Supplement of Hydrol. Earth Syst. Sci., 29, 6181–6200, 2025
https://doi.org/10.5194/hess-29-6181-2025-supplement

Hydrology and Earth System Sciences
Open Access

Supplement of

Integrating historical archives and geospatial data to revise flood estimation equations for Philippine rivers

Trevor B. Hoey et al.

Correspondence to: Pamela Louise M. Tolentino (pammie.tolentino@glasgow.ac.uk, plmtolentino.ac@gmail.com)

SUPPLEMENTARY INFORMATION

Table S1. Q_2 equations for different groups of regions. Meigh (1995) did not include regions 13 or CAR, so the total number of sites in the three groups is 431.

Group	Regions in group	Number of sites	Equation	R^2	se
Growth curve					
A	1,13,CAR	65	$Q_2 = 0.0675(A.RMED)^{0.765}$	0.74	0.287
B	2,3,4A,6,11,12	241	$Q_2 = 0.0218(A.RMED)^{0.836}$	0.68	0.387
C	4B,5,7,10	126	$Q_2 = 0.139(A.RMED)^{0.662}$	0.45	0.443
D	8,9	34	$Q_2 = 0.145(A.RMED)^{0.685}$	0.75	0.203
K-means clustering of regional regression equations					
E	1,6,7,8,11	142	$Q_2 = 0.0222(A.RMED)^{0.851}$	0.73	0.336
F	2,3,4A,CAR	167	$Q_2 = 0.0172(A.RMED)^{0.864}$	0.71	0.391
G	4B,9,10,12,13	103	$Q_2 = 0.148(A.RMED)^{0.659}$	0.59	0.376
H	5	54	$Q_2 = 0.693(A.RMED)^{0.500}$	0.30	0.457
Meigh (1995) contiguous regional groups					
I	1,2	86	$Q_2 = 0.0513(A.RMED)^{0.779}$	0.63	0.369
J	3,4A,4B,5,6,7,8	264	$Q_2 = 0.0583(A.RMED)^{0.759}$	0.61	0.408
K	9,10,11,12	817	$Q_2 = 0.0723(A.RMED)^{0.696}$	0.57	0.342

Table S2. Q_{100} equations for different groups of regions. Meigh (1995) did not include regions 13 or CAR, so the total number of sites in the three groups is 431.

Group	Regions in group	Number of sites	Equation	R^2	se
Growth curve					
A	1,13,CAR	65	$Q_{100} = 0.366(A.RMED)^{0.750}$	0.77	0.255
B	2,3,4A,6,11,12	241	$Q_{100} = 0.248(A.RMED)^{0.751}$	0.59	0.425
C	4B,5,7,10	126	$Q_{100} = 6.026(A.RMED)^{0.447}$	0.27	0.447
D	8,9	34	$Q_{100} = 2.023(A.RMED)^{0.537}$	0.60	0.228
K-means clustering of regional regression equations					
E	1,6,7,8,11	142	$Q_{100} = 0.244(A.RMED)^{0.772}$	0.72	0.312
F	2,3,4A,CAR	167	$Q_{100} = 0.184(A.RMED)^{0.787}$	0.63	0.421
G	4B,9,10,12,13	103	$Q_{100} = 4.198(A.RMED)^{0.477}$	0.41	0.393
H	5	54	$Q_{100} = 24.95(A.RMED)^{0.298}$	0.09	0.520
Meigh (1995) contiguous regional groups					
I	1,2	86	$Q_{100} = 0.317(A.RMED)^{0.756}$	0.58	0.399
J	3,4A,4B,5,6,7,8	264	$Q_{100} = 1.122(A.RMED)^{0.621}$	0.48	0.427
K	9,10,11,12	81	$Q_{100} = 4.111(A.RMED)^{0.446}$	0.33	0.359

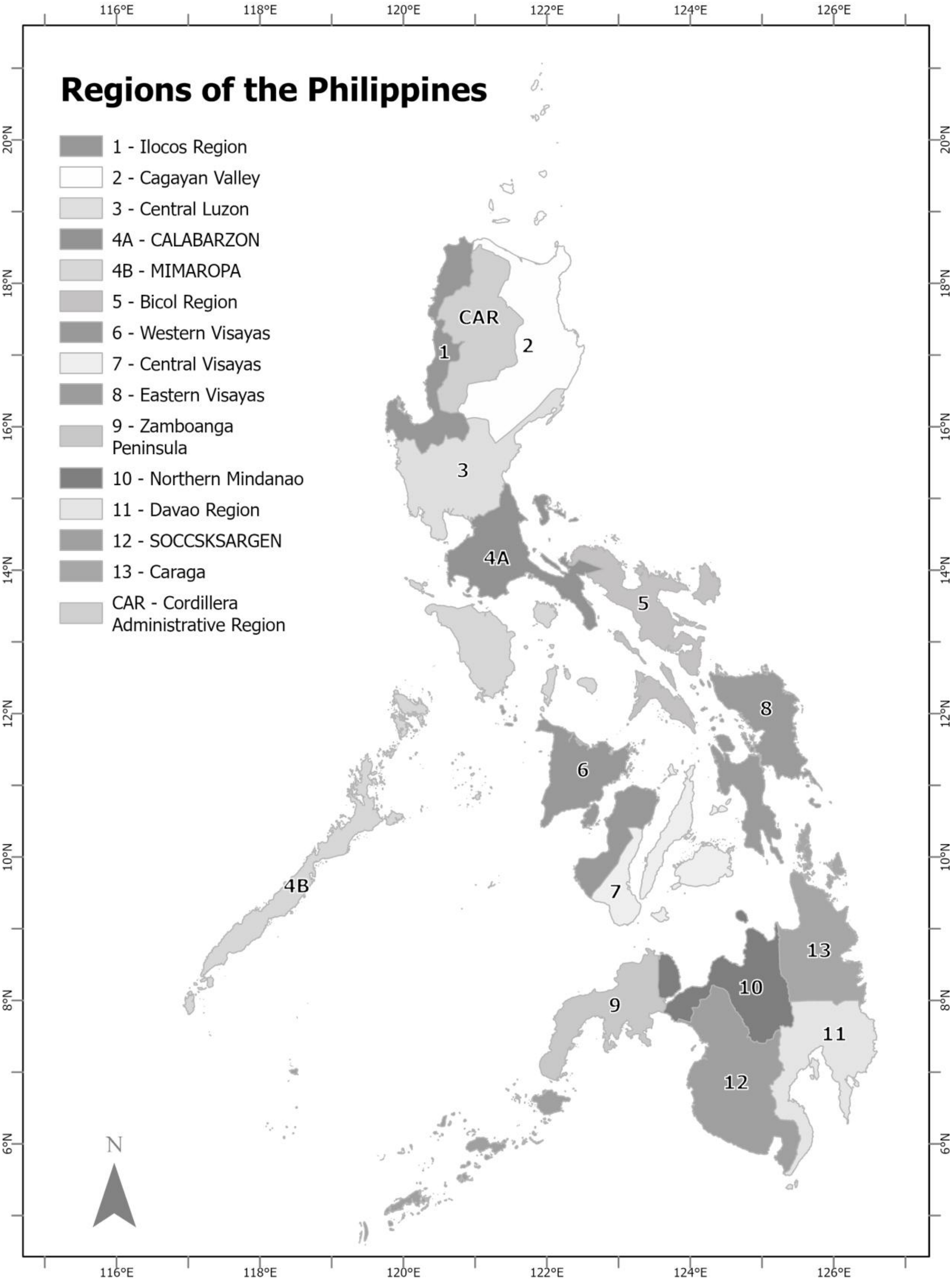


Figure S1. Administrative regions of the Philippines.

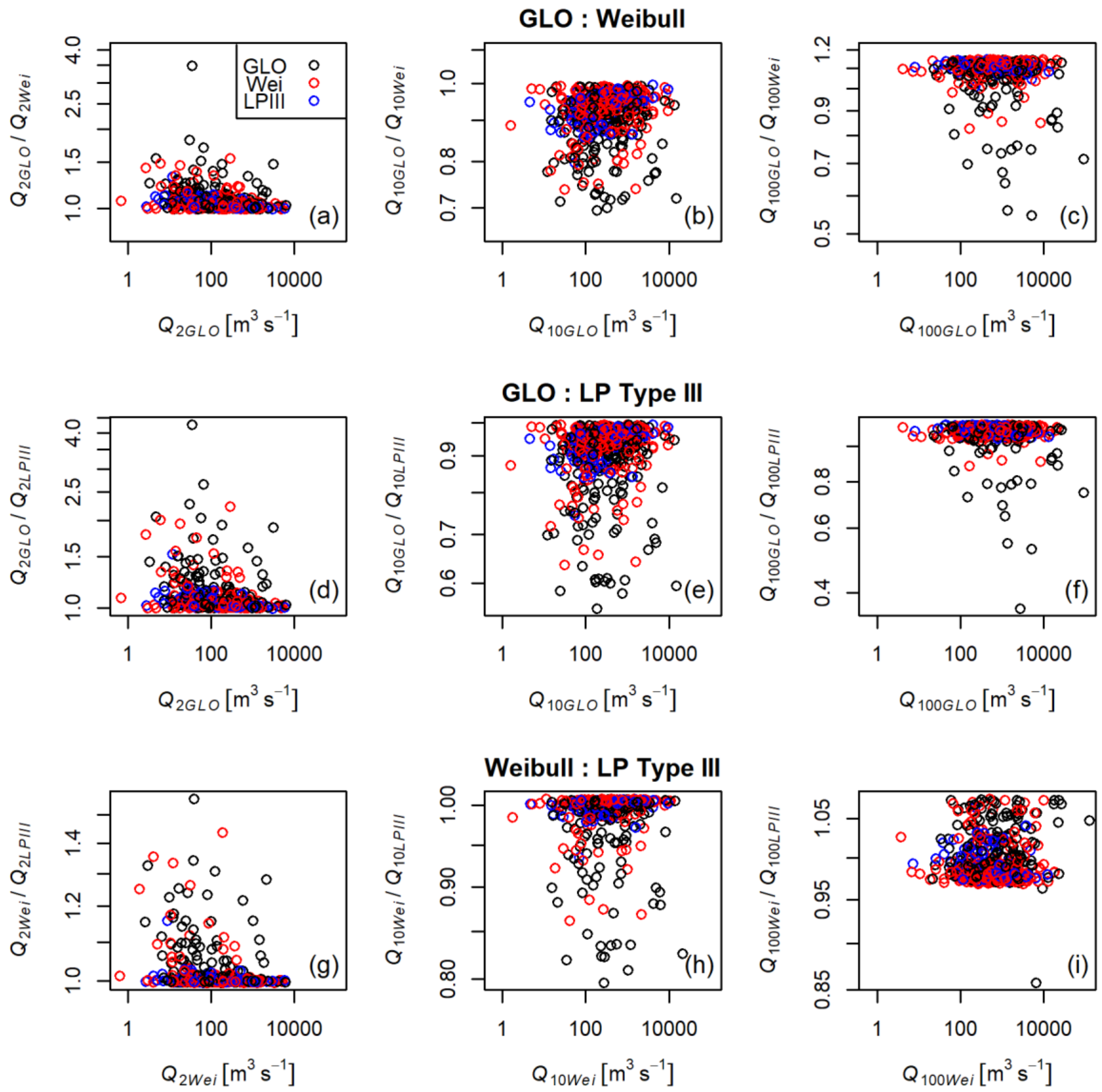


Figure S2. Comparison between flood estimates for 2, 10 and 100 year return periods determined from the three curve fitting methods. GLO = Generalised Logistic Distribution; Wei = Weibull distribution; LPIII = Log-Pearson Type III. x-axes are flood estimates from the GLO (upper two rows) and Weibull distributions, and y-axes are ratios between the estimates obtained from two of the methods, as indicated for each row. Data are colour-coded according to the best-fit curve for each site.

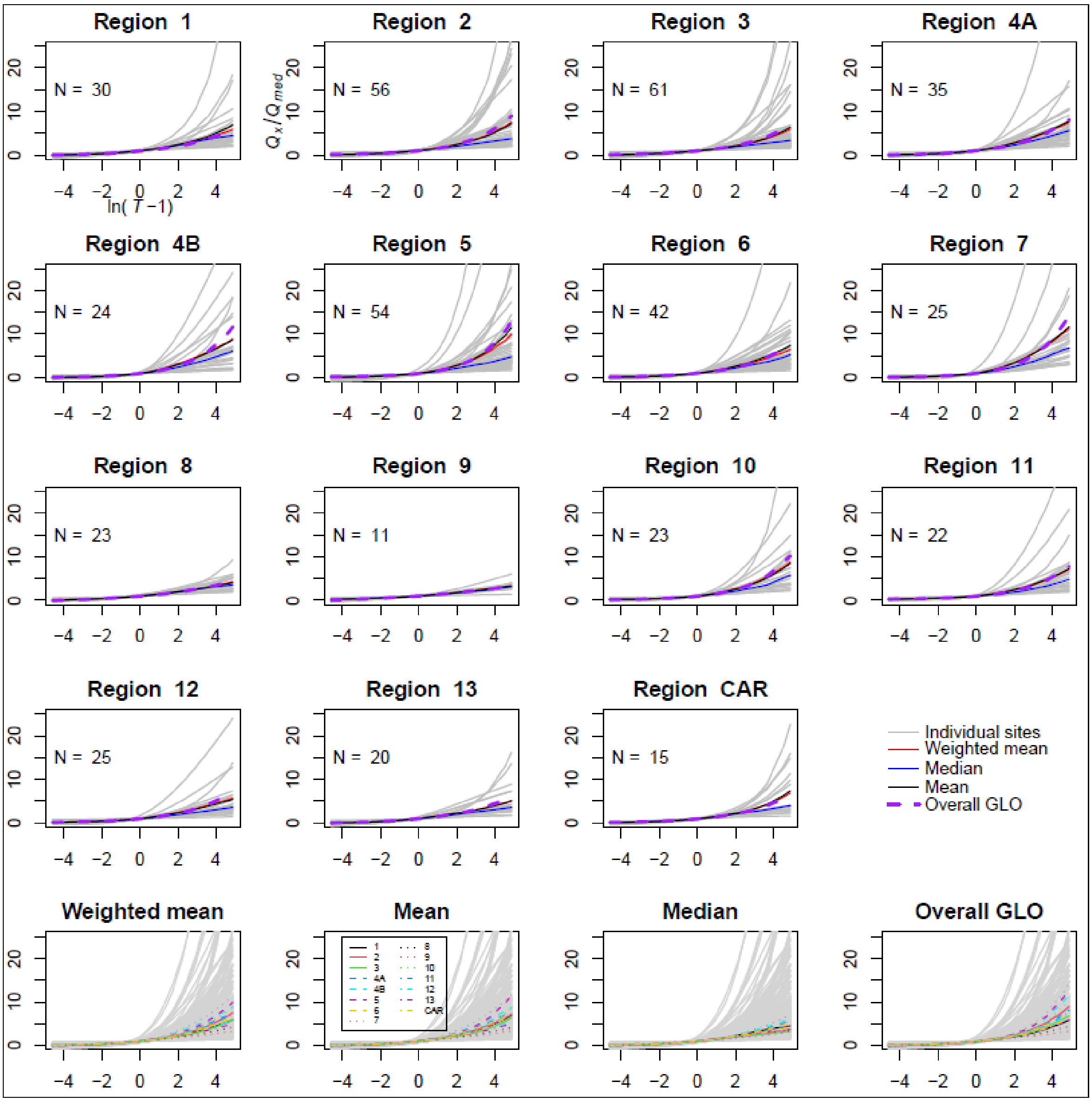


Figure S3. Growth curves for all sites in each region of the Philippines (grey lines) with four methods for calculating a single growth curve for each region. Black line – mean of growth curves for all sites; red line – mean of growth curves for all sites weighted by length of record; blue line – median of growth curves for all sites; and, purple dashed line – GLO curve fitted amalgamated data from all sites within the region. Bottom row shows comparison between the four methods of calculating single growth curves for each region.

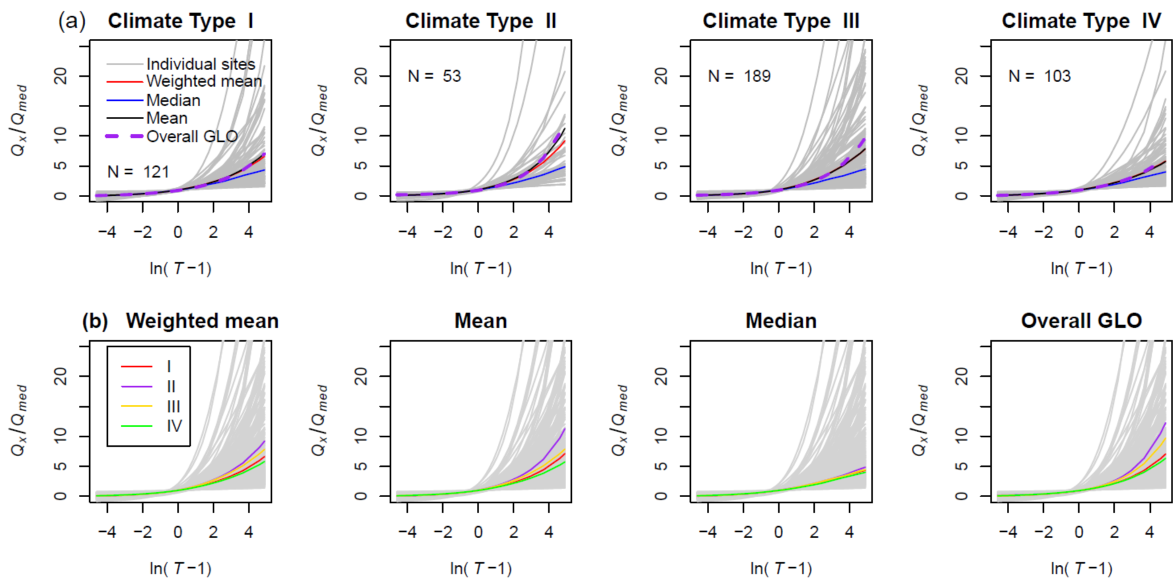


Figure S4. (a) Growth curves for all sites in each climate type of the Philippines (grey lines) with four methods for calculating a single growth curve for each climate type (N is the number of sites of each climate type). Black line – mean of growth curves for all sites; red line – mean of growth curves for all sites weighted by length of record; blue line – median of growth curves for all sites; and, purple dashed line – GLO curve fitted amalgamated data from all sites within each climate type. (b) Comparison between the four methods of calculating single growth curves for catchments within each climate type. Grey lines are growth curves for all individual sites.

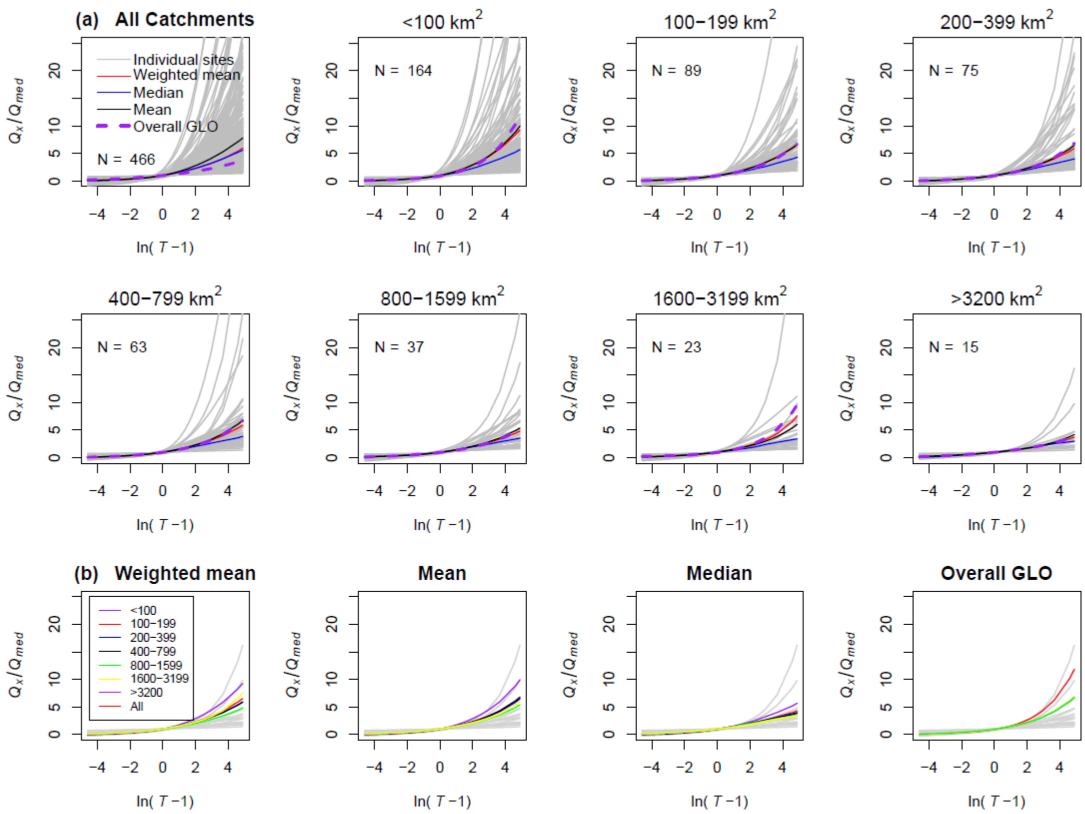


Figure S5. (a) Growth curves for all catchments within catchment area bins indicated (grey lines) with four methods for calculating a single growth curve for each region. Black line – mean of growth curves for all sites; red line – mean of growth curves for all sites weighted by length of record; blue line – median of growth curves for all sites; and, purple dashed line – GLO curve fitted amalgamated data from all sites within each catchment area bin. (b) Comparison between the four methods of calculating single growth curves for within each catchment area bin.

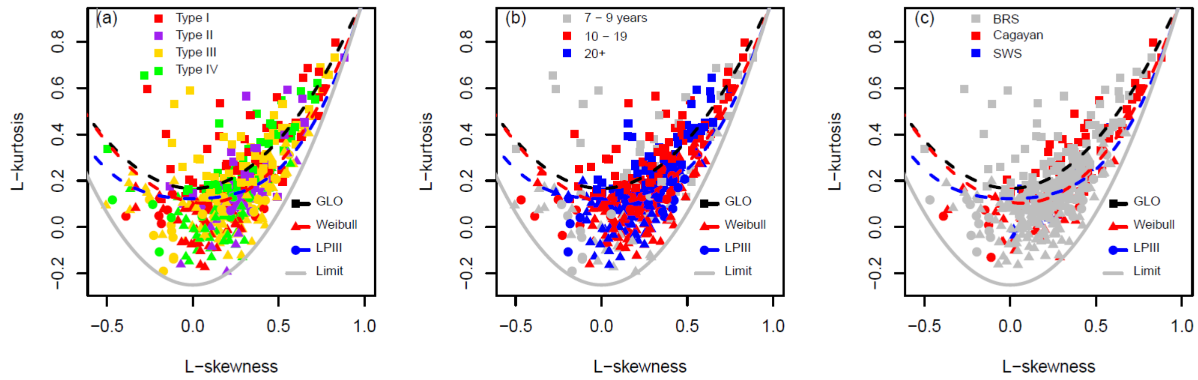
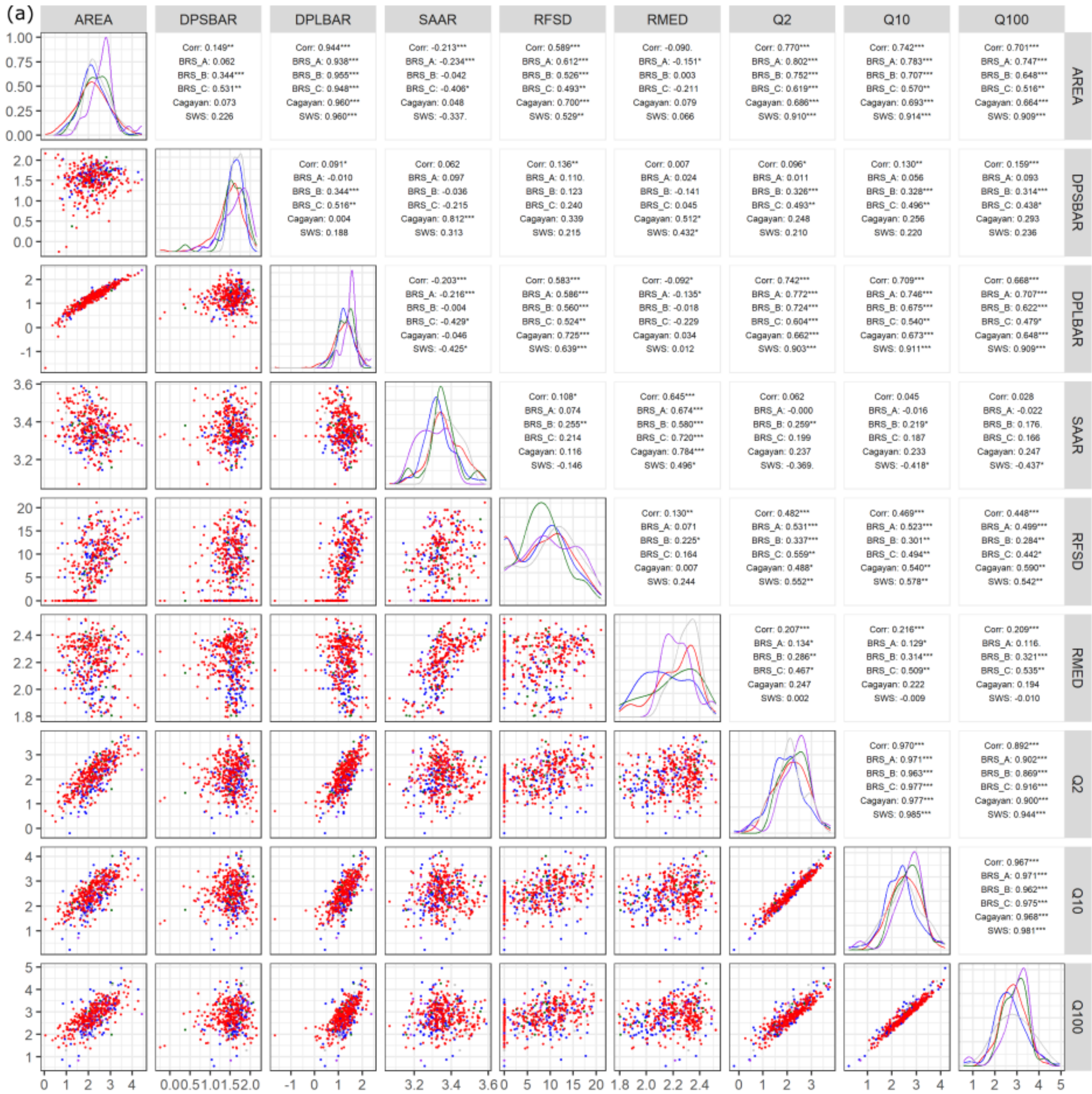
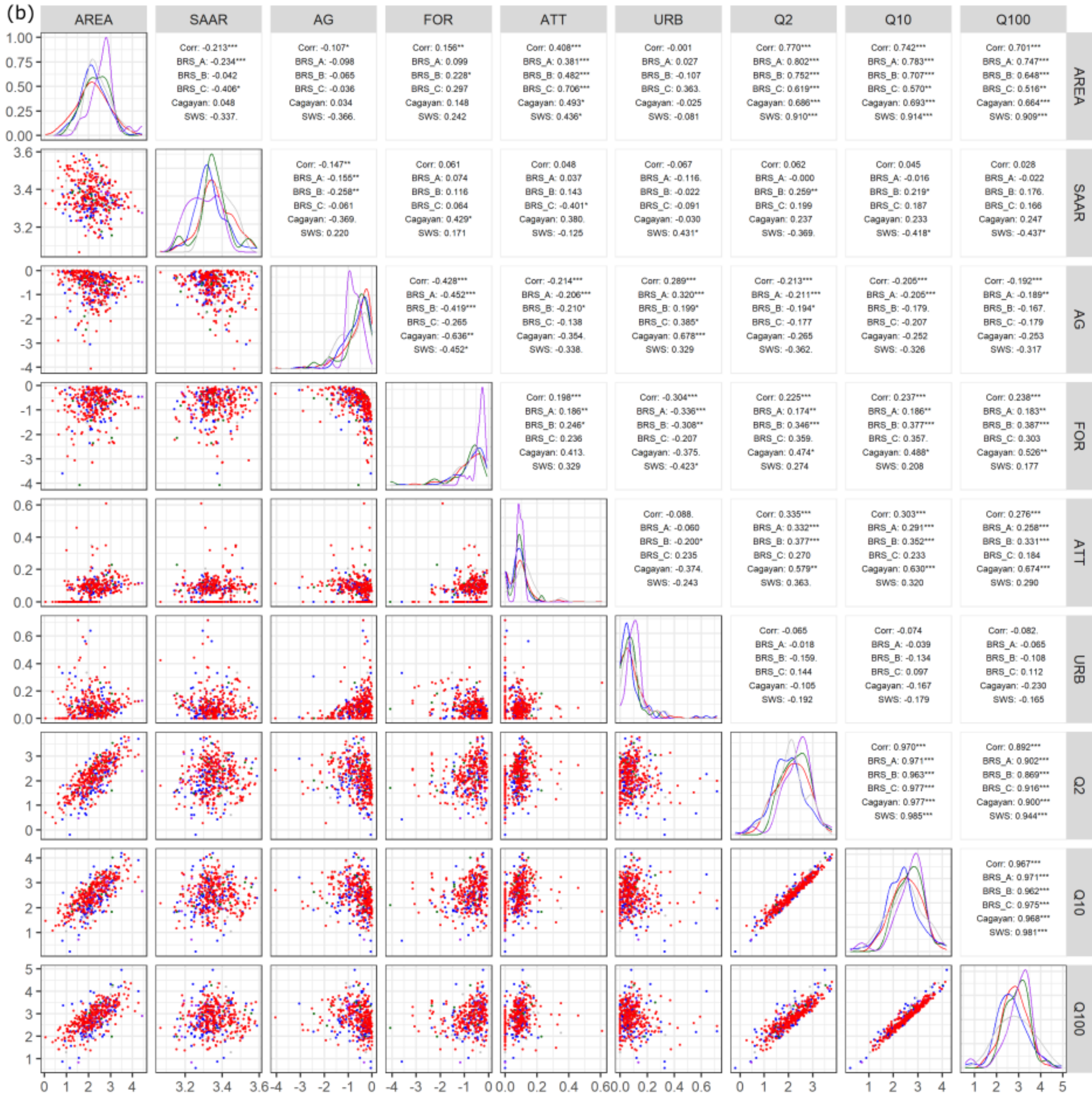


Figure S6. Relationships between L-skewness and L-kurtosis compared with theoretical curves (Hosking and Wallis, 1997). Data classified by: (a) climate zone; (b) length of data record; and, (c) data source. In all cases there is overlap between the best-fit curve type and the classification variable with no obvious clustering of catchments according to climate type, record length or data source. Data points are coloured according to the classification variable, with symbol shape indicating the best-fit curve.


(a)
AREA
DPSBAR
DPLBAR
SAAR
RFSD
RMED
Q2
Q10
Q100



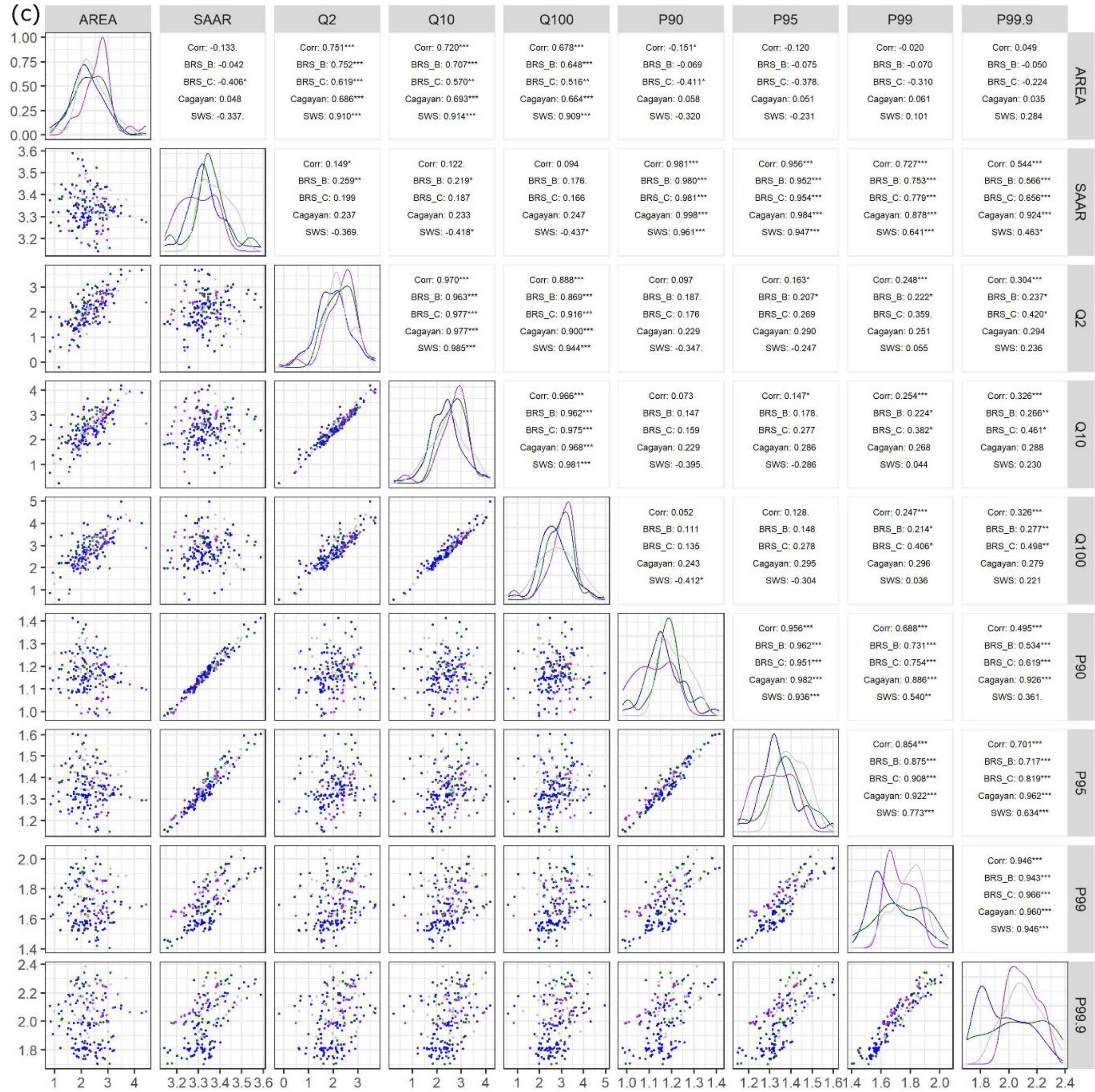


Figure S7. Cross-correlations between all variables used in the analysis. Note that several variables have been transformed prior to plotting, either Log10 (*AREA*, *DPLBAR*, *SAAR*, Q_{MED}, Q_2, Q_{10}, Q_{100}) or square root (*ATT*, *RFSD*, *URB*). See Table 4 for details. (a) topographic variables; (b) land use variables; (c) precipitation descriptors. Points are colour coded by data source: red - BRS_A; blue – BRS_B; green – BRS_C; purple – Cagayan; grey - SWS Plots on the diagonal are probability density functions of the variables. Numbers on the upper right of each figure are correlation coefficients with significance levels indicated (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). On (a), catchment slope (DPSBAR) was converted from $m.m^{-1}$ to $m.km^{-1}$ format before plotting. Rainfall probabilities (P90 to P99.9) were not calculated for the BRS_A sites, hence the absence of red points and lines in (c).

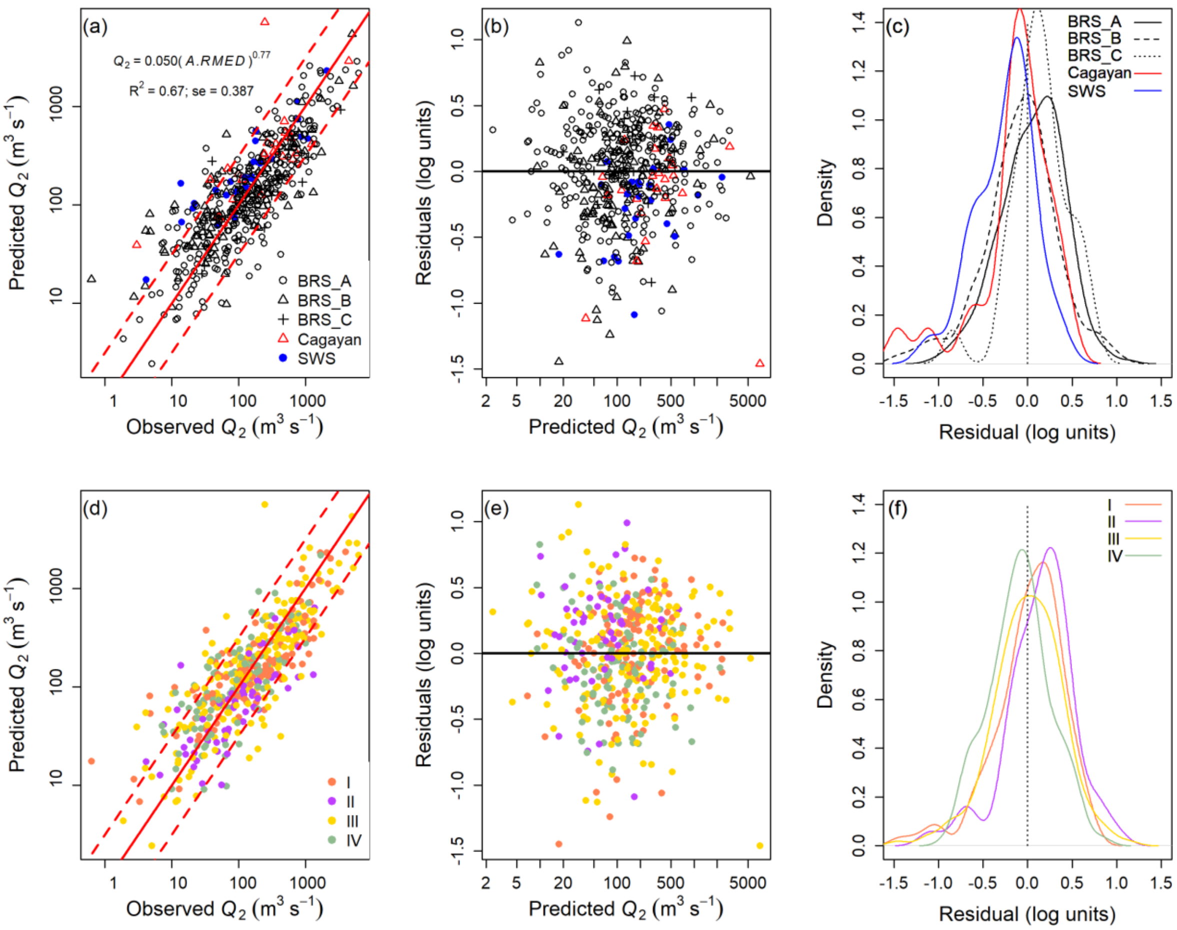


Figure S8. Observed values, prediction and residuals for Q_2 as a function of catchment area (A) multiplied by median daily maximum rainfall ($RMED$). (a)-(c) stratified by data source, (d)-(f) by climate type. (a), (d) are predicted vs. observed values, with 1:1 (solid), 1:2 and 2:1 (dashed) lines shown. Residuals (b) and (e) are normally distributed and show no systematic variation with predicted Q_2. Density plots of residuals (c), (f) confirm the absence of systematic variation with data source and climate type.

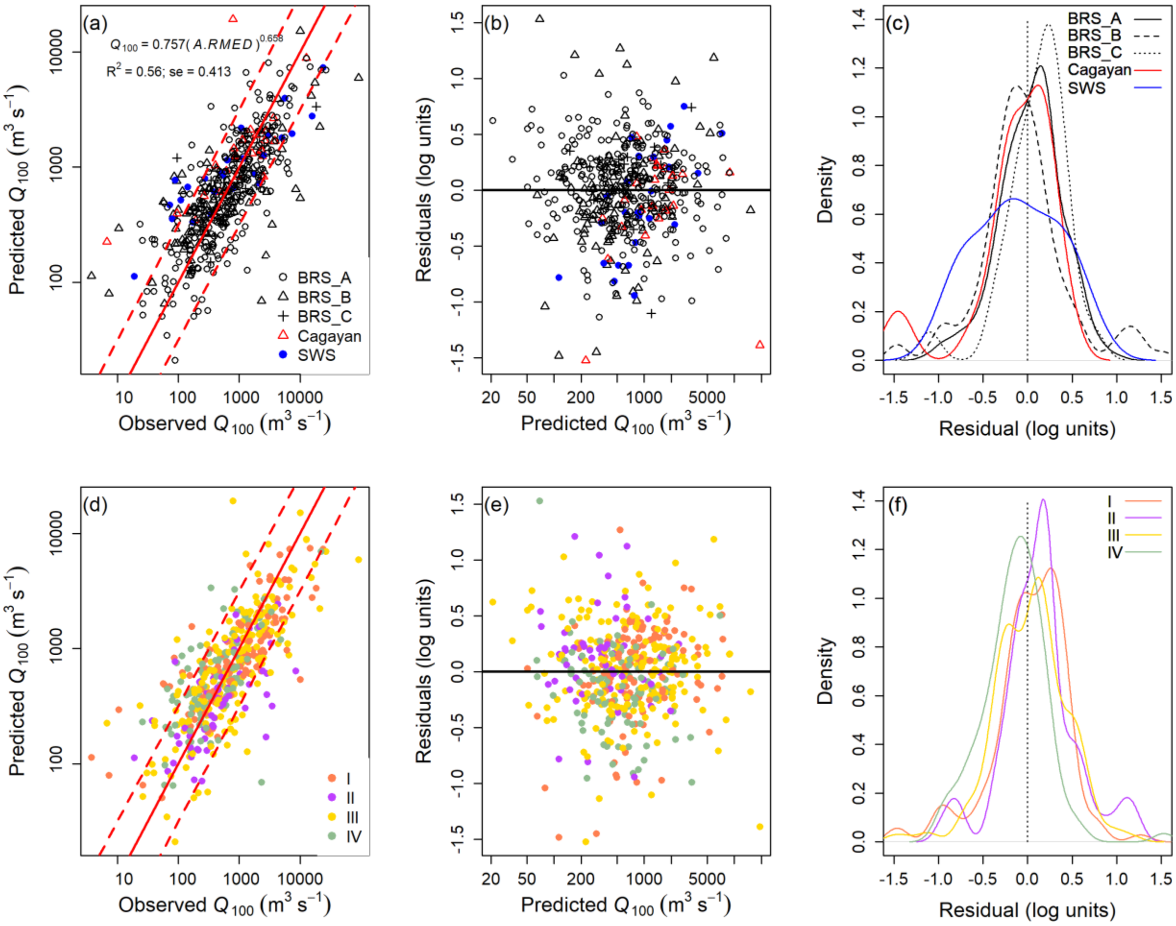


Figure S9. Observed values, prediction and residuals for Q_{100} as a function of catchment area (A) multiplied by median daily maximum rainfall ($RMED$). (a)-(c) stratified by data source, (d)-(f) by climate type. (a), (d) are predicted vs. observed values, with 1:1 (solid), 1:2 and 2:1 (dashed) lines shown. Residuals (b) and (e) are normally distributed and show no systematic variation with predicted Q_{100}. Density plots of residuals (c), (f) confirm the absence of systematic variation with data source and climate type.

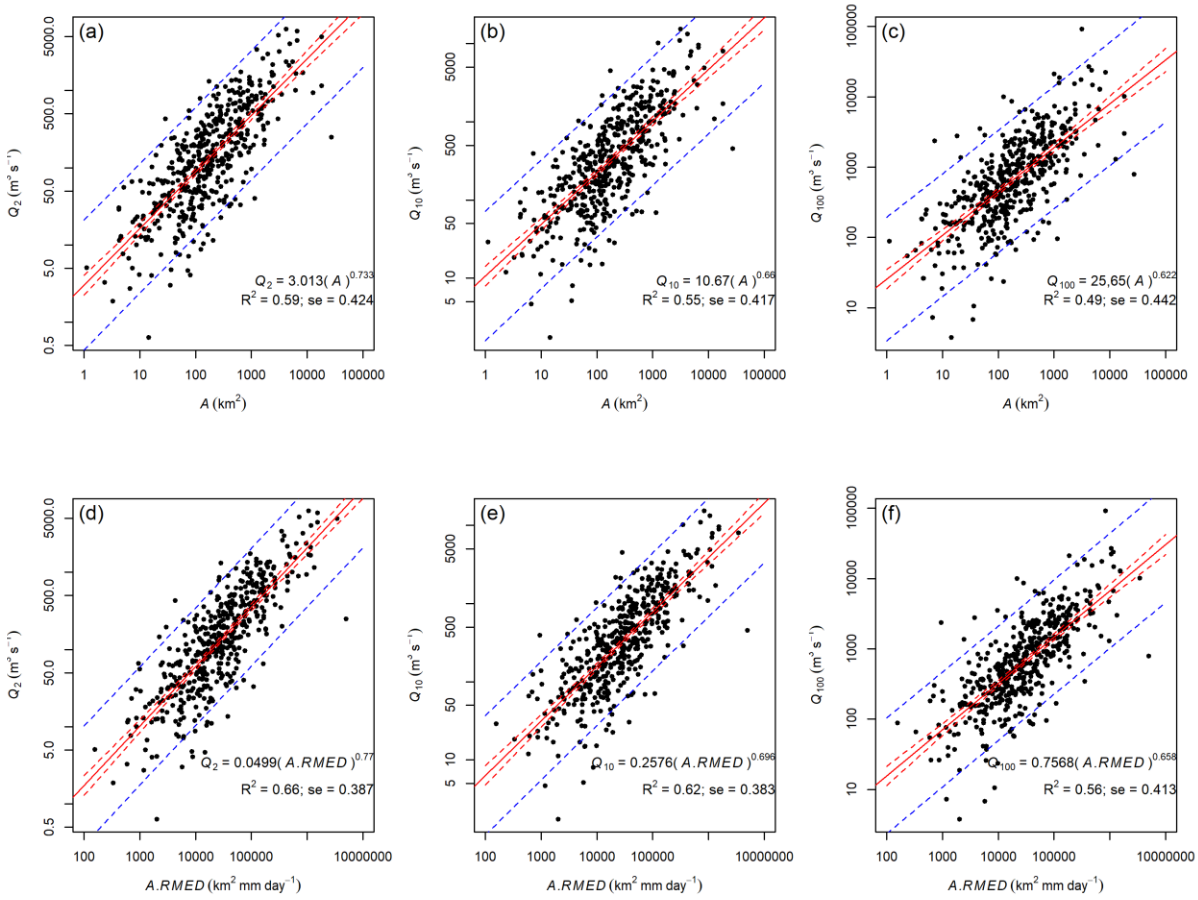


Figure S10. Predictive equations from Table 5 and all 466 data points used in the analysis. Solid red lines are best-fit log-log equations, with 95% confidence (red dashed lines) and prediction (blue dashed lines) intervals shown. (a) – (c) are equations using only catchment area, *A*, and (d)-(f) show results with *A.RMED* as the independent variable.

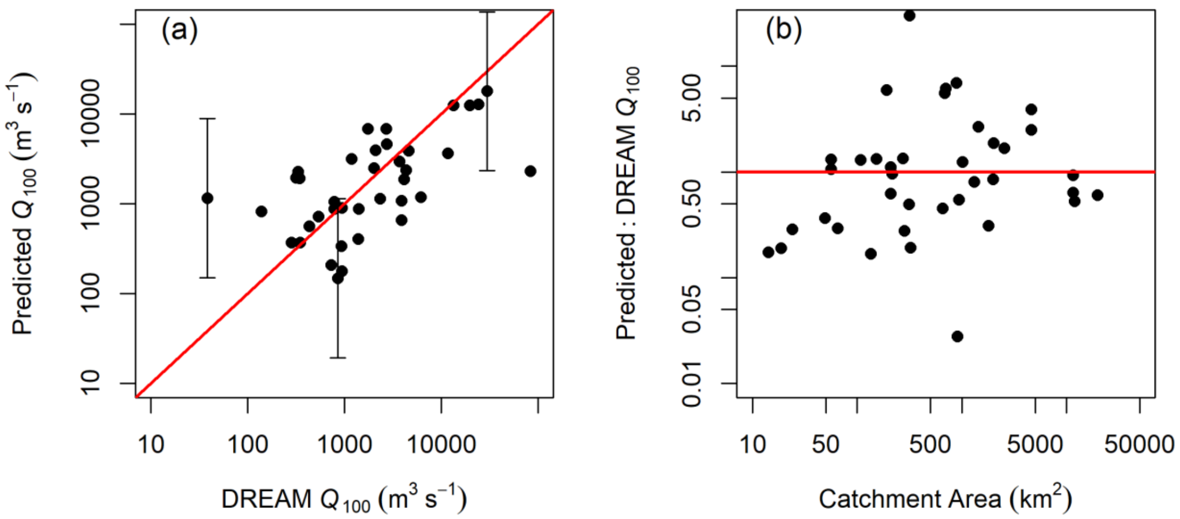


Figure S11. (a) Comparison between Q_{100} estimates based on catchment area (Table 5) and HEC-HMS estimates from the DREAM project. Red line is 1:1 equivalence. (b) Effect of catchment area on the ratio between Q_{100}

values from this paper and the DREAM HEC-HMS modelling. Red line shows equal Q_{100} values from both methods. DREAM estimates are instantaneous peak flows whereas the estimates herein are daily means. As catchment area increases, equivalence between the two methods would show the Q_{100} ratio increasing towards 1.0 as catchment area increases, with lower values in smaller catchments in which flood peaks are of much less than one day duration. 95% prediction intervals are shown for selected points on (a) to indicate the magnitude of statistical uncertainty in the predictions. These are approximated as ± 2s.e., where s.e. is the regression standard error given in Table 5.