Supporting Information for

Advancing Regionalization of IDF Curves in Mainland China: Evaluating Machine Learning versus Spatial Interpolation

Yuantian Jiang[1], Wenting Wang[1*], Andrew T. Fullhart[2], Bofu Yu[3], Bo Chen[4]

[1] Department of Geographic Science, Faculty of Arts and Sciences, Beijing Normal University, Zhuhai 519087, China;

[2] School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, USA;

[3]Australian Rivers Institute, School of Engineering and Built Environment, Griffith University, Brisbane, QLD 4111, Australia;

[4] State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

Contents of this file

Introduction

Figure S1 illustrates the spatial distribution of 2,122 hourly observation stations overlaid on elevation data, while Figure S2 shows their distribution in relation to annual average precipitation patterns, providing context for the representativeness of the observation network. Relative errors of observations and predictions for Kriging with External Drift using annual average precipitation (KED_AP, the best interpolation method) and Gradient Boosting (GB, the best machine learning method) are presented in Figure S3, providing insights into the prediction accuracy across different IDF cases. Residual histograms for KED_AP and GB are shown in Figure S4.

In addition to CHM_PRE (used in this study), we also explored two gridded precipitation datasets covering mainland China: CGDPP (1957-2015) with a daily temporal resolution and a spatial resolution of 0.5°, and CHIRPS (1981-2022) with a daily temporal resolution and a spatial resolution of 0.05°. The accuracy metrics for these datasets were calculated across four cases (Nash-Sutcliffe Efficiency (NSE) and Kling-Gupta Efficiency (KGE) are dimensionless, Root Mean Square Error (RMSE) is in mm/h, and Percent Bias

(PBIAS) is in percentage), and the results are shown in Tables S1 to S12. Table S13 provides the names of 10 different machine learning methods applied to four test cases derived from the IDF curves. Tables S14 to S17 present the hyperparameter grids for machine learning models, including Random Forest (RF), Gradient Boosting (GB), ExtraTrees (ET), and Multilayer Perceptron (MLP). Table S18 summarizes the names of optimal regionalization methods for the developed IDF curves with a spatial resolution of 0.1°×0.1° based on CHM_PRE. Tables S19 to S22 display the performance metrics for the developed IDF curves based on CHM_PRE, including NSE, PBIAS, RMSE, and KGE. Although with a lower accuracy, another IDF curves product for mainland China with a spatial resolution of 0.5°×0.5° based on CGDPP, is also provided as an alternative. Names of optimal regionalization methods and performance metrics for the IDF curves based on CGDPP are shown in Tables S23 to S27. Table S28 summarizes statistical characteristics for four representative station-level IDF cases, including mean, median, standard deviation, and percentile values.

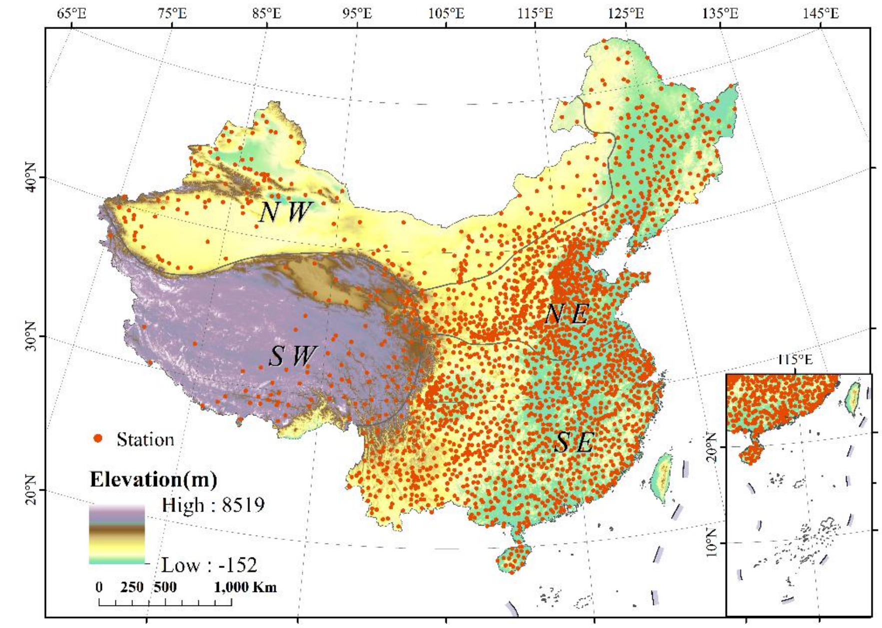

Figure S1. The spatial distribution of 2,122 hourly observation stations and elevation across the study area.

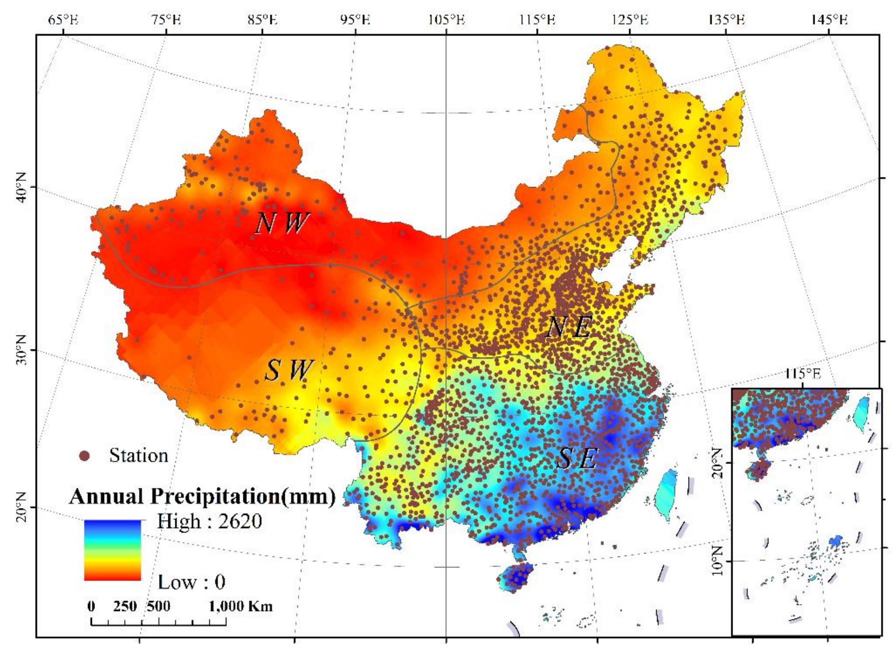

Figure S2. The spatial distribution of 2,122 hourly observation stations and annual average precipitation across the study area.

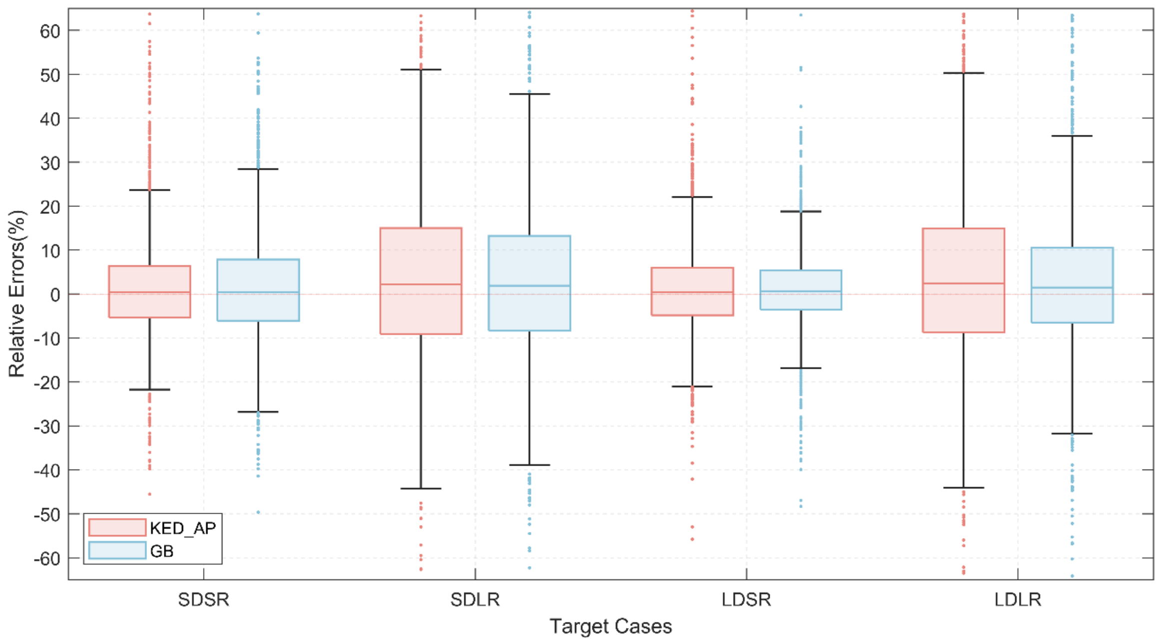

Figure S3. Relative errors of observations and predictions for KED_AP (the best interpolation method) and GB (the best machine learning method). The red line represents the relative errors equal to zero.

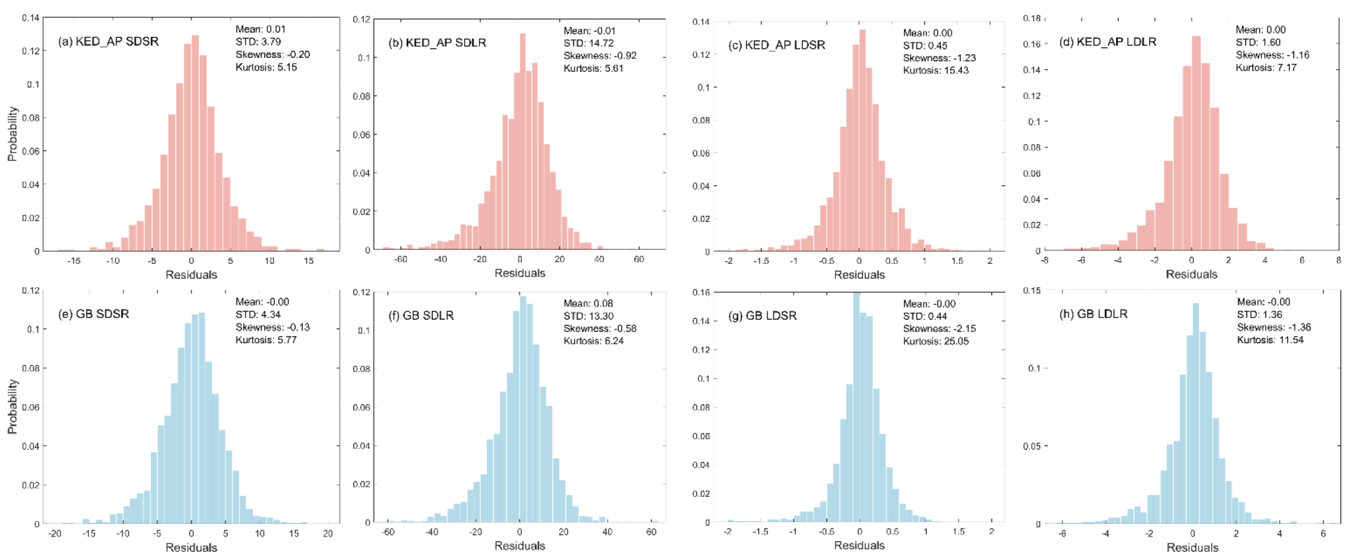

Figure S4. Residual histograms for KED_AP (the best interpolation method, (a)-(d)) and GB (the best machine learning method, (e)-(h)). The normal distribution is characterized by a skewness of 0 and kurtosis of 3.

Table S1. Accuracy metrics of machine learning methods for SDSR (1-hour rainfall intensity for a 5-year return period) using CGDPP.

	NSE	PBIAS	RMSE	KGE
AdaBoost Regression	0.88	1.24	5.36	0.87
Bayesian Ridge Regression	0.87	0.03	5.69	0.90
ExtraTrees Regression	0.91	0.43	4.61	0.94
Gradient Boosting	0.91	0.05	4.60	0.93
K-Nearest Neighbors Regression	0.89	0.92	5.10	0.92
Linear Regression	0.84	0.16	6.24	0.90
Multilayer Perceptron	0.86	-0.30	5.85	0.92
Random Forest	0.92	0.33	4.32	0.94
Ridge Regression	0.87	0.02	5.70	0.90
Support Vector Regression	0.83	-0.44	6.39	0.83

Table S2. Accuracy metrics of machine learning methods for SDLR (1-hour rainfall intensity for a 100-year return period) using CGDPP.

	NSE	PBIAS	RMSE	KGE
AdaBoost Regression	0.70	8.13	17.86	0.74
Bayesian Ridge Regression	0.73	0.03	16.84	0.80
ExtraTrees Regression	0.73	0.75	17.04	0.84
Gradient Boosting	0.78	0.02	15.38	0.83
K-Nearest Neighbors Regression	0.75	0.90	16.38	0.82
Linear Regression	0.63	0.27	19.69	0.78
Multilayer Perceptron	0.71	-0.97	17.50	0.82
Random Forest	0.76	0.74	15.89	0.84
Ridge Regression	0.73	0.03	16.83	0.80
Support Vector Regression	0.66	-3.29	18.89	0.68

Table S3. Accuracy metrics of machine learning methods for LDSR (24-hour rainfall intensity for a 5-year return period) using CGDPP.

	NSE	PBIAS	RMSE	KGE
AdaBoost Regression	0.91	2.82	0.59	0.88
Bayesian Ridge Regression	0.93	0.00	0.51	0.95
ExtraTrees Regression	0.92	0.17	0.54	0.95
Gradient Boosting	0.93	-0.07	0.50	0.95
K-Nearest Neighbors Regression	0.92	0.37	0.54	0.94
Linear Regression	0.89	0.19	0.66	0.94
Multilayer Perceptron	0.92	0.28	0.55	0.95
Random Forest	0.93	0.09	0.51	0.95
Ridge Regression	0.93	0.00	0.51	0.95

Support Vector Regression	0.92	-0.40	0.55	0.93

Table S4. Accuracy metrics of machine learning methods for LDLR (24-hour rainfall intensity for a 100-year return period) using CGDPP.

	NSE	PBIAS	RMSE	KGE
AdaBoost Regression	0.75	9.93	1.96	0.74
Bayesian Ridge Regression	0.82	0.01	1.63	0.87
ExtraTrees Regression	0.77	0.53	1.86	0.87
Gradient Boosting	0.82	0.16	1.63	0.87
K-Nearest Neighbors Regression	0.81	0.38	1.70	0.86
Linear Regression	0.68	0.33	2.22	0.83
Multilayer Perceptron	0.82	0.09	1.65	0.88
Random Forest	0.80	0.51	1.72	0.87
Ridge Regression	0.83	0.00	1.63	0.87
Support Vector Regression	0.80	-2.40	1.74	0.84

Table S5. Accuracy metrics of machine learning methods for SDSR (1-hour rainfall intensity for a 5-year return period) using CHIRPS.

	NSE	PBIAS	RMSE	KGE
AdaBoost Regression	0.87	3.36	5.60	0.85
Bayesian Ridge Regression	0.80	-0.01	6.93	0.85
ExtraTrees Regression	0.92	0.26	4.45	0.93
Gradient Boosting	0.90	0.10	4.86	0.92
K-Nearest Neighbors Regression	0.86	0.62	5.83	0.88
Linear Regression	0.80	-0.12	7.02	0.86
Multilayer Perceptron	0.84	-0.17	6.15	0.90
Random Forest	0.92	0.12	4.43	0.93
Ridge Regression	0.80	0.00	6.99	0.85
Support Vector Regression	0.80	-0.85	7.02	0.78

Table S6. Accuracy metrics of machine learning methods for SDLR (1-hour rainfall intensity for a 100-year return period) using CHIRPS.

	NSE	PBIAS	RMSE	KGE
AdaBoost Regression	0.65	9.89	19.22	0.68
Bayesian Ridge Regression	0.63	0.01	19.87	0.71

ExtraTrees Regression	0.76	0.95	15.78	0.82
Gradient Boosting	0.74	0.33	16.54	0.80
K-Nearest Neighbors Regression	0.68	1.10	18.46	0.77
Linear Regression	0.63	-0.03	19.87	0.71
Multilayer Perceptron	0.67	-0.59	18.67	0.77
Random Forest	0.76	0.63	15.93	0.82
Ridge Regression	0.62	0.03	19.91	0.70
Support Vector Regression	0.57	-4.68	21.28	0.60

Table S7. Accuracy metrics of machine learning methods for LDSR (24-hour rainfall intensity for a 5-year return period) using CHIRPS.

	NSE	**PBIAS**	**RMSE**	**KGE**
AdaBoost Regression	0.83	6.55	0.81	0.80
Bayesian Ridge Regression	0.81	0.00	0.85	0.86
ExtraTrees Regression	0.91	0.09	0.57	0.92
Gradient Boosting	0.89	-0.01	0.64	0.92
K-Nearest Neighbors Regression	0.85	-0.44	0.76	0.87
Linear Regression	0.81	-0.08	0.85	0.86
Multilayer Perceptron	0.87	0.04	0.71	0.90
Random Forest	0.91	-0.02	0.58	0.92
Ridge Regression	0.81	0.01	0.84	0.86
Support Vector Regression	0.84	-1.83	0.78	0.85

Table S8. Accuracy metrics of machine learning methods for LDLR (24-hour rainfall intensity for a 100-year return period) using CHIRPS.

	NSE	**PBIAS**	**RMSE**	**KGE**
AdaBoost Regression	0.64	12.29	2.34	0.65
Bayesian Ridge Regression	0.67	0.05	2.25	0.74
ExtraTrees Regression	0.79	0.65	1.78	0.84
Gradient Boosting	0.76	0.28	1.89	0.82
K-Nearest Neighbors Regression	0.71	0.02	2.11	0.78
Linear Regression	0.66	0.02	2.25	0.74
Multilayer Perceptron	0.72	-0.21	2.05	0.79
Random Forest	0.79	0.29	1.79	0.84
Ridge Regression	0.67	0.05	2.25	0.74
Support Vector Regression	0.69	-3.78	2.18	0.73

Table S9. Accuracy metrics of machine learning methods for SDSR (1-hour rainfall intensity for a 5-year return period) using CHM_PRE.

	NSE	**PBIAS**	**RMSE**	**KGE**
AdaBoost Regression	0.89	2.04	5.09	0.87
Bayesian Ridge Regression	0.89	0.00	5.23	0.92

ExtraTrees Regression	0.92	0.10	4.50	0.93
Gradient Boosting	0.92	-0.02	4.34	0.94
K-Nearest Neighbors Regression	0.90	0.13	5.00	0.92
Linear Regression	0.89	0.00	5.24	0.92
Multilayer Perceptron	0.88	-0.50	5.36	0.93
Random Forest	0.92	0.16	4.34	0.94
Ridge Regression	0.89	0.00	5.25	0.92
Support Vector Regression	0.85	-0.89	6.12	0.85

Table S10. Accuracy metrics of machine learning methods for SDLR (1-hour rainfall intensity for a 100-year return period) using CHM_PRE.

	NSE	PBIAS	RMSE	KGE
AdaBoost Regression	0.76	6.48	15.76	0.74
Bayesian Ridge Regression	0.81	-0.01	14.33	0.86
ExtraTrees Regression	0.82	0.41	13.93	0.87
Gradient Boosting	0.83	0.11	13.29	0.87
K-Nearest Neighbors Regression	0.80	0.00	14.69	0.85
Linear Regression	0.80	0.05	14.56	0.86
Multilayer Perceptron	0.78	-1.21	15.10	0.87
Random Forest	0.83	0.37	13.49	0.87
Ridge Regression	0.81	-0.01	14.33	0.86
Support Vector Regression	0.72	-3.36	17.24	0.73

Table S11. Accuracy metrics of machine learning methods for LDSR (24-hour rainfall intensity for a 5-year return period) using CHM_PRE.

	NSE	PBIAS	RMSE	KGE
AdaBoost Regression	0.88	8.06	0.68	0.82
Bayesian Ridge Regression	0.93	-0.02	0.50	0.95
ExtraTrees Regression	0.93	0.01	0.50	0.95
Gradient Boosting	0.95	-0.07	0.44	0.96
K-Nearest Neighbors Regression	0.93	-0.60	0.53	0.94
Linear Regression	0.93	-0.04	0.50	0.95
Multilayer Perceptron	0.95	-0.14	0.45	0.97
Random Forest	0.94	-0.03	0.47	0.96
Ridge Regression	0.93	-0.02	0.50	0.95
Support Vector Regression	0.92	-1.36	0.55	0.93

Table S12. Accuracy metrics of machine learning methods for LDLR (24-hour rainfall intensity for a 100-year return period) using CHM_PRE.

	NSE	PBIAS	RMSE	KGE
AdaBoost Regression	0.78	10.97	1.83	0.72

Bayesian Ridge Regression	0.87	-0.03	1.40	0.90
ExtraTrees Regression	0.86	0.18	1.45	0.90
Gradient Boosting	0.88	-0.05	1.36	0.91
K-Nearest Neighbors Regression	0.85	-0.72	1.50	0.89
Linear Regression	0.87	0.00	1.41	0.91
Multilayer Perceptron	0.87	0.18	1.37	0.91
Random Forest	0.87	0.22	1.37	0.91
Ridge Regression	0.87	-0.03	1.40	0.90
Support Vector Regression	0.84	-2.06	1.55	0.87

Table S13. 10 different Machine learning methods used for 4 test cases from the IDF curves. The details can be found in the user guide of scikit-learn (https://scikit-learn.org/stable/user_guide.html).

Machine Learning Method	Description
AdaBoost Regression	A boosting technique that adjusts the weights of training samples based on previous errors, focusing more on difficult-to-predict instances in subsequent weak learners.
Bayesian Ridge Regression	A Bayesian approach to linear regression that includes a probabilistic model and regularization, offering a distribution over model parameters rather than point estimates.
ExtraTrees Regression	Similar to Random Forest, but with more randomization in selecting splits, leading to reduced variance and often faster computation.
Gradient Boosting	An ensemble technique that builds decision trees sequentially, with each tree trying to correct the residual errors of the previous trees, optimizing the overall loss function. In each stage a regression tree is fit on the negative gradient of the given loss function.
K-Nearest Neighbors Regression	A non-parametric method that predicts the target value based on the average or weighted average of the k-nearest data points in the feature space.
Linear Regression	A basic linear approach that models the relationship between input features and the target as a linear combination of the input variables.
Multilayer Perceptron	A type of feedforward neural network with one or more hidden layers, capable of learning complex non-linear relationships between inputs and outputs.

Random Forest	A bagging-based ensemble method that builds multiple decision trees and merges their results to improve accuracy and prevent overfitting.
Ridge Regression	A variant of linear regression that includes a regularization term (L2 penalty) to prevent overfitting by shrinking the coefficients of less important features.
Support Vector Regression	A regression variant of support vector machine that attempts to find a hyperplane that best fits the data within a certain margin of tolerance, used for both linear and non-linear regression.

Table S14. Hyperparameter grid for random forest (RF). The default hyperparameters in the scikit-learn library are highlighted in red.

Hyperparameter	**Values**
n_estimators	50, 100, 150, 200, 500
min_samples_split	2, 5, 10, 20, 30
max_features	0.3, 0.5, None, "sqrt", "log2"

Table S15. Hyperparameter grid for gradient boosting (GB). The default hyperparameters in the scikit-learn library are highlighted in red.

Hyperparameter	**Values**
n_estimators	50, 100, 150, 200, 500
max_depth	3, 4, 5, 10
max_features	0.3, 0.5, None, "sqrt", "log2"
subsample	0.2, 0.5, 1
learning_rate	0.05, 0.1, 0.2, 0.5

Table S16. Hyperparameter grid for extratrees (ET). The default hyperparameters in the scikit-learn library are highlighted in red.

Hyperparameter	**Values**
n_estimators	50, 100, 150, 200, 500
min_samples_split	2, 5, 10, 20, 30
max_features	0.3, 0.5, None, "sqrt", "log2"

Table S17. Hyperparameter grid for multilayer perceptron (MLP). The default hyperparameters in the scikit-learn library are highlighted in red.

Hyperparameter	**Values**

hidden_layer_sizes	(50), (100), (500), (750), (1000), (150, 150), (200, 200), (250, 250), (15, 30, 45), (30, 45, 30), (45, 30, 15), (10, 10, 10, 10), (15, 15, 15, 15)
activation	'logistic', 'tanh', 'relu', 'identity'
alpha	0.0001, 0.001, 0.01, 0.1

Table S18. Optimal regionalization methods for developed IDF curves with a spatial resolution of 0.1°×0.1° based on CHM_PRE.

	2 years	**5 years**	**10 years**	**20 years**	**50 years**	**100 years**	**200 years**	**500 years**	**1000 years**
1h	KED_AP	KED_AP	KED_DEM+AP	KED_AP	GB	GB	GB	GB	GB
2h	KED_AP	KED_AP	KED_AP	GB	GB	GB	GB	GB	GB
3h	KED_DEM+AP	KED_AP	KED_AP	GB	GB	GB	GB	GB	GB
4h	KED_DEM+AP	KED_AP	GB	GB	GB	GB	GB	GB	GB
5h	KED_AP	GB	GB	GB	GB	GB	GB	GB	LR
6h	KED_AP	KED_AP	GB	GB	GB	GB	GB	GB	LR
7h	KED_AP	GB	GB	GB	GB	GB	GB	GB	GB
8h	KED_AP	GB	GB	GB	GB	GB	GB	LR	LR
9h	KED_AP	GB	GB	GB	GB	GB	GB	LR	MLP
10h	KED_AP	KED_AP	GB	GB	GB	GB	GB	GB	LR
11h	GB	GB	GB	GB	GB	GB	LR	LR	GB
12h	KED_AP	KED_AP	GB	GB	GB	GB	GB	LR	GB
13h	KED_AP	GB	GB	GB	GB	GB	LR	LR	MLP
14h	GB	GB	GB	GB	GB	GB	GB	MLP	LR
15h	KED_AP	GB	GB	GB	GB	GB	GB	GB	GB
16h	GB	GB	GB	GB	GB	GB	MLP	GB	MLP
17h	GB	GB	GB	GB	GB	GB	GB	LR	MLP
18h	GB	GB	GB	GB	GB	GB	GB	GB	MLP
19h	GB	KED_AP	GB	MLP	GB	GB	LR	GB	MLP
20h	KED_AP	GB	GB	GB	GB	GB	GB	MLP	GB
21h	KED_AP	KED_AP	GB	GB	GB	GB	GB	LR	MLP
22h	KED_DEM+AP	GB	GB	GB	GB	GB	MLP	GB	GB
23h	KED_DEM+AP	GB	MLP	MLP	GB	GB	GB	MLP	GB
24h	KED_AP	GB	MLP	MLP	GB	GB	GB	GB	MLP
36h	KED_AP	MLP	GB	MLP	GB	GB	MLP	MLP	MLP
48h	KED_AP	RF	MLP	MLP	GB	GB	GB	MLP	MLP
72h	KED_AP	KED_AP	RF	MLP	MLP	GB	MLP	GB	MLP

Table S19. Nash-Sutcliffe Efficiency for developed IDF curves with a spatial resolution of 0.1°×0.1° based on CHM_PRE.

	2 years	5 years	10 years	20 years	50 years	100 years	200 years	500 years	1000 years
1h	0.95	0.94	0.93	0.90	0.87	0.83	0.79	0.71	0.66
2h	0.95	0.94	0.94	0.91	0.88	0.85	0.81	0.74	0.67
3h	0.95	0.94	0.93	0.92	0.89	0.86	0.81	0.74	0.68
4h	0.95	0.94	0.93	0.92	0.89	0.86	0.82	0.75	0.68
5h	0.95	0.95	0.94	0.93	0.90	0.86	0.82	0.75	0.67
6h	0.95	0.94	0.94	0.93	0.90	0.86	0.82	0.75	0.67
7h	0.95	0.95	0.94	0.93	0.90	0.86	0.82	0.75	0.68
8h	0.95	0.95	0.94	0.93	0.90	0.86	0.82	0.74	0.68
9h	0.95	0.95	0.94	0.93	0.90	0.87	0.82	0.75	0.70
10h	0.95	0.95	0.94	0.93	0.90	0.87	0.82	0.75	0.68
11h	0.95	0.95	0.94	0.93	0.90	0.87	0.82	0.75	0.69
12h	0.95	0.95	0.94	0.93	0.90	0.87	0.83	0.75	0.69
13h	0.95	0.95	0.94	0.93	0.90	0.87	0.82	0.75	0.71
14h	0.96	0.95	0.94	0.93	0.90	0.87	0.83	0.76	0.69
15h	0.95	0.95	0.94	0.93	0.90	0.87	0.83	0.76	0.69
16h	0.96	0.95	0.94	0.93	0.91	0.87	0.83	0.76	0.71
17h	0.95	0.95	0.94	0.93	0.91	0.87	0.83	0.75	0.71
18h	0.96	0.95	0.94	0.93	0.91	0.87	0.83	0.76	0.71
19h	0.95	0.95	0.94	0.93	0.91	0.87	0.83	0.76	0.71
20h	0.95	0.95	0.94	0.93	0.91	0.87	0.83	0.77	0.69
21h	0.95	0.95	0.94	0.93	0.91	0.87	0.83	0.76	0.71
22h	0.95	0.95	0.94	0.93	0.91	0.87	0.84	0.76	0.69
23h	0.95	0.95	0.94	0.93	0.91	0.88	0.83	0.77	0.69
24h	0.95	0.95	0.94	0.93	0.90	0.88	0.83	0.76	0.71
36h	0.95	0.95	0.94	0.93	0.91	0.88	0.84	0.78	0.72
48h	0.95	0.95	0.94	0.93	0.91	0.87	0.84	0.78	0.72
72h	0.95	0.94	0.94	0.93	0.91	0.88	0.85	0.77	0.73

Table S20. Percent Bias for developed IDF curves with a spatial resolution of 0.1°×0.1° based on CHM_PRE.

	2 years	5 years	10 years	20 years	50 years	100 years	200 years	500 years	1000 years
1h	0.06	0.03	-0.07	-0.02	0.10	0.11	0.06	0.20	0.05
2h	-0.02	0.05	-0.07	-0.01	0.04	0.00	0.04	0.23	0.36
3h	0.07	0.05	0.01	-0.01	0.01	-0.02	0.01	0.12	0.22
4h	0.09	0.07	-0.04	-0.06	-0.02	-0.03	-0.02	0.11	0.20
5h	-0.05	-0.09	-0.09	-0.05	-0.02	-0.01	-0.04	0.12	0.12
6h	0.00	0.03	-0.10	-0.09	-0.03	-0.01	0.07	0.05	0.12
7h	-0.01	-0.09	-0.04	-0.08	-0.05	-0.03	0.03	0.04	0.12
8h	0.03	-0.12	-0.01	-0.05	-0.02	-0.04	0.02	0.07	0.11

9h	0.02	-0.03	-0.10	-0.07	-0.03	0.04	-0.01	0.07	-0.24
10h	-0.03	-0.01	-0.07	-0.07	-0.02	-0.01	0.06	0.10	0.11
11h	-0.11	-0.07	-0.10	-0.08	-0.08	0.03	0.03	0.07	0.09
12h	0.02	0.00	-0.09	-0.07	-0.08	-0.01	0.03	0.07	0.03
13h	0.01	-0.10	-0.12	-0.10	-0.06	0.04	0.03	0.06	-0.06
14h	-0.07	-0.04	-0.14	-0.06	-0.06	0.04	-0.03	-0.06	0.10
15h	0.03	-0.05	-0.13	-0.10	-0.07	0.03	0.04	0.01	-0.01
16h	-0.05	-0.05	-0.15	-0.12	-0.05	-0.01	-0.08	0.05	-0.24
17h	-0.08	-0.08	-0.07	-0.09	0.00	-0.01	0.03	0.06	-0.27
18h	-0.06	-0.08	-0.14	-0.10	-0.09	-0.01	0.00	0.05	-0.12
19h	-0.07	0.00	-0.11	0.01	-0.03	-0.05	0.02	-0.02	0.16
20h	0.05	-0.10	-0.07	-0.07	-0.07	-0.03	0.04	-0.01	-0.02
21h	0.03	0.02	-0.10	-0.08	-0.04	0.00	0.03	0.05	-0.11
22h	-0.03	-0.07	-0.06	-0.10	0.01	-0.02	0.09	-0.01	0.04
23h	0.01	-0.07	0.04	-0.03	-0.03	0.00	-0.03	-0.10	0.02
24h	0.01	-0.07	0.02	0.02	-0.02	-0.05	-0.01	-0.03	-0.23
36h	0.02	0.02	-0.10	0.12	-0.04	-0.04	-0.10	-0.23	-0.12
48h	0.08	0.01	-0.25	-0.35	-0.04	-0.06	-0.04	-0.16	-0.24
72h	0.05	0.06	0.00	0.00	-0.08	-0.02	-0.14	-0.05	0.12

Table S21. Root Mean Square Error for developed IDF curves with a spatial resolution of 0.1°×0.1° based on CHM_PRE.

	2 years	**5 years**	**10 years**	**20 years**	**50 years**	**100 years**	**200 years**	**500 years**	**1000 years**
1h	2.57	3.79	4.99	6.94	10.02	13.29	17.71	26.10	34.68
2h	1.68	2.46	3.22	4.39	6.32	8.49	11.32	16.79	22.71
3h	1.36	1.92	2.53	3.30	4.79	6.43	8.62	12.76	17.20
4h	1.10	1.60	2.09	2.70	3.91	5.25	7.05	10.43	14.06
5h	0.94	1.38	1.76	2.27	3.31	4.47	6.04	8.91	12.27
6h	0.84	1.22	1.54	1.98	2.90	3.91	5.30	7.82	10.71
7h	0.75	1.07	1.36	1.76	2.59	3.50	4.70	6.98	9.41
8h	0.68	0.97	1.24	1.59	2.35	3.17	4.25	6.40	8.61
9h	0.62	0.89	1.13	1.45	2.13	2.90	3.91	5.85	7.62
10h	0.58	0.83	1.05	1.34	1.99	2.69	3.62	5.32	7.26
11h	0.53	0.77	0.97	1.24	1.85	2.49	3.40	5.01	6.69
12h	0.50	0.74	0.91	1.17	1.72	2.33	3.14	4.69	6.27
13h	0.47	0.68	0.85	1.11	1.62	2.19	2.99	4.41	5.73
14h	0.44	0.64	0.81	1.05	1.54	2.06	2.79	4.07	5.61
15h	0.42	0.61	0.77	1.00	1.47	1.96	2.65	3.92	5.26
16h	0.40	0.59	0.73	0.96	1.38	1.87	2.52	3.72	4.88
17h	0.39	0.56	0.70	0.90	1.31	1.79	2.41	3.59	4.67
18h	0.37	0.54	0.67	0.87	1.26	1.70	2.30	3.42	4.45
19h	0.36	0.52	0.65	0.87	1.21	1.63	2.24	3.27	4.28
20h	0.35	0.50	0.63	0.81	1.16	1.59	2.12	3.07	4.23

21h	0.34	0.49	0.61	0.78	1.13	1.52	2.04	3.05	3.96
22h	0.33	0.47	0.59	0.75	1.08	1.46	1.95	2.93	3.95
23h	0.32	0.46	0.59	0.74	1.05	1.41	1.91	2.77	3.82
24h	0.31	0.44	0.56	0.72	1.03	1.36	1.85	2.74	3.56
36h	0.23	0.33	0.42	0.53	0.75	1.01	1.34	1.96	2.61
48h	0.19	0.28	0.33	0.43	0.61	0.82	1.09	1.56	2.09
72h	0.14	0.21	0.26	0.32	0.45	0.61	0.78	1.18	1.53

Table S22. Kling-Gupta Efficiency for developed IDF curves with a spatial resolution of 0.1°×0.1° based on CHM_PRE.

	2 years	**5 years**	**10 years**	**20 years**	**50 years**	**100 years**	**200 years**	**500 years**	**1000 years**
1h	0.96	0.96	0.95	0.93	0.90	0.87	0.84	0.78	0.73
2h	0.97	0.96	0.95	0.93	0.91	0.89	0.85	0.80	0.75
3h	0.97	0.96	0.95	0.94	0.92	0.89	0.86	0.81	0.76
4h	0.96	0.96	0.95	0.94	0.92	0.90	0.86	0.81	0.76
5h	0.97	0.96	0.95	0.94	0.92	0.90	0.86	0.81	0.76
6h	0.96	0.96	0.95	0.94	0.92	0.90	0.87	0.81	0.77
7h	0.96	0.96	0.95	0.94	0.92	0.90	0.87	0.82	0.77
8h	0.97	0.96	0.96	0.95	0.92	0.90	0.87	0.82	0.77
9h	0.96	0.96	0.96	0.95	0.93	0.90	0.87	0.82	0.78
10h	0.96	0.96	0.96	0.95	0.93	0.90	0.87	0.82	0.77
11h	0.97	0.96	0.96	0.95	0.93	0.90	0.87	0.82	0.77
12h	0.97	0.96	0.96	0.95	0.93	0.91	0.87	0.82	0.77
13h	0.97	0.96	0.96	0.95	0.93	0.91	0.87	0.82	0.78
14h	0.97	0.96	0.96	0.95	0.93	0.91	0.87	0.84	0.78
15h	0.97	0.96	0.96	0.95	0.93	0.91	0.88	0.82	0.77
16h	0.97	0.96	0.96	0.95	0.93	0.91	0.88	0.82	0.79
17h	0.97	0.96	0.96	0.95	0.93	0.91	0.88	0.83	0.79
18h	0.97	0.96	0.96	0.95	0.93	0.91	0.88	0.82	0.79
19h	0.97	0.96	0.96	0.95	0.93	0.91	0.88	0.82	0.79
20h	0.97	0.96	0.96	0.95	0.93	0.91	0.88	0.83	0.78
21h	0.96	0.96	0.96	0.95	0.93	0.91	0.88	0.83	0.79
22h	0.96	0.96	0.96	0.95	0.93	0.91	0.89	0.83	0.78
23h	0.97	0.96	0.96	0.95	0.93	0.91	0.88	0.84	0.78
24h	0.96	0.96	0.96	0.94	0.93	0.91	0.88	0.83	0.79
36h	0.97	0.96	0.96	0.95	0.93	0.91	0.89	0.85	0.79
48h	0.97	0.96	0.97	0.96	0.93	0.91	0.88	0.84	0.79
72h	0.96	0.96	0.95	0.94	0.93	0.91	0.90	0.83	0.80

Table S23. Optimal regionalization methods for developed IDF curves with a spatial resolution of 0.5°×0.5° based on CGDPP.

	2 years	5 years	10 years	20 years	50 years	100 years	200 years	500 years	1000 years
1h	KED_AP	KED_AP	KED_DEM+AP	KED_AP	KED_AP	KED_AP	KED_DEM+AP	KED_DEM+AP	GB
2h	KED_AP	KED_AP	KED_AP	KED_AP	KED_AP	KED_AP	KED_DEM+AP	GB	GB
3h	KED_DEM+AP	KED_AP	KED_AP	KED_DEM+AP	KED_AP	KED_AP	KED_DEM+AP	GB	GB
4h	KED_DEM+AP	KED_AP	KED_AP	KED_AP	KED_AP	KED_AP	GB	GB	GB
5h	KED_AP	KED_DEM+AP	KED_DEM+AP	KED_AP	KED_AP	KED_AP	GB	GB	GB
6h	KED_AP	KED_AP	KED_DEM+AP	KED_DEM+AP	KED_AP	GB	GB	GB	GB
7h	KED_AP	KED_AP	KED_DEM+AP	KED_AP	KED_AP	GB	GB	MLP	MLP
8h	KED_AP	KED_AP	KED_DEM+AP	KED_DEM+AP	KED_AP	KED_AP	GB	MLP	MLP
9h	KED_AP	KED_DEM+AP	KED_AP	KED_AP	KED_AP	KED_DEM+AP	GB	MLP	GB
10h	KED_AP	KED_AP	KED_AP	KED_DEM+AP	KED_AP	GB	GB	MLP	MLP
11h	KED_DEM+AP	KED_DEM+AP	KED_AP	KED_DEM+AP	GB	KED_AP	MLP	MLP	MLP
12h	KED_AP	KED_AP	KED_AP	KED_AP	KED_AP	KED_AP	MLP	MLP	MLP
13h	KED_AP	KED_DEM+AP	KED_AP	KED_AP	KED_AP	KED_AP	GB	MLP	GB
14h	KED_AP	KED_AP	KED_DEM+AP	KED_DEM+AP	KED_AP	KED_AP	MLP	MLP	MLP
15h	KED_AP	KED_DEM+AP	KED_AP	KED_AP	KED_DEM+AP	GB	MLP	MLP	MLP
16h	KED_AP	KED_AP	KED_AP	KED_AP	KED_AP	GB	MLP	MLP	MLP
17h	KED_AP	KED_AP	KED_AP	KED_AP	KED_AP	KED_AP	MLP	MLP	MLP
18h	KED_AP	KED_AP	KED_AP	KED_AP	KED_AP	KED_AP	MLP	MLP	MLP
19h	KED_AP	KED_AP	KED_AP	KED_DEM+AP	MLP	GB	MLP	MLP	MLP
20h	KED_AP	KED_AP	KED_AP	KED_AP	MLP	MLP	GB	MLP	MLP
21h	KED_AP	KED_AP	KED_AP	KED_DEM+AP	KED_DEM+AP	MLP	MLP	MLP	MLP
22h	KED_DEM+AP	KED_DEM+AP	KED_AP	KED_AP	KED_AP	GB	MLP	MLP	MLP
23h	KED_DEM+AP	KED_DEM+AP	KED_AP	KED_AP	MLP	MLP	GB	MLP	MLP
24h	KED_AP	KED_AP	KED_DEM+AP	KED_AP	KED_AP	GB	MLP	MLP	MLP
36h	KED_AP	KED_DEM+AP	KED_DEM+AP	KED_AP	MLP	MLP	MLP	MLP	MLP
48h	KED_AP	KED_AP	KED_AP	KED_AP	KED_AP	GB	GB	MLP	MLP
72h	KED_AP	KED_AP	KED_DEM+AP	KED_AP	MLP	MLP	MLP	MLP	MLP

Table S24. Nash-Sutcliffe Efficiency for developed IDF curves with a spatial resolution of 0.5°×0.5° based on CGDPP.

	2 years	5 years	10 years	20 years	50 years	100 years	200 years	500 years	1000 years
1h	0.95	0.94	0.93	0.90	0.85	0.79	0.71	0.61	0.54
2h	0.95	0.94	0.94	0.91	0.86	0.80	0.73	0.64	0.55
3h	0.95	0.94	0.93	0.91	0.86	0.80	0.73	0.65	0.56
4h	0.95	0.94	0.93	0.91	0.86	0.81	0.74	0.65	0.57
5h	0.95	0.94	0.93	0.91	0.86	0.81	0.75	0.65	0.57
6h	0.95	0.94	0.93	0.91	0.87	0.81	0.75	0.65	0.58
7h	0.95	0.94	0.93	0.91	0.87	0.81	0.75	0.65	0.57
8h	0.95	0.94	0.93	0.91	0.86	0.82	0.75	0.65	0.57
9h	0.95	0.94	0.93	0.91	0.87	0.81	0.75	0.65	0.59
10h	0.95	0.95	0.93	0.91	0.87	0.81	0.75	0.65	0.58
11h	0.95	0.94	0.93	0.91	0.86	0.82	0.75	0.66	0.58
12h	0.95	0.95	0.93	0.91	0.87	0.82	0.75	0.66	0.59
13h	0.95	0.94	0.93	0.91	0.87	0.82	0.76	0.66	0.59

14h	0.95	0.95	0.93	0.91	0.87	0.82	0.75	0.66	0.59
15h	0.95	0.94	0.94	0.92	0.86	0.82	0.76	0.66	0.59
16h	0.95	0.95	0.93	0.92	0.87	0.82	0.76	0.67	0.59
17h	0.95	0.95	0.94	0.92	0.87	0.82	0.76	0.67	0.59
18h	0.95	0.95	0.93	0.92	0.87	0.82	0.76	0.67	0.60
19h	0.95	0.95	0.94	0.91	0.86	0.82	0.76	0.67	0.60
20h	0.95	0.94	0.94	0.92	0.86	0.82	0.76	0.67	0.60
21h	0.95	0.95	0.93	0.91	0.87	0.82	0.76	0.68	0.60
22h	0.95	0.94	0.94	0.92	0.87	0.82	0.76	0.67	0.60
23h	0.95	0.94	0.93	0.92	0.86	0.82	0.77	0.68	0.60
24h	0.95	0.95	0.93	0.92	0.87	0.82	0.77	0.68	0.60
36h	0.95	0.94	0.93	0.92	0.87	0.83	0.78	0.69	0.61
48h	0.95	0.94	0.94	0.92	0.87	0.83	0.78	0.69	0.62
72h	0.95	0.94	0.92	0.92	0.88	0.83	0.78	0.70	0.62

Table S25. Percent Bias for developed IDF curves with a spatial resolution of 0.5°×0.5° based on CGDPP.

	2 years	**5 years**	**10 years**	**20 years**	**50 years**	**100 years**	**200 years**	**500 years**	**1000 years**
1h	0.06	0.03	-0.07	-0.02	0.09	-0.01	0.18	0.19	0.49
2h	-0.02	0.05	-0.07	0.07	-0.02	-0.04	0.17	0.32	0.45
3h	0.07	0.05	0.01	0.06	0.07	-0.01	0.11	0.27	0.32
4h	0.09	0.07	-0.02	0.05	0.02	-0.01	0.18	0.28	0.31
5h	-0.05	0.05	0.02	0.01	0.04	0.08	0.12	0.24	0.33
6h	0.00	0.03	0.00	-0.04	-0.04	0.02	0.14	0.28	0.37
7h	-0.01	0.06	0.02	0.04	0.02	0.08	0.17	-0.21	-0.37
8h	0.03	-0.01	0.05	0.07	0.05	-0.04	0.20	-0.45	-0.26
9h	0.02	0.05	-0.06	-0.03	-0.09	-0.01	0.26	-0.29	0.43
10h	-0.03	-0.01	0.01	-0.01	0.06	0.15	0.23	-0.29	-0.29
11h	0.06	-0.02	0.09	-0.04	0.05	0.02	-0.18	0.04	-0.35
12h	0.02	0.00	0.05	0.09	-0.05	0.00	0.06	-0.23	-0.44
13h	0.01	0.02	0.01	0.13	0.07	-0.03	0.28	-0.45	0.51
14h	0.06	0.07	-0.02	0.04	0.00	0.00	0.04	-0.10	-0.21
15h	0.03	0.06	0.06	0.04	0.04	0.13	-0.21	0.03	-0.13
16h	0.00	-0.06	-0.03	-0.03	0.09	0.15	-0.10	-0.12	-0.23
17h	0.10	0.13	-0.02	0.05	-0.03	-0.02	-0.14	-0.38	-0.19
18h	0.04	-0.02	0.07	0.03	-0.01	-0.03	-0.01	0.09	-0.23
19h	0.04	0.00	0.09	0.05	-0.12	0.14	-0.08	-0.19	-0.08
20h	0.05	0.00	-0.03	0.11	-0.07	-0.11	0.30	-0.23	-0.27
21h	0.03	0.02	-0.02	0.02	0.01	0.14	-0.11	-0.03	-0.37
22h	-0.03	0.01	0.02	0.06	0.11	0.19	-0.12	-0.19	0.11
23h	0.01	0.00	0.06	0.07	-0.04	0.03	0.30	-0.20	-0.21
24h	0.01	0.07	-0.02	0.05	0.04	0.15	0.05	-0.15	0.17
36h	0.02	0.01	0.09	-0.03	-0.08	0.10	-0.15	0.24	0.11

48h	0.08	0.10	0.06	0.05	0.01	0.19	0.27	0.00	-0.25
72h	0.05	0.06	0.02	0.08	0.01	-0.12	0.09	0.23	-0.27

Table S26. Root Mean Square Error for developed IDF curves with a spatial resolution of 0.5°×0.5° based on CGDPP.

	2 years	**5 years**	**10 years**	**20 years**	**50 years**	**100 years**	**200 years**	**500 years**	**1000 years**
1h	2.57	3.79	4.99	6.94	10.58	14.72	20.61	30.20	40.07
2h	1.68	2.46	3.22	4.53	6.92	9.62	13.27	19.73	26.34
3h	1.36	1.92	2.53	3.47	5.28	7.42	10.30	14.99	20.03
4h	1.10	1.60	2.12	2.83	4.36	6.06	8.39	12.31	16.39
5h	0.94	1.41	1.84	2.50	3.76	5.13	7.13	10.47	13.97
6h	0.84	1.22	1.62	2.18	3.26	4.64	6.27	9.20	12.16
7h	0.75	1.10	1.47	1.91	2.93	4.16	5.58	8.30	10.94
8h	0.68	1.00	1.31	1.76	2.68	3.63	5.03	7.47	9.92
9h	0.62	0.94	1.19	1.63	2.45	3.44	4.63	6.86	8.96
10h	0.58	0.83	1.11	1.55	2.27	3.18	4.28	6.33	8.30
11h	0.54	0.80	1.03	1.42	2.21	2.88	4.02	5.86	7.76
12h	0.50	0.74	0.96	1.31	1.96	2.69	3.81	5.45	7.23
13h	0.47	0.71	0.92	1.24	1.85	2.53	3.51	5.15	6.81
14h	0.45	0.66	0.86	1.19	1.75	2.41	3.35	4.89	6.45
15h	0.42	0.63	0.81	1.10	1.71	2.35	3.17	4.62	6.08
16h	0.41	0.60	0.79	1.05	1.58	2.23	3.00	4.37	5.79
17h	0.39	0.57	0.74	1.00	1.52	2.08	2.90	4.20	5.52
18h	0.37	0.55	0.72	0.96	1.46	1.99	2.75	4.01	5.29
19h	0.36	0.52	0.68	0.95	1.47	1.96	2.64	3.87	5.08
20h	0.35	0.51	0.67	0.89	1.42	1.89	2.52	3.70	4.88
21h	0.34	0.49	0.65	0.87	1.33	1.84	2.44	3.54	4.66
22h	0.33	0.48	0.62	0.82	1.25	1.74	2.37	3.43	4.54
23h	0.32	0.48	0.61	0.79	1.29	1.70	2.25	3.30	4.39
24h	0.31	0.45	0.59	0.77	1.16	1.63	2.19	3.18	4.24
36h	0.23	0.35	0.45	0.58	0.89	1.20	1.59	2.32	3.08
48h	0.19	0.28	0.35	0.47	0.70	0.96	1.29	1.86	2.46
72h	0.14	0.21	0.29	0.36	0.53	0.71	0.94	1.36	1.81

Table S27. Kling-Gupta Efficiency for developed IDF curves with a spatial resolution of 0.5°×0.5° based on CGDPP.

	2 years	**5 years**	**10 years**	**20 years**	**50 years**	**100 years**	**200 years**	**500 years**	**1000 years**
1h	0.96	0.96	0.95	0.93	0.89	0.84	0.80	0.71	0.64
2h	0.97	0.96	0.95	0.93	0.89	0.85	0.81	0.72	0.66
3h	0.97	0.96	0.95	0.93	0.90	0.85	0.81	0.73	0.66
4h	0.96	0.96	0.95	0.93	0.90	0.85	0.81	0.73	0.67
5h	0.97	0.96	0.95	0.93	0.90	0.86	0.81	0.74	0.67

6h	0.96	0.96	0.95	0.93	0.90	0.86	0.81	0.74	0.68
7h	0.96	0.96	0.95	0.93	0.90	0.86	0.82	0.75	0.68
8h	0.97	0.96	0.95	0.93	0.90	0.86	0.82	0.75	0.69
9h	0.96	0.96	0.95	0.93	0.90	0.86	0.82	0.75	0.69
10h	0.96	0.96	0.95	0.93	0.90	0.87	0.82	0.75	0.68
11h	0.96	0.96	0.95	0.94	0.90	0.86	0.82	0.75	0.69
12h	0.97	0.96	0.95	0.93	0.90	0.86	0.83	0.76	0.69
13h	0.97	0.96	0.95	0.94	0.90	0.86	0.83	0.75	0.69
14h	0.96	0.96	0.95	0.93	0.90	0.86	0.83	0.76	0.69
15h	0.97	0.96	0.95	0.93	0.90	0.87	0.83	0.76	0.69
16h	0.96	0.96	0.95	0.93	0.90	0.87	0.83	0.76	0.69
17h	0.97	0.96	0.95	0.94	0.90	0.86	0.83	0.76	0.70
18h	0.96	0.96	0.95	0.93	0.90	0.86	0.83	0.76	0.69
19h	0.97	0.96	0.95	0.94	0.91	0.87	0.84	0.76	0.70
20h	0.97	0.96	0.95	0.94	0.91	0.88	0.83	0.76	0.70
21h	0.96	0.96	0.95	0.94	0.91	0.88	0.84	0.77	0.70
22h	0.96	0.96	0.95	0.94	0.90	0.87	0.84	0.76	0.70
23h	0.97	0.96	0.95	0.93	0.91	0.87	0.83	0.76	0.70
24h	0.96	0.96	0.95	0.94	0.90	0.87	0.84	0.77	0.70
36h	0.97	0.96	0.95	0.93	0.91	0.88	0.84	0.76	0.71
48h	0.97	0.96	0.95	0.94	0.90	0.88	0.84	0.77	0.71
72h	0.96	0.96	0.95	0.93	0.92	0.90	0.85	0.77	0.73

Table S28. Statistical characteristics of 4 representative observed station-level IDF cases(mm/h).

	Mean	Median	STD*	Minimum	5^{th} Percentile	95^{th} Percentile	Maximum
SDSR	40.68	44.37	15.59	2.63	12.19	61.92	94.54
SDLR	76.38	77.05	32.49	4.52	21.5	129.03	221.95
LDSR	4.5	4.62	1.94	0.17	1.39	7.61	13.34
LDLR	8.42	8.31	3.87	0.36	2.42	14.78	27.17

*STD represents the standard deviation.