ungauged” catchments

187 are catchments that have virtual stations of Envisat-derived water level but mostly have no observations of streamflow (except 3
188 catchments that are located in “geopolitically ungauged” catchments). Among 17 Envisat-“ungauged” catchments, 3 of them are
189 located in “geopolitically ungauged” (shown in bold font), which can be validated with the actual streamflow observations.



"Geopolitically ungauged" stations:
 20 - Chiang Sean 26 - Thakhek
 21 - Luang Prabang 27 - Savannakhek (b)
 22 - Vientiane 28 - Mukdakhon
 23 - Chiang Khan 29 - Khong Chiam
 24 - Nong Khai 30 - Pakse



190
 191 **Figure 2.** Spatial distribution of virtual stations (red triangles) and “geopolitically ungauged” stations (white circles)
 192 employed in the study (Figure 2a). The black lines denote the ground tracks of Envisat altimetry. Figure 2b shows the
 193 names of “geopolitically ungauged” stations. Time series of river elevation at three VS’s are shown in the panels on
 194 the right (Figure 2c, 2d, 2e) (time series of all locations are not presented for reason of brevity).

195 **3. Methods**

196 **3.1 The multi-catchment hydrological model HYPE**

197 HYPE is a process-oriented semi-distributed open-source model that is developed and used
 198 operationally to deliver high-resolution model predictions of water and nutrients (Lindstrom et al.,
 199 2010; Arheimer and Lindstrom, 2013). Initially developed for use in Sweden, it has more recently
 200 been used in applications in, for example, India (Pechlivanidis and Arheimer, 2015), Europe

201 (Donnelly et al., 2016), and across the globe (Arheimer et al., 2019). The HYPE model code has
202 been developed since 2005 with a flexible approach to start with simple process descriptions and
203 further refine and increase complexity when necessary (Lindstrom et al., 2005; Bergstrom 1991;
204 Beven 2001). The model structure is based on a multi-catchment approach allowing simultaneous
205 modeling of multiple river basins, with each river basin divided into multiple subbasins and each
206 subbasin further divided into hydrologic response units (HRUs). Each HRU can be divided
207 vertically into three maximum distinct soil layers (normally the top layer has a thickness of around
208 25 cm, the second of 1-2 meters and the third can be deeper to account for ground water) (Bui et
209 al., 2011). The model is forced by precipitation and temperature at either daily or hourly temporal
210 resolution, and its calculation starts at HRUs and is then aggregated to subbasin level. HYPE
211 calculates flow paths in the soil based on snow melt, evapotranspiration, surface runoff,
212 infiltration, percolation, macropore flow, tile drainage and outflow to the stream from soil layers
213 when water content is above field capacity. Different algorithms are provided to calculate snow
214 melt, evapotranspiration, and infiltration according to the physical characteristics of the modeled
215 catchments. The runoff from the land classes is then routed through the network of rivers and lakes
216 to generate river flow, which could be dampened due to effect of lakes and reservoirs. HYPE can
217 also simulate the effect of floodplains, which is crucial for large river systems and their deltas
218 (Andersson et al., 2017), and it can also simulate the transport and concentration of nutrients in
219 both soil, rivers and lakes (Lindstrom et al., 2010). In addition to natural dynamics, the model can
220 simulate simplified water management schemes, such as regulated reservoirs (hydropower), and
221 irrigation. There are several parameters used in HYPE that can be constrained in a stepwise manner
222 using different types of observed data (Arheimer and Lindstrom 2013). The parameters may be
223 soil type dependent (e.g., field capacity), land cover dependent (e.g., evapotranspiration

224 coefficient) or general across the domain (e.g., river routing parameters). Parameters, which are
225 linked to physiography and/or climate rather than to a specific catchment, are thus assumed to be
226 transferable to ungauged sites. More details on the HYPE model, including visual schematic
227 diagram, can be found in the web-based documentation (<http://www.smhi.net/hype/wiki/>) and
228 Lindstrom et al. (2010).

229 The HYPE model has the explicit lake routing, including two types of lakes, which are
230 local lakes and outlet lakes. Local lakes, which are located inside the subbasin, only receive a
231 portion of local surface runoff and then flows to main river of the same subbasin. Outlet lakes,
232 which are located near the main river, receive both local runoff (after it has passed local lakes) and
233 the river flow from upstream subbasins. Each lake can be set with an individually defined depth.
234 The outflow from lakes (when water level is above a defined threshold) can be either determined
235 by a general rating curve or a specific rating curve.

236 The rating curve for a lake outlet is written as:

$$237 \quad q = k(w - w_0)^p \quad (1)$$

238 Thus, water level can be also seen as a transformation of streamflow:

$$239 \quad (w - w_0) = (q/k)^{1/p} \quad (2)$$

240 where q is the outflow or streamflow (m³/s), w is water level (m), w_0 a threshold (m) and p is an
241 exponent (Lindstrom, 2016). When w_0 is known, $(w - w_0)$ is equal to water depth.

242 Lake water levels, which are easier and cheaper to be measured, are mainly used with the
243 purpose of estimating streamflow through established rating curves. Meanwhile, there are a lot of
244 basins that have no observations of streamflow. Therefore, Lindstrom (2016) tested if HYPE can
245 be calibrated using water level data instead of streamflow. The study found that water levels could
246 be useful for calibration of hydrological models without measuring streamflow by establishing a

247 traditional rating curve but using a constant rating curve exponent. His suggestion of using $p = 2$
248 while adjusting k and w_0 appropriately for all lakes resulted in a reasonable agreement with
249 observed daily water level records based on the assumption of parabolic lake outlets, which agreed
250 with the previous study by Maidment (1992).

251 Accordingly, to integrate river elevation derived from Envisat altimetry into HYPE,
252 modeled streamflow must be converted to water level. Using outlet lake routine from HYPE,
253 negligibly small outlet lakes, which have inflow equal to outflow (no storage capacity to affect
254 streamflow) were added in subbasin, where there is either in-situ observations of streamflow and
255 water level or Envisat-derived river elevation. To reduce uncertainty of estimating water depth, in
256 addition to constant p , constant $k = 100$ and $w_0 = 0$ were used, so equation (2) becomes $w =$
257 $(q/100)^{1/2}$. Since the proposed method (water level based flow correlation, explained in section
258 3.3) emphasized the temporal dynamics rather than the true magnitude of a variable, it was not
259 necessary to estimate the exact water depth. Simulated w would be compared with either in-situ
260 observations of water level at “geopolitically ungauged” catchments or Envisat-derived water
261 elevation at Envisat-“ungauged” catchments. Envisat-“ungauged” catchments are catchments that
262 have virtual stations of Envisat-derived water level but mostly have no observations of streamflow
263 (except 3 catchments that are located in “geopolitically ungauged” catchments) (Table 2).

264 **3.2. Grouping catchments using climatic indexes**

265 Similar seasonal water balance patterns between catchments, which could be explored
266 based on three climatic indices alone, i.e., climatic aridity, timing of seasonal precipitation, and a
267 temperature-based measure of snowiness, was found to provide a useful backdrop to the signatures
268 of streamflow variability over various time scales (daily to decadal) and states (low flow to floods)

269 (Berghuijs et al., 2014). This study applied Berghuijs et al. (2014)'s approach to robustly group
 270 catchments based on their similarity in climatic characteristics. Accordingly, two dimensionless
 271 indices that account for similar water balances among catchments were calculated, namely the
 272 aridity index and the seasonality index (snowiness is not considered in this study since there is
 273 almost no snow impact in the study area).

274 Proposed by Budyko (1974), the aridity index is defined as:

$$275 \quad \varphi = \frac{\bar{E}}{\bar{P}} \quad (3)$$

276 where \bar{E} is the average potential evaporation rate (mm/day) and \bar{P} is the average precipitation rate
 277 (mm/day). This average is calculated from 2002-2009, the same time period used to calibrate the
 278 model. φ can range from 0 to infinity (in theory) with higher values associated with more arid
 279 climate.

280 Here, it is assumed that the seasonal variability of precipitation and air temperature can be
 281 modeled as simple sine curves (Milly, 1994; Potter et al., 2005; Woods, 2009) as follows:

$$282 \quad P(t) = \bar{P} \left[1 + \delta_P \sin \left(\frac{2\pi(t - s_P)}{\tau_P} \right) \right] \quad (4)$$

$$283 \quad T(t) = \bar{T} + \Delta_T \left[\sin \left(\frac{2\pi(t - s_T)}{\tau_T} \right) \right] \quad (5)$$

284 where t is the time (days), s is a phase shift (days), τ is the duration of the cycle under
 285 consideration (here, 365 days), \bar{P} is the average precipitation (mm/day), \bar{T} is the average
 286 temperature ($^{\circ}\text{C}/\text{day}$) over same period 2002-2009, δ_P and Δ_T are dimensionless seasonal
 287 amplitudes, and the subscripts P and T stand for precipitation (mm/day) and temperature ($^{\circ}\text{C}/\text{day}$)
 288 respectively. $P(t)$ is the precipitation rate (mm/day) and $T(t)$ is temperature ($^{\circ}\text{C}/\text{day}$) as a function

289 of t . Using a least squares optimization, δ_P and Δ_T were obtained for all individual 1120
290 catchments in the HYPE study domain.

291 Then, the seasonality index δ_P^* was calculated using Woods (2009):

$$292 \quad \delta_P^* = \delta_P \cdot \text{sgn}(\Delta_T) \cdot \cos\left(\frac{2\pi(s_P - s_T)}{\tau}\right) \quad (6)$$

293 where δ_P^* indicates whether precipitation is in phase with the potential evaporation and temperature
294 regimes. The parameter δ_P^* can range from -1 to +1, with the former representing strongly winter-
295 dominant precipitation (P out of phase with T) and the latter showing strongly summer-dominant
296 precipitation (P in phase with T). $\delta_P^* = 0$ indicates the uniform precipitation throughout the year.

297 **3.3. Water level based flow correlation between gauged and “ungauged” catchments**

298 A measured correlation matrix (Pearson’s r correlation coefficient) between daily in-situ
299 water level of “geopolitically ungauged” catchments and daily streamflow of gauged catchments
300 was calculated to find the most highly correlated reference gauged catchments to the study
301 “ungauged” catchments. Similarly, a measured correlation matrix between Envisat-derived water
302 level of “ungauged” catchments outside of Vietnam and the daily streamflow of gauged
303 catchments was also computed. To examine the assumption that the correlation between two daily
304 streamflow series was similar to water level based flow correlation between daily water level
305 (either insitu observations of water level or Envisat-derived water level) and streamflow, a
306 corresponding correlation matrix between daily streamflow of “geopolitically ungauged”
307 catchments and daily streamflow of gauged catchments was made. From previous studies on flow
308 correlation, correlation coefficients larger than 0.9 were recommended to consider as being highly
309 correlated catchments (Archfield and Vogel, 2010; Betterle et al., 2019). Because there were less

310 catchments considered in the study, $r \geq 0.7$ was selected as the threshold correlation coefficient.
311 Because r was smaller in the study, only catchments in the same climatic group (section 3.2) were
312 examined, to ensure they have similar climate characteristics.

313 **3.4. Step-wise physiography and climate based regionalization at gauged basins**

314 For data-sparse regions, step-wise calibration approach was shown to be a useful method
315 to reduce the problem of equifinality of the final model output (Stromqvist et al., 2012; Arheimer
316 and Lindstrom, 2013; Donnelly et al., 2016; Andersson et al., 2017). At each key process, lumped
317 calibration was carried out simultaneously for sub-groups of gauged basins (representative gauged
318 basins - RGBs) with upstream areas dominated by a specific land-use or soil type. When calibration
319 for a specific group of RGBs is deemed satisfactory, the parameters for that responding land-use
320 or soil type can be kept constant and the next parameters for another group can be calibrated using
321 another set of RGBs. The step-wise separation followed the hydrological pathways through the
322 landscape, starting with climate inputs (precipitation, evapotranspiration), then subsequently
323 moving downstream to soils (infiltration, storage, runoff), then the rivers and lakes (routing and
324 storage). After each step, evaluation of model performance was undertaken for all 19 gauged
325 stations and the best performance parameter set was used in the next step of the model refinement.
326 The period 2002 - 2009 was selected as the calibrated period to analyze errors and refine the model.
327 This period was chosen because it aligned with the availability of Moderate Resolution Imaging-
328 Spectroradiometer (MODIS) - derived potential evapotranspiration (PET) and Envisat-derived
329 water level. The earlier part of the simulation period (1991-2001) was retained for independent
330 validation at the same stations.

331 A key objective in calibrating the Greater Mekong-HYPE model was to represent the main
332 hydrological processes of all river basins. Therefore, model evaluation and refinements primarily
333 focused on achieving satisfactory performance across the whole basin using consistent descriptions
334 rather than excellent performance at few locations. The streamflow signatures to be evaluated in
335 the study were the daily and monthly specific streamflow (mm/day and mm/month), high flow (5th
336 percentile of daily specific flow in mm/day), low flow (95th percentile of daily specific flow in
337 mm/day) and medium flow (50th percentile of daily specific flow in mm/day). These signatures
338 were selected because they are the most important and widely used signatures of catchment runoff
339 response to be applied in water resources planning and environmental studies (Arheimer et al.,
340 2019; Donnelly et al., 2016) (Table 3).

341 The entire domain-scale performance was quantified by first calculating key performance
342 criteria for each of the above flow signatures at each of the 19 streamflow gauges available inside
343 Vietnam (Figure 1), and then computing summary statistics to describe model performance across
344 all locations. The model's ability to simulate daily and monthly streamflow at each gauge was
345 quantified with standard metrics, including the Kling-Gupta Efficiency (KGE) and its components
346 r , β , α , which are directly linked with Pearson's correlation coefficient, relative error (RE) and
347 relative error of standard deviation (RES_D, variability ratio) respectively (Gupta et al., 2009)
348 (Table 4). For constraining PET parameter values, absolute value of RE was used to find the best
349 agreements between modeled PET and MODIS-derived PET.

350 Both automatic and manual calibration approach were employed to take advantage of
351 strengths of both methods. The advantage of the former is power and speed of computation and
352 objective parameter constraints. Nevertheless, it is unlikely to provide physically acceptable
353 parameter estimates, which are mostly addressed by highly labor-intensive manual calibration

354 (Boyle et al., 2000). The automatic approach was the Differential Evolution Markov Chain
355 (DEMC) method (Ter Braak, 2006). DEMC allowed to examine parameter sensitivity, probability
356 based uncertainty estimate and a better convergence towards the global optimum. Two-step DEMC
357 automatic calibration was undertaken. Firstly, short runs (around 400 iterations) were done to
358 examine parameter sensitivity. Secondly, longer runs (with at least 1000 iterations) were
359 undertaken for only sensitive parameters to allow convergence to global optimum values. DEMC
360 automation was then followed by manual checks to ensure the physically acceptable parameter
361 ranges and simulated hydrograph similar to the observed patterns. Table 5 describes the model
362 parameters to be calibrated and lists the initial parameter values for each parameter. [Other
363 parameters were kept as default as the baseline parameters from the first Greater Mekong HYPE
364 model version \(GMHv1\) \(the same roughly calibrated parameter sets of the first Worldwide HYPE
365 model version 1.0 \(WWHv1.0, Arheimer et al., 2019\).](#)

366 Step-wise physiography and climate based regionalization framework for estimating
367 different groups of model parameter values in each step were as follows:

368 (1) For precipitation and temperature, different datasets of precipitation and temperature were
369 used with the baseline parameters from WWHv1.0 (roughly calibrated model at global
370 scale) without undertaking any additional calibration to identify the optimal climate forcing
371 datasets for the region. Daily KGE was used to evaluate this step. This model step after
372 selecting the optimal climate data was named as the Greater Mekong HYPE model version
373 1.0 (GMHv1.0).

374 (2) PET parameter values (lb, kc5, alb: see Table 5 for description of parameter values) were
375 constrained using the absolute value of RE between annually simulated PET and MODIS-
376 derived PET. PET algorithm selected in the study was Food and Agriculture Organization

377 (FAO) Penman-Monteith, which was integrated inside the HYPE model (Allen et al., 1998;
378 Monteith, 1965). This algorithm was selected so that it was more comparable to MODIS-
379 based PET, which was also based on Penman-Monteith logic (Mu et al., 2016). Two-step
380 DEMC automatic calibration was undertaken to obtain the optimal values for each RGB of
381 each land cover type (10 main land cover types were grouped from 36 European Space
382 Agency (ESA) Climate Change Initiative CCI v1.6 data, see details of land cover
383 description and grouping in Arheimer et al., 2019). Thirdly, manual checks of parameter
384 values for each group were made to ensure their acceptable physical meaning. This model
385 step after selecting optimal parameter values was named as the Greater Mekong HYPE
386 model version 1.1 (GMHv1.1).

387 (3) Parameters related to soil storage, flow paths and runoff generation (19 parameters
388 provided in Table 5) were first optimally tuned by two-step DEMC calibration with daily
389 KGE used as the objective function. Because all gauged and ungauged basins were mainly
390 vegetated areas, only parameters for vegetated soils were calibrated. The remaining
391 parameters were kept as default. Following automatic calibration, manual check was done
392 to examine the physical meaning of parameters, and hydrograph simulation of other
393 signatures (daily streamflow, Q95, Q5 and Q50). This model step after selecting the
394 optimal parameter values was named as the Greater Mekong HYPE model version 1.2
395 (GMHv1.2).

396 (4) Each catchment group (section 3.2) was evaluated separately and calibrated using regional
397 correction parameters (cevpcorr, rrscorr: see Table 5 for description of parameter values)
398 (Hundecha et al., 2013). Two-step DEMC with daily KGE as the objective function and
399 manual checks were done for all flow signatures (daily streamflow, Q95, Q5 and Q50).

400 This model step after selecting the optimal regional correction parameter values was named
 401 as the Greater Mekong HYPE model version 1.3 (GMHv1.3).

402 **Table 3**
 403 Flow signatures evaluated in the study (Range estimated from 2002 – 2009 period)

Flow Signatures	Description
MeanDailyQ (QDD)	Mean daily specific flow in mm
MeanMonthlyQ (QMM)	Mean monthly specific flow in mm
Q5	5 th percentile of daily specific flow in mm
Q50	50 th percentile of daily specific flow in mm
Q95	95 th percentile of daily specific flow in mm
MeanDailyW(WDD)	Mean daily water level in m
MeanMonthlyW (WMM)	Mean monthly water level in m

404
 405 **Table 4**
 406 Performance metrics used in the study

Performance metrics	Equation / References	Range	Variables
KGE (Kling-Gupta Efficiency)	$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}$ (Gupta et al., 2009)	Negative Infinity to 1 (the closer to 1, the better simulation)	QDD, QMM
RE (Relative error)	$\beta = \frac{\mu_s}{\mu_o}; RE = (\beta - 1) \cdot 100$ (Gupta et al., 2009)	Infinity to Infinity (the closer to 0, the better simulation)	QDD, QMM, Q5, Q50, Q95
RESD (Relative Error of Standard Deviation)	$\alpha = \frac{\sigma_s}{\sigma_o}; RESD = (\alpha) \cdot 100$ (Gupta et al., 2009)	Infinity to Infinity (the closer to 0, the better simulation)	QDD, QMM
Pearson's r Correlation Coefficient	$r = \frac{cov(x_o, x_s)}{\sigma_s \sigma_o}$	-1 to 1 (the closer to -1 or 1, the better simulation)	QDD, QMM, WDD, WMM
NSEW	$NSEW = NSE + \frac{(\beta - 1)^2}{\sigma_o^2}$ (Lindstrom, 2016)	Negative Infinity to 1 (the closer to 1, the better simulation)	WDD, WMM
NSE _{anom}	$NSE_{anom} = 1 - \frac{\sum_{t=1}^{nt} \{ [x_o(t) - \bar{x}_o] - (x_s(t) - \bar{x}_s) \}^2}{\sum_{t=1}^{nt} [x_o(t) - \bar{x}_o]^2}$ (Getirana, 2010)	Negative Infinity to 1 (the closer to 1, the better simulation)	WDD, WMM

407 Note. x represents the streamflow or water level time series. μ : the mean value of streamflow or water level time series. σ : the
 408 standard deviation of streamflow or water level time series. The sub-indexes o and s are observed and simulated streamflow or
 409 water level time series, respectively. t is the time step (one month for this application), nt is the total number of months.

410 **Table 5**
 411 HYPE Model Parameter Description, Initial parameter range and Posterior parameter values

Hydrological Process	Parameter and description	Initial Parameter Range	Posterior parameter values
Potential evapotranspiration	lb: threshold soil water for activation of PET	0.9	0.9
	kc5: crop coefficient for Penman-Monteith algorithm	[0.9 – 1.4]	[1.2 – 1.9]
	alb: albedo for PET algorithms	[0.12 – 0.23]	[0.12 – 0.23]
Soil water storage and flow path (for vegetated soil and land uses)	rrcs1: recession coefficient for uppermost soil layer	0.3	0.3
	rrcs2: recession coefficient for lowest soil layer	0.03	0.015
	rrcs3: recession coefficient for slope dependent	0.0002	0.0002
	mperc1: maximum percolation capacity from soil layer 1 to soil layer 2	20	20
	mperc2: maximum percolation capacity from soil layer 2 to soil layer 3	20	50
	macrate: fraction for macro-pore/subsurface flow	0.3	0.4
	mactrinf: threshold for macro-pore/subsurface flow	10	6
	mactrsm: threshold soil water for subsurface and surface runoff	0.7	0.1
	srrate: fraction for infiltration excess surface runoff (Horton overland flow)	0.04	0
	wcwp1: wilting point as a fraction for uppermost soil layer	0.2	0.2
	wcwp2: wilting point as a fraction for second soil layer	0.2	0.2
	wcwp3: wilting point as a fraction for lowest soil layer	0.2	0.2
	wcfc1: fraction of soil available for evapotranspiration for uppermost soil layer	0.15	0.15
	wcfc2: fraction of soil available for evapotranspiration for second soil layer	0.15	0.15
	wcfc3: fraction of soil available for evapotranspiration for lowest soil layer	0.15	0.15
	wcep1: effective porosity as a fraction, for uppermost soil layer	0.04	0.015
	wcep2: effective porosity as a fraction, for second soil layer	0.04	0.3
	wcep3: effective porosity as a fraction, for lowest soil layer	0.04	0.4
	srrcs: recession coefficient for saturated surface runoff (Dunne overland flow)	[0.05 – 0.2]	[0 – 0.4]
	Seasonal water balances among catchment groups	cevpcorr: correction factor for PET	0
rrccorr: correction factor for soil recession coefficient		0	[-0.5 – -0.2]

412 Note. Posterior parameter values different from initial parameter range are shown in bold font.

413 3.5. Performance of regionalized parameters at ungauged basins

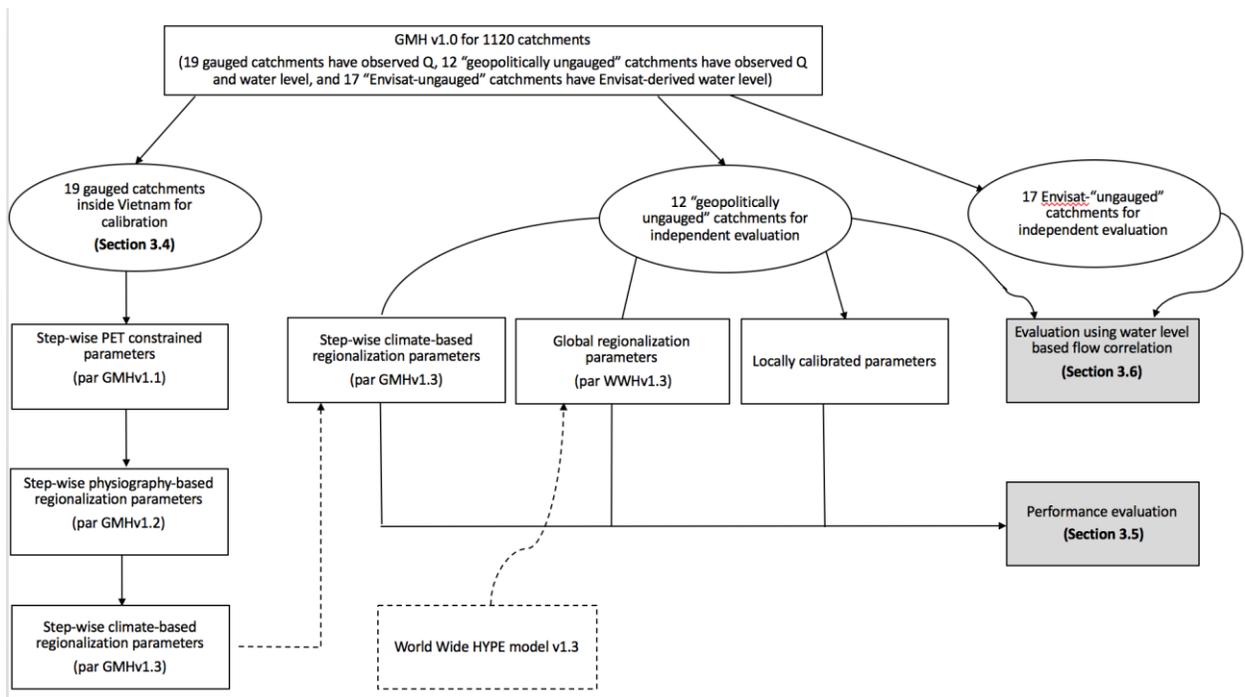
414 The performance of physiography and climate based regionalized parameters was assessed
 415 with the following approach. At 12 “geopolitically ungauged” to be used for independent
 416 evaluation (Figure 3), KGE, RE for daily streamflow, RE for Q95, Q5, Q50 were obtained by
 417 using the following sets of parameters:

418 (1) Step-wise physiography and climate based regionalization parameters transfer from
419 gauged catchments (par GMHv1.3) (section 3.4).

420 (2) Global regionalization parameters from the WWHv1.3 (Arheimer et al., 2019). This
421 parameter set was forced with the same climate data as the WWHv1.3 model (HydroGFD
422 Precipitation and Temperature).

423 (3) Locally calibrated parameters in the ideal situation where observed streamflow were
424 available for calibration (Step one to three of section 3.4 without manual calibration so that
425 selected parameters can be objective).

426 The performance metrics from the three parameter sets were compared to address research
427 objective 2 (section 1) if the proposed regionalization method could help improve prediction of
428 streamflow signatures at ungauged basins.



429
430
431

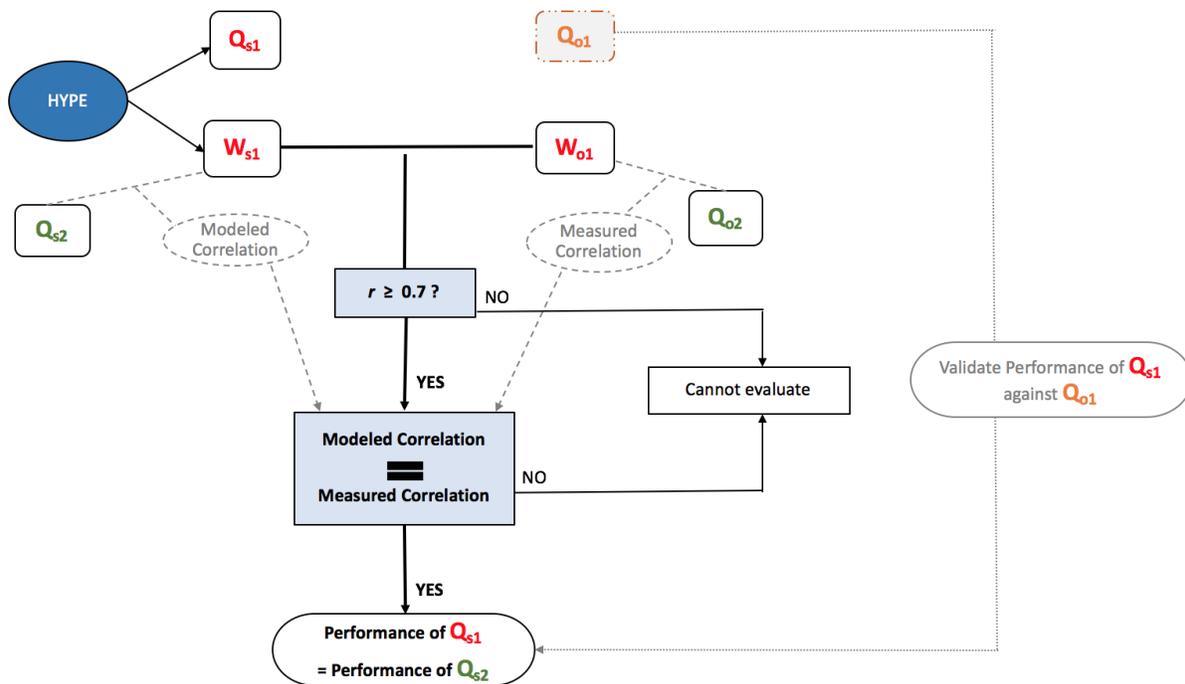
Figure 3. Flow chart summarizing steps of proposed method in the study

432 **3.6. Model evaluation at ungauged basins using water-level based flow correlation**

433 At 12 “geopolitically ungauged” and 17 Envisat-“ungauged” catchments, which were
434 assumed to have no observations of streamflow but only water level, different performance metrics
435 were used to examine if water level can be modelled with a satisfactory level of performance.
436 Pearson’s correlation coefficient was first examined to see if the modeled water level has similar
437 temporal dynamics with the recorded water level, irrespective of their magnitudes. Getirana (2010)
438 proposed Nash-Sutcliffe efficiency for anomalies (NSEanom), which was a modified NSE metrics
439 to eliminate the anomalies or bias caused by different reference water level between modeled and
440 recorded water levels. Similarly, Lindstrom (2016) introduced Nash-Sutcliffe efficiency adjusted
441 for bias (NSEW) to eliminate bias between them (Table 4). However, due to numerical problems,
442 both equations can still yield inaccurate results because magnitudes of variation between the two
443 variables are significantly different (modeled water level with few meters whereas recorded water
444 level with several hundred meters above sea level).

445 Because of numerical problems to evaluate the performance of catchments based on water
446 level only, this study proposed applying hydrologic similarity theory by assuming that the most
447 highly correlated reference gauged catchments (using daily streamflow) also have similar
448 performance to that of the study “ungauged” catchments (using water level). To use this method,
449 first, the modeled water levels of the “ungauged” catchments were evaluated against recorded
450 water levels using Pearson’s correlation coefficient, NSEanom and NSEW. For in-situ water
451 levels, evaluation was undertaken at both daily and monthly time steps. For Envisat-derived water
452 levels, evaluation was performed at any day step that has recorded data (one daily observation
453 every 35 days). NSEanom and NSEW were used only to examine if numerical problems of
454 evaluating models based on water levels existed. When only modeled water level had good

455 correlation with recorded water level ($r \geq 0.7$), following steps were undertaken. This condition
 456 ensured that the temporal variation of modeled water levels against observations was captured.
 457 Secondly, the modeled correlation between modeled water levels of the “ ungauged ” catchments
 458 and modeled streamflow of the reference most highly correlated catchments was computed. If
 459 there was similar result between modeled correlation and measured correlation (modeled
 460 correlation can range from 0.5 to 0.9 compared to measured correlation), performance of reference
 461 gauged catchment was assumed to be the performance of “ ungauged ” catchment. To cross-validate
 462 this assumption, performance of “ ungauged ” catchments against the historical observations of
 463 streamflow, where available, was evaluated and compared with the assumption (Figure 4).



464
 465 **Figure 4.** Model evaluation framework at “ ungauged ” catchments using water level based flow correlation method.
 466 Q, W represents daily or monthly streamflow and water level time series respectively. The subscripts s and o are
 467 simulated and observed time series respectively. The second subscripts 1 and 2 are “ ungauged ” (either “ geopolitical
 468 ungauged ” or Envisat-“ ungauged ”) and their most highly correlated gauged catchments respectively. Qo1 (if available)
 469 is not used in model setup or calibration, but only used to cross-validate the assumption that performance of

470 “ungauged” catchments is similar to that of the reference most highly correlated gauged catchments. The expression
471 “=” is understood as between +/- 0.2 (so modeled correlation can range from 0.5 to 0.9 compared to measured
472 correlation).

473 **4. Results**

474 **4.1. Catchment delineation and characteristics**

475 The World Hydrological Input Set-up Tool (WHIST) developed by SMHI (Swedish
476 Meteorological and Hydrological Institute, developer of HYPE model) was used to delineate
477 catchment borders (Arheimer et al., 2019). Consistent with the WWHv1.3, catchment delineation
478 was defined using the same approach according to the locations of gauging stations in the river
479 network (including 19 “gauged” stations and 12 “geopolitically ungauged”), the outlets of large
480 lakes/reservoirs, and seeking to reach an average catchment size of ~ 1,000 km² (Arheimer et al.,
481 2019). As a result, the Greater Mekong region (~1,2 million km²) was divided into 1,120 sub-
482 catchments with an average size of 1,047 km². Sub-catchments within low-lying areas with
483 extensive floodplains tended to have a larger size (average 3,600 km²), among which the TonleSap
484 basin had the largest size of 10,000 km². The outputs of catchment delineation were quality
485 checked with station metadata (obtained from governmental reports). 100% of the estimated
486 catchment areas were found to fall within +/- 5% of the areas reported by these metadata. For lakes
487 and reservoirs, in total, 15 lakes and 18 reservoirs (only lakes and reservoirs larger than 10 km²
488 recorded by GLWD and GRanD were considered in this version) were identified.

489 Similar to WWHv1.3, HRUs represented a combination of land cover characteristics and
490 elevation, resulting in 169 HRUs (details of HRUs can be found in Arheimer et al., 2019). Different
491 hydrological active soil depths were assigned for the HRUs, based on the variability in vegetation,

492 and elevation they represented as suggested by Troch et al. (2009) and Gao et al. (2014) and
493 currently used in WWHv1.3 (Arheimer et al., 2019). Similar to WWHv1.3, detailed description of
494 soil properties was not included in HYPE model to reduce number of parameters. Nevertheless,
495 five general distinct soil classes (including (i) no soil (water), (ii) urban soil, (iii) rock (no texture),
496 (iv) vegetated soil and (v) irrigated soil) based on impermeable conditions and infiltration of land
497 covers were identified to describe the hydrological processes in the region.

498 **4.2. Grouping catchments using climatic indexes**

499 Across all 1,120 catchments and during the 2002 to 2009 study period, the aridity index
500 ranged from 0.4 to 1.7 whereas seasonality index ranges from -0.3 to 1. Accordingly, consistent
501 with Berghuijs et al. (2014), four catchment groups were made, including group (1): Humidity
502 ($\varphi \leq 0.75$) with Mild seasonality ($\delta_p^* \leq 0.5$); group (2): Humidity ($\varphi \leq 0.75$) with High seasonality
503 ($\delta_p^* > 0.5$); group (3): Sub-humidity ($\varphi > 0.75$) with Mild seasonality ($\delta_p^* \leq 0.5$); and group (4):
504 Sub-humidity ($\varphi > 0.75$) with High seasonality ($\delta_p^* > 0.5$) (Table 6). Figure 5 shows the
505 geographic spread and organization of four catchment groups obtained from this classification
506 approach (See Figure Supplementary 2 for spatial distribution of all catchments based on group
507 classification). Most catchments having historical streamflow observations (both gauged and
508 ungauged) were classified as group 3 or group 4. Catchments of group 1 were mostly located near
509 the coastal area with stronger humidity and more wet-season dominant precipitation. Catchments
510 of group 3 were located mostly in the southwest of the region with less humidity and less seasonal
511 water variability. Catchments of group 4 were located mostly in the northwest of the region with
512 less humidity and more dry-season dominant precipitation.

513 Grouping catchments using climatic indexes could provide a robust reference to further
514 regionalize parameters for each climate group. This study adapted a simple regional calibration
515 approach, following Hundecha et al. (2016). After step-wise calibration of the model for all
516 catchments, we evaluated the model for each catchment group to find out which signatures need
517 to be refined and then performed regional calibration separately by using group-specific correction
518 parameters. It should be noted that only catchment group 1, 3 and 4 could be calibrated and
519 validated whereas catchment group 2 had no validation because there were no gauged stations for
520 this group.

521 **Table 6**
522 Catchment groups using climatic indexes

Group	Description	Total catchments	Available observations within		
			Gauged catchments	“Geopolitically ungauged” catchments	Envisat-“ungauged” catchments
1	Humidity with Mild Seasonality	54	4	0	0
2	Humidity with High Seasonality	68	0	0	1
3	Sub-humidity with Mild Seasonality	322	5	3	9
4	Sub-humidity with High Seasonality	680	10	9	7

523 Note: Humidity ($\varphi \leq 0.75$); Sub-humidity ($\varphi > 0.75$); Mild seasonality ($\delta_p^* \leq 0.5$); High seasonality ($\delta_p^* > 0.5$).

524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550

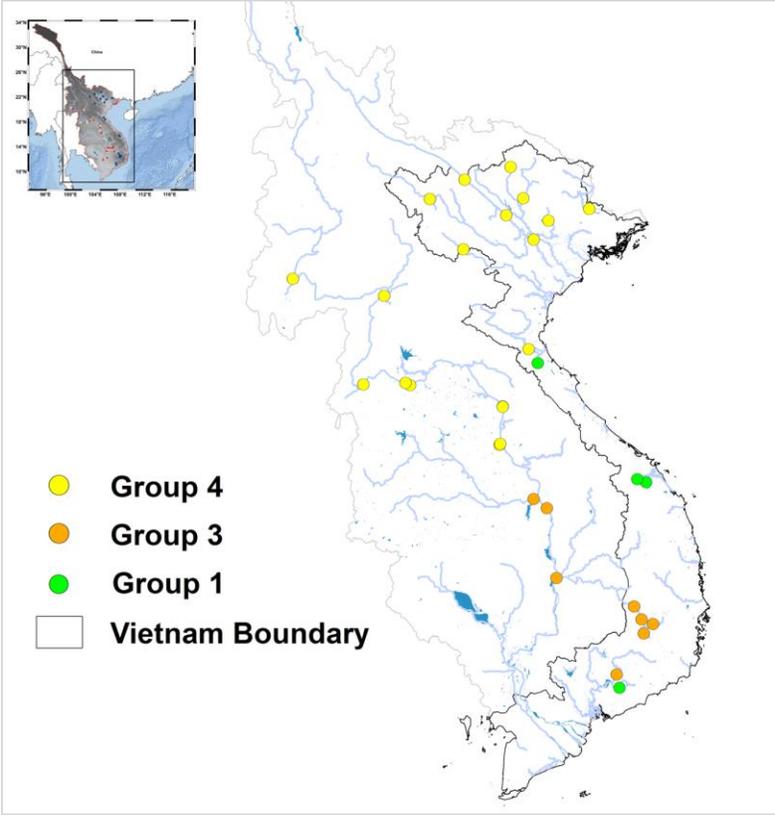
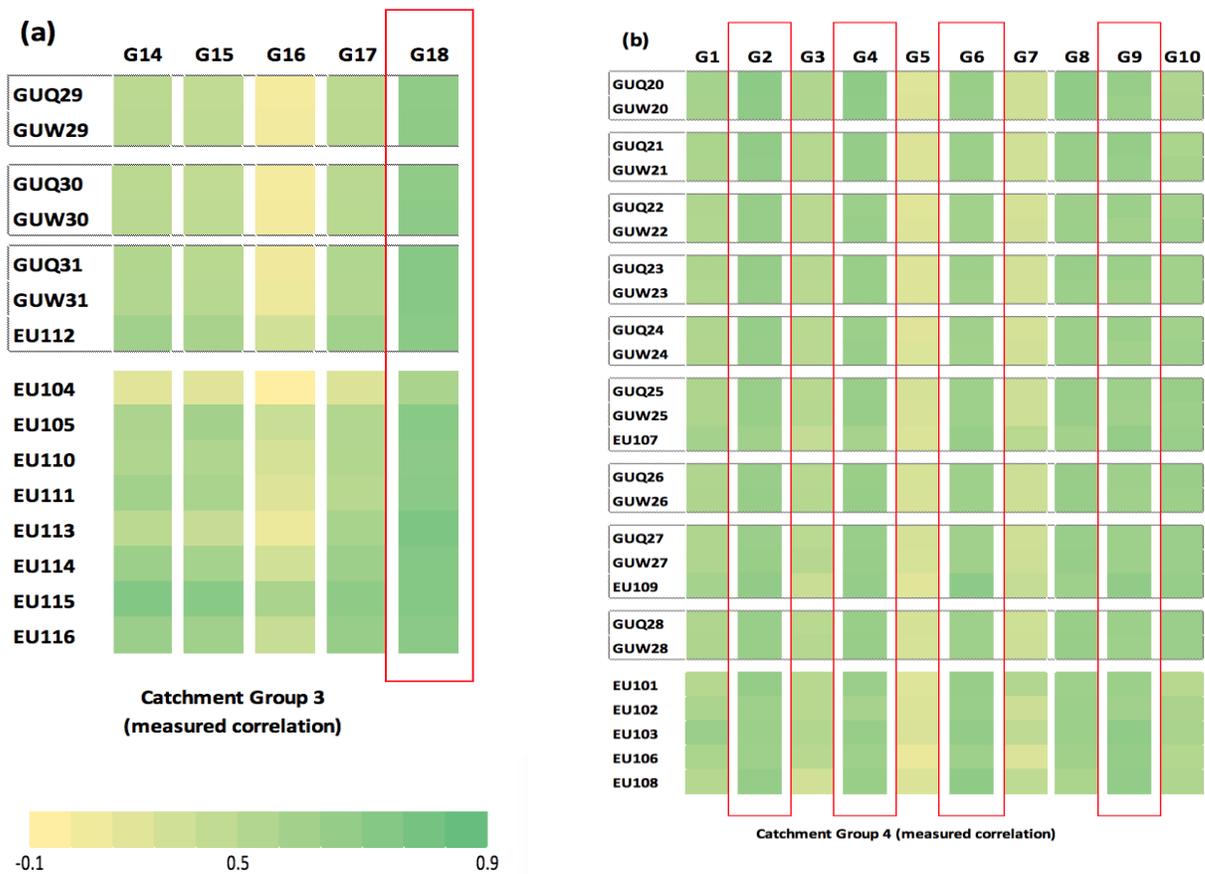


Figure 5. The geographic distribution of all 31 evaluated catchments (gauged catchments inside Vietnam boundary and “geopolitically ungauged” catchments outside of Vietnam) into 4 climatic catchment groups.

551

552 **4.3. Water level based flow correlation between gauged and “ ungauged” catchments**

553 Figure 6 validates the assumption that water level based flow correlation using daily
 554 observed in-situ water levels and streamflow had similar results to the correlation using both daily
 555 observed streamflow observations. In the case of Envisat-derived water levels, because there were
 556 less observations (one daily observation every 35 days), the correlation coefficient became slightly
 557 smaller but the difference was negligible. Accordingly, for the “ geopolitically ungauged” and
 558 Envisat-“ ungauged” catchments in catchment group 3, Talai (G18) was found to be the reference
 559 most highly correlated gauged catchment. For the “ geopolitically ungauged” and Envisat-
 560 “ ungauged” catchments in catchment group 4, LaoCai (G2), LaiChau (G4), YenBai (G6), Xala
 561 (G9) were the reference most highly correlated gauged catchments.



563

564 **Figure 6.** Matrices show Pearson' correlation coefficient between gauged catchments (horizontal positions: G1 → G18
565 using daily streamflow) and “ungauged” catchments (vertical positions, including “geopolitical ungauged”
566 catchments: GUQ20 → GUQ31 using daily streamflow; GUW20 → GUW31 using daily in-situ water levels; and
567 Envisat-“ungauged” catchments EU101 → EU109 using daily Envisat-derived water level). Each dotted box shows
568 the same “ungauged” catchment using different datasets (either streamflow GUQ or in-situ water level GUW or
569 Envisat-derived water level EU) correlated with same gauged catchments. Figure 6a is correlation matrix of
570 “ungauged” catchment group 3 and Figure 6b is correlation matrix of “ungauged” catchment group 4 (Table 6). Red
571 color box highlights the most highly correlated gauged catchments with “ungauged” catchments ($r \geq 0.7$). For details
572 of the location and name of catchments, see Table S1.

573 **4.4. Step-wise physiography and climate-based regionalization at gauged basins**

574 **4.4.1. Baseline model performance (GMHv1.0)**

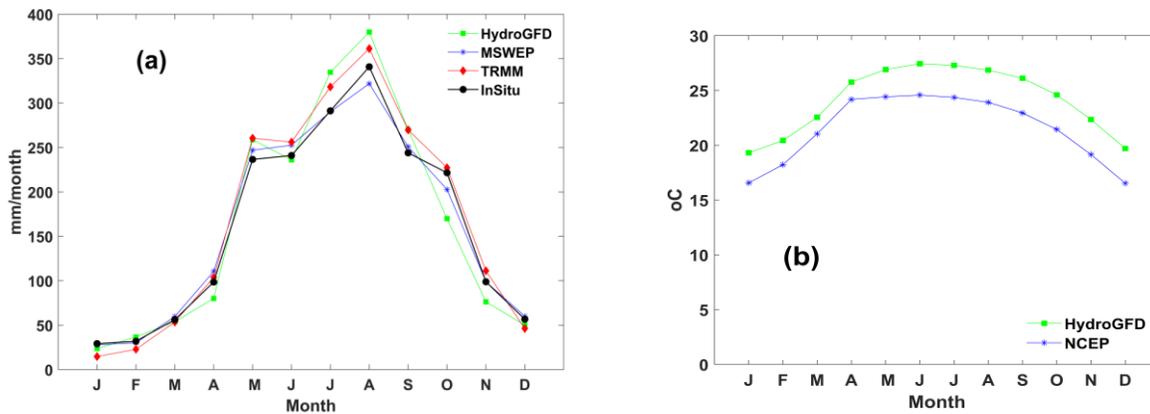
575 Six sets of precipitation and temperature data were used to identify the most appropriate
576 climate inputs for the model. Among them, HydroGFD had the coarsest resolution (0.5° grid)
577 whereas MSWEP, TRMM and NCEP were gridded at had 0.25° resolution. There were 176 in-situ
578 precipitation stations to examine the quality of different climate data inputs of the model. The
579 period 2000-2006 was selected to examine their correlation as it was the period that all datasets
580 were available. In terms of magnitudes, it was found that HydroGFD and TRMM precipitation
581 datasets overestimated during wet months (5% and 7% respectively) and underestimated during
582 dry months (13% and 5% respectively) compared to the in-situ precipitation, resulting in weaker
583 correlation with in-situ precipitation (0.65 and 0.53 respectively) (Table 7). MSWEP had smaller
584 bias (less than 1% for the entire year) and stronger correlation with in-situ precipitation. There
585 was, unfortunately, no in-situ temperature dataset to compare with HydroGFD and NCEP.
586 Monthly average temperature from HydroGFD was larger than monthly average NCEP (Figure 7).

587 Using the initial default parameter set WWHv1.0 with different sets of climate data, no
 588 significant difference in model performance was found between them. Any temperature dataset
 589 combined with the same precipitation dataset resulted in almost similar performance. Among the
 590 precipitation datasets, MSWEP led to the highest model performance, followed by HydroGFD and
 591 TRMM. Since the MSWEP precipitation and NCEP temperature datasets had better resolution
 592 (both at 0.25°), this set of forcing data was selected as the climate input data to be used for the
 593 baseline model (GMHv1.0) (Table 8).

594 **Table 7**
 595 Correlation between different precipitation datasets

Precipitation	HydroGFD	MSWEP	TRMM	In-situ Precipitation
MSWEP	0.78			
TRMM	0.61	0.85		
In-situ Precipitation	0.65	0.75	0.53	

596 Note. Precipitation dataset has the highest correlation with in-situ precipitation are shown in bold font.
 597



599 **Figure 7.** Monthly time series of different climate datasets (Figure 7a: precipitation datasets; Figure 7b: temperature
 600 datasets).

601 **Table 8**
 602 Model performance using different climate datasets

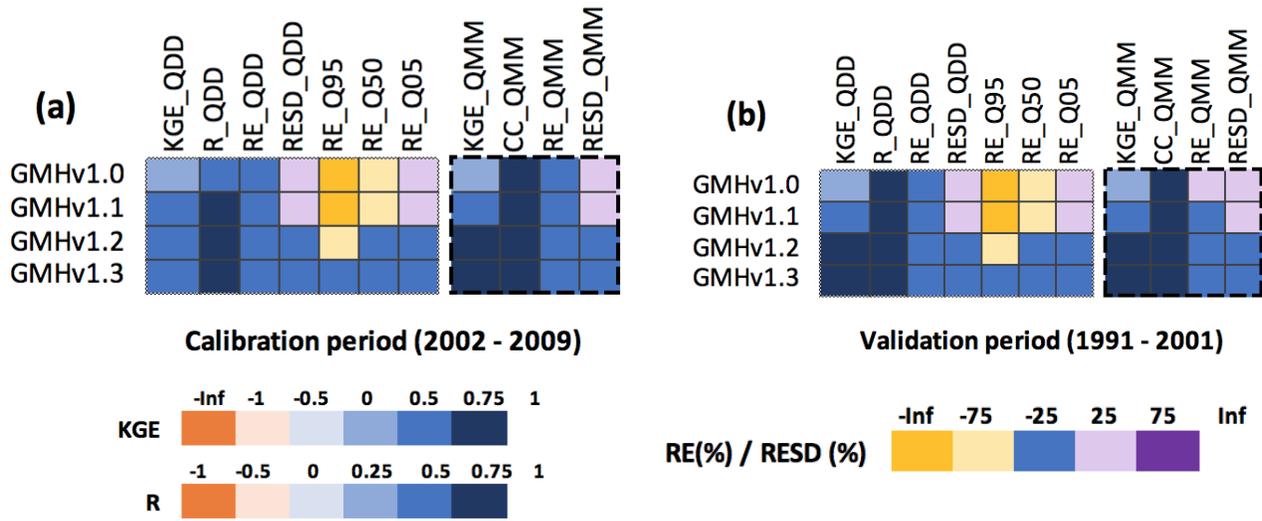
Precipitation	Temperature	KGE	Absolute RE (%)	Absolute RESD (%)	r
HydroGFD	HydroGFD	0.25	16.8	53.55	0.59
HydroGFD	NCEP	0.32	22.61	54.75	0.59
MSWEP	HydroGFD	0.31	21.57	50.16	0.74
MSWEP	NCEP	0.29	19.34	47.59	0.74
TRMM	HydroGFD	0.25	25.27	69.15	0.73
TRMM	NCEP	0.20	25.56	73.64	0.73

603 Note. Table presents median performance metrics for 19 gauged Vietnamese stations. For clarity, in each column, the two best
 604 values are shown in bold font (the highest values (the better) for KGE and r, the smallest values (the better) for absolute RE and
 605 absolute RESD).

606 4.4.2. Refining potential evapotranspiration (GMHv1.1)

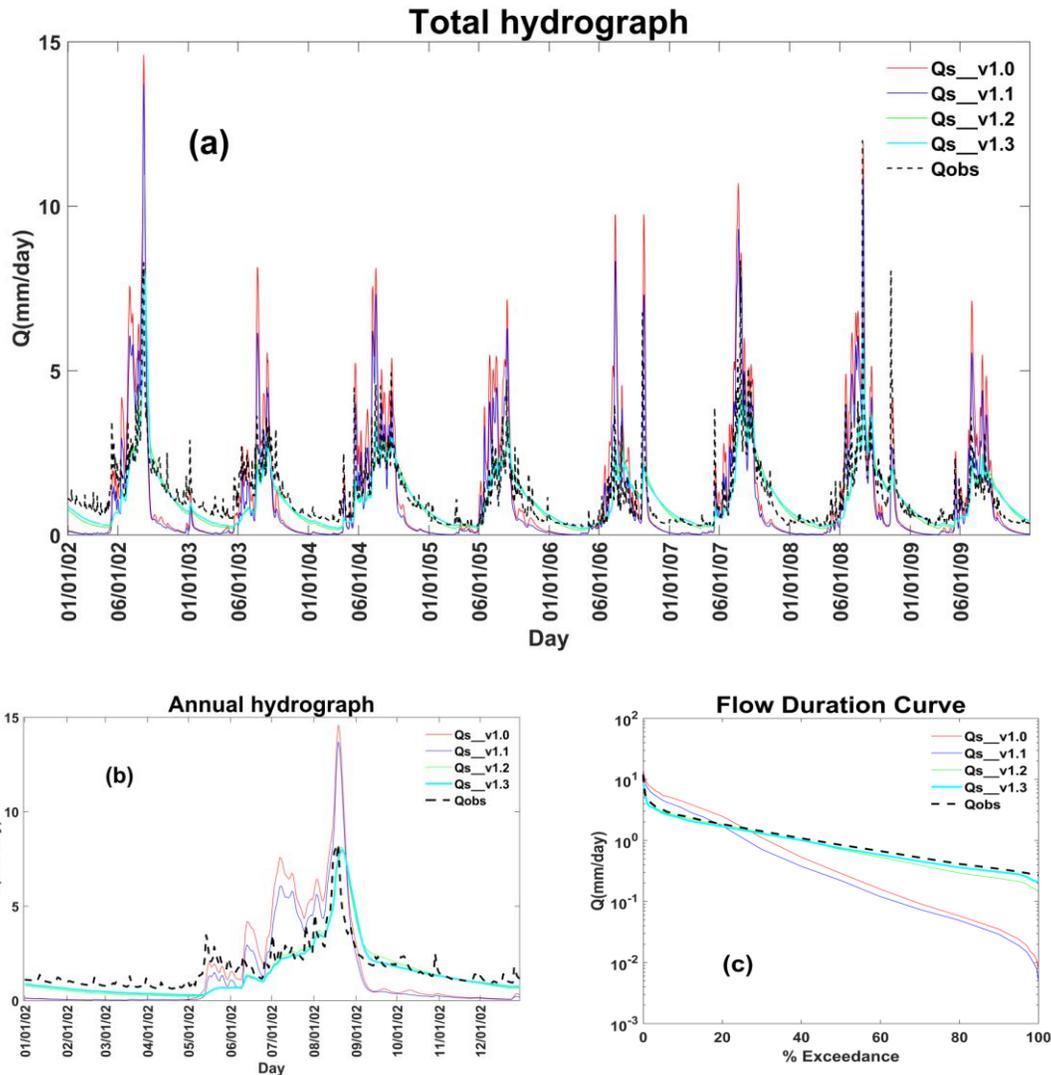
607 Evaporation is a significantly important process in all river basins in Vietnam, accounting
 608 for around 50% of precipitation on average (Nguyen, 2005). Given the importance of evaporation
 609 in the region and large errors of streamflow variability in many locations of the baseline model,
 610 calibration was undertaken to estimate PET – the upper limit of evaporation in the model. Among
 611 three PET related parameters, land use dependent parameter (*kc5*) was found to be sensitive. The
 612 posterior *kc5* was found to reduce relative volumetric errors (RE) between modeled PET and
 613 MODIS-derived PET by 40% compared to initial *kc5* value (Table 5). With this posterior *kc5*
 614 values, model performance for all flow signatures significantly improved over all stations,
 615 particularly for KGE (from 0.3 to 0.47 for daily streamflow). In this model version, nevertheless,
 616 low flows were significantly underestimated while high flow were overestimated, requiring
 617 refinement of the soil storage and flow paths process (Figure 8).

618



620
621
622
623
624
625
626

Figure 8. Model performance (all values are median values) of all model versions at gauged stations in both calibration (Figure 8a) and validation periods (Figure 8b). Dotted box for daily flow signatures and dashed box for monthly flow signatures. Color interpretation of the Figure: blue is good and yellow/red/purple is poor performance.



627

628 **Figure 9.** Total hydrograph of simulated streamflow of all model versions against observed streamflow (Figure 9a),
 629 annual hydrograph (Figure 9b) and flow duration curve of simulated streamflow of all model versions against observed
 630 streamflow (Figure 9c) at one sample gauged location at Lao Cai (located in Red River basin).

631 **4.4.3. Refining soil storage and flow paths (GMHv1.2)**

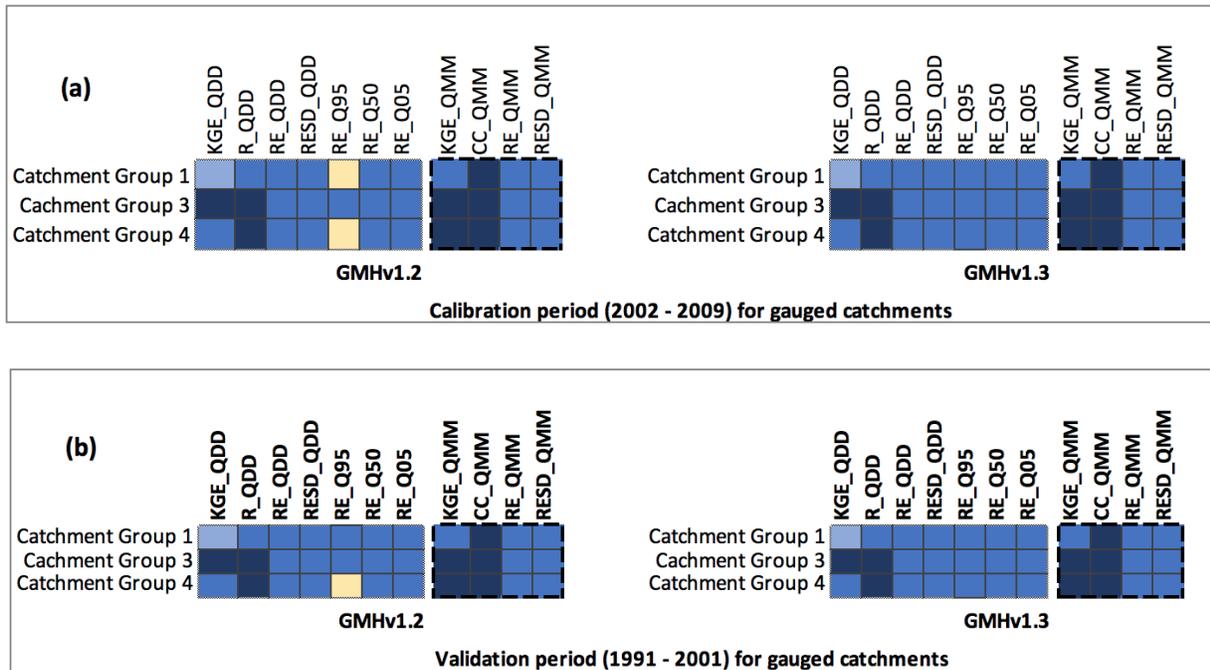
632 The GMHv1.1 model displayed a quick and peaky response of streamflow to rainfall
 633 events, resulting in the underestimation of low flows and the overestimation of high flows (Figure
 634 9). DEMC automation found the sensitive parameters that needed to be calibrated. They were
 635 parameters governing the soil porosities (wcep1, wcep2, wcep3), percolation (mperc2), subsurface

636 runoff and surface runoff (macrate, mactrinf, mactrsm, rrcs2). Parameters of the GMHv1.1 model
637 represented significantly little soil storage, and high recession coefficients. Therefore, soil related
638 parameters were adjusted to increase soil storage capacity, more infiltration and lower recession
639 coefficients for subsurface runoff. In addition, runoff components (Horton overland Flow, Dunne
640 overland flow, subsurface flow) for different soil classes were unreasonable compared to Dunne
641 theory (Dunne, 1978; Li et al., 2014). Accordingly, srrcs (Dunne overland flow related parameter)
642 and srrate (Horton overland flow related parameter) were manually calibrated so that Horton
643 overland flow dominated in urban and bare soil class whereas subsurface runoff and Dunne
644 overland flow dominated in vegetated soil class. Refining these descriptions helped to maintain
645 physical meaning of parameters, whereas significantly improve overall simulated flow signatures
646 for all gauged catchments during both the calibration and validation period. For instance, for the
647 calibration period, compared to the GMHv1.1 model, the KGE for daily streamflow improved
648 from 0.47 to 0.7. On the other hand, volumetric errors of low flows significantly reduced to -27%
649 from -95% and high flow from 35% to 1.7% (Figure 8).

650 **4.4.4. Refining seasonal water balances among catchment groups (GMHv1.3)**

651 The model GMHv1.2 had an overall satisfactory performance for both daily and monthly
652 streamflow time series in both the calibration and validation periods (KGE for daily streamflow
653 was above 0.5). However, the low flow signature (Q95) was underestimated for few stations. In
654 the model GMHv1.2, the global physiography-based parameters, which were based on soil and
655 land cover characteristics of catchments, were used for all catchments. Evaluating the model
656 GMHv1.2 for each catchment group (only 3 groups having gauged stations), the global
657 physiography-based parameters were more suitable for catchment group 3, whereas low flow
658 signatures for both catchment groups 1 and 4 were still underestimated (Figure 10). Catchment

659 group 1 with humidity and mild seasonality has more wet-season dominant storage variation
 660 whereas catchment group 4 with sub-humidity and high seasonality has more dry-season dominant
 661 storage variation (Berghuijs et al., 2014). Accordingly, various correction factors
 662 (evapotranspiration and recession coefficients) were used to simultaneously consider the variety
 663 of climate characteristics between catchments (Hundecha et al., 2016). In a trial and error, the
 664 correction factor for the soil recession coefficient (rrscorr) has resulted in improvement for the
 665 low flow signatures for group 1 and group 4 for both calibration and validation periods whereas
 666 other parameter (cevpcorr) did not result in any improvement.



667



670

671 **Figure 10.** Model performance by different catchment groups for 19 gauged catchments for calibration (Figure 10a)
 672 and validation (Figure 10b). Color interpretation of the Figure: blue is good and yellow/red/purple is poor
 673 performance.

674 **4.5. Performance of regionalized parameters at ungauged basins**

675 Table 9 summarizes the performance in terms of KGE, RE for daily streamflow, RE for
676 Q5, Q95, Q50 obtained in 12 “geopolitically ungauged” evaluation catchments using
677 physiography-based regionalized parameters (par GMHv1.2), physiography and climate based
678 regionalized parameter sets from gauged catchments (par GMHv1.3), global regionalization
679 parameter sets (par WWHv1.3) and locally calibrated parameter sets. Similar to Arheimer et al.
680 (2019), although global regionalization parameters could characterize spatial variability of flow
681 signatures across the globe, they had difficulties in capturing low flows, particularly in tropical
682 catchments. The difference between physiography-based regionalized parameters and
683 physiography and climate-based regionalized parameters was not significant. Nevertheless, the
684 later could significantly reduce volumetric errors of low flow because only one extra parameter
685 was used in the WWHv1.3 compared to previous model version. Compared to locally calibrated
686 parameters, physiography and climate-based regionalized parameters reached nearly 80% in terms
687 of KGE for daily streamflow and was even a slightly better in terms of volumetric errors for low
688 flow, medium and high flow.

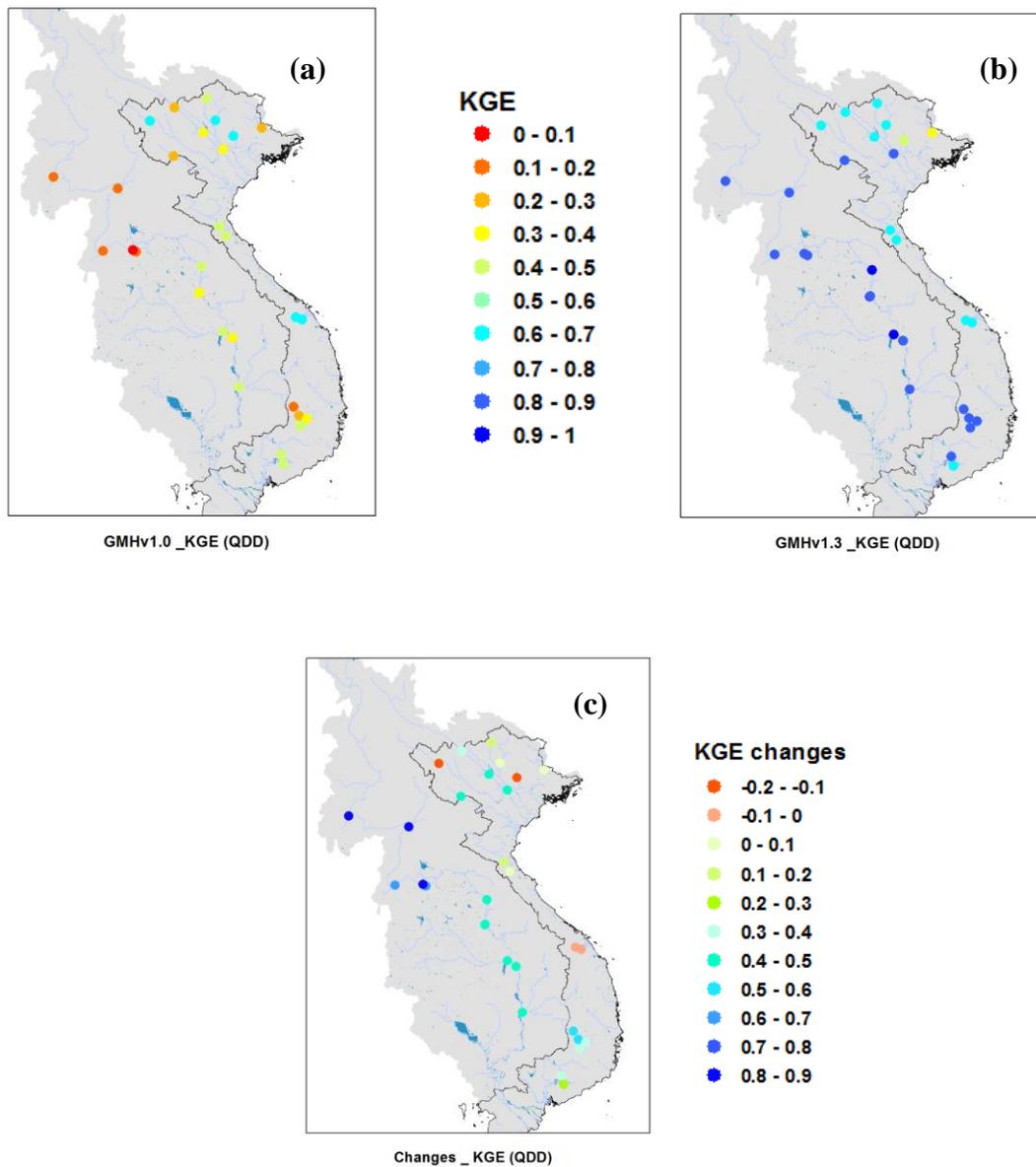
689

690 **Table 9**
 691 Model performance using multiple performance metrics of different flow signatures with various parameter sets for
 692 the 12 “geopolitically ungauged” evaluation catchments for the period 2002 – 2009

Performance metrics	Flow Signatures	Par GMHv1.2 (section 4.4.3)	Par GMHv1.3 (section 4.4.4)	Par WWHv1.3 (global regionalization parameters)	Locally calibrated parameters
KGE	QDD	0.68	0.68	0.32	0.88
RE	QDD	-1.3	-1.7	-61	1
KGE	QMM	0.76	0.76	0.21	0.87
RE	QMM	-1.3	-1.7	-61	0.99
RE	Q95	-32.35	-10.6	-98.17	-14.12
RE	Q5	-4.43	4.06	-48.23	6.45
RE	Q50	1.54	4.51	-80.82	10.61

693 Note. Table presents median performance metrics for 12 “geopolitically ungauged” evaluation catchments. For clarity, in each row,
 694 the two best scores are shown in bold font (the highest values (the better) for KGE, closest values to 0 (the better) for RE).

695 Figure 11 shows performance of daily simulated streamflow of all catchments in terms of
 696 KGE compared to their historical observations of streamflow during validation period (1991 –
 697 2001) for two model versions, including baseline GMHv1.0 and final GMHv1.3. The catchments
 698 presented in Figure 11 include both gauged catchments located inside black Vietnamese boundary
 699 and “geopolitically ungauged” catchments located outside of black Vietnamese boundary. Most of
 700 simulated catchments using the final model version have captured better hydrological processes of
 701 the region, resulting in a substantial improvement (mostly blue dots in Figure 11b). In both model
 702 versions, streamflow in Lang Son (located in Bang Giang Ky Cung basin, Northeast of Vietnam,
 703 the only yellow dot in Figure 11b) was not well simulated. The reason could be the
 704 underrepresented spatial variation of precipitation in the catchment owing to the its small size (the
 705 smallest size 1,500 km² in all evaluated catchments in the study). Future research could be further
 706 improved by using the average of the nearest precipitation grids or higher resolution precipitation
 707 datasets like NASA Global Precipitation Measure integrated multi-satellite retrievals with 0.1°
 708 resolution (GPM IMERG-F-V6) (Le et al., 2020).



709 **Figure 11.** Spatial overview of the model performance for GMHv1.0 (Figure 11a), GMHv1.3 (Figure 11b) and their
 710 changes from GMHv1.0 to GMHv1.3 (Figure 11c) in terms of KGE for daily streamflow time series. Model
 711 performance for both gauged (inside boundary of Vietnam) and “ungauged” catchments for validation periods (1991
 712 – 2001). See Figure Supplementary 3 for model performance of simulated low flows and high flows.
 713

714 **4.6. Model evaluation at ungauged basins using water level based flow correlation**

715 To evaluate the performance of model at ungauged catchments that have observed water
716 levels, evaluation framework using water level based flow correlation (Figure 4) and existing
717 performance metrics of simulated water levels was used. This section applied the framework and
718 performance metrics for both baseline and final model versions to examine if this method and/or
719 performance metrics can work for both scenarios (Figure 12 and Figure 13). Accordingly, firstly,
720 the daily and/or monthly simulated water levels were evaluated against the recorded water levels
721 using the existing performance metrics for simulated water levels, including Pearson's correlation
722 coefficients, NSEanom and NSEW. From Figure 12 and Figure 13, conflicting performance results
723 of simulated water levels compared to observed water levels were found. In any row of both
724 figures, inconsistent colors between r, NSEanom and NSEW for simulated water levels (especially
725 for Envisat-derived water level) were shown. For example, from Figure 12d, at EU_107 station,
726 Pearson's correlation coefficient (dark blue color – good result) showed that simulated water level
727 had good temporal correlation with observed water levels; NSEanom (blue color – acceptable
728 result) informed that they had acceptable magnitude bias; NSEW (orange color – bad result)
729 advised that they had significantly high magnitude bias. Comparing simulated streamflow of this
730 station with observed streamflow (blue color of KGE – acceptable result), the model was found to
731 simulate daily flow at acceptable level but could not capture high flows (light purple color of Q5
732 – overestimated) and low flows (yellow color of Q95 – highly underestimated). This finding
733 confirmed previous studies that model performance using on only water levels could yield
734 inaccurate results in modelling streamflow signatures (Lindstrom, 2016; Jian et al., 2017).
735 Additionally, unlike water levels that have limited performance metrics and derived hydrological
736 signatures, there is a high variety of performance metrics to evaluate various signatures of

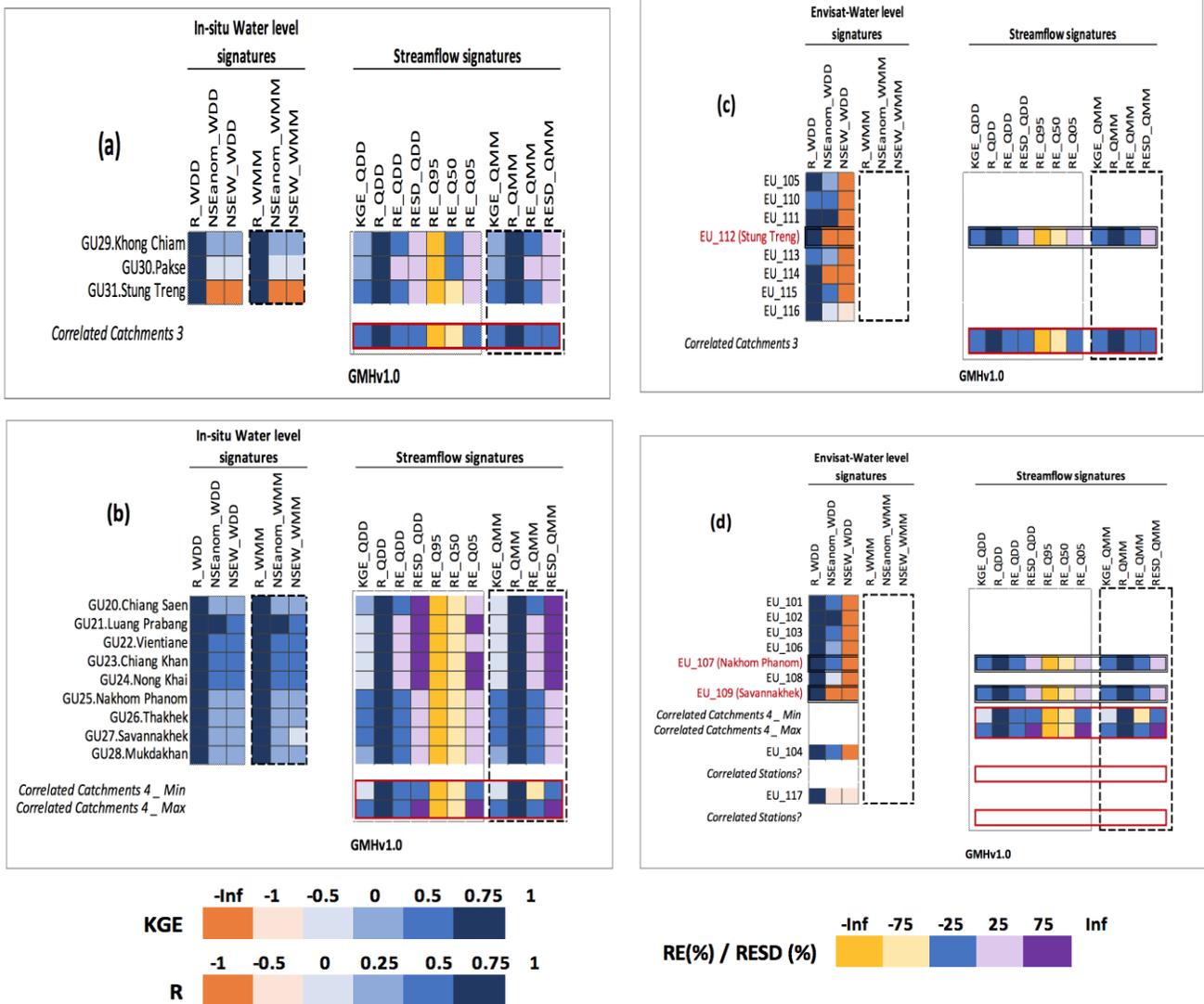
737 streamflow that could help diagnose model problems and inform where to improve. For example,
738 KGE metric can inform whether temporal pattern or variation or magnitude of daily flow is not
739 good (Gupta et al., 2009) whereas relative volumetric errors of flow signatures from flow duration
740 curve (high, low and medium) can inform which part of runoff (surface or subsurface runoff) is
741 not well represented (Yokoo and Sivapalan, 2011). Accordingly, it raised a question how to have
742 extra important model diagnostic information if only observations of water levels are available.

743 Water level based flow correlation was found to possibly address the above question. Using
744 water level based flow correlation evaluation framework, firstly, temporal patterns of simulated
745 water levels were examined against observations using correlation coefficients, which were all
746 above 0.7 for both model versions (Figure 12, Figure 13). Secondly, modeled correlation between
747 simulated water levels of “ungauged” catchments and simulated streamflow of gauged catchments
748 were compared against measured correlation between observed water levels of “ungauged”
749 catchments and observed streamflow of gauged catchments. The difference between them was
750 within +/- 0.2 for both model versions, thus the performance of “ungauged” catchments was
751 similar to the performance of the most highly correlated gauged catchments (see Figure
752 Supplementary 4). Accordingly, in both baseline and final model versions, performance of
753 “ungauged” catchments were similar to performance of the most highly correlated gauged
754 catchments for all flow signatures. It was then validated with any “ungauged” catchments that have
755 historical observed streamflow to cross-validate the hypothesis. Consistent results were found for
756 all flow signatures between the reference most highly correlated gauged catchments and
757 “ungauged” catchments having observed streamflow (where available for cross-validation) to
758 accept the hypothesis. For catchments having only Envisat-derived water level without observed
759 streamflow, its performance cannot be validated. Nevertheless, since this method worked for both

760 in-situ water levels and 3 catchments having both Envisat-derived water level and observed
761 streamflow, performance of the remaining 14 catchments with Envisat-derived water level could
762 be evaluated using the reference most highly correlated gauged catchments.

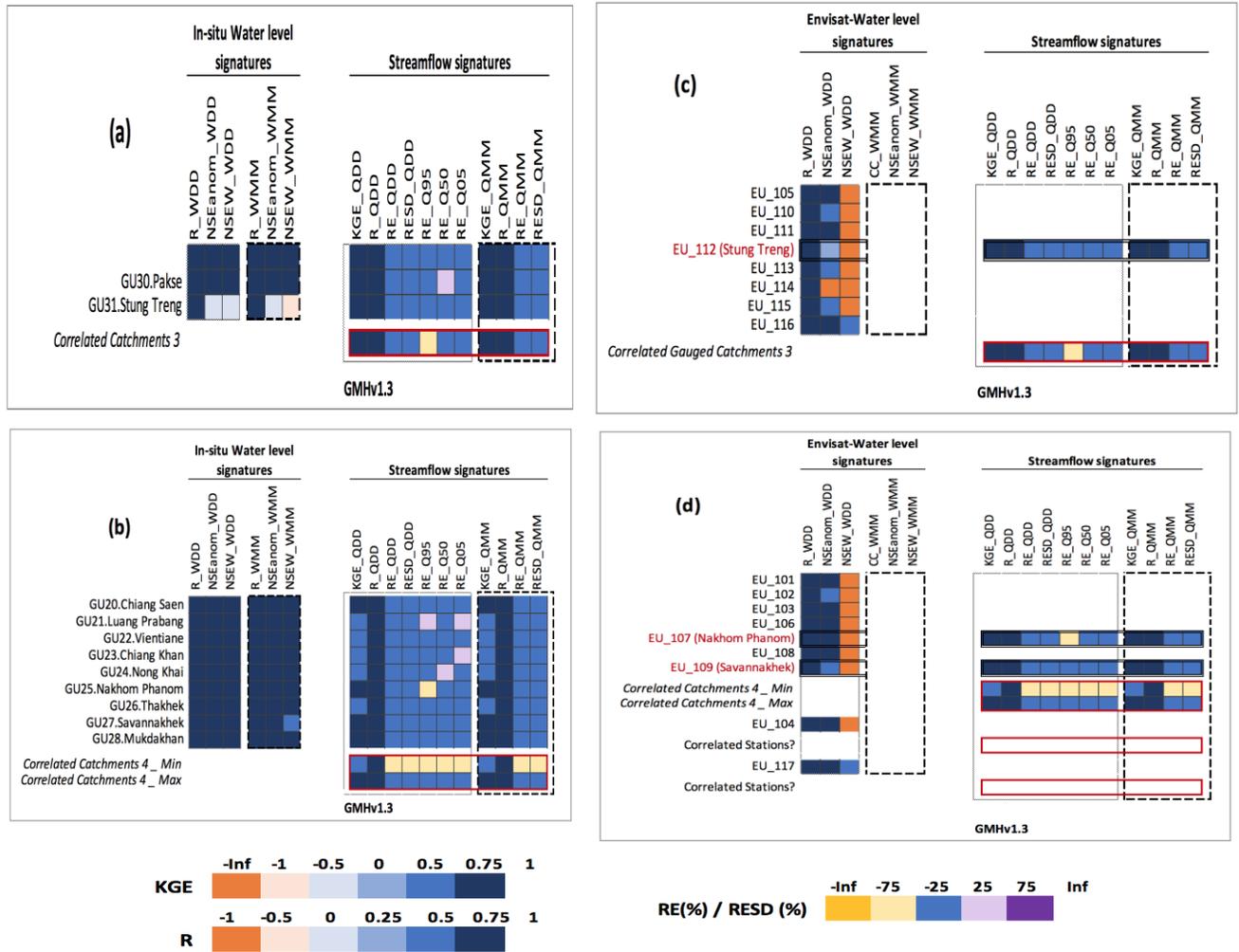
763 Accordingly, it showed that water level based flow correlation method could be used to
764 evaluate the model performance at ungauged catchments having only observations of water levels.
765 Furthermore, compared to previous studies that used water levels to evaluate model performance,
766 this approach can not only overcome numerical problems of existing performance metrics for
767 water levels but also provide important model diagnostic information on how to improve model
768 performance without streamflow observations.

769



770

771 **Figure 12.** Evaluation of the baseline model GMHv1.0 for “geopolitically ungauged” catchment group 3 (Figure 12a),
 772 “geopolitically ungauged” catchment group 4 (Figure 12b), Envisat-“ungauged” catchment group 3 (Figure 12c) and
 773 Envisat-“ungauged” catchment group 4 (Figure 12d). Simulated water levels were evaluated against in-situ water level
 774 (left images) and Envisat-derived water level (right images). Since modeled correlation was similar to measured
 775 correlation, simulated streamflow of “geopolitically ungauged” or Envisat-“ungauged” catchments were similar to
 776 that of the reference most highly correlated gauged catchments (red highlight box). This simulation was then validated
 777 against observed streamflow for any “ungauged” catchments that have historical observations (observed streamflow
 778 of “ungauged” catchments were only used for cross-validation, not used in calibration). For the reference most highly
 779 correlated gauged catchment 4, there were 4 catchments, thus both minimum and maximum performance metrics were
 780 presented



781
 782 **Figure 13.** Evaluation of the final model GMHv1.3 for “geopolitically ungauged” catchment group 3 (Figure 13a),
 783 “geopolitically ungauged” catchment group 4 (Figure 13b), Envisat-“ungauged” catchment group 3 (Figure 13c) and
 784 Envisat-“ungauged” catchment group 4 (Figure 13d). Simulated water levels were evaluated against in-situ water level
 785 (left images) and Envisat-derived water level (right images). Since modeled correlation was similar to measured
 786 correlation, simulated streamflow of “geopolitically ungauged” or Envisat-“ungauged” catchments were similar to
 787 that of the reference most highly correlated gauged catchments (red highlight box). This simulation was then validated
 788 against observed streamflow for any “ungauged” catchments that have historical observations (observed streamflow
 789 of “ungauged” catchments were only used for cross-validation, not used in calibration). For the reference most highly
 790 correlated gauged catchment 4, there were 4 catchments, thus both minimum and maximum performance metrics were
 791 presented.

792 **5. Discussion**

793 **5.1. Step-wise physiography and climate-based regionalization at gauged basins**

794 Catchment models are important tools to support decision makers in sustainable planning
795 of water resources. The Greater Mekong region is the top global biodiversity hotspot but
796 increasingly facing urgent socio-economic development and climate change impacts. Accordingly,
797 it is imperative to have a multi-national and multi-catchment model to support river basin
798 authorities. It could thus help predict river flows across administrative borders and allocate water
799 resources among water users in a harmonized manner. For the first time, a multi-national and
800 multi-catchment Greater Mekong HYPE was set up in this important region. The analysis of the
801 final model GMHv1.3 version (KGE of daily and monthly streamflow is 0.7 and 0.8 respectively)
802 shows that the model is useful for water authorities in managing water related issues. The model
803 has been setup on the foundation of the Worldwide HYPE model and successfully refined to
804 capture the hydrological processes for the region. It shows that global hydrological model, in this
805 case the worldwide HYPE model, could be a useful starting point as a time-saving alternative for
806 other regions to further refine it with local expert knowledge, so that it could be useful in
807 supporting decision makers for water management. Additionally, further refining an existing
808 model would allow critical knowledge and experiences shared between research groups and
809 practitioners, thus increasing full transparency in the research process, further understanding of
810 general hydrological patterns, process and functions between catchments. It can thus ultimately
811 advance hydrological sciences toward a unified theory of hydrology at catchment scale (Sivapalan,
812 2005) and better predict flow signatures at ungauged basins (Bloschl et al., 2013).

813 The approach of sequentially and iteratively (both automatically and manually) refining
814 inadequately described hydrological processes, together with local knowledge can substantially

815 improve the appropriateness of model application in a new region. Calibration is inevitable in
816 physically distributed model because of impossibility to measure all required model parameters at
817 the model simulation scale (Beven, 1989; Bloschl and Sivapalan, 1995). This study combined both
818 automatic and manual calibration to combine the strengths of both methods to achieve more
819 physically acceptable parameters at a timely efficient manner at each step in hydrological
820 processes. The study area has various seasonal water variability due to its substantial precipitation
821 and evaporation variability from tropical monsoon effect. Therefore, using climatic indexes
822 (aridity index and seasonality index) is a useful approach to group catchments so that all
823 catchments can be simulated in the same modeling domain. Adding one simple step (step 4 in the
824 step-wise calibration approach in section 3.4) into the common step-wise physiography-based
825 parameters helped reduce the underestimation of low flow of two catchment groups. In this study,
826 simple regionalized parameter approach (correction parameters) were used. More substantial
827 model improvement could be made if other regionalization approaches could be employed, such
828 as linear parameter estimation based on catchment descriptors (Hundecha et al., 2016). Future
829 studies could examine this hypothesis.

830 **5.2. Performance of regionalized parameters at ungauged basins**

831 HYPE with physiography and climate based regionalized parameters appears to perform
832 as good as locally calibrated parameters and outperform global regionalization parameters in all
833 flow signatures. This result confirms findings of the previous studies that similarity in catchment
834 characteristics and climate characteristics can lead to similarity in rainfall-runoff responses (Beck
835 et al., 2016; Berghuijs et al., 2014). Climatic indexes based on observations of precipitation and
836 temperature during the same period with observations of streamflow could provide more
837 dynamically agreeing characteristics of each catchment rather than using Koppen climate

838 classification, which has different timeline with streamflow observation (Kottek et al., 2006). The
839 evaluated catchments are all vegetated (either forest or agricultural lands) catchments so the
840 difference in physiography is not significant. Nevertheless, physiography-based regionalized
841 parameters are demonstrated to predict well flow signatures in ungauged basins across Europe
842 (Donnelly et al., 2016). This approach could be helpful for existing model using physiography-
843 based regionalized parameters to be further improved without altering the current parameter sets.
844 Since this is a poorly-gauged region, obtaining more streamflow observations would be
845 challenging. Therefore, it is important to develop more approaches to validate the simulated
846 streamflow from model for ungauged catchments. The next section is one of those attempts.
847 Another approach to cross-validate simulated streamflow for ungauged catchments could be using
848 ensemble learning regression combining satellite altimetry data and a hydrologic model, which
849 could be HYPE model in this case (Kim et al., 2019c).

850 **5.3. Model evaluation at ungauged basins using water level based flow correlation**

851 To evaluate model performance at ungauged basins, both existing performance metrics of
852 water levels and proposed water level based flow correlation were adopted. Inconsistent and even
853 conflicting performance results using different performance metrics happened for both baseline
854 and final models, which make diagnosing and evaluating the model at ungauged basins difficult
855 (Figure 12 and 13). Meanwhile, using water level based flow correlation method (both in-situ and
856 Envisat derived water levels) can provide more details regarding model diagnostics of which
857 signature needs to be further refined. For instance, in the baseline model (Figure 12), using the
858 performance of the reference most highly correlated gauged catchments, it informed that model
859 could not capture low flow for both catchment group 3 and catchment group 4. For this study, the
860 threshold for identifying the most correlation catchments were only 0.7 because of limited ground

861 observations. Data access in this region is particularly arduous, therefore correlation threshold was
862 lower than other studies (Archfield and Vogel, 2010; Betterle et al., 2017; Betterle et al., 2019).
863 Lower correlation threshold could have reduced the matching performance between the reference
864 most highly correlated gauged catchments and “ungauged” catchments although the difference is
865 not significant. Future researches could further examine this hypothesis. Findings show that flow
866 correlation method with in-situ water level can be used to evaluate the performance of ungauged
867 catchments through the most highly correlated gauged catchments. For Envisat-derived water
868 level, since there are only 3 virtual stations located in catchments having ground observations, 3
869 out of 17 Envisat-“ungauged” catchments have been validated. Nevertheless, since this method
870 was found to work with both in-situ water level and 3 Envisat virtual stations, it is assumed that
871 the remaining 14 Envisat-“ungauged” catchments could have similar satisfactory simulation to the
872 reference gauged catchments.

873 It is expected that not only sub-continental multi-catchment hydrological models but also
874 multi-continental multi-catchment hydrological models would be benefited from this approach if
875 water level-based flow correlation was found between altimetry-derived water level in ungauged
876 catchments of a poorly gauged continent and streamflow in gauged catchments of another
877 excessively gauged continent. In this study, expanding possibility of study area of GMv1.3, along
878 with water level based flow correlation, could further validate the performance of the current non-
879 validated catchment group 2 of model. Meanwhile, global-scale model could more satisfactorily
880 capture the full range of variability of hydrological regimes that actually exist within their large
881 domains. Thus, it can further increase the ability of hydrological models to be employed routinely
882 and with confidence to ungauged basins. More altimetry satellite missions with denser coverage

883 in the future could further advance this approach to improve predictions of flow regimes in
884 ungauged basins.

885 **6. Conclusion**

886 The study uses a novel approach to combine regionalization and satellite observations of
887 various hydrological variable to improve prediction of streamflow signatures at “geopolitically
888 ungauged” basins. Using the proposed step-wise physiography and climate-based regionalization
889 approach, the model performance at ungauged basins reached 80% of performance of the ideal
890 situation, where observed streamflow data were available for calibration, and significantly
891 outperformed the global regionalization parameters using the Koppen climate classification. This
892 approach would be helpful for both new model setup and existing physically distributed models
893 because it is flexible and does not change the current parameter values of existing models.
894 Additionally, the proposed water level based flow correlation was found to help diagnose models
895 and outperform the existing performance metrics of simulated water levels at ungauged basins. It
896 is expected that more satellite altimetry missions with a denser coverage in the future, together
897 with macroscale hydrological model, either at continental scale or global scale with a wide variety
898 of observed streamflow patterns (Alemaw and Chaoka, 2003; Arheimer et al., 2019; Doll et al.,
899 2013; Beck et al., 2016) could be benefited from this approach to further evaluate model
900 performance in ungauged basins.

901 The study also helps to setup the first multi-national, multi-catchment hydrological model
902 in the Greater Mekong region, the top global biodiversity and major disaster risk hotspot in the
903 world. This model version would be useful for water authorities to monitor and plan sustainable
904 use of water resources across administrative boundaries under rapid changing development

905 activities and climate impacts. Using a common hydrological model concept and setup approach
906 compared to the global hydrological model would allow critical sharing of knowledge and
907 experiences to advance toward a unified theory of hydrology at catchment scale and better predict
908 flow signatures at ungauged basins. Nevertheless, knowledge gaps in aquifers, floodplain effect,
909 and water extraction by human have not been addressed. Future model version could be further
910 improved, such as using average of the nearest precipitation grids (for better reproducing regimes
911 in small catchments), incorporating other hydrological data (e.g. groundwater level, total terrestrial
912 storage change, soil moisture), and adding water management modules (e.g. regulated reservoirs,
913 irrigation, water quality) to explore impacts of various changing scenarios from climate and human
914 activities on the vital water, food and energy security in the region. Web data portal could be
915 developed to allow more data accesses and knowledge sharing of water status in this important
916 region (McDonald et al., 2019; Biswas and Faisal, 2018).

917

918 **Acknowledgement**

919

920 This study is supported by NASA's Applied Sciences Program for SWOT Science Team
921 (NNX16AQ33G); GEO Program (80NSSC18K0423); Vingroup Innovation Foundation
922 (VINIF.2019.DA17); the Greater Mekong HYPE Study Subcontract (No. SC1) of the ECMWF
923 COPERNICUS project (C3S_422_Lot1_SMHI); SERVIR Program (80NSSC20K0152); Vietnam
924 National Foundation for Science and Technology Development (NAFOSTED) and the United
925 Kingdom's Natural Environment Research Council (NERC) (NE/S002847/1); USAID's PEER
926 Project (AID-OAA-A-11-00012); and South Korea Ministry of Environment's Demand
927 Responsive Water Supply Service Program (2019002650004). Especially, the authors would like
928 to express our gratitude to SMHI hydrological research unit and water monitoring and forecasting
929 team at NAWAPI, where various prior common works over multiple years on the regional
930 modelling platform were done to make this study possible. We would like to thank Dr. Phil Graham
931 for supporting with HYPE model setup, Kel Markert for helping with Google Earth Engine coding,
932 and data providers listed in Table 1 for providing us important resources and data to undertake this
933 work.

934 **Credit Author Statement**

935

936 Tien L.T. Du: Conceptualization, Methodology, Investigation, Writing – Original Draft, Writing
937 – Review & Editing; Hyongki Lee: Supervision & Resources; Duong D. Bui: Supervision &
938 Resources; Berit Arheimer: Resources, Writing – Review & Editing; Hong-Yi Li: Writing –
939 Review & Editing; Jonas Olsson: Writing – Review & Editing; Stephen E. Darby: Writing –
940 Review & Editing; Justin Sheffield: Writing – Review & Editing; Donghwan Kim: Review &
941 Editing; Euiho Hwang: Writing – Review & Editing

942

943

944 **References**

- 945 Abbaspour, K.C., Rouholahnejad, E., Vaghefi, Srinivasan, R., Yang, H. and Kløve, B., 2015. A continental-scale
 946 hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale
 947 SWAT model. *Journal of Hydrology*, 524, pp.733-752.
- 948 Ad Hoc Group, Vörösmarty, C., Askew, A., Grabs, W., Barry, R.G., Birkett, C., Döll, P., Goodison, B., Hall, A.,
 949 Jenne, R. and Kitaev, L., 2001. Global water data: A newly endangered species. *Eos, Transactions American
 950 Geophysical Union*, 82(5), pp.54-58.
- 951 Alemaw, B.F. and Chaoka, T.R., 2003. A continental scale water balance model: a GIS-approach for Southern
 952 Africa. *Physics and Chemistry of the Earth, Parts A/B/C*, 28(20-27), pp.957-966.
- 953 Allen, R.G., Pereira, L.S., Raes, D. and Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop
 954 water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome*, 300(9), p.D05109.
- 955 Andersson, J.C., Arheimer, B., Traoré, F., Gustafsson, D. and Ali, A., 2017. Process refinements improve a
 956 hydrological model concept applied to the Niger River basin. *Hydrological processes*, 31(25), pp.4540-4554.
- 957 Archfield, S.A. and Vogel, R.M., 2010. Map correlation method: Selection of a reference streamgage to estimate daily
 958 streamflow at ungauged catchments. *Water Resources Research*, 46(10).
- 959 Arheimer, B. and Lindström, L., 2013. Implementing the EU Water Framework Directive in Sweden. Chapter 11.20
 960 in: Bloeschl, G., Sivapalan, M., Wagener, T., Viglione, A. and Savenije, H. (Eds). *Runoff Predictions in
 961 Ungauged Basins – Synthesis across Processes, Places and Scales*. Cambridge University Press, Cambridge,
 962 UK. (p. 465) pp. 353-359.
- 963 Arheimer, B., Pimentel, R., Isberg, K., Crochemore, L., Andersson, J. C. M., Hasan, A., and Pineda, L., 2019 Global
 964 catchment modelling using World-Wide HYPE (WWH), open data and stepwise parameter estimation,
 965 *Hydrol. Earth Syst. Sci. Discuss.*, <https://doi.org/10.5194/hess-2019-111>, in review.
- 966 Beck, H.E., van Dijk, A.I., De Roo, A., Miralles, D.G., McVicar, T.R., Schellekens, J. and Bruijnzeel, L.A., 2016.
 967 Global-scale regionalization of hydrologic model parameters. *Water Resources Research*, 52(5), pp.3599-
 968 3622.
- 969 Beck, H.E., Van Dijk, A.I., Levizzani, V., Schellekens, J., Gonzalez Miralles, D., Martens, B. and De Roo, A., 2017.
 970 MSWEP: 3-hourly 0.25 global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis
 971 data. *Hydrology and Earth System Sciences*, 21(1), pp.589-615.
- 972 Berg, P., Donnelly, C. and Gustafsson, D., 2018. Near-real-time adjusted reanalysis forcing data for
 973 hydrology. *Hydrology and Earth System Sciences*, 22(2), pp.989-1000.
- 974 Berghuijs, W.R., Sivapalan, M., Woods, R.A. and Savenije, H.H., 2014. Patterns of similarity of seasonal water
 975 balances: A window into streamflow variability over a range of time scales. *Water Resources
 976 Research*, 50(7), pp.5638-5661.
- 977 Bergström, S., 1991. Principles and confidence in hydrological modelling. *Hydrology Research*, 22(2), pp.123-136.
- 978 Betterle, A., Radny, D., Schirmer, M. and Botter, G., 2017. What do they have in common? Drivers of streamflow
 979 spatial correlation and prediction of flow regimes in ungauged locations. *Water Resources Research*, 53(12),
 980 pp.10354-10373.
- 981 Betterle, A., Schirmer, M. and Botter, G., 2019. Flow dynamics at the continental scale: Streamflow correlation and
 982 hydrological similarity. *Hydrological processes*, 33(4), pp.627-646.
- 983 Beven, K.J., 2011. *Rainfall-runoff modelling: the primer*. John Wiley & Sons.
- 984 Beven, K., 1989. Changing ideas in hydrology—the case of physically-based models. *Journal of hydrology*, 105(1-
 985 2), pp.157-172.
- 986 Biswas, N.K. and Hossain, F., 2018. A scalable open-source web-analytic framework to improve satellite-based
 987 operational water management in developing countries. *Journal of Hydroinformatics*, 20(1), pp.49-68.
- 988 Blöschl, G., Sivapalan, M., Savenije, H., Wagener, T. and Viglione, A. eds., 2013. *Runoff prediction in ungauged
 989 basins: synthesis across processes, places and scales*. Cambridge University Press.
- 990 Blöschl, G. and Sivapalan, M., 1995. Scale issues in hydrological modelling: a review. *Hydrological processes*, 9(3-
 991 4), pp.251-290.
- 992 Boyle, D.P., Gupta, H.V. and Sorooshian, S., 2000. Toward improved calibration of hydrologic models: Combining
 993 the strengths of manual and automatic methods. *Water Resources Research*, 36(12), pp.3663-3674.
- 994 Budyko, M.I., 1974. *Climate and life*. New York: Academic press.

995 Bui, D.D., Kawamura, A., Tong, T.N., Amaguchi, H., Nakagawa, N. and Iseri, Y., 2011. Identification of aquifer
996 system in the whole Red River Delta, Vietnam. *Geosciences Journal*, 15(3), p.323.

997 Chang, C.H., Lee, H., Hossain, F., Basnayake, S., Jayasinghe, S., Chishtie, F., Saah, D., Yu, H., Sothea, K. and Du
998 Bui, D., 2019. A model-aided satellite-altimetry-based flood forecasting system for the Mekong
999 River. *Environmental modelling & software*, 112, pp.112-127.

1000 CTOH_ENVISAT_2014_01. Doi:10.6096

1001 Dilley, M., Chen, R.S., Deichmann, U., Lerner-Lam, A.L. and Arnold, M., 2005. *Natural disaster hotspots: a global
1002 risk analysis*. The World Bank.

1003 Döll, P., Kaspar, F. and Lehner, B., 2003. A global hydrological model for deriving water availability indicators:
1004 model tuning and validation. *Journal of Hydrology*, 270(1-2), pp.105-134.

1005 Donnelly, C., Andersson, J.C. and Arheimer, B., 2016. Using flow signatures and catchment similarities to evaluate
1006 the E-HYPE multi-basin model across Europe. *Hydrological Sciences Journal*, 61(2), pp.255-273.

1007 Du, T.L.T., Bui, D.D., Quach, X.T. and Robins, L., 2016. Water governance in a changing era: Perspectives on
1008 Vietnam. *Water governance dynamics in the Mekong region*, pp.241-248.

1009 Du, T., Bui, D., Nguyen, M. and Lee, H., 2018. Satellite-Based, Multi-Indices for Evaluation of Agricultural Droughts
1010 in a Highly Dynamic Tropical Catchment, Central Vietnam. *Water*, 10(5), p.659.

1011 Dunne, T., 1978. Field studies of hillslope flow processes. *Hillslope hydrology*, pp.227-293.

1012 Gao, H., Hrachowitz, M., Schymanski, S.J., Fenicia, F., Sriwongsitanon, N. and Savenije, H.H.G., 2014. Climate
1013 controls how ecosystems size the root zone storage capacity at catchment scale. *Geophysical Research
1014 Letters*, 41(22), pp.7916-7923.

1015 Fischhendler, I., 2008. Ambiguity in transboundary environmental dispute resolution: The Israeli—Jordanian water
1016 agreement. *Journal of Peace Research*, 45(1), pp.91-109.

1017 Garambois, P.A., Roux, H., Larnier, K., Labat, D. and Dartus, D., 2015. Parameter regionalization for a process-
1018 oriented distributed model dedicated to flash floods. *Journal of Hydrology*, 525, pp.383-399.

1019 Gerlak, A.K., Lautze, J. and Giordano, M., 2011. Water resources data and information exchange in transboundary
1020 water treaties. *International Environmental Agreements: Politics, Law and Economics*, 11(2), pp.179-199.

1021 Getirana, A.C., 2010. Integrating spatial altimetry data into the automatic calibration of hydrological models. *Journal
1022 of Hydrology*, 387(3-4), pp.244-255.

1023 Gupta, H.V., Kling, H., Yilmaz, K.K. and Martinez, G.F., 2009. Decomposition of the mean squared error and NSE
1024 performance criteria: Implications for improving hydrological modelling. *Journal of hydrology*, 377(1-2),
1025 pp.80-91.

1026 He, Y., Bárdossy, A. and Zehe, E., 2011. A review of regionalisation for continuous streamflow simulation. *Hydrology
1027 and Earth System Sciences*, 15(11), pp.3539-3553.

1028 Hossain, F., Sikder, S., Biswas, N., Bonnema, M., Lee, H., Luong, N.D., Hiep, N.H., Du Duong, B. and Long, D.,
1029 2017. Predicting water availability of the regulated Mekong river basin using satellite observations and a
1030 physical model. *Asian Journal of Water, Environment and Pollution*, 14(3), pp.39-48.

1031 Hrachowitz, M., Savenije, H.H.G., Blöschl, G., McDonnell, J.J., Sivapalan, M., Pomeroy, J.W., Arheimer, B., Blume,
1032 T., Clark, M.P., Ehret, U. and Fenicia, F., 2013. A decade of Predictions in Ungauged Basins (PUB)—a
1033 review. *Hydrological sciences journal*, 58(6), pp.1198-1255.

1034 Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Stocker, E. and Wolff, D.B., 2006. The
1035 TRMM Multi-satellite Precipitation Analysis (TMPA): quasi-global precipitation estimates at fine scales.

1036 Hundecha, Y., Arheimer, B., Donnelly, C. and Pechlivanidis, I., 2016. A regional parameter estimation scheme for a
1037 pan-European multi-basin model. *Journal of Hydrology: Regional Studies*, 6, pp.90-111.

1038 Jian, J., Ryu, D., Costelloe, J.F. and Su, C.H., 2017. Towards hydrological model calibration using river level
1039 measurements. *Journal of Hydrology: Regional Studies*, 10, pp.95-109.

1040 Kibler, K.M., Biswas, R.K. and Juarez Lucas, A.M., 2014. Hydrologic data as a human right? Equitable access to
1041 information as a resource for disaster risk reduction in transboundary river basins. *Water Policy*, 16(S2),
1042 pp.36-58.

1043 Kim, D., Yu, H., Lee, H., Beighley, E., Durand, M., Alsdorf, D.E. and Hwang, E., 2019a. Ensemble learning regression
1044 for estimating river discharges using satellite altimetry data: Central Congo River as a Test-bed. *Remote
1045 sensing of environment*, 221, pp.741-755.

1046 Kim, D., Lee, H., Chang, C.H., Bui, D.D., Jayasinghe, S., Basnayake, S., Chishtie, F. and Hwang, E., 2019b. Daily
1047 River Discharge Estimation Using Multi-Mission Radar Altimetry Data and Ensemble Learning Regression
1048 in the Lower Mekong River Basin. *Remote Sensing*, 11(22), p.2684.

- 1049 Kim, D., Lee, H., Beighley, E. and Tshimanga, R.M., 2019c. Estimating discharges for poorly gauged river basin
1050 using ensemble learning regression with satellite altimetry data and a hydrologic model. *Advances in Space*
1051 *Research*.
- 1052 Kim, U. and Kaluarachchi, J.J., 2008. Application of parameter estimation and regionalization methodologies to
1053 ungauged basins of the Upper Blue Nile River Basin, Ethiopia. *Journal of Hydrology*, 362(1-2), pp.39-56.
- 1054 Kottek, M., Grieser, J., Beck, C., Rudolf, B. and Rubel, F., 2006. World map of the Köppen-Geiger climate
1055 classification updated. *Meteorologische Zeitschrift*, 15(3), pp.259-263.
- 1056 Le, M.H., Lakshmi, V., Bolten, J. and Du Bui, D., 2020. Adequacy of Satellite-derived Precipitation Estimate for
1057 Hydrological modeling in Vietnam Basins. *Journal of Hydrology*, p.124820.
- 1058 Lee, H., Shum, C.K., Yi, Y., Ibaraki, M., Kim, J.W., Braun, A., Kuo, C.Y. and Lu, Z., 2009. Louisiana wetland water
1059 level monitoring using retracked TOPEX/POSEIDON altimetry. *Marine Geodesy*, 32(3), pp.284-302.
- 1060 Lehner, B. and Döll, P., 2004. Development and validation of a global database of lakes, reservoirs and
1061 wetlands. *Journal of Hydrology*, 296(1-4), pp.1-22.
- 1062 Lehner, B., Liermann, C.R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K.,
1063 Magome, J. and Nilsson10, C., 2011. Global reservoir and dam (grand) database. *Technical Documentation*,
1064 *Version, 1*, pp.1-14.
- 1065 Li, H.Y., Sivapalan, M., Tian, F. and Harman, C., 2014. Functional approach to exploring climatic and landscape
1066 controls of runoff generation: 1. Behavioral constraints on runoff volume. *Water Resources*
1067 *Research*, 50(12), pp.9300-9322.
- 1068 Lindström, G., Rosberg, J. and Arheimer, B., 2005. Parameter precision in the HBV-NP model and impacts on
1069 nitrogen scenario simulations in the Rönneå River, Southern Sweden. *AMBIO: A Journal of the Human*
1070 *Environment*, 34(7), pp.533-538.
- 1071 Lindström, G., Pers, C., Rosberg, J., Strömqvist, J. and Arheimer, B., 2010. Development and testing of the HYPE
1072 (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrology*
1073 *research*, 41(3-4), pp.295-319.
- 1074 Lindström, G., 2016. Lake water levels for calibration of the S-HYPE model. *Hydrology Research*, 47(4), pp.672-
1075 682.
- 1076 Maidment, D.R., 1992. *Handbook of hydrology*. New York: McGraw-Hill.
- 1077 McDonald, S., Mohammed, I.N., Bolten, J.D., Pulla, S., Meechaiya, C., Markert, A., Nelson, E.J., Srinivasan, R. and
1078 Lakshmi, V., 2019. Web-based decision support system tools: The Soil and Water Assessment Tool Online
1079 visualization and analyses (SWATOnline) and NASA earth observation data downloading and reformatting
1080 tool (NASAaccess). *Environmental Modelling & Software*, 120, p.104499.
- 1081 Milly, P.C.D., 1994. Climate, soil water storage, and the average annual water balance. *Water Resources*
1082 *Research*, 30(7), pp.2143-2156.
- 1083 Mohammed, I.N., Bolten, J.D., Srinivasan, R. and Lakshmi, V., 2018. Satellite observations and modeling to
1084 understand the Lower Mekong River Basin streamflow variability. *Journal of hydrology*, 564, pp.559-573.
- 1085 Monteith, J.L., 1965. Evaporation and environment. In *Symposia of the society for experimental biology* (Vol. 19, pp.
1086 205-234). Cambridge University Press (CUP) Cambridge.
- 1087 Mu, Q., Zhao, M. and Running, S.W., 2011. Improvements to a MODIS global terrestrial evapotranspiration
1088 algorithm. *Remote Sensing of Environment*, 115(8), pp.1781-1800.
- 1089 Okeowo, M.A., Lee, H., Hossain, F. and Getirana, A., 2017. Automated generation of lakes and reservoirs water
1090 elevation changes from satellite radar altimetry. *IEEE Journal of Selected Topics in Applied Earth*
1091 *Observations and Remote Sensing*, 10(8), pp.3465-3481.
- 1092 Parajka, J., Viglione, A., Rogger, M., Salinas, J.L., Sivapalan, M. and Blöschl, G., 2013. Comparative assessment of
1093 predictions in ungauged basins–Part 1: Runoff-hydrograph studies. *Hydrology and Earth System*
1094 *Sciences*, 17(5), pp.1783-1795.
- 1095 Parajka, J., Blöschl, G. and Merz, R., 2007. Regional calibration of catchment models: Potential for ungauged
1096 catchments. *Water Resources Research*, 43(6).
- 1097 Pechlivanidis, I. and Arheimer, B., 2015. Large-scale hydrological modelling by using modified PUB
1098 recommendations: the India-HYPE case. *Hydrology and Earth System Sciences*, 19(11), pp.4559-4579.
- 1099 Potter, N.J., Zhang, L., Milly, P.C.D., McMahon, T.A. and Jakeman, A.J., 2005. Effects of rainfall seasonality and
1100 soil moisture capacity on mean annual water balance for Australian catchments. *Water Resources*
1101 *Research*, 41(6).
- 1102 Razavi, T. and Coulibaly, P., 2012. Streamflow prediction in ungauged basins: review of regionalization
1103 methods. *Journal of hydrologic engineering*, 18(8), pp.958-975.

1104 Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.T., Chuang, H.Y., Iredell, M.
1105 and Ek, M., 2011. NCEP climate forecast system version 2 (CFSv2) 6-hourly products. *Research Data*
1106 *Archive at the National Center for Atmospheric Research, Computational and Information Systems*
1107 *Laboratory*.

1108 Sellami, H., La Jeunesse, I., Benabdallah, S., Baghdadi, N. and Vanclooster, M., 2014. Uncertainty analysis in model
1109 parameters regionalization: a case study involving the SWAT model in Mediterranean catchments (Southern
1110 France).

1111 Schwatke, C., Dettmering, D., Bosch, W. and Seitz, F., 2015. DAHITI—an innovative approach for estimating water
1112 level time series over inland waters using multi-mission satellite altimetry. *Hydrology and Earth System*
1113 *Sciences*, 19(10), pp.4345-4364.

1114 Shiklomanov, A.I., Lammers, R.B. and Vörösmarty, C.J., 2002. Widespread decline in hydrological monitoring
1115 threatens pan-Arctic research. *Eos, Transactions American Geophysical Union*, 83(2), pp.13-17.

1116 Sivapalan, M., 2005. Pattern, process and function: elements of a unified theory of hydrology at the catchment
1117 scale. *Encyclopedia of hydrological sciences*.

1118 Strömqvist, J., Arheimer, B., Dahné, J., Donnelly, C. and Lindström, G., 2012. Water and nutrient predictions in
1119 ungauged basins: set-up and evaluation of a model at the national scale. *Hydrological Sciences*
1120 *Journal*, 57(2), pp.229-247.

1121 Sun, W., Ishidaira, H. and Bastola, S., 2012. Calibration of hydrological models in ungauged basins based on satellite
1122 radar altimetry observations of river water level. *Hydrological Processes*, 26(23), pp.3524-3537.

1123 Tang, X., Zhang, J., Gao, C., Ruben, G.B. and Wang, G., 2019. Assessing the Uncertainties of Four Precipitation
1124 Products for Swat Modeling in Mekong River Basin. *Remote Sensing*, 11(3), p.304.

1125 Tordoff, A.W., Bezuijen, M.R., Duckworth, J.W., Fellowes, J.R., Koenig, K., Pollard, E.H.B., Royo, A.G. and ZUN,
1126 P., 2012. Ecosystem profile: Indo-Burma biodiversity hotspot 2011 update. *Critical Ecosystem Partnership*
1127 *Fund, Arlington, Virginia*.

1128 Ter Braak, C.J., 2006. A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy
1129 Bayesian computing for real parameter spaces. *Statistics and Computing*, 16(3), pp.239-249.

1130 Troch, P.A., Martinez, G.F., Pauwels, V.R., Durcik, M., Sivapalan, M., Harman, C., Brooks, P.D., Gupta, H. and
1131 Huxman, T., 2009. Climate and vegetation water use efficiency at catchment scales. *Hydrological Processes:*
1132 *An International Journal*, 23(16), pp.2409-2414.

1133 Turton, A., Ashton, P. and Cloete, E., 2003. An introduction to the hydrological drivers in the Okavango River
1134 basin. *Transboundary Rivers, Sovereignty and Development: Hydrological Drivers in the Okavango River*
1135 *Basin. Pretoria & Geneva: AWIRU & Green Cross International*.

1136 UNEP-DHI, U.N.E.P., 2016. Transboundary river basins: status and trends (summary for policy makers). *Nairobi:*
1137 *United Nations Environment Programme (UNEP)*.

1138 Woods, R.A., 2009. Analytical model of seasonal climate impacts on snow hydrology: Continuous
1139 snowpacks. *Advances in water resources*, 32(10), pp.1465-1481.

1140 World Bank, 2019. Vietnam: Toward a safe, clean and resilient water system. Washington DC: The World Bank

1141 Yamazaki, D., O'Loughlin, F., Trigg, M.A., Miller, Z.F., Pavelsky, T.M. and Bates, P.D., 2014. Development of the
1142 global width database for large rivers. *Water Resources Research*, 50(4), pp.3467-3480.

1143 Yokoo, Y. and Sivapalan, M., 2011. Towards reconstruction of the flow duration curve: development of a conceptual
1144 framework with a physical basis. *Hydrology & Earth System Sciences*, 15(9).

1145