# Technical Note: Evaluation and bias correction of an observationsbased global runoff dataset using streamflow observations from small tropical catchments in the Philippines

Daniel E. Ibarra<sup>1,2</sup>, Carlos Primo C. David<sup>3</sup>, Pamela Louise M. Tolentino<sup>3</sup>

5

10

<sup>1</sup> Department of Earth and Planetary Science, University of California, Berkeley, California 94720 USA

<sup>2</sup> Institute at Brown for Environment and Society and the Department of Earth, Environmental and Planetary Science, Brown University, Providence, Rhode Island 02912 USA

<sup>3</sup> National Institute of Geological Sciences, University of the Philippines, Diliman, Quezon City, Philippines 1101 *Correspondence to*: Daniel E. Ibarra (dibarra@berkeley.edu) and Carlos Primo C. David (cpdavid@nigs.upd.edu.ph)

## Abstract

Even in relatively wet tropical regions, seasonal fluctuations in the water cycle affect the consistent and reliable supply of water for urban, industrial and agricultural uses. <u>Historic</u> streamflow monitoring datasets are crucial in assessing our ability to model and subsequently plan for future hydrologic changes. In this technical note, we evaluate a new <u>observations-based</u> global product of monthly runoff (GRUN; Ghiggi et al., 2019) for <u>55</u> small tropical catchments in the Philippines with at least 10 years of data, extending back to 1946 in some cases. Since GRUN didn't use discharge data from the Philippines to train/calibrate their models, the data presented in this study, 11,915 monthly data points, provide an

- 20 independent evaluation of this product. We demonstrate across all observations a significant but weak correlation (r<sup>2</sup> = 0.372) between the GRUN predicted values and observed river discharge, and somewhat skilful prediction (Volumetric Efficiency = 0.363 and log(Nash-Sutcliff Efficiency) = 0.453). GRUN performs best among catchments located in Climate Types III (no pronounced maximum rainfall with short dry season) and IV (evenly distributed rainfall, no dry season). There was a weak negative correlation between volumetric efficiency and catchment area, and a positive correlation between
- 25 volumetric efficiency and mean observed runoff. Further, analysis for individual rivers demonstrates systematic biases (over and under estimation) of baseflow during the dry season, and under-prediction of peak flow during some wet months for most catchments. To correct for underprediction during wet months, we applied a log-transform bias correction which greatly improves the nationwide Root Mean Square Error between GRUN and the observations by an order of magnitude (2.648 vs. 0.292 mm/day). This technical note demonstrates the importance of performing such corrections when
- 30 <u>determining</u> the proportional contribution of smaller catchments <u>or</u> tropical <u>islands</u> such as the Philippines to global tabulations of discharge. These results <u>also</u> demonstrate the potential use of GRUN and future data products of this nature after consideration and correction of systematic biases to: 1) assess trends in regional scale runoff over the past century, 2)

Deleted: In

Deleted: complicate

Deleted: Importantly, historic

**Deleted:** . In particular, we evaluate the observations-based monthly runoff product using archived monthly streamflow data from 55 catchments with at least 10 years of data

Deleted: We also found

Deleted: land

Deleted: ) in	
Deleted: Finally, to	
Deleted: perform	
Deleted: accounting for	

validate hydrologic models for un-monitored catchments in the Philippines, and 3) assess the impact of hydrometeorological phenomena to seasonal water supply in this wet but drought prone archipelago.

#### **1** Introduction

	The global water crisis affects an estimated two-thirds of the world's population and is considered one of the three	
50	biggest global issues that we need to contend with (Kummu et al., 2016; WEF, 2018). The most important source of	
	freshwater in terms of use is surface water. It is the primary resource for irrigation, industrial use and provides the bulk of	1
	water supply for many large cities. Long term streamflow datasets are useful for resource management and infrastructure	
	planning (e.g., Evaristo and McDonnell, 2019). Such data is even more critical in areas that rely on run-of-the-river and do	
	not use storage structures, such as dams and impoundments. Further, a robust, long term dataset is crucial in the face of	

- 55 increased variability in stream discharge due to land use change, increased occurrence of mesoscale disturbances and climate change (e.g., Abon et al., 2016; David, et al., 2017; Kumar et al., 2018). In the absence of long-term streamflow datasets for most locations in the world, several researchers have compiled datasets worldwide that are used to estimate streamflow in <u>un</u>-gauged areas (Maybeck et al., 2013, Gudmundsson et al., 2018, Do et al., 2018; Alfieri et al., 2020; Harrrigan et al., 2020). Several global hydrological models have also been created to project variations in streamflow and extend present-day
- 60 measurements to the future (Hagemann et al., 2011; Davie et al., 2013; Winsemius et al., 2016). A recent contribution to modelled global runoff products is the Global Runoff Reconstruction (GRUN) (Ghiggi et al., 2019). GRUN is a global gridded reconstruction of monthly runoff for the period 1902-2014 at 0.5 degree (~50km by 50km) spatial resolution. It used global streamflow data from 7,264 river basins that to train a machine learning algorithm that inferred runoff generation processes from precipitation and temperature data.

65

70

There is a disparity in the availability of long-term gauged rivers datasets between continental areas and smaller island nations, which have more dynamic hydrometeorologic system owing to the size of the catchments and proximity to the ocean (e.g., Abon et al., 2011; Paronda et al., 2019). <u>As a result, information from island nations is usually not used to train or evaluate global model</u>. The Philippines offers a unique example where manual stream gauging programs have started in 1904 and, while spotty at times, have continued to today. This island nation on the western side of the Pacific Ocean is characterized by a very dynamic hydrologic system because it is affected by tropical cyclones, seasonal monsoon rains, sub-decadal cycles such as the El Nino Southern Oscillation (ENSO) and climate change (Abon et al., 2016; David, et al., 2017; Kumar et al., 2018). The impact of climate change on the hydrological cycle can be observed the most for tropical island nations, including the Philippines (Nurse et al., 2014).

This technical note evaluates the accuracy of the GRUN dataset (GRUN\_v1) as applied to the hydrodynamically-75 active smaller river basins in the Philippines. Additionally, it explores the possible hydrologic parameters that may need to be considered and/or optimized <u>so that</u> such global datasets <u>can</u> predict runoff in smaller, ungauged basins more accurately. Deleted: phenomenon

<b>Deleted:</b> ; it affects an estimated two-thirds of the world's population
Deleted: Among the sources of freshwater, the
Deleted: compartment
Deleted: supply

(	Deleted: which
······	Deleted: extrapolate
······(	Deleted: non
	Deleted: The latest

Deleted: which learn the

**Deleted:** However, the impact of climate change on the hydrological cycle can be observed the most for tropical island nations (Nurse et al., 2014).As a result, information from island

Deleted: that Deleted: are able to

## 2 Dataset and Methods

# 95

2.1\_Climate Types Moved (insertion) [1] The Philippines has four Climate Types (see also Abon et al., 2016; Tolentino et al., 2017; Figure 1): Type I Climate on the western seaboard of the Philippines is characterized by distinct wet (May to October) and dry (November to April) seasons; Type II Climate on the eastern seaboard has no distinct dry period with maximum rainfall occurring from 100 November to February; Type III inland climate experiences less annual rainfall with a short dry season (December to May) and a less pronounced wet season (June to November); and, Type IV southeast inland climate receives less rainfall and is characterized by an evenly distributed rainfall pattern throughout the year. 2.2 Historical streamflow data 105 In this contribution we analyse monthly observations of discharge from 55 manually observed streamflow stations Moved (insertion) [2] from three Philippine datasets. The observations span the period between 1946 to 2016, although only data through 2014 are used due to the time period included in GRUN (see section 2.3) All datasets needed to include at least 10 years of data. The

## 110 2.2.1 Bureau of Research and Standards (BRS) Dataset

location of all streamflow stations is shown on Figure 1 and listed in Table 1.

The discharge data was originally acquired from the Bureau of Research Standards (BRS) under the Department of Public Works and Highways (DPWH). The records keeping was transferred to the Bureau of Design, also under DPWH, which continues to record gage data from some rivers up to today. The majority of the reprocessed BRS data used in this analysis come from Tolentino et al. (2016), some of the datasets were updated using data available from the Department of

115 Public Works and Highways. A discussion of the accuracy of this data based on comparison to manual daily discharge measurements can be found in Tolentino et al. (2016).

Moved (insertion) [3]

Moved (insertion) [4]



#### 2.2.2 Global Runoff Data Centre (GRDC) Reference Dataset

 Data
 from
 ten
 catchments
 from
 the
 GRDC
 Reference
 Dataset

 125
 (https://www.bafg.de/GRDC/EN/04\_spcldtbss/43\_GRfN/refDataset\_node.html; requested July 2019) were analysed. Over
 45 sites from the Philippines are available in the GRDC data; however, almost none fulfilled our criteria of having over 10

 years of data. Four of these catchments match or extended the BRS datasets, and one extends a GSIM dataset (see below).
 Notably four of the time series available from GRDC are available back to the 1940s (Table 1).

## 130 2.2.3 Global Streamflow Indices and Metadata Archive (GSIM) Reference Dataset

Only two of the available five GSIM time series for the Philippines (Gudmundsson et al., 2016, Do et al., 2018) contain more than 10 years of data.

#### 2.3 GRUN observations-based global gridded (0.5°x0.5°) runoff dataset

135 GRUN is a recently published global reconstruction of monthly runoff time series for the period 1902 to 2014. It was created using a machine learning algorithm based on temperature and precipitation data from the Global Soil Wetness Project Phase 3 (GSWP3; Kim et al., 2017; <u>http://hydro.iis.u-tokyo.ac.jp/GSWP3/index.html</u>) using the Global Streamflow Indices and Metadata Archive (GISM) (Ghiggi et al., 2019). In this contribution we analysed GRUN v1 (<u>https://figshare.com/articles/GRUN\_Global\_Runoff\_Reconstruction/9228176;</u> accessed September 9<sup>th</sup> 2019) which was

- 140 trained on a selection of catchments with an area between 10 and 2500 km<sup>2</sup> GSIM (Do et al., 2018; Gudmundsson et al., 2018) and validated using 379 large (>50,000 km<sup>2</sup>) monthly river discharge datasets from the Global Runoff Data Centre (GRDC) Reference Dataset (<u>https://www.bafg.de/GRDC/EN/04\_spcldtbss/43\_GRfN/refDataset\_node.html</u>). Additionally, due to the criteria for training data, GRUN's calibration is biased towards the northern hemisphere mid-latitudes; <u>discharge</u> data are available for only a few sites in the tropics in Africa and southeast Asia (Mulligan, 2013). Ghiggi et al. (2019)
- 145 discuss that because of the dataset training technique, uncertainty scales with the magnitude of runoff. GRUN is likely to have high prediction uncertainty in regions with less dense runoff observations <u>and high discharge</u> such as in tropical southeast Asia. However, <u>Ghiggi et al. (2019)</u> show for southeast Asia an increase in runoff and a strong correlation of runoff with ENSO for the period of analysis (1902 to 2014). We refer the reader to Ghiggi et al. (2019) for more information but note that because of the catchment size filtering criteria, none of the GISM and GRDC data from the Philippines were
- 150 used in calibration or evaluation (*Personal Communications*, G. Ghiggi, 2019). As such, we view our analysis as a completely independent test of the GRUN runoff prediction for small tropical catchments.

#### 2.4 Catchment Area and Pairing with GRUN Gridcells

All catchment areas were verified using the digital elevation model from the 2013 Interferometric Synthetic Aperture Radar (IfSAR) data. All runoff datasets were normalized (mm/yr), i.e., 'specific discharge'. We only considered streamflow stations where the published and verified areas agreed. Catchment areas span 4 orders of magnitude (8.93 to

Moved (in	sertion) [7]
-----------	--------------

Moved (insertion) [8]

Deleted: because

Deleted. mcy

Formatted: Font: Bold

## Deleted: 2.2 Historical streamflow observations

**Moved up [2]:** In this contribution we analyse monthly observations of discharge from 55 manually observed streamflow stations from three Philippine datasets. The observations span the period between 1946 to 2016, although only data through 2014 are used due to the time period included in GRUN.

### Deleted:

**Moved up [3]:** All datasets needed to include at least 10 years of data. The location of all streamflow stations is shown on Figure 1 and listed in Table 1.

#### 2.2.1 Bureau of Research and Standards (BRS) Dataset The discharge data was originally acquired from the Bureau of

Research Standards (BRS) under the Department of Public Works and Highways (DPWH). The records keeping was transferred to the Bureau of Design, also under DPWH, which continues to record gage data from some rivers up to today. A majority of the

**Deleted:** A majority of the reprocessed BRS data used in this analysis come from Tolentino et al. (2016), however, some of the datasets were subsequently updated using

**Moved up [4]:** updated using data available from the Department of Public Works and Highways. A discussion of the accuracy of this data based on comparison to manual daily discharge measurements can be found in Tolentino et al. (2016).<sup>¶</sup>

#### Deleted: overlaid on

**Moved up [5]:** climatic type (as in Tolentino et al., 2016; Kintanar, 1984; Jose and Cruz, 1999). Note that no

**Deleted:** Note that no publicly available long-term stations are available for Palawan.

## Moved up [6]: ¶

2.2.2 Global Runoff Data Centre (GRDC) Reference Dataset Deleted: Ten catchments from the GRDC Reference Dataset

(https://www.bafg.de/GRDC/EN/04\_spcldtbss/43\_GRfN/refDataset

BRS datasets, and one extends a GSIM dataset (see below). Notably

Deleted: 2.2.3 Global Runoff Data Centre (GSIM) Reference Dataset ¶ .... [2]

**Moved up [8]:** (Gudmundsson et al., 2016, Do et al., 2018) contain more than 10 years of data.

#### Deleted:

2.3 Criteria for Inclusion of Datasets

6,487 km<sup>2</sup>) and cover the majority of the Philippines excluding Palawan (see Figure 1). The location of catchments was paired to GRUN grid cells (0.5° by 0.5°) for the analysis. Instead of computing the weighted area runoff over the catchment, we employed the nearest neighbour interpolation between the catchment outlet location and the GRUN gridded product (0.5° by 0.5° resolution). All but one catchment is smaller than the area of the GRUN grid cells (~2,500 km<sup>2</sup>), thus, we view this pairing as sufficient for validation purposes. This assumption was tested by interpolating the GRUN grid to the gauging

location as well as the watershed centroids, but this did not lead to a significant difference in correlation.

## 2,5 Comparison of GRUN Estimates and Observations

To assess the performance of GRUN, we use a suite of metrics commonly used to assess model performance in hydrologic studies. Given the emphasis on a country scale evaluation of GRUN we primarily focus below on results in aggregate grouped by climate type or for all catchments. These metrics are calculated for each individual catchment (n=55) and in aggregate for each climate type (n=4; see below) shown in Figure 1. GRUN was not intended to be used for estimating discharge for single small catchments, therefore we focus on the aggregated data but also report the range for the result for the individual catchments.

Firstly, we use the commonly used coefficient of determination ( $r^2$  or r-squared). This bivariate correlation metric measures the linear correlation between two variables. In this case the predicted monthly values from GRUN and the observed monthly values from the streamflow datasets. It varies from 0 (no linear correlation) to 1 (perfect correlation). The use of r-squared does not account for systematic over- or under-prediction in runoff because it only accounts for correlation

among the observed and predicted values (see Krause et al. (2005) for further discussion of the use of r-squared in 240 hydrological model assessment).

The second metric used here is the Volumetric Efficiency (VE) (Criss and Winston, 2008), utilized previously by Tolentino et al. (2016) on a subset of the BRS catchments analysed here. VE is defined as:

245

250

225

230

235

 $VE = 1 - \frac{\sum |Q_P - Q_O|}{\sum Q_O}$ (1)

where Q is the monthly discharge and subscript P is used for the modelled/predicted values and O the observed runoff values. A value of 1 indicates a perfect score. Because we are interested in the performance of GRUN over the period of each streamflow record, we calculate VE using all paired monthly observed and simulated values rather than the monthly medians that were used in Tolentino et al. (2016). This results in lower VE scores than those reported by Tolentino et al. (2016).

Further, we use the linear Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe (1970):

$$NSE = 1 - \frac{\sum (Q_P - Q_O)^2}{\sum (Q_O - mean(Q_O))^2}$$
6

Deleted: 4 Statistical Performance Metrics Deleted: Tools

Moved (insertion) [9]

Deleted: square

(2)

Deleted: is

Deleted: both

Deleted: and logged

Deleted:  $\frac{\sum Q_P - Q_O}{\sum Q_O}$ 

The NSE can vary between -∞ and 1 (perfect fit). NSE values are useful because values less than zero indicate that the model is no better than using the mean value of the observed data as a predictor. The logged NSE was also calculated using logarithmic values of runoff to reduce the influence of a mismatch during peak flow and to increase the influence of low flow values (see further discussion in Krause et al., 2005).

To evaluate a possible strategy for performing a bias correction of the GRUN simulated values at a countrywide scale, we use the Root Mean Square Error (RMSE) in units of runoff (i.e., mm/day). The RMSE <u>vas</u> applied to the raw GRUN simulated values and the observation based bias corrected GRUN values at the country, climate type (see below) and individual catchment level.

Finally, to evaluate distributions of flow duration using flow duration curves by catchment and aggregated by climate type we use the 'fdc' function in the R package 'hydroTSM' (Zambrano-Bigiarini, 2020), and include in our comparison GRUN predicted values only for months which observations are available.

7

275

Deleted: is

Deleted: is

## Moved up [1]: Climate Types

The Philippines has four Climate Types (see also Abon et al., 2016; Tolentino et al., 2017; Figure 1): Type I Climate on the western seaboard of the Philippines is characterized by distinct wet (May to October) and dry (November to April) seasons; Type II Climate on the eastern seaboard has no distinct dry period with maximum rainfall occurring from November to February; Type III inland climate experiences less annual rainfall with a short dry season (December to May) and a less pronounced wet season (June to November); and, Type IV southeast inland climate experiencing

## Deleted: 2.5

**Deleted:** experiencing depressed rainfall and is characterized by an evenly distributed rainfall pattern throughout the year.



**Figure 2.** Example timeseries of GRUN predicted (red lines) and observed (black lines) runoff values, and cross-plots (log-scale) with VE,  $r^2$  and RSME values for the worst (top) and best (bottom) performing river basins within Climate Types I, II, III and IV (panels A-D, respectively).

300

#### **3** Results and Discussion

Figure 2 and the supplemental figures (Figure A1) show comparisons between the time series of the GRUN runoff values and runoff (area normalized discharge). Statistics performance metrics across all data as well as by climate types I to IV are listed in Table 2 and Table A1. In the following sections, we break down the comparison between the streamflow observations and GRUN by reporting summary statistics, comparing runoff distributions and extreme values at the individual basin level and by Climate Type, analyse flow duration curves, and finally look at several correlations of VE to watershed characteristics. Following this we calculate bias correction regressions and provide an outlook for future work.

3.1 Comparison of runoff distributions

305

Average runoff values among all catchments are somewhat well predicted by GRUN. Across all observations the rsquared of the correlation between GRUN prediction and observation was of 0.372 and VE was 0.363 (Table 2). Using log(runoff) values (following Criss and Winston, 2008) the r-squared improved to 0.546 and a VE to 0.733, suggesting reasonable utility in the GRUN product at the country scale for the Philippines, even though no training data from the Philippines were used in the creation of GRUN. The RMSE across the dataset was 2.648 mm/day (Table 2). NSE and NSE-

310 log10 values, ranged, from -10.70 to 0.68 and -11.53 to 0.76, for individual catchment comparisons, with median values of 0.02 and 0.24, respectively. For more than half of the catchments, (29 of 55) NSE values were greater than 0, but for, only 5 catchments it was higher than 0.5. Similarly, for 32 of 55 catchments the NSE-log10 values were greater than 0, and for 12 catchments greater than 0.5.

The VE of the median values of runoff (black dots in Figure 3A) was 0.509 across all catchments; the average (mean) difference between the observed and simulated median runoff values was +16% (Figure 3A). The median and interquartile ranges (IQR, 25% to 75%) of the runoff for the individual catchments for the GRUN and the observations, overlap (Figure 2E). For two large catchments and three relatively small catchments the IQR of the observations does not overlap with the GRUN runoff IQR. The three small catchments are located in Climate Type III (yellow) and the two large catchments are Climate Type IV. For two catchments of moderate size, the GRUN IQR is greater than the observed IQR 320 runoff range.

Looking at extreme monthly values (maximum and minimum) over the period of observation demonstrates significant underprediction <u>during the</u> wettest conditions (orange dots in Figure 3A and 3D). For almost all catchments' the maximum observations <u>plot</u> above the 1:1 line, <u>The VE for the maximum runoff is lower than for</u> the median runoff (0.194). The minimum values plot around the 1:1 line and are more evenly distributed, however the VE score <u>for the minimum runoff</u>

(0.154) is similarly low due to greater spread than for the median values.

Regardless of Climate Type, there is a general underestimation of the model is seen for the highest runoff but this is especially evident for catchments in Climate Types I and II with pronounced wet seasons (as also shown by their lower rsquared values; Figure 4) and lower VE values. For Climate Type II the RMSE value of 4.55 mm/day (compared to an

Deleted: time series comparisons between the time series of the[3]

**Moved up [9]:** Given the emphasis on a country scale evaluation of GRUN we primarily focus below on results in aggregate grouped by climate type or for all catchments. In the following

Deleted: In the following sections

Deleted: ofwas 0.363 (Table 2). Using log(runoff) values [4]

Deleted: In Figure 3A ... he VE of the median values of runoff [5]

Deleted: induring the wettest conditions (orange dots in Figure 16]

Deleted: months, looking at the distributions by basin. This...ut[7]

average observed flow of 9.03 mm/day is also the highest. Climate Types III and IV have comparable r-squared and VE values and skewness towards underprediction during the highest runoff months is still evident, particularly for Climate Type IV. These patterns are particularly evident looking at Figure 5, which shows flow duration curves (FDC) by climate type for individual catchments groups (see Figure A2 for individual catchment comparisons). Such an analysis allows for inspection of runoff distributions and biases across the range of observed and predicted values. At low flows (high exceedance probability, >80%) there is reasonable agreement in the shape and magnitude of the distributions between GRUN and the observations for Climate Types I and IV (bottom right of the FDC plots). For Climate Type III there is a consistent bias across all runoff values with an exceedance probability 590%. At high flow (low exceedance probability, <20%), as also noted above, runoff for all Climate Types is underestimated by GRUN, with the greatest discrepancy for Climate Types I

## Deleted: )

1	
(	Deleted: though
1	Balatada a star
	Deleted: catchments
	Deleted: flow
	Deleteu: now
	Deleted: for Climate Types Land IV
. (	Deleted. for Chinate Types I and Tv
(	Deleted: demonstrates
(	Deleted: less than
. /	× · · · · · · · · · · · · · · · · · · ·
	Deleted: of ~
(	Deleted: Climate Types Land II showing
	Deleter Chinate Types Land It showing

395 and II (top left of each FDC plot).

390



Figure 3. Comparison of runoff ranges and distributions. (A) Comparison of median and extreme (maximum and minimum) monthly runoff for the observations and GRUN in log space. (B-D) As in (A) for the minimum, median and maximum 400

Deleted: values between

monthly values, respectively, in linear space. (E) Distribution of runoff (observations (coloured) and GRUN (white)) for the individual catchments (ordered by size). Plots show the median (line), interquartile range (box) and maximum/minimum values (whiskers). The boxplots for GRUN also only includes months for which observations are available.

Deleted: between

Deleted: include

Deleted: Cross plots

Climate Type I Climate Type II 1e+02 1e+02 1e+01 1e+01 Observed (mm/day) 1e+00 1e+00 20 le-01 e-01  $r^2 = 0.211$ = 0.409 VE = 0.252 VE = 0.354 e-02 e-02 1e-02 1e-01 1e+00 1e+01 1e+02 1e-02 1e-01 1e+00 1e+01 1e+02 GRUN Predicted (mm/day GRUN Predicted (mm/day) Climate Type IV Climate Type III 1e+02 1e+02 1e+01 1e+01 Observed (mm/day) 1e+00 e+00 1e-01 e-01  $r^2 = 0.471$  $r^2 = 0.535$ VE = 0.404 VE = 0.403 le-02 -02 1e-02 1e-01 1e+00 1e+02 1e-02 1e+01 1e+02 1e+01 1e-01 1e+00 GRUN Predicted (mm/day) GRUN Predicted (mm/day)

### 415

Figure 4: <u>Plots</u> of GRUN predicted vs. observed monthly runoff by climate type (see Figure 1 for climate type distributions). Grey dots represent all data, colored dots represent data points from that region. The squared pearson correlation coefficient (r<sup>2</sup>) and volumetric efficiency (VE) metrics are listed for each panel.



## 3.2 Correlation and trends with watershed characteristics

The biases noted above are likely due to the high uncertainty and underprediction in monsoonal precipitation that is used for input into GRUN. There is a significant negative correlation (at p<0.01,  $r^2 = 0.391$ ) between log values of maximum runoff difference (observed minus predicted) and catchment area (not shown). This suggests two possibilities: first, that particularly for small catchments which may have steeper average slope, GRUN underpredicts monthly runoff values during the wet season due to the bias in the precipitation datasets used to create GRUN; and second model-data agreement improves with catchment size. In this section we explore these possibilities.

There was a weak positive correlation (r<sup>2</sup> = 0.041, p = 0.137) between VE and log(catchment size) (Figure 6) and a stronger negative correlation with mean runoff (r<sup>2</sup> = 0.182, p < 0.01). However, for catchments with a low mean runoff there is significant spread in VE score, driven primarily by Climate Type I catchments (red box and whisker plot in Figure 6C). These catchments experience distinct wet and dry seasons and are predominantly located in the northwest Philippines. The positive correlation with catchment size is likely primarily due to the extreme wet months. The bias of the GRUN data at high and moderate flow conditions is particularly evident in the flow duration curves (Figure 5). This is also in line with the Nash-Sutcliff Efficiency (NSE) than the log(Nash-Sutcliff Efficiency) (NSE-log10) scores (Table 2 and Table A1) because NSE puts more weight on high flow (Criss and Winston, 2008) as well as the higher VE scores for log10 runoff values for all

the runoff data across the entire dataset (0.363 vs. 0.733; Table 2). The physical significance of these results could be that for large basins the time of concentration of any given flood event will be much longer, thus flood peaks will be wider and subdued due to infiltration into the shallow aquifers. However, a more likely explanation is that the average rainfall intensity is too low in the GSWP3 precipitation data used in producing GRUN. This is likely due to downscaling from averages over

- larger areas with topographic complexity. While GSWP3 uses downscaled 20<sup>th</sup> century reanalysis products (Kim et al., 2017) at T248 resolution (~0.5°, the same resolution as GRUN), the topographic complexity of tropical islands such as the Philippines on the sub-0.5° scale would likely results in smoothing of the variability and lowering of the absolute magnitude of the precipitation fields, particularly <u>for small catchments and</u> during the wet season and during large synoptic precipitation
- 455 events during the monsoon season. Further, because of the monthly (rather than daily or sub-daily) output of GRUN and the comparison carried out here, the biases observed in our analysis could be <u>due to</u> averaging of the input products and/or discharge datasets to monthly values, which would mask extreme values influencing the observational dataset since the <u>runoff distribution of daily or sub-daily runoff is heavier tailed than that of monthly distributions</u>. Finally, while the GSWP3 precipitation inputs are bias-corrected using the Global Precipitation Climatology Centre precipitation network, previous work has highlighted data quality issues with some historical data from the Philippines (Schneider et al., 2014).

The underestimation of runoff values during extreme rain events is most likey due to an underestimation in the wet season rainfall products from GSWP3 used to create GRUN as described above. Alternatively, it may be a result of the fast saturation of the overlying soil and shallow aquifers filling up during the wet season, and potentially high amounts of direct runoff (e.g., Tarasova et al., 2018). On the other hand, the underestimation of flow during low flow events may be a result of not accurately accounting for stream baseflow which is fed by shallow aquifers and influenced by land use and surface

Deleted: inputted

Deleted: versus

Formatted: Space Before: 0 pt

Deleted: at

Deleted:	dominantly
Deleted:	This is
Deleted:	from
Deleted:	and
Deleted:	in
Deleted:	. Since
Deleted:	large
Deleted:	), it is not surprising that our NSE-
Deleted: weighted lo	scores are better because extremely wet months are ess than using the raw runoff
Deleted: VE scores	compared to raw NSE scores. It is also notable that the using log10 valuesfor all the runoff data
Deleted:	is are better
Deleted:	observations
Deleted: particularly	A bias at high and moderate flow conditions is y evident in the flow duration curves (Figure 5) and This

Deleted: because of

Moved (insertion) [10]

13

properties. These effects may be buffered in larger catchments leading the increase in that model-data agreement with catchment size (Figure 6B).

Previous studies have investigated the correlation between runoff and catchment size (Mayor et al., 2011), and the different hydrologic and geologic factors that cause non-linear relationships between these two variables (Rodriguez et al.,

490

2014). Recently, Zhang et al. (2019) point out that runoff coefficients increase logarithmically as catchment size decreases. Moreover, the same <u>study</u> reports <u>that</u> the effects of vegetation cover, slope, and land use are larger <u>for smaller catchments</u> than for larger catchments. This implies that <u>prediction</u> of basin runoff for smaller catchments <u>is</u> more difficult due to the <u>variantions</u> in the compounding factors mentioned above. We hypothesize that <u>the</u> effects proposed by Zhang et al. (2019) also influence the Philippines streamflow dataset <u>used</u> in this study. Moved (insertion) [11]

### Deleted: paper

Deleted: for smaller catchments

## Deleted: compared to

Deleted: predictability

### Deleted: are

Deleted: variances

Deleted: these

#### Deleted: use

**Deleted:** As such, we suggest that GRUN, is a useful new tool for studying trends, seasonality and average runoff from tropical catchments in the Philippines...

**Moved down [12]:** (such as in previous work: e.g., Merz et al., 2011; Wanders and Wada, 2015). However, we qualify this finding

Deleted: ). However, we qualify this finding by noting, based on the limitations discussed above, that GRUN is not suitable for extreme value analyses associated with major tropical storms during the wet seasons unless suitable bias corrections (see next section) can be effectively carried out.





Figure 6: Diagnostic plots of volumetric efficiency (VE) results and comparison across metrics. Cross plots show the
 correlation of VE with (A) catchment area and (B) mean runoff. (C) Box and whisker plots show data from distribution of
 VE by climate type. Box and whisker plots show the median, interquartile range and 95% confidence intervals and outliers
 (dots). The regression in (A) is between the VE scores and ln(Catchment Area). (D-F) As in (C) for r<sup>2</sup>, NSE and NSE-log10.

#### 3.3 Bias Correction and Outlook

Overall, GRUN underestimates the actual observed runoff for <u>the</u> Philippine basins. The GRUN dataset predicts a range of 0 to 10mm/day for most basins and up to 20mm/day for larger basins. The observed maximum <u>monthly</u> runoff values are on average higher and exceed 50mm/day during months with extreme rain events (Figure 4). <u>The bias is largest</u> for the high flows (Figure 3a and Figure 5). Furthermore, the GRUN dataset also appears to underestimate minimum flow in streams from highly seasonal catchments (e.g. Types I and II).

525

520

Given the biases and in particular the clear underprediction of streamflow in GRUN during the wettest months we perform a bias correction of the GRUN dataset at a nationwide level using all the available data used in our analysis. We do so in a two-step process to both correct the mean offset and stretch the wettest months to higher values with all transformations occurring in log-transform space (i.e., as displayed in cross plots in Figures 2 and 4). Thus, we first add the mean log10(runoff) difference between the observations and the predicted values (0.117  $\pm$  0.045). Following this, using the

- 530 *Im* function in R, we fit a linear regression between the observations and the GRUN predicted values (log10(runoff, observed) =  $m \times log10(runoff, predicted) + b$ ) and correct the predicted values using the slope (m=0.774 ± 0.058) and intercept (b=0.099 ± 0.030) derived from this regression. Uncertainties reported here are 68% confidence intervals and were assessed by bootstrap resampling observation and prediction pairs from 20 catchments (vertical line in Figure 7) without replacement 10,000 times. In Figure 7 we show the influence of including an increasing number of catchments from the
- 535 dataset in our bootstrap resampling to assess how the mean value of the coefficients asymptote as more catchments are included. While the mean log10(runoff) offset is relatively unaffected, the slope and intercept of the bias correction do not reach a stable number until more than 10 catchments are included in the analysis. A leave-one-out approach (i.e., calculating the coefficients with 54 of 55 catchments) indicates 68% confidence uncertainties of ± 0.005, ± 0.006, and ± 0.003 for the log10(runoff) offset, the slope and intercept, lower than those reported above, as expected.
- 540 By carrying out these calculations in log-transform space the highest GRUN runoff values are the most affected, which are the data points that were most underpredicted (Figures 3A and 4). Because these corrections were carried out in log10 space statistical bias in the form of underestimation is possible (Ferguson, 1986). Following Ferguson (1986) we <u>calculated</u> the unbiased estimate of the variance (notated as 's') as 0.0686 mm/day which gives a correction factor (calculated as exp(2.65s<sup>2</sup>) of 1.0126. This correction factor, a multiplier, can be applied to the bias corrected values to adjust for possible the bias due to the log10 space regression we have implemented.

To assess this bias correction, we calculated RMSE values at a catchment, climate type and countrywide level (Figure 8 and Tables 2 and A1). The log-transform bias correction improves the nationwide RMSE value by an order of magnitude (2.648 vs. 0.292) and most significantly improves catchments in Climate Types III and IV (Figure 8; 2.285 vs. 0.432 and 2.398 vs. 0.131, respectively; Table 2). Interestingly, the median RMSE value for Climate Type I and II

550 catchments <u>was</u> not notably improved, however, the RMSE range for both <u>was</u> reduced (red and blue boxes in Figure 8, respectively).

Moved up [10]: direct runoff (e.g., Tarasova et al., 2018).

**Deleted:** The underestimation of runoff values during extreme rain events may be a result of the fast saturation of the overlying soil and exceedance of rates of infiltration partly as a result of shallow aquifers filling up and consequently the conversion of excess rainfall into

Moved up [11]: These effects may be buffered in larger catchments leading the increase in that model-data agreement with catchment size (Figure 6B). Alternatively, this could be due to an

**Deleted:** On the other hand, the underestimation of flow during low flow events may be a result of not accurately accounting for stream baseflow which is fed by shallow aquifers, as well as other effects such as land use and surface properties....

Deleted: Alternatively, this could be due to an underestimation in the wet season rainfall products from GSWP3 used to create GRUN as described above.

-(	Deleted:	asymptote
-(	Deleted:	Finally, a

Deleted: calculate

Deleted: are
Deleted: have been

This analysis and the improvement of RMSE values, as well as some of performance metrics such as NSE (see scores tabulated in Table 2), using a simple log-transform based bias correction demonstrates the importance of either: 1) 575 including smaller catchments in future products such as GRUN, or 2) performing similar bias corrections on a country, region or even catchment scale as needed. This is particularly important because if taken at face value the proportional contribution of relatively small tropical land areas to global discharge (e.g., Dai and Trenberth, 2002) would be underestimated without such corrections.

> 1.0 68% C. 0.9 Slope (m) 0.8 0.7 Mean 0.6 0.20 0.15 Intercept (b) 0.10 0.05 15 10 20 5 No. of Catchments

580

Figure 7. Bootstrap uncertainty of bias correction. Plotted versus the number of catchments is the mean and confidence interval (68%) log10(runoff) offset, the slope and intercept of the bias correction determined via random sampling at the catchment level without replacement. The vertical line at 20 catchments is the uncertainty reported in the text where all three coefficients asymptote to the mean values determined using the entire dataset.

585



Deleted: appropriate



Figure 8. Box and whisker plots of the Root Mean Square Error (RMSE) for catchments grouped by climate type of observed values versus raw GRUN values (light-coloured boxes) and bias-corrected GRUN values (bold-coloured boxes).
590 For bias correction equation and country-wide results see Table 2.

## 4 Conclusion

Based on monthly runoff observations from catchments in the Philippines with more than 10 years of data between 1946 and 2014, there is a significant but weak correlation ( $r^2 = 0.372$ ) between the GRUN-predicted runoff values and actual observations. The results indicate a somewhat skilful prediction for monthly runoff (Volumetric Efficiency = 0.363 and Deleted: Using

Deleted: we

log(Nash-Sutcliff Efficiency) = 0.453) when all data are pooled. Looking at different hydrometeorological regimes, we demonstrated that GRUN performs <u>better for</u> low rainfall catchments located in climate types III and IV. There was a weak negative positive correlation between volumetric efficiency and catchment area. Further, we found that particularly for smaller catchments, maximum wet season <u>runoff</u> values are grossly underpredicted by GRUN. The application of a nationwide bias correction to stretch high runoff values using log<u>transformed</u> runoff values greatly improved the RMSE of the predicted values. Global databases such as GRUN <u>can be</u> applicable for aggregated stream discharge estimates and to investigate general trends in the hydrologic characteristics of a region (such as in previous work: e.g., Merz et al., 2011; <u>Wanders and Wada, 2015</u>) but bias correction is needed when applying them to smaller catchments or regions for which data were not used in the development of the dataset. GRUN was not intended to be used for estimating discharge for single small catchments, however it is applicable for use in regional and country scale analyses provided that proper statistical comparison of modelled versus actual gauged data are performed. <u>The recommended bias correction presented here will</u>

610 <u>likely improve such estimates and analysis for the Philippines.</u> We thus propose that the use of the GRUN dataset can be extended to other ungauged tropical regions with smaller catchments after at least applying a similar correction as described in this study.

(	Deleted:	)

## Deleted: best among

 Deleted: transform
 Deleted: are
 Moved (insertion) [12]
 Deleted: .
 <b>Moved down [13]:</b> The recommended bias correction presented here will likely improve such estimates and analysis for the philipping.

Moved (insertion) [13]

#### 615

### Acknowledgements

We thank the Republic of the Philippines Department of Science and Technology Philippine Council for Industry, Energy, and Emerging Technology Research and Development (DOST-PCIEERD) Balik Scientist program for facilitating collaboration between the authors in 2019. This project was partially supported through the DOST-PCIEERD project entitled, Catchment Susceptibility to Hydrometeorological Events supporting Tolentino and David, and a DOST-PCIEERD Balik Scientist award to Ibarra. Ibarra is funded by the UC Berkeley Miller Institute and UC President's Postdoctoral Fellowship. We acknowledge the help of Mart Geronia, Allen Marvin Yutuc, Iris Joy Gonzales, Jia Domingo and EJ Terciano Jr. for help with dataset compilation, and Tyler Kukla for discussing the statistical methods. We thank DPWH-BRS, GISM and GRDC for datasets availability. Finally, we thank Jaivime Evaristo and an anonymous referee for thorough and helpful reviews and edits, the editor, Wouter Buytaert, for handling our paper, and Gionata Ghiggi for ensuring accurate

representation of the GRUN product and comments on the manuscript.

630

Data availability

640 Data was compiled from the DPWH-BRS, GISM and GRDC datasets (see links in text) and is made available as a supplemental file. The supplemental file "PhilippinesRiverDischarge\_Ibarra\_HESS.xlsx" contains individual tabs for each catchment in the same order as Figures A1 and A2. The runoff time series start with January of the first year that data is available. Blank cells indicate no measurement. The first three rows include the major river basin, the river name and the location of the station. Further metadata including location (latitude and longitude), area and data source can be found in Table 1.

## Author contributions

DEI and CPCD designed the study, DEI and PLMT carried out the dataset compilation and screening, PLMT verified catchment areas, DEI and PLMT carried out the analysis, and DEI and CPCD prepared the manuscript with contributions from PLMT.

### **Competing interests**

The authors declare that they have no conflict of interest.

#### References

655 Abon, C. C., David, C. P. C., and Bellejera, N. E. B.: Reconstructing the Tropical Storm Ketsana flood event in Marikina River, Philippines, Hydrol. Earth Syst. Sci., 15, 1283–1289, doi:10.5194/hess-15-1283-2011, 2011.

 Abon, C. C., Kneis, D., Crisologo, I., Bronstert, A., David, C. P. C., Heistermann, M.; Evaluating the potential of radarbased rainfall estimates for streamflow and flood simulations in the Philippines, Geomat. Nat. Haz. Risk., 7, 1390-1405, doi:10.1080/19475705.2015.1058862, 2016.

Alfieri, L., Lorini, V., Hirpa, F. A., Harrigan, S., Zsoter, E., Prudhomme, C., and Salamon, P.: A global streamflow reanalysis for 1980–2018. Journal of Hydrology X, 6, 100049, doi: 10.1016/j.hydroa.2019.100049, 2020.

665 Criss, R. E. and Winston, W. E.: Do Nash values have value? Discussion and alternate proposals. Hydrol. Process., 22(14), 2723-2725, doi:10.1002/hyp.7072, 2008.

Dai, A., and Trenberth, K. E.: Estimates of Freshwater Discharge from Conntinents: Latitudinal and Seasonal Variations, Journal of Hydrometeorology, 3 (6), 660-687, 2002.

670

David, C. P., Cruz, R. V. O., Pulhin, J. M., and Uy, N. M., Freshwater Resources and Their Management, In, Philippine Climate Change Assessment Report WG2: Impacts, Vulnerabilities and Adaptation, OML Foundation, 34-54, 2017.

 Davie, J.C., Falloon, P.D., Kahana, R., Dankers, R., Betts, R., Portmann, F.T., Wisser, D., Clark, D.B., Ito, A., Masaki, Y.
 and Nishina, K.: Comparing projections of future changes in runoff from hydrological and biome models in ISI-MIP, Earth System Dynamics, 4(2), 359-374, doi: 10.5194/esd-4-359-2013, 2013

Do, H. X., Gudmundsson, L., Leonard, M., and Westra, S.: The Global Streamflow Indices and Metadata Archive (GSIM) – Part 1: The production of a daily streamflow archive and metadata, Earth Syst. Sci. Data, 10, 765–785, https://doi.org/10.5194/essd10-765-2018, 2018.

Evaristo, J., and McDonnell, J. J.: Global analysis of streamflow response to forest management: Nature, 570, 455-461, doi:10.1038/s41586-019-1306-0, 2019.

685 Ferguson, R. I.: River Loads Underestimated by Rating Curves, Water Resources Research, 22, 74-76, doi; 10.1029/WR022i001p00074, 1986.

Ghiggi, G., Humphrey, V., Seneviratne, S.I., Gudmundsson, L. 2019. GRUN: An observations-based global gridded runoff dataset from 1902 to 2014. Earth Sys. Sci., 11, 1655-1674, doi:10.5194/essd-11-1655-2019, 2019. Dataset analysed:
https://figshare.com/articles/GRUN Global Runoff Reconstruction/9228176

Gudmundsson, L., Do, H. X., Leonard, M., and Westra, S.: The Global Streamflow Indices and Metadata Archive (GSIM) – Part 2: Quality control, time-series indices and homogeneity assessment, Earth Syst. Sci. Data, 10, 787–804, https://doi.org/10.5194/essd-10-787-2018, 2018.

695

Hagemann, S., Chen C., Haerter, J. O., Heinke, J., Gerten, D., and Piani, C.: Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models, Journal of Hydrometerology, 12 (4), 556-578, doi: 10.1175/2011JHM1336.1, 2011.

700 Harrigan, S., Zsoter, E., Alfieri, L., Prudhomme, C., Salamon, P., Wetterhall, F., Barnard, C., Cloke, H., and Pappenberger, F.: GloFAS-ERA5 operational global river discharge reanalysis 1979-present. Earth System Science Data, doi: 10.5194/essd-2019-232, 2020

Jose, A. M., and Cruz, N. A.: Climate change impacts and responses in the Philippines: water resources, Clim Res., 12 (2-3), 705 77–84, 1999.

Kim, H., Watanabe, S., Chang, E. C., Yoshimura, K., Hirabayashi, J., Famiglietti, J., and Oki, T.: Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1) [Data set], Data Integration and Analysis System (DIAS), doi:10.20783/DIAS.501, 2017.

710

Kintanar R. L.: Climate of the Philippines, PASAGA report, 38, 1984.

Kumar, P., Masago, Y., Mishra, B. K., and fukushi, K., Evaluating future stress due to combined effects of climate change and rapid urbanization for Pasig Marikina river, Manila, Groundwater Sustain. Develop., 6, 227-34, doi:10.1016/j.gsd.2018.01.004, 2018.

Kummu, M., Guillaume, J. H. A., de Moel, H., Eisner, S., Flörke, M., Porkka, M., Siebert, S., Veldkamp, T. I. E., and Ward, P. J.: The world's road to water scarcity: shortage and stress in the 20th century and pathways towards sustainability, Nat. Sci. Rep., 6, 38495, https://doi.org/10.1038/srep38495, 2016.

720

Meybeck, M., Kummu, M, and Dürr, H. H.: Global hydrobelts and hydroregions: improved reporting scale for water-related issues, Hydrol. Earth Syst. Sci., 17, 1093–1111, doi: 10.5194/hess-17-1093-2013, 2013.

Mayor, A. G., Bautista, S., and Bellot, J.: Scale-dependent variation in runoff and sediment yield in a semiarid 725 Mediterranean catchment, J. Hydrol., 397 (1-2), 128-135, doi:10.1016/j.jydrol.2010.11.039, 2010.

Merz, R., Parajka, J., and Bloschl, G.: Time stability of catchment model parameters: Implications for climate impact analyses, Water Resources Research, 47 (2), doi: 10.1029/2010WR009505, 2011.

730 Mulligan, M.: WaterWorld: a self-parameterising, physically based model for application in data-poor but problem-rich environments globally, Hydrology Research, 44 (5), 748-769, doi: 10.2166/nh.2012.217, 2013.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282-290, doi:10.1016/0022-1694(70)90255-6, 1970.

735

Nurse, L. A., McLean, R. F., Agard, J. Briguglio, L.P., Duvat-Magnan, V., Pelesikoti, N., Tompkins, E. and Webb, A.: Small islands. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Barros, V.R., et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, Ch. 29, 1613-1654, 2014.

740

Paronda, G. R. A., David, C. P. C., and Apodaca, D. C.: River flow patterns and heavy metals concentrations in Pasig River, Philippines as affected by varying seasons and astronomical tides, IOP Conf. Series: Earth and Environmental Science 344, 012049, doi: 10.1088/1755-1315/344/1/012049, 2019.

- 745 Rodríguez-Caballero, E., Cantón, Y., Lazaro, R., and Sole-Benet, A.: Cross-scale interactions between surface components and rainfall properties. Non-linearities in the hydrological and erosive behavior of semiarid catchments. Journal of Hydrology 517, 815–825, doi: 10.1016/j.jhydrol.2014.06.018, 2014.
- 750 Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., and Rudolf, B.: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theoretical and Applied Climatology 115, 15-40, doi: 10.1007/s00704-013-0860-x, 2014.

 Tarasova, L., Basso, S., Zink, M. and Merz, R.: Exploring Controls on Rainfall-Runoff Events: 1. Time Series-Based Event Separation and Temporal Dynamics of Event Runoff Response in Germany, Water Resources Research, 54 (10), 7711-7732, doi:10.1029/2018WR022587, 2018.

Tolentino, P. L. M., poortinga, A., Kanamaru, H., Keesstra, S., Maroulis, J., David, C. P. C., Ritsema, C. J.: Projected Impact of Climate Change on Hydrological Regimes in the Philippines, PLoS ONE, 11(10), e0163941, doi: 10.1371/journal.pone.0163941, 2016.

760

Wanders, N. and Wada, Y.: Decadal predictability of river discharge with climate oscillations over the 20th and early 21st century, Geophys. Res. Lett., 42, 10689–10695, https://doi.org/10.1002/2015GL066929, 2015.

Winsemius, H.C., Aerts, J.C., Van Beek, L.P., Bierkens, M.F., Bouwman, A., Jongman, B., Kwadijk, J.C., Ligtvoet, W.,
 Lucas, P.L., Van Vuuren, D.P. and Ward, P.J.: Global drivers of future river flood risk. Nature Climate Change, 6(4), 381-385, doi: 10.1038/nclimate2893, 2016.

WEF: The Global Risks Report 2018, available at: http://reports. weforum.org/global-risks-2018/.

770 Zhang, Q., Liu, J., Yu, X. and chen, L.: Scale effects on runoff and a decomposition analysis of the main driving factors in the Haihe Basin mountainous area. Sci. Tot. Env., 690, 1089-1099, doi: 10.1016/j.scitotenv.2019.06.540, 2019

Table 1: List of stations	used in this analysis including full station names, upd	lated catchment	areas, years of c	overage, divisio	n and climate ty	pe.		
					Years of	Catchment		
River Name	Station Name	Latitude	Longitude	Coverage	Coverage	Area (km²)	Dataset	Climate Type
Sinalang River	Penarrubia, Abra	17.622	120.715	1984-2015	32	136.128	BRS	1
Antequera River	Sto. Rosario, Antequera, Bohol	9.493	123.890	1984-2016	33	54	BRS	4
Amparo River	Brgy. Mabini, Macrohon, So. Leyte	10.042	124.018	1985-2007	23	74	BRS	4
Hira-an River	Upper Hiraan, Rarigara, Leyte	11.258	124.672	1986-2010	27	8.93	BRS	4
Leyte River	San Joaquin, Capocan, Leyte	11.880	124.829	1985-2007	23	29.15	BRS	3
Surigao River	Surigao City	9.796	125.808	1986-2010	25	85	BRS	4
Bais River	Cabanlutan, Bais City, Negros Oriental	9.876	124.140	1989-2015	27	41	BRS	2
Lingayaon River	Lingayon, Alang-Alang, Leyte	11.192	124.863	1957-1991	35	18	BRS	4
Sapiniton River	Libton, San Miguel, Leyte	11.188	124.795	1984-2010	27	277.3	BRS	4
Laoag River	Poblacion, Laoag City, Ilocos Norte	18.203	120.590	1984-2016	33	1355	BRS; updated from Tolentino et al. (2016)	1
Pared River	Baybayog, Alcala, Cagayan	17.682	121.270	1985-1996	14	966	BRS; used in Tolentino et al. (2016)	1
Ganano River	Ipil, Echague, Isabela	16.812	121.211	1986-2001	16	9//	BRS; used in Tolentino et al. (2016)	3
Magat River	Baretbel, Bagabag, Nueva Vizcaya	16.992	121.073	1986-2002	17	2199	BRS; used in Tolentino et al. (2016)	3
Caming River	Poblacion, Mayanioc, Tariac	13.018	120.303	1985-2017	33	288	BRS; updated from Totentino et al. (2016)	1
Dia China Dimo	Sta. Cruz, Lubao, Pampanga	14.900	120.441	1985-2001	22	570	BRS; used in Tolentino et al. (2016)	1
Son Juan Pittar	Sto. Kosario, Zaragosa, Nueva Ecija Bonzo, Colombo, Loguno,	13.038	120.088	1986-1999	14	165	BRS; used in Tolentino et al. (2016) BRS: used in Tolentino et al. (2016)	1
Dan salara Disas	Porecher Disconducer Orientel Mindeer	19.976	121.207	1080 1000	14	20	DRG, used in Tolentino et al. (2010)	-
Pangaiaan River	Sta Nino II Hinungangan Lauta	10.295	121.260	1987-2007	21	50	BRS; used in Tolentino et al. (2016) BRS: used in Tolentino et al. (2016)	3
Tuloren River	Tinotonoon, Tuluungangan, Leyte	7 627	122.020	1986-2009	24	147	BRS, used in Tolentino et al. (2016) BRS: used in Tolentino et al. (2016)	2
Hijo River	Anokan Tagum Dayao del Norte	7.812	125.050	1986-2016	31	634	BRS: undated from Tolentino et al. (2016)	4
Coorner do Oro Pissor	Cabula Cagavan da Ora City Micamis Oriantal	9 216	124.811	1991-2004	14	1079	BRS: usad in Tolentino et al. (2016)	4
Davao River	Tiestto Davao City	7 320	125.634	1984-1999	14	1683	BRS: used in Tolentino et al. (2016)	4
Allah River	Immao Jaulan Sultan Kudarat	6 568	124.085	1980-1994	15	1231	BRS: used in Tolentino et al. (2016)	3
Anucan Canyon River	Camp Philing Manolo Fortich Bukidnon	8 296	124.665	1986-2004	19	48	BRS: undated from Tolentino et al. (2016)	3
Wawa River	Wawa Bayaoan Aousan del Sur	8 261	125 501	1981-2010	30	396	BRS: undated from Tolentino et al. (2016)	4
Buayan River	Malandag Malungon South Cotabato	6 317	125.749	1986-2004	19	207	BRS: used in Tolentino et al. (2016)	4
Gaseas River	Manalnac Solsona	18 080	120.830	1978-1988	11	73	GISM	1
Jalaur River	Calvan Pototan Iloilo	10.930	122.670	1976-1988	13	1499	GISM and GRDC	3
Padsan River	Baneay	18 080	120 700	1946-1979	34	534	GRDC	1
Pamnanea River	San Agustin	15 170	120.780	1946-1977	32	6487	GRDC	1
Sinocot River	Sabano	13.810	122,990	1946-1970	25	447	GRDC	2
Mambusao River	Tumalalud	11.260	122.570	1950-1978	29	307	GRDC	3
Padada River	Lapulabao	6.660	125.280	1949-1978	30	821	GRDC	4
Aloran River	Juan Bacay, Aloran, Misamis Occ.	8.420	123.820	1978-2003	26	30	GRDC + BRS	3
Cabacanan River	Baduane, Pagudpud	18,580	120.800	1979-2017	39	60	GRDC + BRS	1
Maragayap River	Sta. Rita, Bacnotan, La Union	16.750	120.374	2004-2017	14	40	BRS	1
Abacan River	San Juan, Mexico, Pampanga	15.118	120.703	2004-2017	14	217	BRS	1
Hibayog River	La Victoria, Carmen, Bohol	9.876	124.141	2004-2017	14	41	BRS	4
Manaba River	Calma, Garcia-Hernandez, Bohol	9.631	124.131	2001-2016	16	98	BRS	4
Gabayan River	Canawa, Candijay, Bohol	9.848	124.450	2001-2017	17	48.5	BRS	4
				1984-1990;				
Bangkerohan River	Brgy. Tagaytay, Bato, Leyte	10.342	124.834	2000-2009	17	168	BRS	4
Borongan River	Brgy. San Mateo, Borongan City	11.628	125.403	1990-2008	19	111	BRS	2
Loom River	Brgy. Calico-an, Borongan City	10.594	125.404	1986-2004	19	42	BRS	2
Pagbanganan River	Brgy. Makinhas, Baybay City	10.637	124.865	1984-2008	25	128	BRS	4
Rizal River	Brgy. Rizal, Babatngon, Leyte	11.389	124.908	1990-2008	18	15	BRS	4
Tenani River	Brgy. Tenani, Paranas (Wright), Samar	11.806	125.127	1985-2001	17	394	BRS	2
Dischar Disco	Disks Maska Zaskana dd V	0.400	122.046	1985-1991;		100	DBC	2
Disakan River	Disakan, Manukan, Zamboanga del Norte	8.480	123.048	2002 2011	11	109	BRS	3
Kapasalan River	Danker, Kabasalan, Sibugay, Province	/.851	122.778	1000-2001	10	145	DRS	3
Sindangan River	Dicoyong, Sindangan, Zamboanga del Norte	8.217	123.057	1990-2003	14	590.5	BKS	3
Aubijid River	Alumija, Misamis Oriental Thumis Koodow Down del Norte	8.570	124.476	2004 2017	19	124	DRS	3
Ripaliko River	rioucia, Kapaiong, Davao dei None	7.002	123.081	1087-1005	15	147	DRJ	-4
Banaue River	Poblacion, Banaue, Ifugao	16.915	121.061	2005-2010	15	15	BRS	3
Aciga River	Santiago, Agusan del Norte	9.269	125.570	2002-2015	14	80	BRS	4
				1982; 1984-				
				1987; 1989-				
Agusan River	Sta. Josefa, Agusan del Sur	7.993	126.036	2010	27	1633	BRS	4

Table 2. Results of Annual agreement between OKON	aggregated by Children	Type and for the chine	united (see Those Ser It	Cinter Calciments	c steams monthly tank	1 Tanks ac 3.23, 7.03,	4.77 and 5.51 million	Nash-Sutcliff	,
				Nash-Sutcliff		Volumetric	Nash-Sutcliff	Efficiency (NSE-	Root Mean Square
		Volumetric	Nash-Sutcliff	Efficiency (NSE-	Root Mean Square	Efficiency (VE) Bias	Efficiency (NSE)	log10) Bias	Error Bias Corrected
	Pearsons Coeff (r2)*	Efficiency (VE)	Efficiency (NSE)	log10)	Error	Corrected **	Bias Corrected **	Corrected **	**
Entire Dataset	0.372	0.363	0.091	0.453	2.648	0.323	0.182	0.385	0.292
Entire Dataset log10(runoff)	0.546	0.733	n/a	n/a	n/a	1.067	n/a	n/a	n/a
Climate Type 1 (n=12)	0.211	0.252	0.062	0.538	2.476	0.168	0.111	0.432	0.298
Climate Type 2 (n=6)	0.409	0.354	0.05	0.49	4.554	0.349	0.188	0.457	0.544
Climate Type 3 (n=15)	0.471	0.404	0.026	0.23	2.285	0.345	0.011	0.188	0.432
Climate Type 4 (n=22)	0.535	0.403	0.159	0.414	2.398	0.377	0.323	0.36	0.131
Entire Dataset, Individual Catchment Median	0.658	0.404	0.02	0.239	2.072	0.322	-0.158	0.051	1.840
Entire Dataset, Individual Catchment Range	0.123 to 0.859	-2.079 to 0.633	-10.697 to 0.678	-11.525 to 0.761	0.038 to 14.681	-5.082 to 0.579	-53.842 to 0.650	-10.534 to 0.779	0.035 to 12.602
Climate Type 1 (n=12), Individual Catchment Median	0.444	0.236	0.112	0.485	2.232	0.258	0.027	0.330	2.303
Climate Type 1 (n=12), Individual Catchment Range	0.123 to 0.833	-0.782 to 0.567	-1.481 to 0.567	-1.472 to 0.761	0.513 to 7.684	-2.316 to 2.316	-6.917 to 0.389	-2.172 to 0.779	0.325 to 5.448
Climate Type 2 (n=6), Individual Catchment Median	0.574	0.348	-0.021	0.312	4.319	0.339	0.007	0.158	3.742
Climate Type 2 (n=6), Individual Catchment Range	0.329 to 0.760	-0.322 to 0.633	-0.555 to 0.509	-0.338 to 0.696	0.395 to 13.389	-1.083 to 0.533	-3.937 to 0.516	-0.346 to 0.646	1.783 to 9.622
Climate Type 3 (n=15), Individual Catchment Median	0.697	0.443	-0.069	0.163	1.491	0.360	-0.614	-0.002	0.911
Climate Type 3 (n=15), Individual Catchment Range	0.495 to 0.855	0.155 to 0.612	-0.835 to 0.678	-4.577 to 0.753	0.057 to 14.681	-0.478 to 0.540	-3.079 to 0.126	-3.921 to 0.557	0.035 to 12.602
Climate Type 4 (n=22), Individual Catchment Median	0.681	0.414	-0.022	0.11	1.959	0.357	-0.064	-0.185	1.588
Climate Type 4 (n=22), Individual Catchment Range	0.261 to 0.859	-2.079 to 0.609	-10.697 to 0.607	-11.525 to 0.640	0.038 to 8.728	-5.082 to 0.579	-53.842 to 0.650	-10.534 to 0.627	0.561 to 5.855

		Volumetric	Nasł
	Pearsons Coeff (r <sup>2</sup> )*	Efficiency (VE)	Effic
Entire Dataset	0.372	0.363	0.093
Entire Dataset log10(runoff)	0.546	0.733	n/a
Climate Type 1 (n=12)	0.211	0.252	0.062
Climate Type 2 (n=6)	0.409	0.354	0.05
Climate Type 3 (n=15)	0.471	0.404	0.02
Climate Type 4 (n=22)	0.535	0.403	0.15
Notes			
* For represeions forced through i	ntercent of 0		

Nets 1 or programming of the set of the set

Page 5: [1] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 5: [2] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
		4/2/21 12:40:00 111
·		
Page 9: [3] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [3] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
		4,2,22 12:10:00 11
·		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
	Daniel Enrique Ibarra	4/2/21 12:45:00 114
•		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
	Daniel Enrique Ibarra	4/2/21 12:45:00 114
·		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
		., _,
*		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
· · · · · · · · · · · · · · · · · · ·		-, _,
•		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
V		-, _,
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
<b>V</b>		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM

**V**...

Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
Χ		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
V	······	· · · · · · · · · · · · · · · · · · ·
D	Devid Fusions theme	4/2/21 12:45:00 PM
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
· · · · · · · · · · · · · · · · · · ·		
Page 9: [4] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
Tuge 5. [5] Deleteu	Daniel Enrique Ibarta	+, 2, 21 12.+5.00 TH
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
<b>V</b>		
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
V	•	· · ·
Dama O. [E] Dalatad	Deniel Fusieure Therma	4/2/21 12:45:00 DM
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
<b>V</b>		
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
		7/2/21 22.73.00 FM
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
▼		
Page 9: [5] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM

Dage () [6] Deleted	Daniel Envigue Thanks	4/2/21 12:45:00 PM
Page 9: [6] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
<i>I</i>		
Page 9: [6] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
Υ		
Page 9: [6] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
۲		
Page 9: [6] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
Υ		
Page 9: [6] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
۲		
Page 9: [6] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
Χ		
Page 9: [6] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
Χ		
Page 9: [7] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
۷		
Page 9: [7] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
۷		
Page 9: [7] Deleted	Daniel Enrique Ibarra	4/2/21 12:45:00 PM
۷		
Page Q. [7] Deleted	Daniel Envigue Theres	4/2/24 42-45-00 DM