Supplement of Hydrol. Earth Syst. Sci., 21, 6461–6483, 2017
https://doi.org/10.5194/hess-21-6461-2017-supplement

Hydrology and Earth System Sciences
Open Access
EGU

Supplement of

Does nonstationarity in rainfall require nonstationary intensity–duration–frequency curves?

Poulomi Ganguli and Paulin Coulibaly

Correspondence to: Poulomi Ganguli (poulomi.ganguli@alumnimail.iitkgp.ac.in, gangulip@mcmaster.ca)

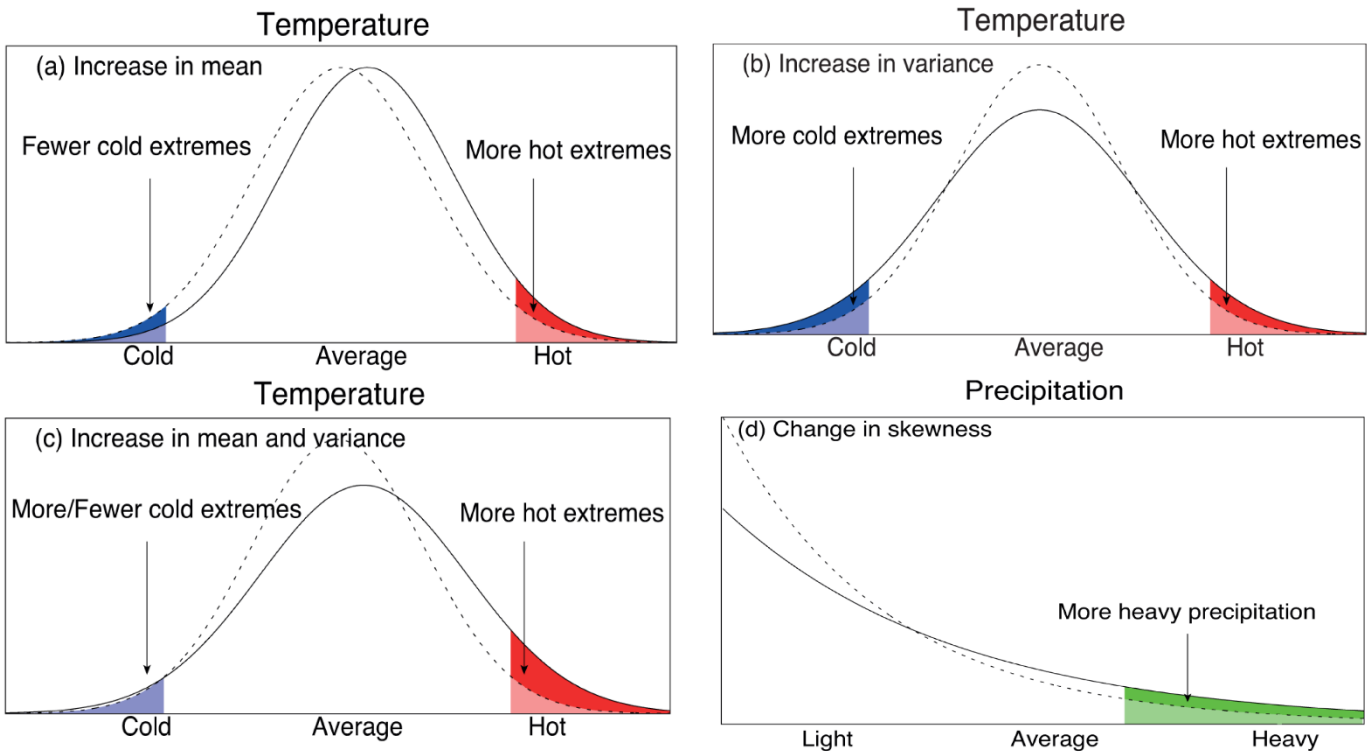

Figure S1. IPCC AR5 conceptual representation of changes in probability density functions of daily temperature (**a** – **c**) and precipitation. The previous and the new distributions are marked by the solid and the dashed lines respectively. The frequency (probability of occurrence) of extremes is denoted by the shaded areas (Source: IPCC AR5 Working Group I report, Figure 1.8, page no. 134).

SI 1. Infilling Missing AMP Record: Method and Results

Multiplicative Random Cascade (MRC) Models for Temporal Disaggregation of Rainfall

Multiplicative Random Cascades were first developed for studies of turbulence (Mandelbrot, 1999; Yaglom, 1966) with a motivation to have mathematical models, which produce time series that have statistically scale-invariant properties. In general, random cascade model for rainfall assumes a division of known rainfall total R_L occurring over an interval of time among a number of smaller intervals of fixed size, which implies a successive fine graining process that starts from an original, large-scale resolution R_L and continues till a target small-scale resolution is reached. The approach is based on scaling laws, which describe the scale-invariant properties or relationships that connect the statistical properties of rainfall for different timescales (Willems, 2012). The number of subintervals is defined by the branching number *b*, is set to 2, which is a redistribution of total rainfall in period *i* at a resolution r, $R_{i,r}$, between the amount associated with the first and last half respectively.

Here we implement a micro-canonical (exact conservation of mass in each cascade branching) cascade-based temporal disaggregation model as proposed by (Olsson, 1998), in which daily rainfall is disaggregated using a uniformly distributed generator, dependent on rainfall intensity and position of the rain sequence. The technique was later successfully implemented by (Güntner et al., 2001; Jebari et al., 2012; Rana et al., 2013) for temporal disaggregation of point rainfall and the development of IDF-curves from short-duration rainfall extremes. In the disaggregation process, each time interval (box) at a given resolution (for example 1 day) is split into two half of the original length (1/2 day). The procedure is continued as a cascade until the desired time resolution is reached, i.e., to ¼ day, then to 1/8 of a day and so on. Each step is termed as a cascade step, with cascade step 0 as the longest time period with only one box (*i.e.*, a day). The distribution of the volume between two sub-intervals (or smaller boxes) is computed by multiplication with the cascade weights ($W_{i,r}$), $0 \le W_{i,r} \le 1$ that assigns $W_{i,r} \bullet R_{i,r}$ to the first half of the period and $\left(1 - W_{i,r}\right) \bullet R_{i,r}$ to the next half. In each branching two possibilities exist: (1) $W_1 = 0$, $W_1 = 1$ (2) $0 < W_1 < 1$. The occurrence of (1) and (2) may be expressed in terms of probabilities, $P_{01} = P_{(1/0)}$ or $P_{(0/1)}$ = P($W_1 = 0$ or $W_1 = 1$) and $P_{xx} = P_{(x/x)}$ = P($0 < W_1 < 1$) = 1 - P_{01}.

Depending on the range of resolution involve, P_{01} either be assumed as resolution independent or parameterized as a scaling law: $Pr_{01}(r) = c_1 r^{c_2}$ where c_1 and c_2 are constants. The distribution of $W_{i,r}$ is termed as cascade generator, assumed to follow 1-parameter beta distribution (Olsson, 2012). Following (Olsson, 1998), the probabilities *P*, the probability distribution of cascade generator are assumed to be related to (1) position in rainfall sequence, and (2) rainfall volume. The wet boxes, with a rainfall volume $V > 0$, can be characterized by their position in the rainfall series: (1) the starting box, box preceded by a dry box ($V = 0$) and succeeded by a wet box ($V > 0$); (2) the enclosed box, box preceded and succeeded by wet boxes; (3) the ending box, box preceded by a wet box and succeeded by a dry box, and (4) the isolated box, box preceded and succeeded by dry boxes. On the other hand, based on volume dependence, if the volume is large then it is more likely that both halves of the subintervals

contribute to nonzero volume than if the volume is small. Following (Olsson, 1998), a partition into three volume classes (v_c = 1, 2, 3) was used, separated by percentiles 33rd and 67th of the values at the cascade step. Next, the variation of $P_{(x/x)}$ with volume is parameterised as, $P_{(x/x)} = a + b_m \bullet v_c$, where a is the intercept at v_c = 0, b_m is the mean slope of linear regression obtained from all cascade steps and v_c is volume class. For details about theory and implementation issues of MRC-based disaggregation tool, interested readers are requested to refer (Olsson, 1998).

For calibration and application of the disaggregated model from an original resolution R_l, two situations are considered: (1) when representative data at target resolution R_s is available; such as in this case, disaggregation from daily to hourly time steps, in which hourly data were available. Hence, parameters are calibrated over the actual resolution interval $R_l < r < R_s$ using 5 cascade steps. This implies, 5 successive "halving" from one day to generate 45-minute (2700 seconds) data. (2) When no representative high-resolution data are available, then parameters are estimated by coarse graining from lower resolution R_l to a higher resolution R_s by successive disaggregation steps. In this study, it is the disaggregation from daily to minute scales (or sub-hourly time steps), in which no sub-hourly data were available for any of the representative sites. In such case, parameters are calibrated from daily rainfall data using 7 cascade steps. This implies disaggregating by halving from 1-day to 11-minute 15 seconds (675 seconds) data. After calibration, Monte Carlo simulation is performed to gradually fine-grain the data and generate realizations at desired resolution, R_s. Since R_s in these cases are not directly achieved by exact resolution doubling from R_l, the target resolutions are obtained by geometric interpolation of the disaggregated model output at the final time step. In the next sub-sections, we demonstrate the performance of the MRC-based disaggregation tool using two different sets of observations.

S 1.1 Performance Evaluation of MRC-based disaggregation Tools for McMaster Weather Station Data

McMaster weather station is situated on a rooftop of McMaster University campus (43.26° N latitude and 79.92° W longitude, 114 m above sea level). We obtain daily, hourly and sub-hourly (15-min) rainfall information from the year 2010 to 2013, archived at McMaster Weather Station (MUWS; http://geomedia.mcmaster.ca/muws/weatherstation.html) website. The time slice was chosen based on data completeness and quality of available records.

We investigate two cases: (1) First, we calibrate the model using 1-year (2012 – 2013) hourly data, estimated model parameters, and then assess the performance of disaggregated model output using 2010 – 2013 observed data (2) Next, the observed 15-min time series was first aggregated to an hourly time step and then calibrated the model using aggregated hourly data. Then we compare disaggregated versus observed data at a 15-min temporal resolution for validation. Since we do not have 15-min daily rainfall information available for any of rain gauge locations, in both cases we calibrate the model using hourly data, which in turn gives us the opportunity to evaluate the performance of the disaggregation algorithm. In all cases, disaggregation was performed from daily time scales. Table S1 shows results of disaggregation experiments. We find a satisfactory performance between

observed and simulated model output, especially for the simulation of the percentage of zero values. However, we find a slight underestimation for simulated standard deviations of event volume and duration, whereas variance of mean inter-arrival time is overestimated.

Table S1. Comparison between observed and disaggregated 15-min and 1-hour time series for McMaster Weather Station Data

Time scale	Metrics	Observed	Simulated
Hourly (1-hr)	Zero values (%)	92	91.5
	Individual Rainfall volume (mean ± SD) mm	1.60 ± 2.95	1.45 ± 2.89
	Event volume (mean ± SD) mm	4.52 ± 10.40	4.50 ± 8.82
	Event duration (mean ± SD) hour	2.83 ± 3.51	3.09 ± 2.69
	Mean inter-arrival time between event, hour	36.82 ± 64.70	37.33 ± 65.10
	Mean annual maxima (mean ± SD) mm	39.97 ± 21.42	32.97 ± 12.33
Minute (15-min)	Zero values (%)	94.63	95.35
	Individual Rainfall volume (mean ± SD) mm	0.032 ± 1.20	0.031 ± 1.28
	Event volume (mean ± SD) mm	1.87 ± 6.28	1.96 ± 4.64
	Event duration (mean ± SD) hour	3.15 ± 5.77	2.94 ± 3.47
	Mean inter-arrival time between event, hour	59.54 ± 177.33	64.12 ± 182.01
	Mean annual maxima (mean ± SD) mm	20.4 ± 10.22	23.4 ± 12.70

S 1.2 Performance Evaluation of MRC-based disaggregation Tools for the Nine Locations in Southern Ontario

Figures S2 – S8 display disaggregation fit of the MRC-based tools for Toronto International Airport. Figure S2 presents variations of probability with volume classes, which often show substantial differences between the classes. The figure shows a linear relationship between precipitation, P and volume class, v_c changes with cascade steps. The mean regression line $a_p + b_m * vc$, (where a_p is the intercept and b_m is the slope) is shown as a dashed line with squares (Figure S2). Figures S3 and S4 show variations of intercepts with cascade steps. The probabilities P(x/x) and P(0/1) for daily to minute-scale disaggregation of four different type of boxes are shown in Tables S2 and S3. P(x/x) can be modelled assuming linear dependence on the cascade step, which is given as, $P(x/x) = a_p + b_m * vc$ with a_p is estimated as, $a_p = c_1 + c_2 * C_s$, where c_1 and c_2 are the slopes and the intercept of the linear regression. Since P(0/1) [or P(1/0)] is relatively independent of cascade step, it can be modeled using $a_p + b_m * vc$. Finally, P(0/1) [or P(1/0)] can be estimated as, P(0/1)=1-(P(x/x)+P(1/0) [or P(1/0)=1-(P(x/x)+P(0/1)]. The empirical histograms (the observed, shown in bars) and the fitted beta distributions (shown in lines) [Figures S5 and S6] show a good agreement in the overall fit. However, at higher cascade steps, i.e. finer time scales, the number of bins in the histogram is small, and the fits appear naturally uncertain.

Figure S7 shows time series of disaggregated versus observed annual maxima (the 15-min and 1-hour) for Toronto International Airport. The Quantile Mapping (QM) is employed to adjust occasional overestimation due to disaggregation. Except for a few extremes, we find a close agreement between the observed versus bias-corrected disaggregated annual maxima time series. For example, the algorithm overestimates the wettest event in Toronto (137.4 mm of rainfall) during Hurricane Hazel (1954) at hourly disaggregation time steps. However, due to lack of observation, we could not validate 15-min disaggregation model performance. Figure S8 shows 1-hour disaggregated model performance for the nine sites. Although we find evidence of occasional overestimations in disaggregated model output, the adjustment of extremes by QM could correct biases to some extent. However, the algorithm fails to correct extreme wet biases for London International Airport (during 1954), Trenton Airport (2000), Stratford WWTP (2002) and Fergus Shand Dam (2004 and 2006).

Table S2. Probabilities $P(x/x)$ as function of volume class, cascade step and type of box for daily to minute scale disaggregation

Volume Class/Cascade Step	**1**	**2**	**3**	**4**	**5**	**6**	**7**	Mean
1. Isolated Box (b_m = 0.223)								
1	0.001	0.030	0.051	0.048	0.073	0.082	0.031	0.045
2	0.244	0.254	0.266	0.299	0.249	0.289	0.267	0.27
3	0.509	0.505	0.510	0.513	0.512	0.470	0.420	0.49
2. Starting Box (b_m = 0.244)								
1	0.054	0.122	0.081	0.064	0.101	0.130	0.095	0.092
2	0.399	0.415	0.345	0.290	0.275	0.283	0.437	0.349
3	0.635	0.612	0.610	0.527	0.487	0.570	0.628	0.581
3. Enclosed Box (b_m = 0.263)								
1	0.423	0.357	0.272	0.157	0.097	0.156	0.194	0.236
2	0.862	0.817	0.703	0.510	0.335	0.345	0.463	0.576
3	0.951	0.938	0.875	0.736	0.585	0.562	0.685	0.762
4. Ending Box (b_m = 0.247)								
1	0.063	0.033	0.064	0.087	0.084	0.109	0.114	0.079
2	0.381	0.372	0.323	0.253	0.303	0.237	0.439	0.330
3	0.651	0.639	0.591	0.564	0.474	0.496	0.593	0.572

Table S3. Probabilities $P(0/1)$ as function of volume class, cascade step and type of box for daily to minute scale disaggregation

Volume Class/Cascade Step	1	2	3	4	5	6	7	Mean
1. Isolated Box (b_m = -0.113)								
1	0.446	0.449	0.494	0.466	0.427	0.482	0.50	0.466
2	0.369	0.388	0.370	0.330	0.400	0.369	0.30	0.361
3	0.257	0.248	0.236	0.242	0.241	0.259	0.193	0.239
2. Starting Box (b_m = -0.175)								
1	0.747	0.713	0.771	0.711	0.681	0.565	0.526	0.673
2	0.535	0.521	0.583	0.593	0.558	0.409	0.295	0.499
3	0.320	0.333	0.347	0.408	0.394	0.274	0.195	0.324
3. Enclosed Box (b_m = -0.113)								
1	0.272	0.283	0.330	0.387	0.400	0.405	0.374	0.350
2	0.078	0.103	0.159	0.220	0.350	0.348	0.253	0.216
3	0.027	0.031	0.077	0.154	0.204	0.224	0.158	0.125
4. Ending Box (b_m = -0.068)								
1	0.215	0.211	0.167	0.202	0.210	0.330	0.298	0.233
2	0.093	0.083	0.098	0.144	0.187	0.318	0.263	0.169
3	0.040	0.054	0.068	0.072	0.120	0.174	0.159	0.098

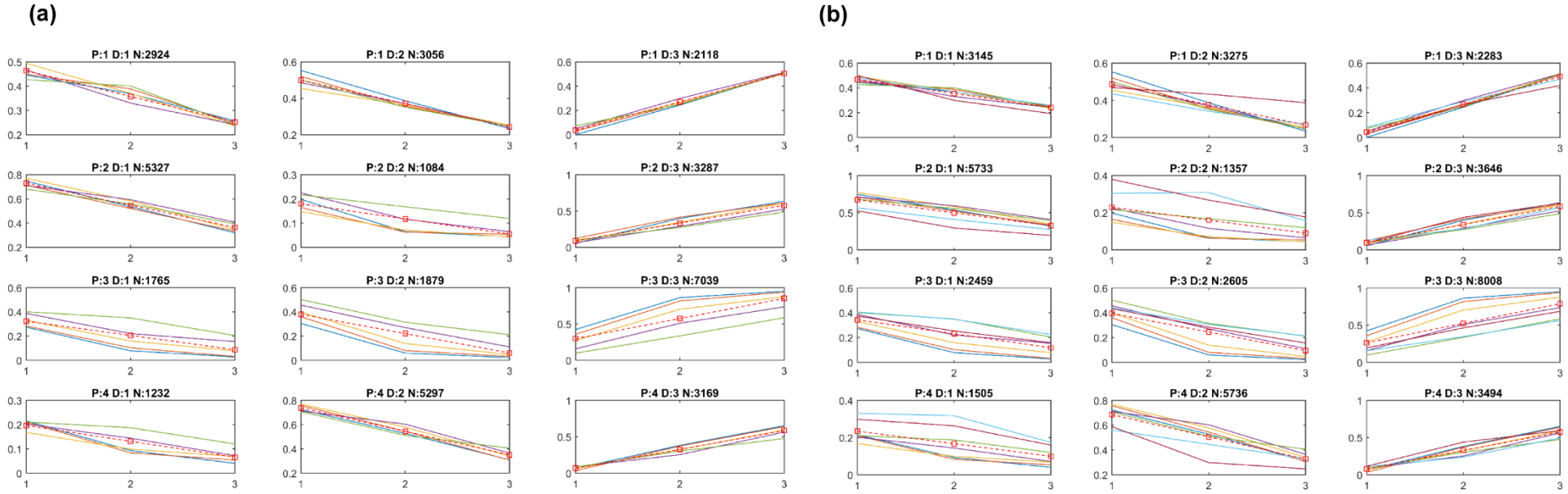

Figure S2. Variation of probabilities with volume class, v_c for Toronto International Airport at (***a***) daily to hourly-, and (***b***) daily to minute-time step disaggregation. X-axes show, v_c: 1 – small, 2 – medium, and 3 – large; Y-axes show, probabilities. *P* – denotes position type, 1: Isolated, 2: Starting, 3: Enclosed, and 4: Ending, *D* denotes division type (1: 0/1, 2: 1/0, 3: x/x), and N denotes the total number of periods for associated position and division type. The solid lines marked with different colors indicate different cascade steps. The dashed line with squares represents the mean of all cascade levels.

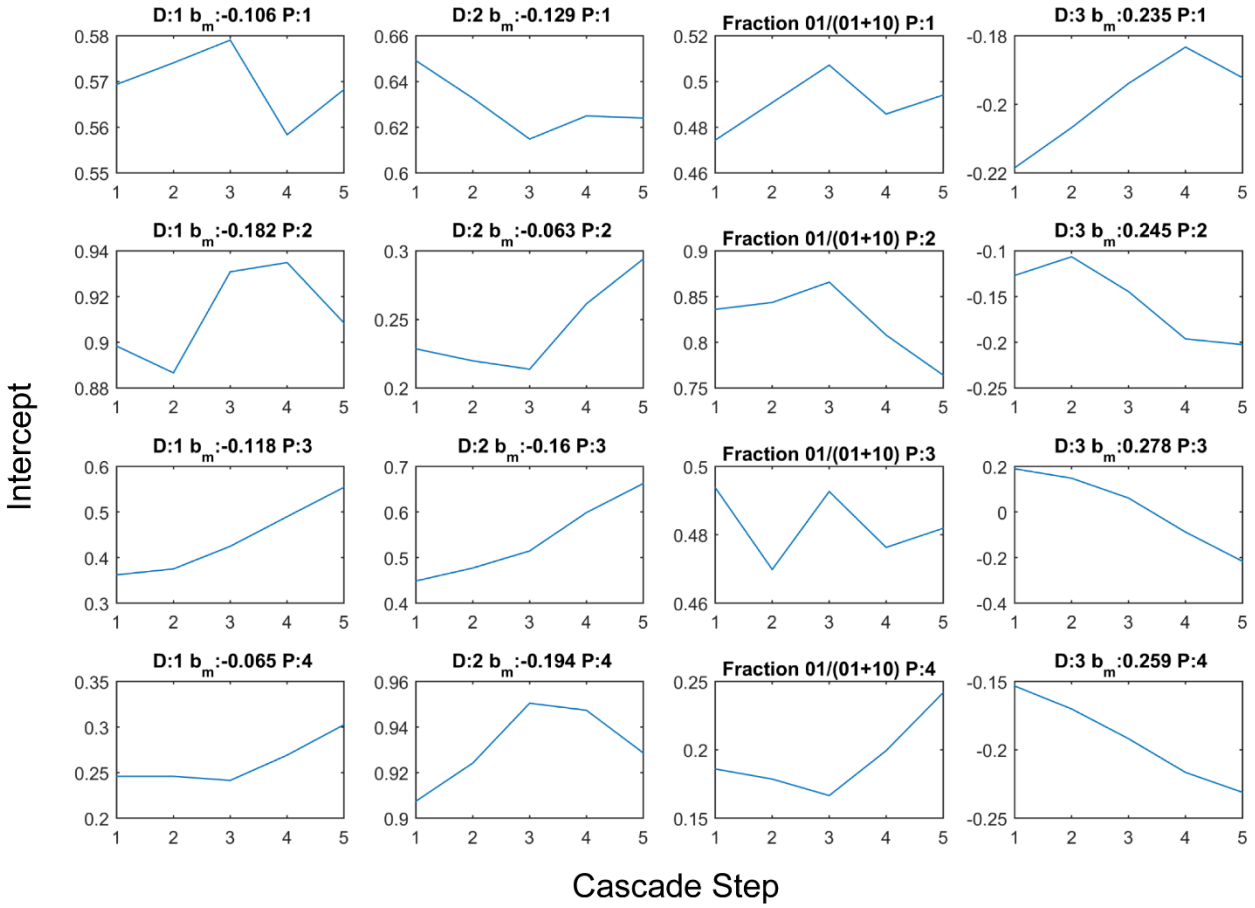

Figure S3. Variations of intercept, a_p with cascade steps for daily to hourly- time step disaggregation at Toronto International Airport. *P*- denotes position types, 1: Isolated, 2: Starting, 3: Enclosed and 4: Ending. *D* denotes division type (1: 0/1, 2: 1/0, 3: x/x), and b_m indicates the mean slope for corresponding position and division type, estimated from the fitted mean lines as in Figure S3. The third column indicates the fraction of 0/1-divisions of all "non-x/x-divisions" summed over volume classes.

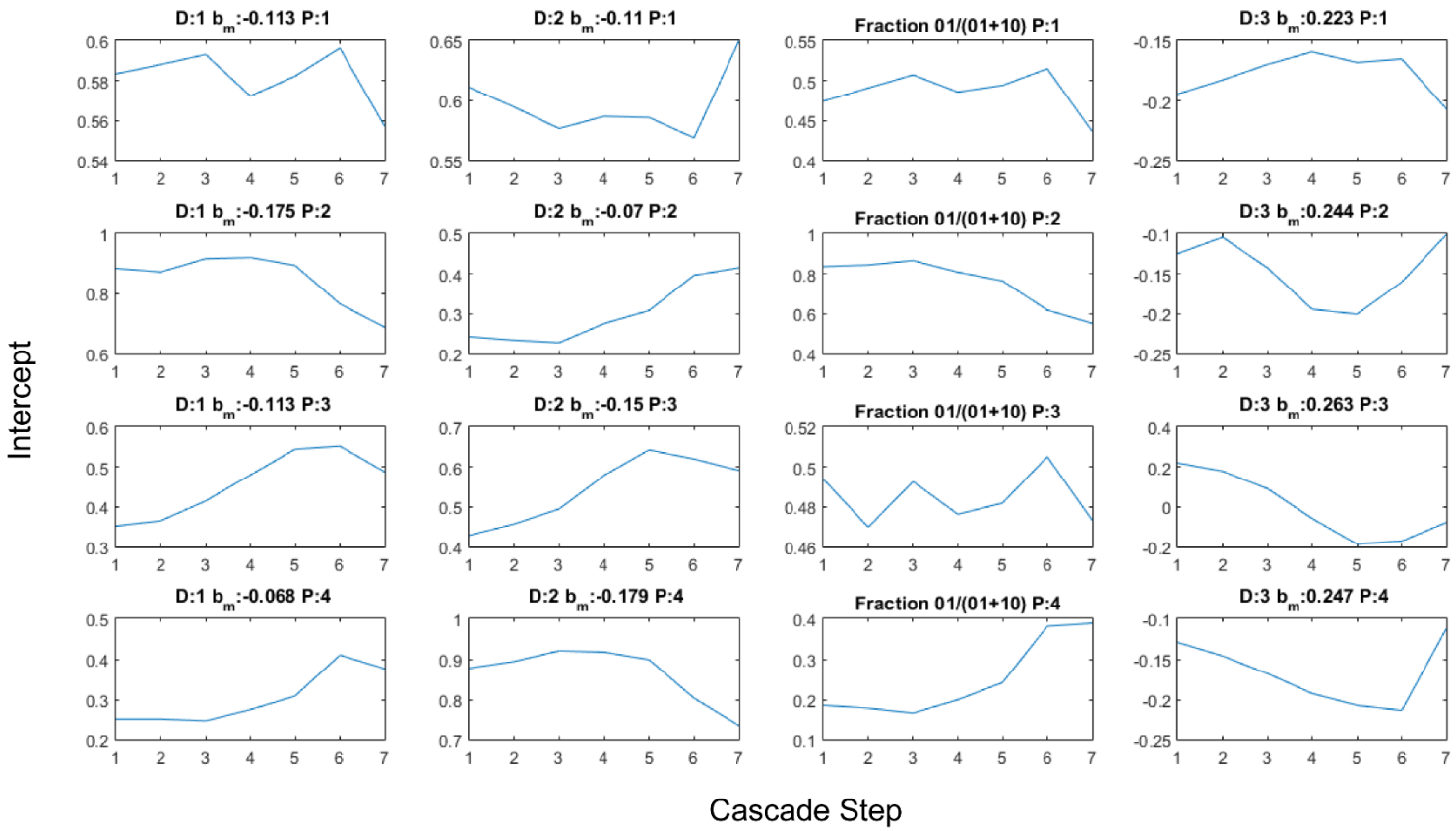

Figure S4. Same as in Figure S3 but for daily to minute- time step disaggregation.

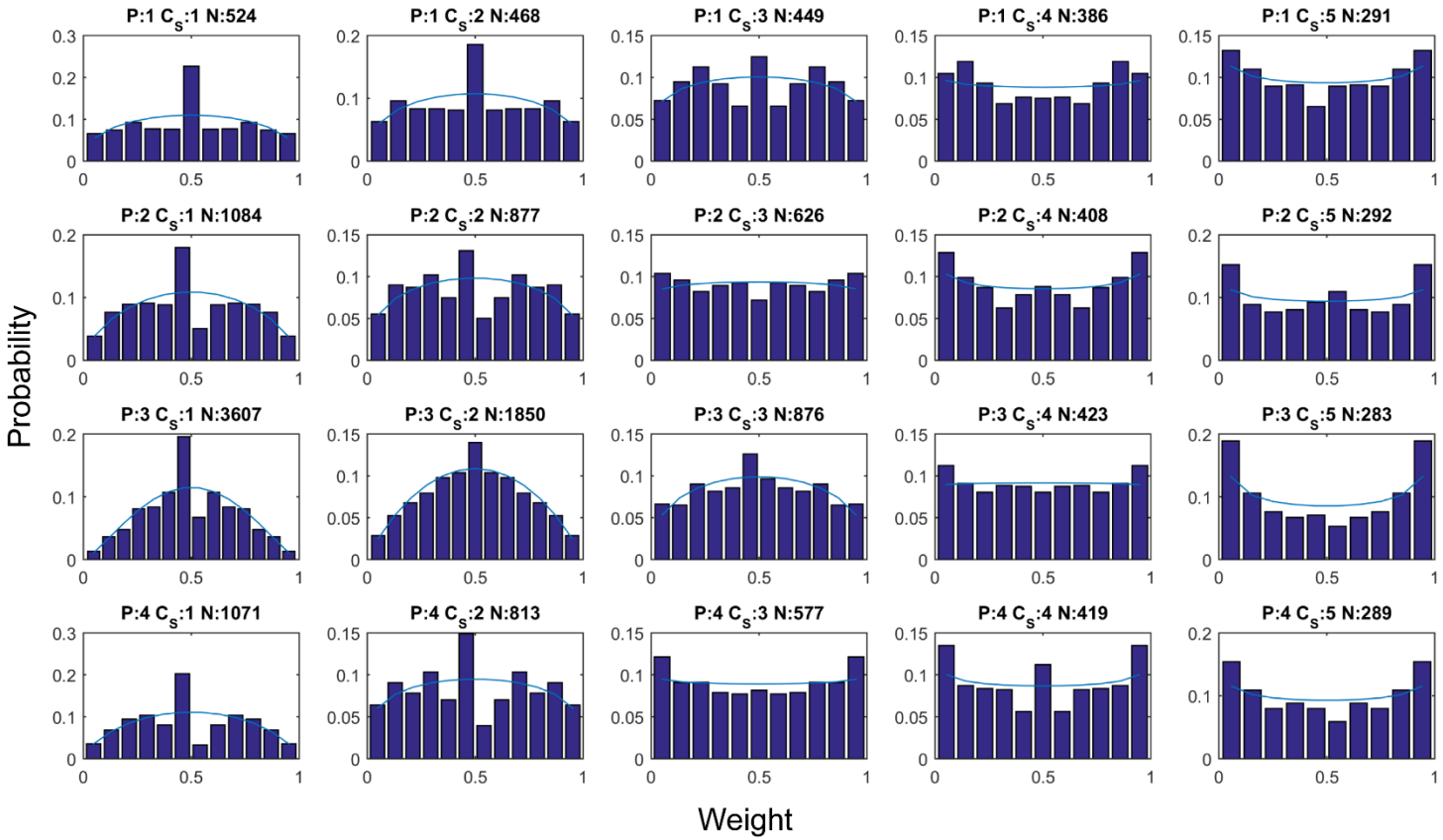

Figure S5. Variations of empirical *x/x*-distributions with cascade steps (bars) and fitted beta distributions (lines) for daily to hourly- time step disaggregation at Toronto International Airport. *P*-denotes position types, 1: Isolated, 2: Starting, 3: Enclosed and 4: Ending; C_s denotes cascade steps, for example, step 1 shows cascading from 32 to 16 minute, step 2 indicates from 16 to minute and so on. *N* indicates the total number of *x/x*- divisions for this position type and cascade level.

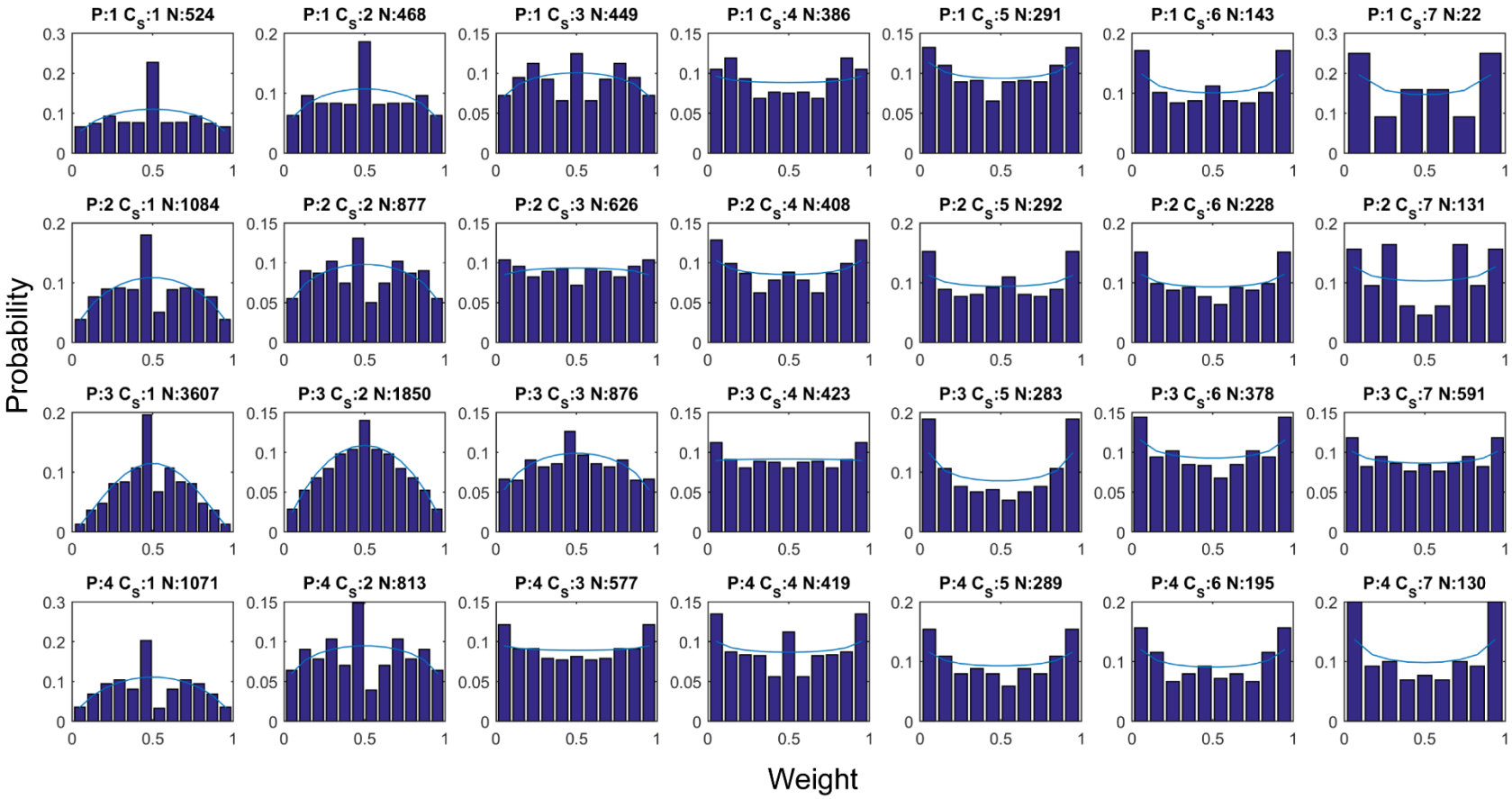

Figure S6. Same as in Figure S5 but for daily to minute- time step disaggregation.

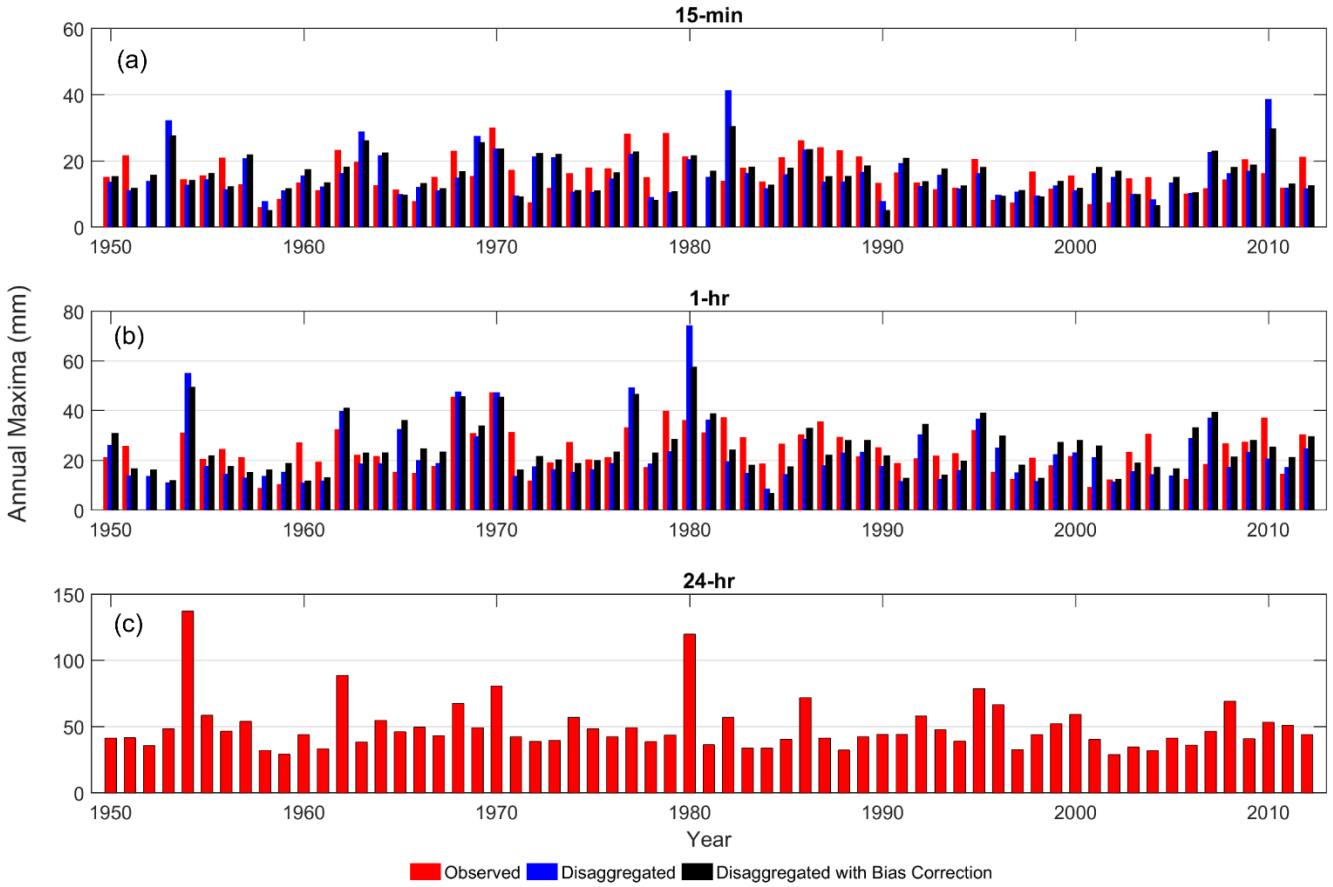

Figure S7. Comparison of observed versus disaggregated (**a**) sub-hourly and (**b)** hourly Annual Maximum (AM) rainfall values (Observed – red, disaggregated – blue and disaggregated and bias corrected - black) for Toronto International Airport. (**c**) The bottom panel shows AM daily rainfall values. A multiplicative random cascade-based disaggregation tool is used to disaggregate daily rainfall time series into a sub-hourly and hourly time step.

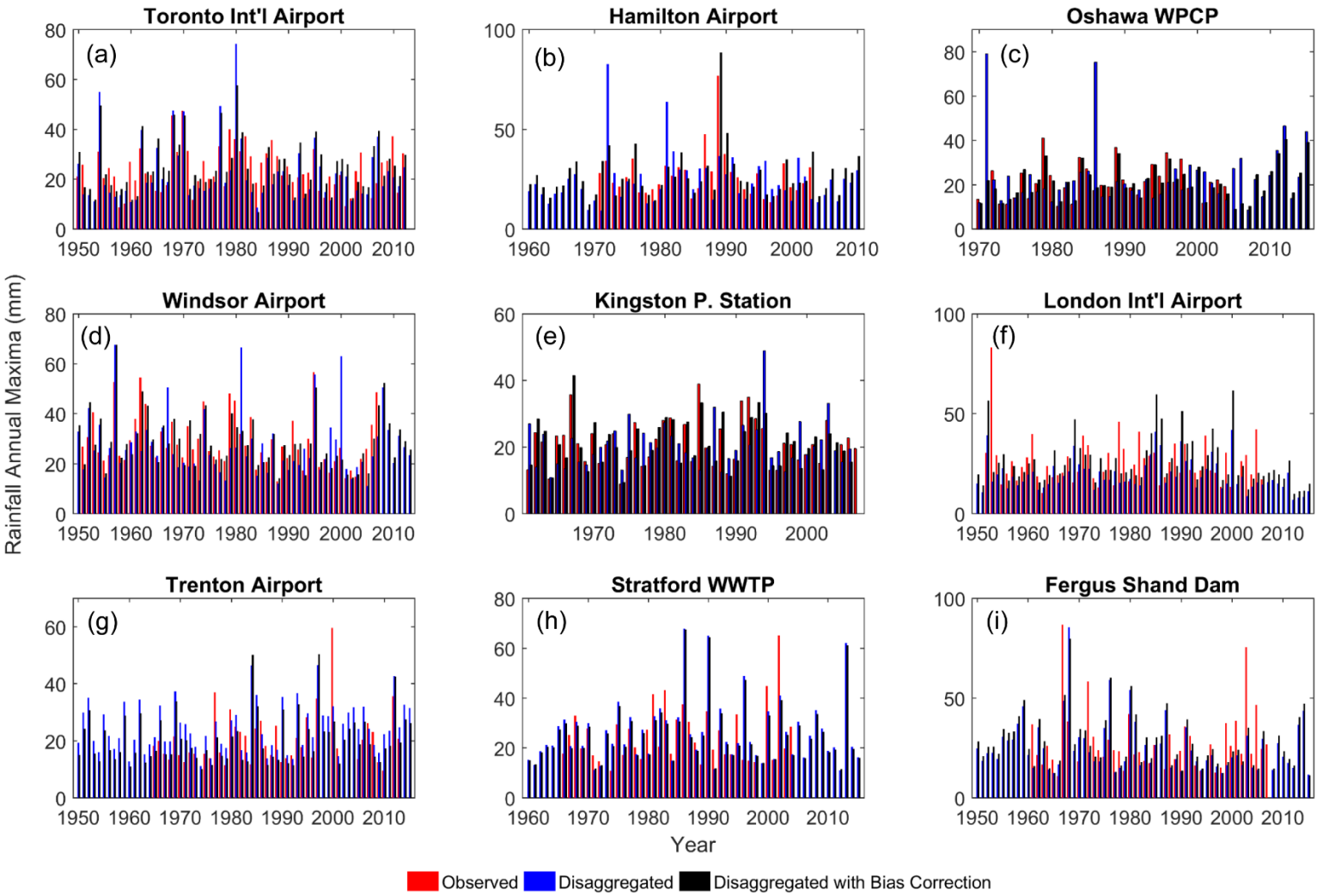

Figure S8. Comparison of observed versus disaggregated time series for 1-hour AMP across nine locations (**a-g**)

SI 2 Trend and Change point, Autocorrelation: Methods and Results

SI 2.1 Detection of Monotonic trend: Mann-Kendall Test Statistics

The null hypothesis H_0 of the test assumes that no temporal trend exists in the data and the alternate hypothesis H_1 assumes that a significant temporal trend (upward or downward) exists. The test statistic Z_{MK} is computed as (Hirsch et al., 1982)

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{Var(S)}} \; if \; S > 0; \\ 0, \quad if \; S = 0; \\ \dfrac{S+1}{\sqrt{Var(S)}} \; if \; S < 0 \end{cases} \tag{2.1}$$

where S is defined by,

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}\left(x_j - x_k\right) \qquad \operatorname{sgn}\left(x_j - x_k\right) = \begin{cases} 1, & if \; (x_j - x_k) > 0 \\ 0, & if \; (x_j - x_k) = 0 \\ -1, & if \; (x_j - x_k) < 0 \end{cases} \tag{2.2}$$

where x_j and x_k are the data points in time periods j and k ($j > k$) respectively, and n is number of observed data points. For $n \geq 10$, the test statistic S is approximately normally distributed with the mean of E(S) =0, and the variance of,

$$Var(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{i=1}^{g} t_i.(i).(i-1).(2i+5) \right] \tag{2.3}$$

where g is the number of tied groups and t_i is the number of data points in the i^{th} group.

SI 2.2 Detection of Abrupt or Step Change

SI 2.2.1 General Formulation of Sequential Change Point Test

When a sequence of random variables is divided into two segments represented by $x_1, ..., x_{t_0}$ and $x_{t_0+1}, x_{t_0+2}, ..., x_{t_0}$, if each segment has distribution functions, $F_1(x)$ and $F_2(x)$, where $F_1(x) \neq F_2(x)$, then change point is identified at t_0. Thus the null hypothesis of the test is "no change", $H_0 : \tau = T$ against the alternative of "change" $H_1 : 1 \leq \tau < T$. Suppose t points in the sequence of observations and we wish to test whether a change point has occurred at some point in the historical observation. For any fixed $k < t$ the hypothesis that a change point occurs at the k^{th} observation can be written as (Ross et al., 2011)

$$H_o \forall i \quad X_i \sim F_0, \quad H_1 : X_i \sim \begin{cases} F_0 & if\ i < k \\ F_1 & if\ i \geq k \end{cases} \tag{2.4}$$

Now a two-sample hypothesis test can be used to test for a change point at a location *k* using a test statistic $N_{k,t}$. The test statistic $N_{k,t}$ can be evaluated for all values of $1 < k < t$, and then using the maximum value, *i.e.*,

$$N_{\max,t} = \max_k N_{k,t} \tag{2.5}$$

Then, the hypothesis that no change has occurred before the t^{th} observation is rejected if $N_{\max,t} > h_t$ at a threshold h_t . The value τ of the change point is then the value of *k* which maximizes $N_{\max,t}$

Change Point in Location

- **Pettitt Change Point Test**

The test is based on following statistic (Serinaldi and Kilsby, 2016; Xie et al., 2014)

$$K_T = \max_{1 \leq t \leq T} \left| U_{t,T} \right| \tag{2.6}$$

Where $U_{t,T}$ the ranked sample as is defined in Eq. 2.2. The statistical significance of Pettitt test quantified by *p*-value is approximately evaluated as (Xie et al., 2014), $p = 2\exp\left(\frac{-6K_T^2}{T^2 + T^3}\right)$. Given a certain significance level α , if $p < \alpha$, we reject the null hypothesis and conclude that x_τ is a significant change point at level α . The analysis was performed using R statistical software with add-on package "trend".

- **Mann-Whitney Change Point Test**

Mann-Whitney (MW) test statistics detect shifts in location parameter in the time series. The test is based on the following observation: if there are *n* points spread over two samples *S* and *T*, containing n_s and n_t points respectively, $n = n_s + n_t$. The MW U statistic is then defined as (Ross, 2013)

$$U' = r_s - \frac{n_s\left(n_s + 1\right)}{2} \tag{2.7}$$

Where r_s is the sum of the ranks of the points in sample *S*. Under the null hypothesis that *S* and *T* have equal location parameter, the distribution of of *U* is independent of the distribution of the observations in two samples and its first two moments are

$$\mu_{U'} = n_s n_t / 2\,,\ \sigma_{U'}^2 = n_s n_t \left(1 + n_s + n_t\right)/12 \tag{2.8}$$

The standardized MW statistic, $U = \left|\left(U' - \mu_{U'}\right)/\sigma_{U'}\right|$ (2.9)

A nonparametric change point for location shifts is then defined as by replacing $N_{\max,t}$ with this *U* statistic.

Change Point in Scale: **Mood Test**

The Mood Test uses a test statistic which measures the extent to which the rank of each point deviates from its expected value (Ross et al., 2011)

$$M' = \sum_{x_i \in S} \left(r\left(x_i\right) - \left(n+1\right)/2 \right)^2 \tag{2.10}$$

Where $r\left(x_i\right)$ denotes the rank of x_i in the pooled sample. Similar to MW statistic, the distribution of Mood statistic is independent of the underlying random variables. The mean and variance of the Mood statistic is given as

$$\mu_{M'} = n_s\left(n^2-1\right)/12\,,\ \sigma^2_{M'} = n_s n_T\left(n+1\right)\left(n^2-4\right)/180 \tag{2.11}$$

Then the standardized statistic is given as (Ross et al., 2011)

$$M = \left|\left(M' - \mu_{M'}\right)/\sigma_{M'}\right| \tag{2.12}$$

Similar to MW statistic, a nonparametric change point for scale shifts is defined as by replacing $N_{\max,t}$ with this *M* statistic. Both Mann-Whitney and Mood tests were performed using R statistical software with add-on package "cpm".

SI 2.3 Non-parametric Trend Free Pre-Whitening (TPFW) for Correction of Autocorrelation

To correct the autocorrelation present in the data, we followed the non-parametric procedure of trend-free pre-whitening (TPFW) as suggested by the (Petrow and Merz, 2009):

- At first, the trend of annual maxima time series of a particular duration is estimated by the non-parametric trend slope estimator, β as suggested by (Sen, 1968), which is the median of all pair wise slopes (b_i) in the time series of length (*N*):

$$\beta = \begin{cases} b_{(N+1)/2} & if\ N\ is\ odd \\ \dfrac{1}{2}\left(b_{N/2} + b_{(N+1)/2}\right) & if\ N\ is\ even \end{cases} \tag{2.13}$$

Where, $b_i = \dfrac{\left(x_j - x_k\right)}{j-k}$, $i = 1, 2, ..., N\ \ and\ \ j > k$, and x_j and x_k are the data points in time periods *j* and *k* (*j* > *k*) respectively. Hence, if there are *n* values of data in the time series, it results into as many as $N = {}^{n}C_2$ number of b_i values.

- Secondly, the computed trend is removed from the original series:

$$Y_t = X_t - \beta \times t \tag{2.14}$$

Where, X_t is the original time series and *t* is the time.

- Next, the lag1- autocorrelation [ρ , using MATLAB function 'autocorr ()'], is computed from Y_t. If no statistically significant (significance is checked at 5 and 10% significance level) autocorrelation is found, the trend and change-point detection algorithms are directly applied to the original time series. Otherwise, the lag-1 autocorrelation is removed from the time series:

$$Y_t' = Y_t - \rho \times Y_{t-1} \tag{2.15}$$

- Finally, the removed trend in the first step is added back into the time series free from trend and autocorrelation.

$$Y_t'' = Y_t' + \beta \times t \tag{2.16}$$

The resulting time series Y'' includes the original trend but free from autocorrelation.

Table S4. Detection of trends and nonstationarity in Toronto Pearson International Airport

Time Slice	Ljung-Box Test	Augmented Dickey-Fuller Test		KPSS Test		Mann-Kendall Trend Test			Priestley-Subbarao Test	Pettitt Test	Mann-Whitney Change Point Test			Mood Change Point Test		
	p-value	Test statistics	p-value	Test statistics	p-value	MK_Z	β	p-value	p-value	p-value	Year	Test statistics	Threshold	Year	Test statistics	Threshold
15-min	0.014**	-0.326	0.53**	0.120	0.098*	-0.71	-0.23	0.48	0.383	0.141	x	x	x	x	x	x
30-min	0.027**	-0.177	0.58**	0.123	0.092*	-0.52	-0.25	0.61	0.283	0.179	x	x	x	x	x	x
1-hr	0.805	-0.077	0.62**	0.138	0.064*	0.39	0.22	0.69	0.409	0.368	x	x	x	x	x	x
2-hr	0.920	0.043	0.66**	0.129	0.081*	-0.07	-0.05	0.94	0.143	0.564	x	x	x	x	x	x
6-hr	0.959	0.051	0.67**	0.060	> 0.10	-0.61	-0.48	0.54	0.025	0.733	x	x	x	x	x	x
12-hr	0.965	0.106	0.67**	0.065	> 0.10	-0.55	-0.49	0.58	$2.0e^{-4}$**	0.659	x	x	x	x	x	x
24-hr	0.885	0.106	0.69**	0.046	> 0.10	-0.06	-0.06	0.95	0.046**	-†	x	x	x	x	x	x

** and * indicate statistically significant at 5% and 10% significance levels, 'x' denotes no change point is detected using Mann-Whitney and Mood tests respectively. β indicates slope per decade calculated using Theil-Sen method. *P-values* larger than 0.1 in KPSS test indicates test statistics are non-significant, whereas *p-values* smaller than 0.01 are considered to be highly significant. The standardized Mann-Kendall test statistic (MK_Z) is positive (negative) with an increasing (decreasing) trend, and statistically significant at 5% and 10% significance levels when $|MK_Z| > 1.96$ and $|MK_Z| > 1.64$ respectively. Change point tests are performed at 10% significance level. †*p*-values are not reported due to analytical intractability.

Table S4.1 Detection of trends and nonstationarity after performing TFPW in Toronto Pearson International Airport

Time Slice	Ljung-Box Test	Augmented Dickey-Fuller Test		KPSS Test		Mann-Kendall Trend Test			Priestley-Subbarao Test	Pettitt Test	Mann-Whitney Change Point Test			Mood Change Point Test		
	p-value	Test statistics	p-value	Test statistics	p-value	MK_Z	β	p-value	p-value	p-value	Year	Test statistics	Threshold	Year	Test statistics	Threshold
15-min	0.258	-0.60**	0.43	0.089	>0.10	-0.17	-0.04	0.87	0.504	0.531	x	x	x	x	x	x
30-min	0.243	-0.35**	0.52	0.098	>0.10	0.07	0.05	0.94	0.498	0.777	x	x	x	x	x	x

Table S5. Detection of trends and nonstationarity in Hamilton Airport

Time Slice	Ljung-Box Test	Augmented Dickey-Fuller Test		KPSS Test		Mann-Kendall Trend Test			Priestley-Subbarao Test	Pettitt Test	Mann-Whitney Change Point Test			Mood Change Point Test		
	p-value	Test statistics	p-value	Test statistics	p-value	MK_Z	β	p-value	p-value	p-value	Year	Test statistics	Threshold	Year	Test statistics	Threshold
15-min	0.173	-0.454	0.48**	0.033	>0.10	-0.65	-0.24	0.516	0.406	0.438	x	x	x	x	x	x
30-min	0.911	-0.428	0.49**	0.047	>0.10	-1.186	-0.45	0.236	2.56e-5**	0.402	x	x	x	x	x	x
1-hr	0.918	-0.170	0.58**	0.082	>0.10	0.244	0.28	0.807	0.002**	0.870	x	x	x	x	x	x
2-hr	0.735	-0.176	0.58**	0.154	0.043**	1.023	0.75	0.306	8.03e-6**	0.300	x	x	x	x	x	x
6-hr	0.099*	-0.138	0.60**	0.193	0.019**	-0.097	-0.11	0.922	3.49e-9**	0.216	x	x	x	x	x	x
12-hr	0.150	-0.055	0.63**	0.183	0.023**	-0.122	-0.14	0.903	2.78e-10**	0.506	x	x	x	x	x	x
24-hr	0.059*	-0.035	0.63**	0.307	0.010**	0.309	0.39	0.758	1.82e-7**	0.199	x	x	x	x	x	x

Table S5.1 Detection of trends and nonstationarity after performing TFPW in Hamilton Airport

Time Slice	Ljung-Box Test	Augmented Dickey-Fuller Test		KPSS Test		Mann-Kendall Trend Test			Priestley-Subbarao Test	Pettitt Test	Mann-Whitney Change Point Test			Mood Change Point Test		
	p-value	Test statistics	p-value	Test statistics	p-value	MK_Z	β	p-value	p-value	p-value	Year	Test statistics	Threshold	Year	Test statistics	Threshold
6-hr	0.108	-0.211	0.570**	0.197	0.02**	-0.02	-0.03	0.981	2.83e-6**	0.227	x	x	x	x	x	x
24-hr	0.282	-0.202	0.573**	0.282	0.01**	0.318	0.41	0.751	9.78e-4**	0.316	x	x	x	x	x	x

Table S6. Detection of trends and nonstationarity in Oshawa WPCP

Time Slice	Ljung-Box Test	Augmented Dickey-Fuller Test		KPSS Test		Mann-Kendall Trend Test			Priestley-Subbarao Test	Pettitt Test	Mann-Whitney Change Point Test			Mood Change Point Test		
	p-value	Test statistics	p-value	Test statistics	p-value	MK_Z	β	p-value	p-value	p-value	Year	Test statistics	Threshold	Year	Test statistics	Threshold
15-min	0.986	-0.152	0.59**	0.044	>0.10	0.93	2.12	0.35	0.18	0.89	x	x	x	x	x	x
30-min	0.620	-0.251	0.55**	0.045	>0.10	0.73	1.02	0.47	0.92	-	x	x	x	1974	2.62*	2.59
1-hr	0.994	-0.065	0.62**	0.081	>0.10	1.29	1.5	0.20	0.57	0.562	x	x	x	x	x	x
2-hr	0.938	-0.116	0.61**	0.08	>0.10	1.68	0.82	0.09*	0.005**	0.259	2009	2.64*	2.58	x	x	x
6-hr	0.924	-0.311	0.53**	0.075	>0.10	1.49	0.3	0.14	0.027**	0.248	x	x	x	x	x	x
12-hr	0.998	-0.928	0.31**	0.048	>0.10	1.73	0.13	0.08*	$2.4e^{-4}$**	0.321	1975	2.59*	2.67	x	x	x
24-hr	0.990	-1.716	0.08*	0.040	>0.10	2.20	0.14	0.03**	$2.4e^{-3}$**	0.056*	1983	2.90*	2.58	x	x	x

Table S7. Detection of trends and nonstationarity in Windsor Airport

Time Slice	**Ljung-Box Test**	**Augmented Dickey-Fuller Test**		**KPSS Test**		**Mann-Kendall Trend Test**		**Priestley-Subbarao Test**	**Pettitt Test**	**Mann-Whitney Change Point Test**			**Mood Change Point Test**		
	p-value	Test statistics	p-value	Test statistics	p-value	MK_Z	β	p-value	p-value	Year	Test statistics	Threshold	Year	Test statistics	Threshold
15-min	0.595	-0.237	0.56**	0.108	>0.10	-1.29	-0.42	0.198	0.156	0.141	x	x	x	x	x	x
30-min	0.749	-0.363	0.51**	0.074	>0.10	-1.59	-0.77	0.111	0.845	0.093*	x	x	x	x	x	x
1-hr	0.462	-0.366	0.51**	0.083	>0.10	-1.52	-0.9	0.129	0.851	0.103	x	x	x	x	x	x
2-hr	0.614	-0.222	0.56**	0.127	0.085*	-0.82	-0.74	0.414	0.529	0.281	2005	2.99*	2.63	x	x	x
6-hr	0.504	-0.098	0.61**	0.128	0.084*	-0.47	-0.45	0.656	0.256	0.300	2005	2.95*	2.63	x	x	x
12-hr	0.969	-0.238	0.56**	0.065	>0.10	-1.48	-1.4	0.138	0.296	0.141	x	x	x	x	x	x
24-hr	0.414	-0.177	0.58**	0.041	>0.10	-0.035	-0.01	0.972	0.976	0.940	x	x	x	x	x	x

Table S8. Detection of trends and nonstationarity in Kingston P. Station

Time Slice	Ljung-Box Test	Augmented Dickey-Fuller Test		KPSS Test		Mann-Kendall Trend Test			Priestley-Subbarao Test	Pettitt Test	Mann-Whitney Change Point Test			Mood Change Point Test		
	p-value	Test statistics	p-value	Test statistics	p-value	MK_Z	β	p-value	p-value	p-value	Year	Test statistics	Threshold	Year	Test statistics	Threshold
15-min	0.257	-0.502	0.46**	0.024	>0.10	0.63	0.2	0.53	0.040	-	x	x	x	x	x	x
30-min	0.094	-0.485	0.47**	0.027	>0.10	0.18	0.12	0.85	0.322	-	x	x	x	x	x	x
1-hr	0.371	-0.316	0.53**	0.039	>0.10	0.16	0.12	0.88	0.094*	-	x	x	x	1996	2.73*	2.6
2-hr	0.124	-0.306	0.53**	0.084	>0.10	0.05	0.03	0.96	0.333	-	x	x	x	x	x	x
6-hr	0.273	-0.330	0.53**	0.168	0.032**	-0.22	-0.36	0.83	0.0009**	0.397	x	x	x	x	x	x
12-hr	0.886	-0.298	0.54**	0.102	>0.10	0.56	0.39	0.58	$5.53e^{-6}$**	0.347	x	x	x	x	x	x
24-hr	0.346	-0.258	0.55**	0.131	0.078*	0.68	0.75	0.50	$1.06e^{-5}$**	0.514	1965	2.62*	2.60	x	x	x

Table S9. Detection of trends and nonstationarity in London International Airport

Time Slice	Ljung-Box Test	Augmented Dickey-Fuller Test		KPSS Test		Mann-Kendall Trend Test			Priestley-Subbarao Test	Pettitt Test	Mann-Whitney Change Point Test			Mood Change Point Test		
	p-value	Test statistics	p-value	Test statistics	p-value	MK_Z	β	p-value	p-value	p-value	Year	Test statistics	Threshold	Year	Test statistics	Threshold
15-min	0.344	-0.662	0.41**	0.179	0.024**	-0.549	-0.23	0.583	0.092*	0.253	2006	2.94*	2.63	2006	3.09*	2.66
30-min	0.288	-0.598	0.43**	0.106	>0.10	-1.422	-0.64	0.155	0.019**	0.240	2005	2.87*	2.63	x	x	x
1-hr	0.783	-0.487	0.47**	0.117	>0.10	-1.173	-0.66	0.241	0.031**	0.359	2011	3.06*	2.63	2011	2.96*	2.66
2-hr	0.845	-0.552	0.45**	0.135	0.070*	-1.566	-1.03	0.117	0.224	0.216	2011	3.14*	2.63	2011	3.15*	2.66
6-hr	0.700	-0.418	0.49**	0.071	>0.10	-1.893	-1.11	0.058*	0.199	0.161	2011	3.01*	2.63	2011	2.80*	2.66
12-hr	0.587	-0.347	0.52**	0.090	>0.10	-2.142	-1.47	0.032**	0.755	0.093*	2011	2.98*	2.63	x	x	x
24-hr	0.733	-0.414	0.50**	0.223	0.010**	-2.297	-2.45	0.022**	0.909	0.024**	2005	3.46*	2.63	x	x	x

Table S10. Detection of trends and nonstationarity in Trenton Airport

Time Slice	Ljung-Box Test	Augmented Dickey-Fuller Test		KPSS Test		Mann-Kendall Trend Test			Priestley-Subbarao Test	Pettitt Test	Mann-Whitney Change Point Test			Mood Change Point Test		
	p-value	Test statistics	p-value	Test statistics	p-value	MK_Z	β	p-value	p-value	p-value	Year	Test statistics	Threshold	Year	Test statistics	Threshold
15-min	0.263	-0.398	0.50**	0.051	>0.10	0.708	0.16	0.479	$2.65e^{-6}$**	0.893	x	x	x	1979	3.66*	2.66
30-min	0.161	-0.325	0.53**	0.034	>0.10	1.732	0.68	0.083*	0.0033**	0.264	x	x	x	1978	3.16*	2.66
1-hr	0.373	-0.328	0.53**	0.061	>0.10	1.776	0.87	0.076*	0.0013**	0.098*	x	x	x	x	x	x
2-hr	0.423	-0.341	0.52**	0.063	>0.10	2.313	1.27	0.021**	$5.1e^{-5}$**	0.053*	1994	2.81*	2.63	x	x	x
6-hr	0.638	-0.244	0.56**	0.078	>0.10	2.031	1.12	0.042**	0.005**	0.092*	1994	2.66*	2.63	x	x	x
12-hr	0.295	-0.105	0.61**	0.204	0.014**	2.009	1.64	0.045**	0.044**	0.022**	1996	3.25*	2.63	x	x	x
24-hr	0.321	-0.131	0.60**	0.238	0.010**	2.031	1.82	0.042**	0.39	0.014**	1994	3.37*	2.63	x	x	x

Table S11. Detection of trends and nonstationarity in Stratford WWTP

Time Slice	**Ljung-Box Test**	**Augmented Dickey-Fuller Test**		**KPSS Test**		**Mann-Kendall Trend Test**			**Priestley-Subbarao Test**	**Pettitt Test**	**Mann-Whitney Change Point Test**			**Mood Change Point Test**		
	p-value	Test statistics	p-value	Test statistics	p-value	MK_Z	β	p-value	p-value	p-value	Year	Test statistics	Threshold	Year	Test statistics	Threshold
15-min	0.070	-0.321	0.53**	0.068	>0.10	-0.233	-0.07	0.816	0.143	-	1994	3.37*	2.63	x	x	x
30-min	0.473	-0.375	0.51**	0.039	>0.10	-0.297	-0.13	0.767	0.004**	-	x	x	x	2011	3.23*	2.63
1-hr	0.857	-0.230	0.56**	0.077	>0.10	0.417	0.26	0.677	0.013**	-	x	x	x	x	x	x
2-hr	0.149	-0.263	0.55**	0.055	>0.10	0.389	0.3	0.697	0.016**	-	x	x	x	1967	2.73*	2.63
6-hr	0.375	-0.134	0.60**	0.062	>0.10	-0.459	-0.51	0.646	$3.9e^{-4}$**	0.802	x	x	x	x	x	x
12-hr	0.336	-0.186	0.58**	0.058	>0.10	0.82	-0.67	0.412	$8.9e^{-5}$**	0.254	x	x	x	x	x	x
24-hr	0.789	-0.208	0.57**	0.056	>0.10	-1.442	-1.43	0.149	$8.7e^{-4}$**	0.360	x	x	x	x	x	x

Table S12. Detection of trends and nonstationarity in Fergus Shand Dam

Time Slice	Ljung-Box Test	Augmented Dickey-Fuller Test		KPSS Test		Mann-Kendall Trend Test			Priestley-Subbarao Test	Pettitt Test	Mann-Whitney Change Point Test			Mood Change Point Test		
	p-value	Test statistics	p-value	Test statistics	p-value	MK_Z	β	p-value	p-value	p-value	Year	Test statistics	Threshold	Year	Test statistics	Threshold
15-min	0.955	-0.267	0.55^{**}	0.184	0.022^{**}	-0.99	-0.43	0.322	0.049^{**}	0.25	x	x	x	x	x	x
30-min	0.935	-0.204	0.57^{**}	0.172	0.028^{**}	-0.44	-0.27	0.658	0.015^{**}	0.33	x	x	x	x	x	x
1-hr	0.879	-0.494	0.47^{**}	0.068	>0.10	-1.00	-0.72	0.317	0.004^{**}	0.57	x	x	x	x	x	x
2-hr	0.849	-0.540	0.45^{**}	0.055	>0.10	-0.92	-0.69	0.358	0.002^{**}	0.67	x	x	x	x	x	x
6-hr	0.967	-0.318	0.53^{**}	0.104	>0.10	-0.26	-0.29	0.795	0.053^{**}	0.77	x	x	x	x	x	x
12-hr	0.018^{**} (0.033^{**})	-0.116	0.61^{**}	0.035	>0.10	-0.45	-0.28	0.652	0.255	-	x	x	x	x	x	x
24-hr	$1.5e^{-4**}$ (0.011^{**})	-0.218	0.57^{**}	0.035	>0.10	-1.34	-1.2	0.180	0.216	0.53	x	x	x	x	x	x

[1]Values within first Brackets indicate p-values of Ljung-Box Test after performing two-successive TFPWs.

SI 3 GEV Fitting: Methods and Results

SI 3.1 Estimation of GEV Parameters

GEV parameters are estimated using Bayesian Inference. For this, a Bayesian analysis is performed by imposing a prior distribution on the parameters. We estimated parameters using Bayesian Inference (BI) coupled with Differential Evaluation Markov Chain (DE-MC) simulation as in (Cheng and AghaKouchak, 2014; Cheng et al., 2014). This approach combines knowledge from a prior distribution and the observation vector $\boldsymbol{y} = \{y_t\}_{t=1:N}$ (*i.e.*, annual maxima rainfall, where *N* denotes the number of annual maxima in the observation vector $\boldsymbol{Y}$) into the posterior distribution of parameters $\boldsymbol{\omega} = \{\mu, \sigma, \zeta\}$ and $\boldsymbol{\lambda} = \{\mu_1, \mu_0, \sigma_1, \sigma_0, \zeta\}$ assuming stationary and nonstationary conditions respectively. The priors for the location and scale parameters are non-informative normal distributions, whereas prior for the shape parameter is a normal distribution with a standard deviation of 0.3 (Renard et al., 2013) as used in the default option in the NEVA package. The Bayes theorem for estimation of GEV parameters under stationarity assumption can be expressed as (Cheng et al., 2014; Renard et al., 2013)

$$p(\boldsymbol{\omega}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{\omega})p(\boldsymbol{\omega}) = \prod_{t=1}^{N} p(\boldsymbol{y}|\boldsymbol{\omega})p(\boldsymbol{\omega}) \tag{3.1}$$

Under the assumption of nonstationarity, the equation (3.1) is given as (Cheng et al., 2014; Renard et al., 2013),

$$p\left(\boldsymbol{\lambda}|\boldsymbol{y},\boldsymbol{x}\right) \propto p\left(\boldsymbol{y}|\boldsymbol{\lambda},\boldsymbol{x}\right) p\left(\boldsymbol{\lambda}|\boldsymbol{x}\right) \tag{3.2}$$

$$p\left(\boldsymbol{y}|\boldsymbol{\lambda},\boldsymbol{x}\right) = \prod_{t=1}^{N} p\left(\boldsymbol{y}_t|\boldsymbol{\lambda},\boldsymbol{x}(t)\right) = \prod_{t=1}^{N} p\left(\boldsymbol{y}_t|\mu(t),\sigma(t),\zeta\right) \tag{3.3}$$

Where, $\boldsymbol{x}(t)$ denote covariates under nonstationarity. The $p(\boldsymbol{\omega}|\boldsymbol{y})$ and $p\left(\boldsymbol{\lambda}|\boldsymbol{y},\boldsymbol{x}\right)$ in Eqns. 3.1 and 3.2 indicate resulting posterior distributions whereas $p(\boldsymbol{y}|\boldsymbol{\omega})$ and $p\left(\boldsymbol{y}_t|\boldsymbol{\lambda},\boldsymbol{x}(t)\right)$ denote likelihood functions. Since the posterior distributions of model parameters are, in general, analytically intractable, the DE-MC is integrated with BI to generate a large number of realizations (*Cheng et al.*, 2014). DE-MC is an adaptive Monte Carlo Markov Chain (MCMC) algorithm (Ter Braak and Vrugt, 2008; Ter Braak, 2006), in which multiple chains (here, we fix chain length '*n*' as 5) are run in parallel. The resulting MC simulations are then run to an equilibrium. It is a standard practice to discard the initial iterations (often referred to as the *burn-in* period) of simulated samples since they are strongly influenced by starting values and do not provide usable information of the target distribution. Here we run DE-MC simulations for 3000 iterations and kept the 2001-3000th iterations of each chain. The convergence of MC simulation is checked by the "potential scale reduction

factor ($\widehat{R}$)" as in (Gelman et al., 2011), which suggests the value of $\hat{R}$ should remain below the threshold value of 1.1. The post burn-in random draws from posterior distribution are then used to construct predictive distributions. For annual maxima time series of each duration, the mean and associated 95% credibility interval of parameters ($\mu(t), \sigma(t)$) are derived by computing 50th (the median), 2.5th and 97.5th (bounds) percentiles of post *burn-in* random draw (for example, 50th percentile of $\mu(t_1), \ldots, \mu(t_{100})$). The derived model parameters are then used to compute corresponding design rainfall quantiles at *T*-year return period. For more details of BI and parameter estimation by DE-MC simulation, interested readers are requested to refer (Gelman et al., 2014; Renard et al., 2013).

SI 3.2 Model Selection

SI 3.2.1 Bayes Factor

To evaluate the fit of the stationary model (null model, *M*1) relative to the nonstationary model (alternative model, *M*2) Bayes factor is computed based on the posterior distributions of sampled parameters, which is given as (Cheng et al., 2014)

$$\gamma = \frac{\Pr\left(\boldsymbol{X} \mid M_1\right)}{\Pr\left(\boldsymbol{X} \mid M_2\right)} = \frac{\int \Pr\left(\boldsymbol{\omega} \mid M_1\right) \Pr\left(\boldsymbol{X} \mid \boldsymbol{\omega}, M_1\right) d\boldsymbol{\omega}}{\int \Pr\left(\boldsymbol{\lambda} \mid M_2\right) \Pr\left(\boldsymbol{X} \mid \boldsymbol{\lambda}, M_2\right) d\boldsymbol{\lambda}} \tag{3.4}$$

Where $\boldsymbol{X}$ denotes input data, $\boldsymbol{\omega}$ and $\boldsymbol{\lambda}$ denote model parameters as described in the previous section. The term $\Pr\left(\boldsymbol{X} \mid M\right)$ can be expressed using Monte Carlo integration estimation as described in (Kass and Raftery, 1995). A value of $\gamma < 1$ indicates nonstationary model, the nonstationary model, *M*2 fits the data better than the stationary model, *M*1.

SI 3.2.2 Akaike Information Criterion (AIC) for Small Samples

The AIC (Anderson et al., 1994; Bozdogan, 2000) is defined as follows:

$$AIC\left(m\right) = n\log\left(MSE\right) + 2m \tag{3.5}$$

Where *n* is the number of observations, *m* denotes the number of fitted model parameters. *MSE* is the Mean Square Error of the fitted distribution against empirical distribution, which is expressed as (Dawson et al., 2007; Hu, 2007; Karmakar and Simonovic, 2007, 2007),

$$MSE = \frac{1}{n-m}\sum_{i=1}^{n}\left(O_i - F_i\right)^2 \tag{3.6}$$

Where O_i and F_i are empirical (observed) and estimated distributions. The observed distribution (O_X) is computed using rank-based Gringorten's plotting position formula (Yue, 2001)

$$O_X\left(X \le x_i\right) = \frac{i - 0.44}{n + 0.12} \qquad \forall i = 1, 2, ..., n \tag{3.7}$$

Where, *i* is the rank in ascending order and x_i is the i^{th} largest variate in a data series of size *n*.

The *AIC* for small sample sizes ($n/m < \sim 40$) is given by (Hurvich and Tsai, 1995),

$$AIC_c\left(m\right) = AIC + \frac{2m\left(m+1\right)}{n - m - 1} \tag{3.8}$$

As sample size increases, the last term of the second-order AIC statistics (AIC_c) approaches zero, and the AICc tends to give the same conclusion as the traditional AIC (Burnham and Anderson, 2003). Since in the present study, length of data series varies from 46 to 66 years, we used AICc instead of traditional AIC, which is also a widely used method for model selection in hydrology (Caroni and Panagoulia, 2016; Gu et al., 2017; Panagoulia et al., 2014).

Table S13. Performance of stationary and nonstationary models for Oshawa WPCP

Time Slice	Model	Location parameter	Scale parameter	Shape parameter	AIC_c	Bayes-factor	LB (100yr)	UB (100yr)	UB/LB
15-min	GEV_t-0	48.33	15.71	-0.075	-332.55	-	81.65	156.01	1.91
	GEV_t-I	47.22+0.095*t*	17.95	-0.202	**-304.31**	7736.97	86.91	134.98	1.55
	GEV_t-II	51.03-0.074*t*	exp(3.06-0.007*t*)	-0.108	-290.55	5.44	91.48	156.69	1.71
30-min	GEV_t-0	31.67	11.21	-0.21	-306.10	-	47.64	107.8	2.26
	GEV_t-I	27.66+0.135*t*	10.73	-0.098	**-287.17**	0.15	51.74	74.03	1.43
	GEV_t-II	28.55+0.104*t*	exp(2.47-0.005*t*)	-0.155	-279.13	2.13	51.03	94.42	1.85
1-hr	GEV_t-0	18.10	7.33	-0.03	-317.73	-	31.82	104.55	3.28
	GEV_t-I	16.62+0.082*t*	7.28	-0.05	**-320.87**	0.47	39.59	59.46	1.50
	GEV_t-II	16.44+0.07*t*	exp(1.97+0.002*t*)	-0.002	-319.36	0.42	38.23	77.35	2.02
2-hr	GEV_t-0	11.09	3.81	0.13	-333.15	-	21.21	73.71	3.47
	GEV_t-I	10.04+0.05*t*	3.72	0.095	**-316.87**	0.16	23.03	48.89	2.12
	GEV_t-II	9.61+0.07*t*	exp(1.41+0.0034*t*)	0.007	-304.53	0.54	24.97	47.67	1.91
6-hr	GEV_t-0	5.02	1.42	0.094	-315.85	-	9.49	22.66	2.39
	GEV_t-I	4.5+0.02*t*	1.44	0.062	**-302.12**	0.51	10.8	15.46	1.43
	GEV_t-II	4.6+0.02*t*	exp(0.52-0.004*t*)	0.032	-291.60	0.67	10.34	15.94	1.54
12-hr	GEV_t-0	2.83	0.78	0.28	-310.34	-	6.36	20.48	3.22
	GEV_t-I	2.65+0.0073*t*	0.78	0.36	**-319.6**	0.46	9.3	15.85	1.70
	GEV_t-II	2.68+0.0092*t*	exp(-0.24+0.0017*t*)	0.29	-304.38	3.83	8.2	14.16	1.72
24-hr	GEV_t-0	1.65	0.51	0.17	-318.65	-	3.3	10.1	3.06
	GEV_t-I	1.51+0.0056*t*	0.50	0.17	**-316.24**	0.46	3.8	7.63	2.00
	GEV_t-II	1.51+0.0072*t*	exp(-0.68+0.00135*t*)	0.14	-305.97	1.98	3.94	7.18	1.82

*GEV_t-0 is stationary model whereas GEV_t-I and GEV_t-II are nonstationary models with time-variant mean, and both time-variant mean and standard deviation respectively. The selected best fitted nonstationary model is marked in bold letters. Bayes factor, $\gamma < 1$ indicates that the nonstationary model fits better than the stationary model. However, in cases $\gamma > 1$, to compare with stationary model, the nonstationary model is selected following minimum *AICc* criteria. LB and UB indicate lower and upper bounds of DSI at 100-year return period.

Table S14. Performance of stationary and nonstationary models for Kingston Airport

Time Slice	Model	Location parameter	Scale parameter	Shape parameter	AIC_c	Bayes-factor	LB (100yr)	UB (100yr)	UB/LB
15-min	GEV_t-0	44.59	12.70	-0.04	-315.08	-	76	140.88	1.85
	GEV_t-I	41.36+0.13*t*	12.83	-0.04	-311.53	1.93	85.2	118.76	1.40
	GEV_t-II	40.83+0.19*t*	exp(2.92-0.013*t*)	-0.07	**-317.65**	5.19	78.8	142.36	1.81
30-min	GEV_t-0	28.40	9.04	-0.05	-345.71	-	48.1	97.36	2.02
	GEV_t-I	27.61+0.042*t*	9.01	-0.04	**-348.92**	0.91	53.41	98.27	1.84
	GEV_t-II	27.57+0.035*t*	exp(2.11+0.0026*t*)	-0.07	-331.17	37.31	53.12	73.41	1.38
1-hr	GEV_t-0	18.35	5.87	-0.11	-318.56	-	29.45	56.19	1.91
	GEV_t-I	18.4+0.008*t*	6.02	-0.12	**-326.50**	3.76	31.21	61.45	1.97
	GEV_t-II	16.77+0.046*t*	exp(1.84-0.0059*t*)	0.02	-291.35	1.84	37.8	52.93	1.40
2-hr	GEV_t-0	11.32	3.16	0.06	-330.11	-	19.36	49.78	2.57
	GEV_t-I	11.09+0.009*t*	3.19	0.05	-321.46	1.27	23.74	33.6	1.41
	GEV_t-II	11.12+0.0055*t*	exp(1.10+0.0004*t*)	0.15	**-325.34**	4.47	22.89	46.26	2.02
6-hr	GEV_t-0	5.21	1.44	0.10	-327.18	-	10.1	23.64	2.34
	GEV_t-I	5.07+0.0073*t*	1.41	0.13	**-329.47**	0.23	11.17	17.79	1.59
	GEV_t-II	4.88+0.0093*t*	exp(0.38-0.00335*t*)	0.21	-326.79	0.77	12.91	19.57	1.51
12-hr	GEV_t-0	3.13	0.83	0.11	-312.19	-	6.09	13.04	2.14
	GEV_t-I	2.85+0.009*t*	0.78	0.14	**-321.61**	0.57	6.36	10.59	1.66
	GEV_t-II	2.84+0.009*t*	exp(-0.126-0.006*t*)	0.14	-317.09	0.16	6.78	9.5	1.40
24-hr	GEV_t-0	1.77	0.46	0.07	-297.12	-	3.18	6.43	2.02
	GEV_t-I	1.59+0.0071*t*	0.44	0.08	**-299.78**	0.54	6.36	10.59	1.66
	GEV_t-II	1.64+0.005*t*	exp(-0.81-0.0006*t*)	0.07	-296.75	0.97	3.42	5.57	1.63

Table S15. Performance of stationary and nonstationary models for Trenton Airport

Time Slice	Model	Location parameter	Scale parameter	Shape parameter	AIC_c	Bayes-factor	LB (100yr)	UB (100yr)	UB/LB
15-min	GEV_t-0	43.08	12.41	0.24	-485.30	-	95.12	250.55	2.63
	GEV_t-I	43.96-0.05*t*	11.87	0.19	**-474.35**	4.18	98.79	173.25	1.75
	GEV_t-II	40.37+0.16*t*	exp(1.65+0.04*t*)	-0.20	-472.29	7075420	86.12	120.33	1.40
30-min	GEV_t-0	26.94	8.50	0.13	-495.06	-	55.05	133.25	2.42
	GEV_t-I	26.71+0.014*t*	8.28	0.09	**-477.87**	13.50	55.16	111.31	2.02
	GEV_t-II	22.39+0.16*t*	exp(2.34+0.0024)	-0.07	-418.98	77.76	56.23	96.9	1.72
1-hr	GEV_t-0	16.25	5.022	0.15	-448.76	-	34.7	88.1	2.54
	GEV_t-I	15.40 + 0.032*t*	5.44	0.17	**-476.72**	0.73	39.3	73.74	1.88
	GEV_t-II	14.62+0.057*t*	exp(1.31+0.010)	0.13	-463.04	0.06	37.0	81.02	2.19
2-hr	GEV_t-0	10.13	3.32	0.17	-524.45	-	21.72	59.53	2.74
	GEV_t-I	9.03 + 0.03*t*	3.38	0.09	-481.72	0.99	20.46	51.05	2.49
	GEV_t-II	8.63 + 0.05*t*	exp(1.02 + 0.007*t*)	0.09	**-493.82**	0.16	24.76	36.64	1.48
6-hr	GEV_t-0	5.100	1.31	0.14	-419.26	-	10.24	22.05	2.15
	GEV_t-I	4.72 + 0.0124*t*	1.37	0.16	**-429.78**	1.83	12.56	17.14	1.36
	GEV_t-II	4.75 + 0.014*t*	exp(0.33 + 0.0012*t*)	0.15	-415.55	1.11	12.65	17.03	1.35
12-hr	GEV_t-0	3.19	0.94	0.14	-480.19	-	6.09	14.86	2.44
	GEV_t-I	2.85 + 0.0114*t*	0.87	0.09	**-491.15**	0.93	6.74	10.26	1.52
	GEV_t-II	2.84 + 0.0113*t*	exp(-0.16+0.0014*t*)	0.15	-488.52	0.15	7.95	10.58	1.33
24-hr	GEV_t-0	1.90	0.52	-0.04	-499.41	-	3.38	4.99	1.48
	GEV_t-I	1.51 + 0.011*t*	0.50	0.02	-467.34	0.26	3.19	5.84	1.83
	GEV_t-II	1.57 + 0.0096*t*	exp(-0.80 + 0.0033*t*)	-0.02	**-475.74**	0.14	3.38	4.99	1.48

Table S16. Performance of stationary and nonstationary models for Stratford WWTP

Time Slice	Model	Location parameter	Scale parameter	Shape parameter	AIC_c	Bayes-factor	LB (100yr)	UB (100yr)	UB/LB
15-min	GEV_t-0	56.30	15.70	0.04	-423.32	-	100.59	216.41	2.15
	GEV_t-I	57.31-0.05*t*	15.06	0.18	**-414.85**	5.26	105.08	144.59	1.38
	GEV_t-II	60.72-0.13*t*	exp(3.04-0.009*t*)	0.12	-384.65	54.70	134.37	184.4	1.37
30-min	GEV_t-0	33.83	10.81	0.17	-413.45	-	73.35	205.5	2.80
	GEV_t-I	34.76-0.023*t*	11.16	0.28	**-419.35**	6.24	110.13	177.76	1.61
	GEV_t-II	38.75-0.164*t*	exp(2.97-0.015*t*)	0.45	-314.14	68.72	311.05	163.32	1.90
1-hr	GEV_t-0	19.03	6.31	0.21	-388.39	-	44.6	149.69	3.36
	GEV_t-I	18.62-0.0045*t*	6.42	0.36	**-398.93**	3.93	74.66	125.4	1.68
	GEV_t-II	17.43+0.0552*t*	exp(1.5+0.0125*t*)	0.17	-372.02	2.79	49.84	84.84	1.70
2-hr	GEV_t-0	11.71	4.37	0.26	-390.44	-	28.84	118.37	4.10
	GEV_t-I	11.49+0.023*t*	4.99	0.34	**-371.43**	3.77	33.97	158.15	4.65
	GEV_t-II	11.82-0.027*t*	exp(1.70-0.008*t*)	0.53	-342.82	3.35	70.43	147.35	2.09
6-hr	GEV_t-0	5.32	1.80	0.24	-400.01	-	13.34	38.24	2.87
	GEV_t-I	5.76-0.012*t*	1.78	0.21	-392.41	3.97	13.2	31.6	2.40
	GEV_t-II	5.52-0.007*t*	exp(0.59+0.0002*t*)	0.28	**-400.2**	29.5	18.2	28.63	1.57
12-hr	GEV_t-0	3.11	0.89	0.26	-365.66	-	7.45	19.5	2.62
	GEV_t-I	3.38-0.01*t*	0.89	0.34	**-377.72**	0.96	9.75	17.42	1.79
	GEV_t-II	3.44-0.0097*t*	exp(-0.11-6e^{-5}*t*)	0.29	-367.39	0.25	9.88	14.77	1.50
24-hr	GEV_t-0	1.68	0.45	0.26	-376.96	-	3.75	9.93	2.65
	GEV_t-I	1.77-0.004*t*	0.45	0.29	-377.32	0.71	9.75	17.42	1.79
	GEV_t-II	1.82-0.005*t*	exp(-0.80-0.0008*t*)	0.30	**-378.44**	0.41	4.56	8.69	1.90

Table S17. Performance of stationary and nonstationary models for Fergas Shand Dam

Time Slice	Model	Location parameter	Scale parameter	Shape parameter	AIC_c	Bayes-factor	LB (100yr)	UB (100yr)	UB/LB
15-min	GEV_t-0	54.92	15.4	0.13	-538.02	-	116.76	224.08	1.92
	GEV_t-I	61.55-0.09*t*	20.77	0.047	-516.45	3.40	119.9	255.47	2.13
	GEV_t-II	65.49-0.19*t*	exp(2.99-0.00064*t*)	0.03	**-529.81**	1.22	117.27	230.15	1.96
30-min	GEV_t-0	37.63	13.2	0.11	-469.76	-	86.68	185.36	2.14
	GEV_t-I	38.27-0.024*t*	13.6	0.11	**-462.14**	2.98	85.51	200.77	2.35
	GEV_t-II	34.08+0.074*t*	exp(2.12+0.0115*t*)	0.14	-444.82	105.23	93.66	150.24	1.60
1-hr	GEV_t-0	22.10	8.57	0.16	-532.24	-	52.91	137.9	2.61
	GEV_t-I	25.35-0.096*t*	8.58	0.17	**-526.26**	0.12	60.42	115.43	1.91
	GEV_t-II	19.54+0.04*t*	exp(1.59+0.013*t*)	0.28	-476.74	11.4	74.7	112.48	1.50
2-hr	GEV_t-0	12.54	5.26	0.12	-542.06	-	30.02	76.43	2.55
	GEV_t-I	13.75-0.049*t*	5.02	0.23	**-487.49**	1.17	41.93	69.76	1.66
	GEV_t-II	12.89-0.026*t*	exp(1.64-0.00013*t*)	0.32	-462.93	1.95	49.89	89.83	1.80
6-hr	GEV_t-0	5.34	2.31	0.055	-558.5	-	12.11	28.94	2.39
	GEV_t-I	5.30-0.0026*t*	2.35	0.064	**-537.24**	1.35	14.66	21.69	1.48
	GEV_t-II	5.15+0.0011*t*	exp(0.88-0.00035*t*)	0.046	-524.53	1.25	14.86	20.51	1.38

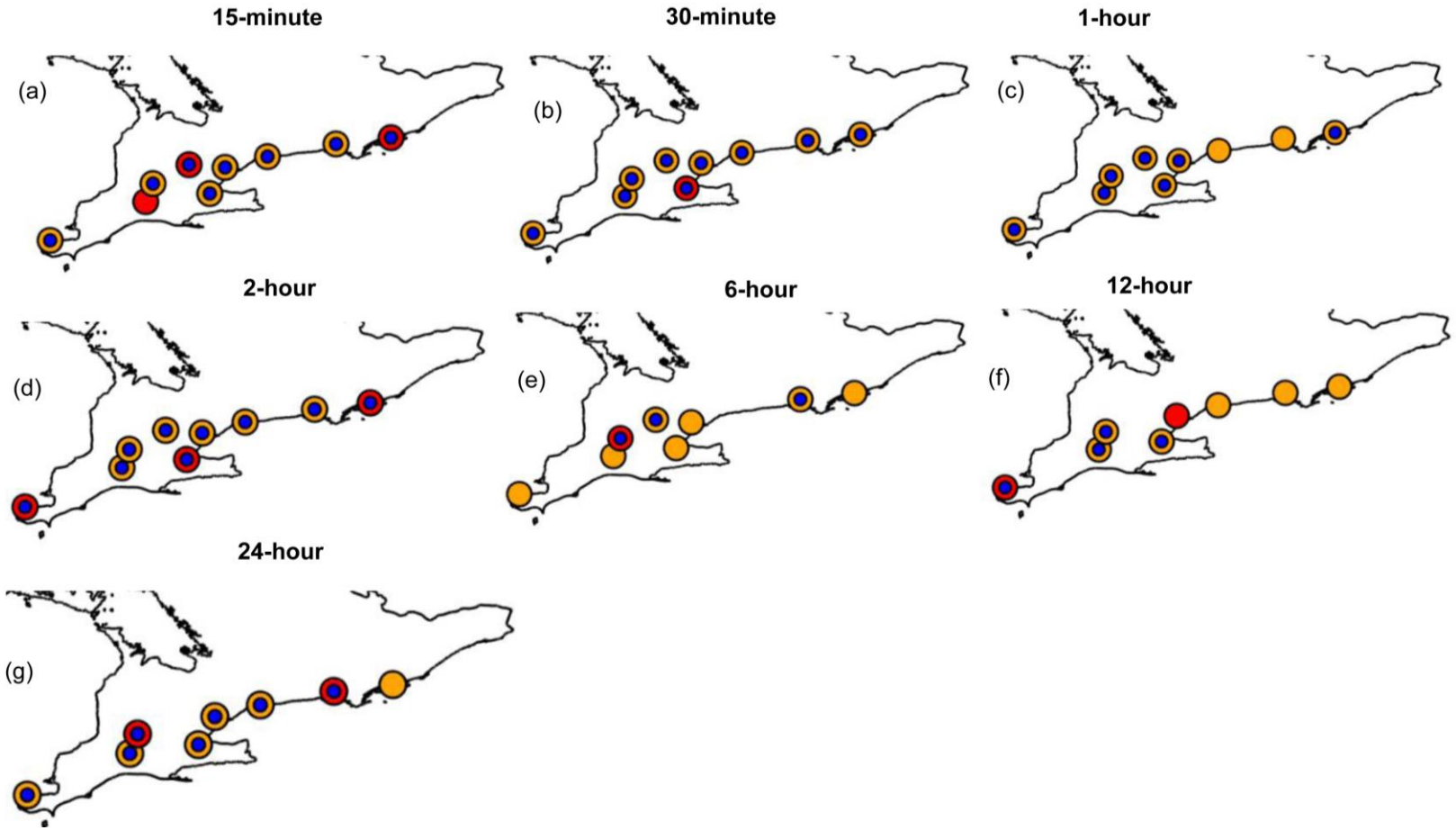

Figure S9. The performance of stationary (Blue) versus nonstationary (Orange - nonstationary with time variant mean; red - nonstationary with time variant mean and standard deviation) GEV models with durations ranging between 15-minute and 24-hour (**a-g**) in nine urbanized sites. Sites with double circles indicate either better or comparable performance of stationary model relative to the nonstationary model. The shades denote the type of distribution.

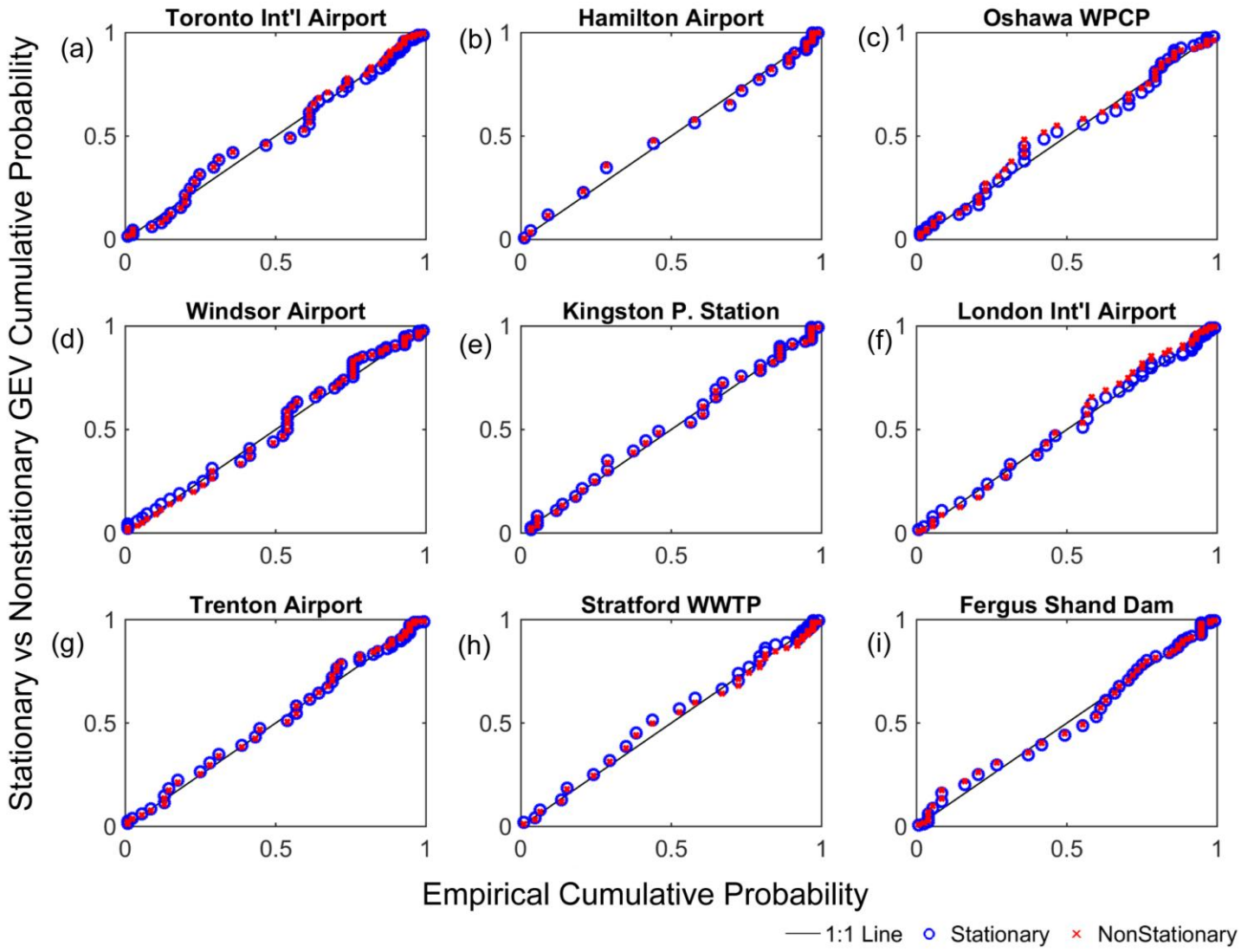

Figure S10. *PP* plots of nonstationary versus stationary GEV distributions for 30-min storm duration in nine sites (**a - i**)

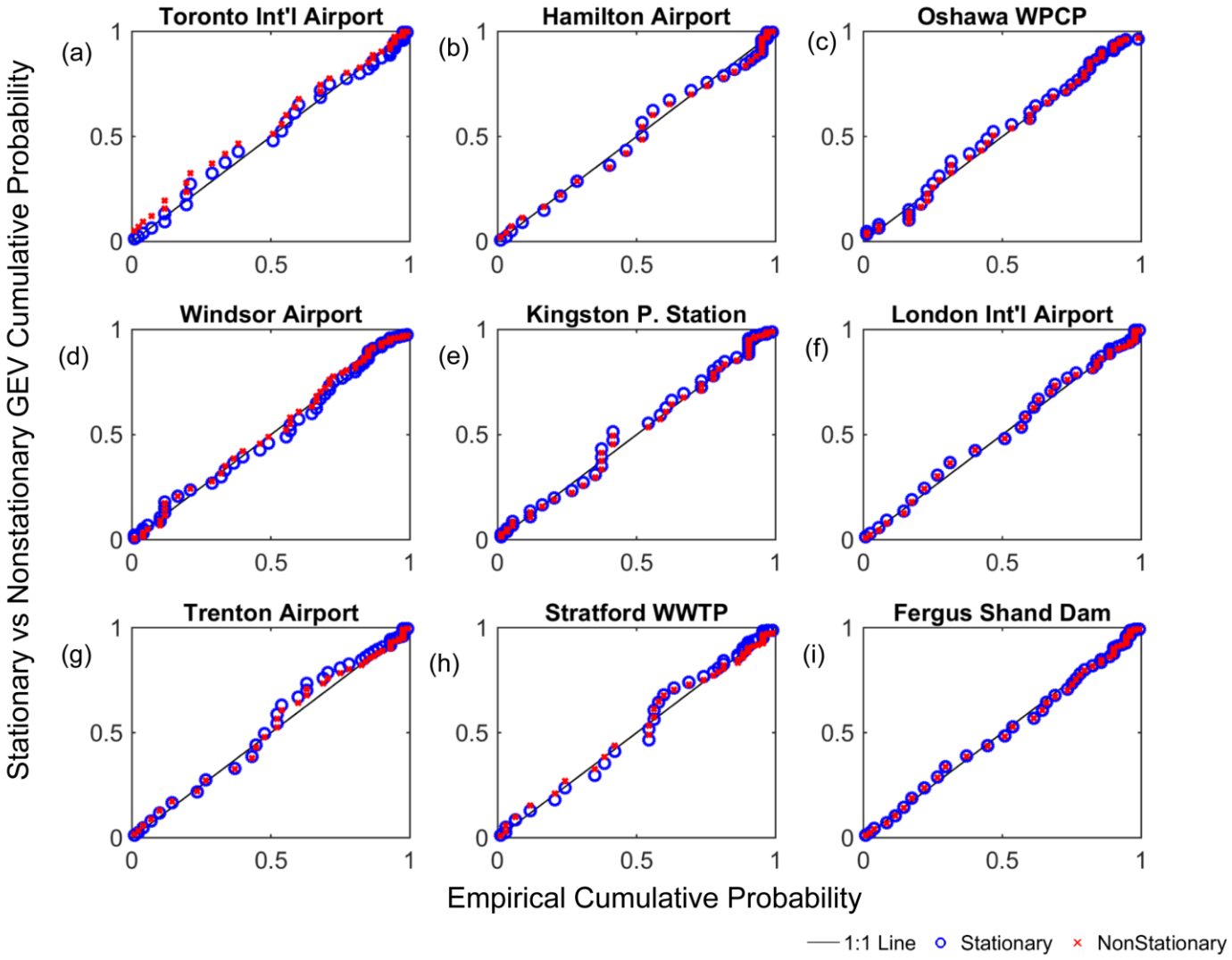

Figure S11. *PP* plots of nonstationary versus stationary GEV distributions for 1-hour storm duration in nine sites (**a - i**).

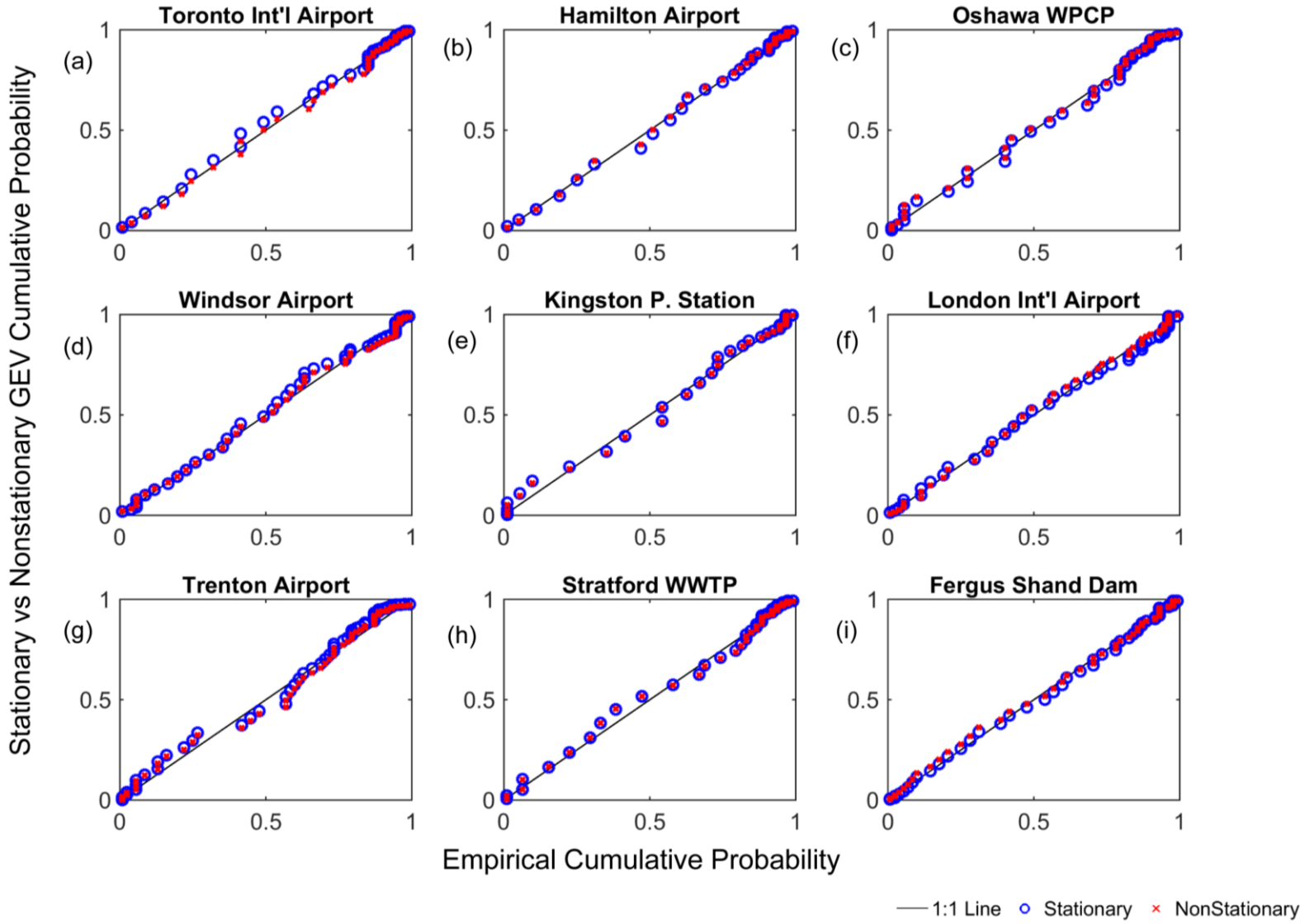

Figure S12. *PP* plots of nonstationary versus stationary GEV distributions for 6-hour duration in nine sites (**a - i**)

4 Estimation of Design Storm Intensity (DSI)

The DSI q_p, often referred as *return level* in the literature is the value which is expected to be exceeded on an average once in every $1/p$ periods, where $1-p$ is the specific probability associated with the quantile (Gilleland and Katz, 2014). We obtain the $1-p$ (*i.e.,* non-exceedance probability of occurrence) quantile of the annual maximum rainfall fitted with stationary GEV distribution using following expression (Coles and Tawn, 1996):

$$q_p = \mu + \frac{\sigma}{\zeta}\left[\left\{-\ln\left(1-p\right)\right\}^{-\zeta} - 1\right] \quad \forall \zeta \neq 0 \tag{4.1}$$

Where, q_p is the DSI, associated with a $T = 1/(1-p)$ - year return period. The nonstationary design intensity is analogous to the standard stationary precipitation intensity with the exception of inclusion of time variant parameters (Cheng et al., 2014)

$$q_p = \bar{\mu} + \frac{\bar{\sigma}}{\zeta}\left[\left\{-\ln\left(1-p\right)\right\}^{-\zeta} - 1\right] \quad \forall \zeta \neq 0 \tag{4.2}$$

Where time-variant parameters, $\bar{\mu}$ and $\bar{\sigma}$ are derived by computing 0.50, 0.025 and 0.975 quantiles of DE-MC sampled $\mu(t)$ and $\sigma(t)$ $\left(\left\{\mu(t_1),....,\mu(t_n)\right\} \text{ and } \left\{\sigma(t_1),....,\sigma(t_n)\right\}\right)$ respectively. In Eq. 4.1, when $\zeta \to 0$, the GEV distribution reduces to Gumbel distribution (or Extreme Value Type I). It should be noted that Gumbel Extreme value distribution has been commonly used to estimate design storm by Environment Canada (CSA, 2010). The Gumbel probability distribution has following form (Wang et al., 2015)

$$q_p = \mu + K_p \sigma \tag{4.3}$$

Where K_p denotes frequency factor depending on the return period *T*, which is obtained using following relationship (Wang et al., 2015)

$$K_p = \frac{-\sqrt{6}}{\pi}\left[0.5772 + \ln\left(\ln\left(\frac{T}{T-1}\right)\right)\right] \tag{4.4}$$

Environment Canada uses this method to estimate rainfall frequency at a given duration and obtain nationwide IDF curves.

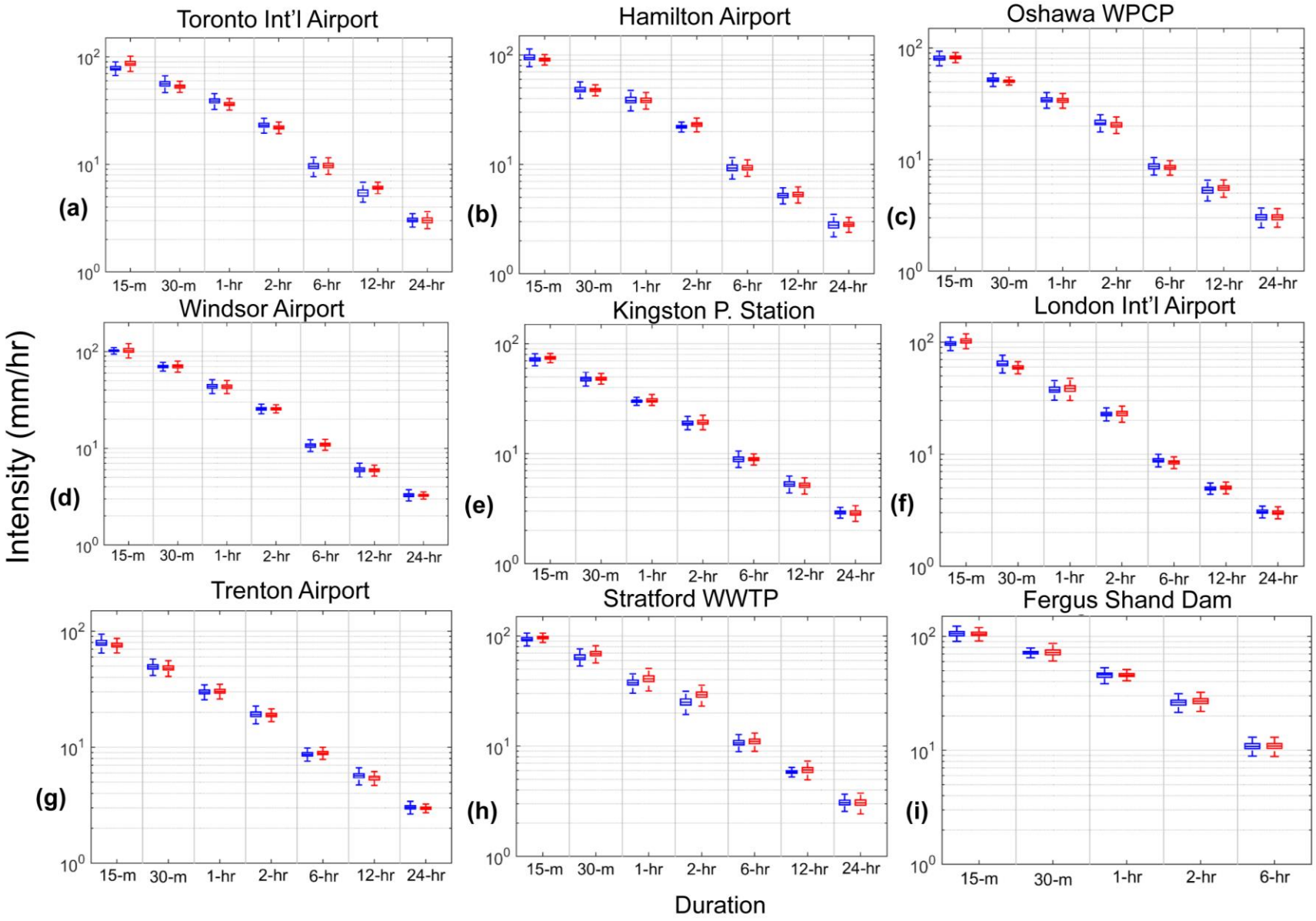

Figure S13. Uncertainty in DSI for 10-year return periods for stationary (blue) versus nonstationary (red) models The boxplots indicate 95% credibility interval in DSI estimate.

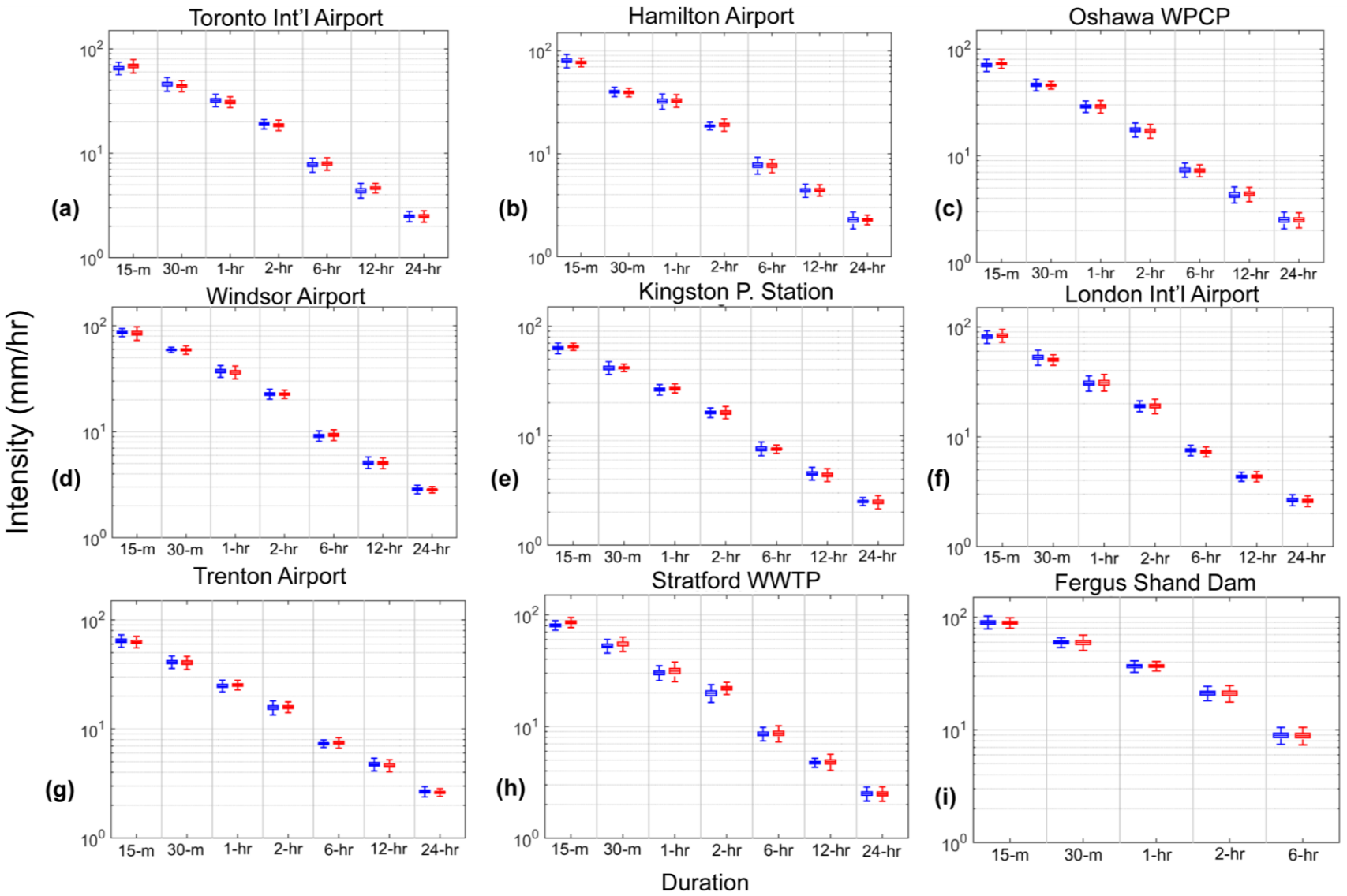

Figure S14. Same as Figure S13 but for 5-year return period.

Table S18.1 Ratio and percentage changes between updated nonstationary versus EC-generated DSI for Toronto International Airport for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	0.72	0.84	0.91	0.99	1.05	1.12
	[-38.6%]	[-19.1%]	[-10.0%]	[-0.8%]	[**5.1%**]	[***10.4***%]
30-min	0.75	0.80	0.81	0.82	0.82	0.82
	[-33.2%]	[-25.1%]	[-22.9%]	[-21.7%]	[-21.6%]	[-21.8%]
1-hour	0.94	0.93	0.92	0.90	0.88	0.86
	[-6.7%]	[-7.3%]	[-8.9%]	[-11.7%]	[-14.1%]	[-16.7%]
2-hour	0.97	0.93	0.91	0.90	0.89	0.88
	[-3.3%]	[-7.5%]	[-9.4%]	[-11.3%]	[-12.6%]	[-13.7%]
6-hour	0.96	0.92	0.94	0.99	1.05	1.13
	[-4.4%]	[-9.17%]	[-6.8%]	[-0.7%]	[**5.2%**]	[***11.7****%]
12-hour	0.94	0.94	1.03	1.23	1.47	1.80
	[-6.4%]	[-6.7%]	[**2.5%]**	[***18.9%***]	[***32.2%***]	[***44.5%***]
24-hour	0.94	0.89	0.90	0.96	1.03	1.11
	[-5.9%]	[-12.1%]	[-10.7%]	[-4.18%]	[**2.8%**]	[***10.1%***]

*Results within brackets indicate percentage changes between updated versus EC-generated DSI. More than 1% increase in DSI is shown in bold, whereas 10% or more is marked with bold italic letters.

Table S18.2 Ratio and percentage changes between updated stationary versus EC-generated DSI for Toronto International Airport for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	0.73	0.80	0.82	0.83	0.83	0.83
	[-37.7%]	[-25.3%]	[-22.1%]	[-20.6%]	[-20.5%]	[-25.6%]
30-min	0.76	0.83	0.86	0.89	0.91	0.93
	[-32.3%]	[-21.0%]	[-16.5%]	[-12.4%]	[-10.0%]	[-8.08%]
1-hour	0.97	0.97	0.98	1.01	1.03	1.05
	[-2.9%]	[-2.9%]	[-1.7%]	[0.6%]	[**2.5%**]	[**4.6**%]
2-hour	0.98	0.96	0.97	0.99	1.02	1.05
	[-1.6%]	[-4.35%]	[-3.4%]	[-0.7%]	[**1.9%**]	[**4.9%**]
6-hour	0.93	0.89	0.91	0.98	1.05	1.13
	[-7.7%]	[-12.6%]	[-9.7%]	[-2.4%]	[**4.4%**]	[***11.8****%]
12-hour	0.93	0.88	0.91	1.0	1.1	1.25
	[-8.0%]	[-13.5%]	[-9.8%]	[0.3%]	[**9.8%**]	[***19.9%***]
24-hour	0.94	0.90	0.91	0.98	1.06	1.15
	[-5.9%]	[-11.6%]	[-9.6%]	[-2.3%]	[**5.3%**]	[***13.2%***]

*More than 10% increase is marked in bold italics.

Table S19.1 Ratio and percentage changes between updated nonstationary versus EC-generated DSI for Hamilton Airport for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	0.96 [-4.7%]	0.93 [-7.1%]	0.94 [-6.1%]	0.97 [-3.2%]	1.0 [-0.3%]	1.03 [**3.1%**]
30-min	0.71 [-40.1%]	0.72 [-39.5%]	0.75 [-34.2%]	0.80 [-24.3%]	0.86 [-15.7%]	0.94 [-6.7%]
1-hour	0.98 [-2.1%]	0.93 [-7.8%]	0.91 [-9.7%]	0.90 [-10.9%]	0.90 [-11.2%]	0.90 [-11.2%]
2-hour	0.94 [-6.1%]	0.89 [-12.3%]	0.90 [-11.7%]	0.93 [-7.8%]	0.97 [-3.5%]	1.02 [**1.5%**]
6-hour	0.90 [-11.6%]	0.85 [-17.7%]	0.87 [-15.2%]	0.93 [-7.8%]	1.0 [-0.3%]	1.09 [**7.9%**]
12-hour	0.91 [-9.2%]	0.88 [-13.8%]	0.89 [-12.07%]	0.94 [-6.5%]	0.99 [-0.76%]	1.06 [**5.59%**]
24-hour	0.75 [-32.9%]	0.80 [-25.8%]	0.84 [-19.0%]	0.92 [-8.9%]	0.99 [-1.1%]	1.07 [**6.5%**]

Table S19.2 Ratio and percentage changes between updated stationary versus EC-generated DSI for Hamilton Airport for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	0.98 [-1.7%]	0.97 [-2.8%]	0.99 [-1.5%]	1.02 [**1.6%**]	1.05 [**4.5%**]	1.08 [**7.7%**]
30-min	0.72 [-38.5%]	0.72 [-38.3%]	0.75 [-34.2%]	0.79 [-26.5%]	0.83 [-19.8%]	0.61 [-63.9%]
1-hour	0.96 [-3.7%]	0.92 [-9.2%]	0.91 [-9.5%]	0.93 [-7.6%]	0.95 [-5.2%]	0.98 [-2.1%]
2-hour	0.94 [-6.4%]	0.87 [-15.4%]	0.85 [-17.5%]	0.85 [-17.5%]	0.86 [-16.3%]	0.87 [-14.4%]
6-hour	0.91 [-10.1%]	0.86 [-16.8%]	0.86 [-15.6%]	0.91 [-10.4%]	0.96 [-4.7%]	1.02 [**1.7%**]
12-hour	0.92 [-8.5%]	0.87 [-14.5%]	0.87 [-14.2%]	0.90 [-10.4%]	0.94 [-6.3%]	0.99 [-1.3%]
24-hour	0.76 [-31.3%]	0.79 [-26.3%]	0.83 [-21.1%]	0.88 [-13.3%]	0.93 [-7.2%]	0.99 [-1.2%]

Table S20.1 Ratio and percentage changes between updated nonstationary versus EC-generated DSI for Oshawa WPCP for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	1.09	1.08	1.06	1.01	0.97	0.94
	[**7.8%**]	[**7.6%**]	[**5.3%**]	[**1.06%**]	[-2.7%]	[-6.8%]
30-min	1.09	1.05	1.0	0.93	0.88	0.83
	[**8.7%**]	[**4.83%**]	[0.1%]	[-7.4%]	[-13.6%]	[-20.3%]
1-hour	1.05	1.07	1.07	1.06	1.05	1.04
	[**4.3%**]	[**6.4%**]	[**6.5%**]	[**5.9%**]	[**5.1%**]	[**4.2%**]
2-hour	1.0	1.02	1.04	1.07	1.10	1.14
	[0.24%]	[**1.81%**]	[**3.7%**]	[**6.9**%]	[**9.4**%]	[***12.0***%]
6-hour	1.0	0.99	1.0	1.01	1.03	1.04
	[0%]	[-0.55%]	[-0.11%]	[**1.3%**]	[**2.5%**]	[**4.0%**]
12-hour	0.96	0.90	0.93	1.03	1.15	1.30
	[-4.2%]	[-11.6%]	[-7.56%]	[**3.04%**]	[***12.8*%**]	[***23.0*%**]
24-hour	0.99	0.95	0.96	0.99	1.03	1.08
	[-1.1%]	[-5.6%]	[-4.6%]	[-1.05%]	[**2.9%**]	[**7.3%**]

Table S20.2 Ratio and percentage changes between updated stationary versus EC-generated DSI for Oshawa WPCP for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	1.05	1.05	1.04	1.03	1.01	0.99
	[**4.9%**]	[**4.9%**]	[**4.1%**]	[**2.5%**]	[1.0%]	[-0.6%]
30-min	1.07	1.06	1.03	0.99	0.95	0.91
	[**6.3%**]	[**5.6%**]	[**3.0%**]	[-1.5%]	[-5.5%]	[-9.9%]
1-hour	1.03	1.06	1.07	1.08	1.08	1.08
	[**2.6%**]	[**5.9%**]	[**6.9%**]	[**7.4%**]	[**7.5%**]	[**7.3%**]
2-hour	1.0	1.04	1.07	1.13	1.18	1.23
	[0.4%]	[**3.5%**]	[**6.7%**]	[***11.4*%**]	[***15.1*%**]	[***18.9***%]
6-hour	1.01	1.01	1.02	1.04	1.06	1.09
	[0.9%]	[0.5%]	[**1.5%**]	[**3.7%**]	[**5.9%**]	[**8.1%**]
12-hour	0.96	0.88	0.88	0.93	1.0	1.08
	[-3.8%]	[-14.2%]	[-13.3%]	[-7.0%]	[0%]	[**7.8%**]
24-hour	0.98	0.95	0.96	0.99	1.03	1.07
	[-1.6%]	[-5.5%]	[-4.6%]	[-1.0%]	[**2.7%**]	[**6.9%**]

Table S21.1 Ratio and percentage changes between updated nonstationary versus EC-generated DSI for Windsor Airport for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	0.94	0.97	1.02	1.11	1.20	1.31
	[-6.1%]	[-3.3%]	[**1.6%**]	[**9.8%**]	[***16.5***%]	[***23.4***%]
30-min	0.99	1.00	1.03	1.07	1.11	1.15
	[-1.5%]	[0.5%]	[**2.9%**]	[**6.8%**]	[**9.9**%]	[***13.2***%]
1-hour	0.98	0.98	1.00	1.04	1.08	1.12
	[-1.7%]	[-1.7%]	[0.21%]	[**3.9%**]	[**7.4%**]	[***10.9%***]
2-hour	1.06	1.06	1.04	1.00	0.98	0.95
	[**5.8%**]	[**5.3%**]	[**3.5%**]	[0.27%]	[-2.5%]	[-5.5%]
6-hour	1.04	1.06	1.08	1.10	1.11	1.13
	[**3.7%**]	[**5.9%**]	[**7.2%**]	[**8.9%**]	[***10.3%***]	[***11.6%***]
12-hour	1.01	1.01	1.03	1.05	1.08	1.11
	[0.5%]	[**1.2%**]	[**2.7%**]	[**5.2%**]	[**7.4**%]	[**9.8**%]
24-hour	1.00	1.00	1.01	1.02	1.03	1.03
	[-0.45%]	[0%]	[0.9%]	[**1.8%**]	[**2.6%**]	[**3.2%**]

Table S21.2 Ratio and percentage changes between updated stationary versus EC-generated DSI for Windsor Airport for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	0.97	0.98	1.00	1.04	1.08	1.12
	[-3.0%]	[-1.94%]	[0.32%]	[**4.2%**]	[**7.5%**]	[***11.1%***]
30-min	0.98	1.01	1.03	1.05	1.07	1.09
	[-1.56%]	[0.79%]	[**2.6%**]	[**4.9%**]	[**6.7%**]	[**8.5%**]
1-hour	1.01	1.01	1.02	1.03	1.04	1.05
	[1.04%]	[0.99%]	[**1.5%**]	[**2.7%**]	[**3.7%**]	[**4.8%**]
2-hour	1.07	1.06	1.03	1.0	0.97	0.94
	[**6.9%**]	[**5.5%**]	[**3.3%**]	[-0.3%]	[-3.4%]	[-6.8%]
6-hour	1.03	1.04	1.05	1.07	1.08	1.09
	[**2.4%**]	[**3.7%**]	[**4.7%**]	[**6.1%**]	[**7.2%**]	[**8.4%**]
12-hour	1.00	1.01	1.04	1.07	1.10	1.14
	[-0.26%]	[**1.4%**]	[**3.5%**]	[**6.8%**]	[**9.5%**]	[***12.3%***]
24-hour	1.00	1.01	1.01	1.02	1.03	1.03
	[0.45%]	[0.7%]	[**1.2%**]	[**2.1%**]	[**2.6%**]	[**3.2%**]

Table S22.1 Ratio and percentage changes between updated nonstationary versus EC-generated DSI for Kingston P. Station for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	1.00	1.00	1.00	0.99	0.98	0.96
	[-0.2%]	[0.5%]	[-0.03%]	[-1.3%]	[-2.5%]	[-3.9%]
30-min	0.98	0.98	0.97	0.95	0.94	0.93
	[-1.8%]	[-2.5%]	[-3.4%]	[-4.78%]	[-5.93%]	[-7.13%]
1-hour	1.00	0.97	0.95	0.91	0.88	0.86
	[0.1%]	[-2.8%]	[-5.68%]	[-9.8%]	[-13.3%]	[-16.8%]
2-hour	0.96	0.95	0.97	1.01	1.05	1.10
	[-4.5%]	[-4.9%]	[-2.84%]	[1.18%]	[4.9%]	[8.82%]
6-hour	0.99	0.98	0.99	1.03	1.06	1.09
	[-1.04%]	[-1.98%]	[-0.56%]	[**2.5%**]	[**5.4%**]	[**8.6%**]
12-hour	0.98	0.96	0.97	1.00	1.03	1.07
	[-2.08%]	[-4.1%]	[-3.1%]	[**-0.3%**]	[**2.8%**]	[**6.1%**]
24-hour	0.98	0.98	0.99	1.01	1.03	1.05
	[-1.5%]	[-1.6%]	[-0.7%]	[1.2%]	[**3.1%**]	[**4.9%**]

Table S22.2 Ratio and percentage changes between updated stationary versus EC-generated DSI for Kingston P. Station for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	0.97	0.97	0.97	0.96	0.95	0.95
	[-2.8%]	[-2.7%]	[-3.1%]	[-4.0%]	[-4.8%]	[-5.7%]
30-min	0.98	0.97	0.96	0.95	0.93	0.92
	[-2.5%]	[-3.2%]	[-4.1%]	[-5.7%]	[-7.1%]	[-8.5%]
1-hour	0.99	0.96	0.93	0.90	0.87	0.85
	[-1.5%]	[-4.5%]	[-7.3%]	[-11.4%]	[-14.7%]	[-18.1%]
2-hour	0.97	0.95	0.95	0.96	0.97	0.98
	[-3.3%]	[-5.1%]	[-4.9%]	[-4.1%]	[-3.0%]	[-1.8%]
6-hour	0.99	0.98	0.99	1.01	1.03	1.06
	[-1.2%]	[-2.2%]	[-1.4%]	[0.8%]	[**2.9%**]	[**5.3%**]
12-hour	1.0	0.98	0.99	1.01	1.04	1.07
	[0.3%]	[-1.5%]	[-0.8%]	[**1.4%**]	[**3.9%**]	[**6.5%**]
24-hour	0.99	1.0	1.01	1.02	1.04	1.06
	[-0.5%]	[-0.4%]	[0.7%]	[**2.3%**]	[**3.9%**]	[**5.6%**]

Table S23.1 Ratio and percentage changes between updated nonstationary versus EC-generated DSI for London International Airport for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	0.92	0.96	1.0	1.08	1.1	1.21
	[-9.2%]	[-4.3%]	[0.45%]	[**7.1%**]	[***12.3%***]	[***17.4%***]
30-min	0.92	0.89	0.90	0.92	0.94	0.97
	[-9.3%]	[-12.3%]	[-11.7%]	[-9.2%]	[-6.4%]	[-3.3%]
1-hour	0.88	0.88	0.91	0.98	1.05	1.12
	[-13.3%]	[-13.3%]	[-9.3%]	[-1.9%]	[**4.3%**]	[***10.8%***]
2-hour	0.92	0.93	0.95	0.98	1.0	1.03
	[-8.9%]	[-7.8%]	[-5.8%]	[-2.5%]	[0.12%]	[**2.9%**]
6-hour	0.93	0.92	0.92	0.92	0.93	0.94
	[-7.2%]	[-9.0%]	[-9.1%]	[-8.4%]	[-7.5%]	[-6.3%]
12-hour	0.95	0.97	0.98	1.01	1.03	1.05
	[-5.6%]	[-3.4%]	[-1.8%]	[0.5%]	[**2.5%**]	[**4.3%**]
24-hour	0.92	0.94	0.94	0.94	0.94	0.93
	[-9.3%]	[-6.2%]	[-5.9%]	[-5.9%]	[-6.7%]	[-7.6%]

Table S23.2 Ratio and percentage changes between updated stationary versus EC-generated DSI for London International Airport for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	0.92	0.94	0.95	0.97	0.98	1.0
	[-8.9%]	[-6.8%]	[-5.2%]	[-3.2%]	[-1.7%]	[-0.1%]
30-min	0.93	0.94	0.97	1.02	1.07	1.13
	[-7.6%]	[-6.2%]	[-3.1%]	[**2.2%**]	[**6.6%**]	[***11.2%***]
1-hour	0.88	0.87	0.89	0.92	0.95	0.99
	[-13.1%]	[-14.8%]	[-12.9%]	[-8.9%]	[-5.1%]	[-1.05%]
2-hour	0.92	0.92	0.93	0.94	0.95	0.97
	[-8.4%]	[-8.3%]	[-7.5%]	[-5.9%]	[-4.8%]	[-3.4%]
6-hour	0.94	0.94	0.96	0.98	1.0	1.02
	[-6.4%]	[-5.8%]	[-4.4%]	[-2.1%]	[0.08%]	[**2.3%**]
12-hour	0.95	0.96	0.96	0.96	0.96	0.96
	[-5.3%]	[-4.6%]	[-4.3%]	[-4.1%]	[-3.8%]	[-3.8%]
24-hour	0.92	0.96	0.96	0.97	0.96	0.96
	[-8.7%]	[-4.5%]	[-3.9%]	[-3.3%]	[-3.8%]	[-4.4%]

Table S24.1 Ratio and percentage changes between updated nonstationary versus EC-generated DSI for Trenton Airport for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	0.99	0.99	1.01	1.07	1.13	1.19
	[-0.67%]	[-1.11%]	[**1.45%**]	[**6.56%**]	[***11.13***%]	[***16.0***%]
30-min	0.99	0.99	1.01	1.04	1.07	1.10
	[-1.32%]	[-0.54%]	[**1.11**%]	[**3.9%**]	[**6.3**%]	[**8.88**%]
1-hour	0.97	0.94	0.96	1.0	1.04	1.09
	[-3.2%]	[-5.93%]	[-4.23%]	[0%]	[**4.0%**]	[**8.5%**]
2-hour	0.99	0.95	0.95	0.96	0.97	0.99
	[-1.5%]	[-5.29%]	[-5.54%]	[-4.3%]	[-2.8%]	[-0.9%]
6-hour	0.99	1.00	1.03	1.08	1.12	1.18
	[-0.53%]	[0.13%]	[**2.6%**]	[**7.1%**]	[***11.1%***]	[***15.3***%]
12-hour	1.00	1.02	1.03	1.06	1.09	1.12
	[0%]	[**1.5%**]	[**3.3%**]	[**6.04%**]	[**8.5%**]	[***10.9***%]
24-hour	1.02	1.03	1.03	1.03	1.03	1.03
	[**1.93%**]	[**3.05%**]	[**3.02%**]	[**2.9%**]	[**2.9%**]	[**2.5%**]

Table S24.2 Ratio and percentage changes between updated stationary versus EC-generated DSI for Trenton Airport for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	1.01	1.03	1.07	1.16	1.25	1.35
	[1.38%]	[**2.78**%]	[**6.8%**]	[***13.9%***]	[***19.9%***]	[***26.06***%]
30-min	0.98	1.01	1.04	1.09	1.13	1.18
	[-1.63%]	[0.8%]	[**3.7%**]	[**8.1%**]	[***11.7%***]	[***15.4%***]
1-hour	0.95	0.91	0.91	0.94	0.97	1.01
	[-4.8%]	[-9.8%]	[-9.4%]	[-6.4%]	[-3.1%]	[0.7%]
2-hour	0.96	0.94	0.96	1.0	1.05	1.10
	[-3.7%]	[-5.95%]	[-4.1%]	[0.45%]	[**4.6%**]	[**9.1%**]
6-hour	0.98	0.98	0.99	1.03	1.06	1.11
	[-1.97%]	[-2.3%]	[-0.82%]	[**2.69%**]	[**6.0%**]	[**9.5%**]
12-hour	1.0	1.05	1.09	1.15	1.21	1.27
	[-0.3%]	[**4.41%**]	[**8.1%**]	[***13.04%***]	[***17.1%***]	[***21.1%***]
24-hour	1.03	1.05	1.05	1.05	1.04	1.04
	[**2.87%**]	[**4.51%**]	[**4.6%**]	[**4.3%**]	[**4.2%**]	[**3.6%**]

Table S25.1 Ratio and percentage changes between updated nonstationary versus EC-generated DSI for Stratford WWTP for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	1.05 [5.03%]	1.08 [**7.02%**]	1.06 [**5.6%**]	1.02 [**2.4%**]	0.99 [-0.76%]	0.96 [*-4.35%*]
30-min	0.96 [-3.99%]	1.00 [0.33%]	1.07 [**6.6%**]	1.19 [***16.2%***]	1.31 [***23.72%***]	1.45 [***31.2%***]
1-hour	0.90 [-10.9%]	0.94 [-6.1%]	1.03 [**2.6%**]	1.19 [***16.2%***]	1.36 [***26.5%***]	1.57 [***36.4%***]
2-hour	0.97 [-3.3%]	0.98 [-1.9%]	1.06 [**5.3%**]	1.21 [***17.4%***]	1.37 [***26.8%***]	1.56 [***36.04%***]
6-hour	0.89 [-12.3%]	0.84 [-19.4%]	0.86 [-16.1%]	0.93 [-7.2%]	1.01 [**1.09**%]	1.11 [**9.8%**]
12-hour	0.90 [-11.5%]	0.83 [-20.7%]	0.84 [-18.9%]	0.90 [-11.04%]	0.97 [-2.9**%**]	1.06 [**5.9%**]
24-hour	0.84 [-19.02%]	0.79 [-27.2%]	0.80 [-24.3%]	0.87 [-15.6%]	0.94 [-6.8%]	1.03 [**2.6%**]

Table S25.2 Ratio and percentage changes between updated stationary versus EC-generated DSI for Stratford WWTP for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	1.0 [-0.1%]	1.01 [**1.4**%]	1.03 [**2.47**%]	1.04 [**3.97%**]	1.05 [**5.1%**]	1.07 [**6.3%**]
30-min	0.95 [-5.25%]	0.96 [-4.5%]	0.99 [-1.45%]	1.04 [**3.89%**]	1.09 [**8.5%**]	1.15 [***13.2%***]
1-hour	0.92 [-8.7%]	0.91 [-10.03%]	0.94 [-6.7%]	1.0 [-0.06%]	1.06 [***5.8%***]	1.14 [***12.05***%]
2-hour	0.92 [-8.7%]	0.88 [-13.6%]	0.91 [-10.2%]	0.98 [-2.18%]	1.05 [***5.1%***]	1.15 [***12.7***%]
6-hour	0.89 [-12.3%]	0.82 [-21.3%]	0.84 [-19.7%]	0.88 [-13.1%]	0.94 [-6.45%]	1.01 [0.8%]
12-hour	0.89 [-12.4%]	0.81 [-22.7%]	0.82 [-21.9%]	0.86 [-15.7%]	0.92 [-8.86%]	0.99 [-1.3%]
24-hour	0.84 [-19.02%]	0.79 [-27.2%]	0.80 [-25.5%]	0.84 [-18.5%]	0.90 [-11.4%]	0.97 [-3.1%]

Table S26.1 Ratio and percentage changes between updated nonstationary versus EC-generated DSI for Fergus Shand dam for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	1.05	1.02	1.01	1.01	1.01	1.02
	[**4.5%**]	[**1.8%**]	[**1.21%**]	[**1.1**%]	[**1.3**%]	[**1.75**%]
30-min	1.02	0.98	0.98	1.0	1.02	1.05
	[**2.1%**]	[-2.05%]	[-1.89%]	[0.06%]	[**2.3**%]	[**4.8**%]
1-hour	0.99	0.94	0.95	0.99	1.03	1.08
	[-0.8%]	[-6.5%]	[-5.4%]	[-1.4%]	[**2.8**%]	[**7.4**%]
2-hour	0.91	0.92	0.96	1.05	1.13	1.23
	[-10.4%]	[-9.0%]	[-3.9%]	[**4.7%**]	[***11.6%***]	[***18.7%***]
6-hour	0.94	0.97	0.99	1.02	1.05	1.07
	[-6.2%]	[-3.0%]	[-0.7%]	[**2.2**%]	[**4.3**%]	[**6.6**%]

Table S26.2 Ratio and percentage changes between updated stationary versus EC-generated DSI for Fergus Shand dam for different durations

Duration	2-year	5-year	10-year	25-year	50-year	100-year
15-min	1.05	1.02	1.02	1.02	1.02	1.03
	[**4.5%**]	[**2.2%**]	[**1.73%**]	[**1.86%**]	[**2.27%**]	[**2.86%**]
30-min	1.03	0.97	0.97	0.98	1.00	1.03
	[**2.4%**]	[-2.74%]	[-3.1%]	[-1.60%]	[0.36%]	[**3.76%**]
1-hour	0.99	0.93	0.94	0.97	1.01	1.05
	[-1.03%]	[-7.1%]	[-6.48%]	[-3.02%]	[0.67%]	[**4.86%**]
2-hour	0.94	0.92	0.94	0.97	1.0	1.03
	[-6.3%]	[-8.44%]	[-6.95%]	[-3.47%]	[-0.31%]	[**3.19%**]
6-hour	0.96	0.97	0.99	1.01	1.03	1.05
	[-4.36%]	[-2.7%]	[-1.1%]	[1.12%]	[**2.79**%]	[**4.53**%]

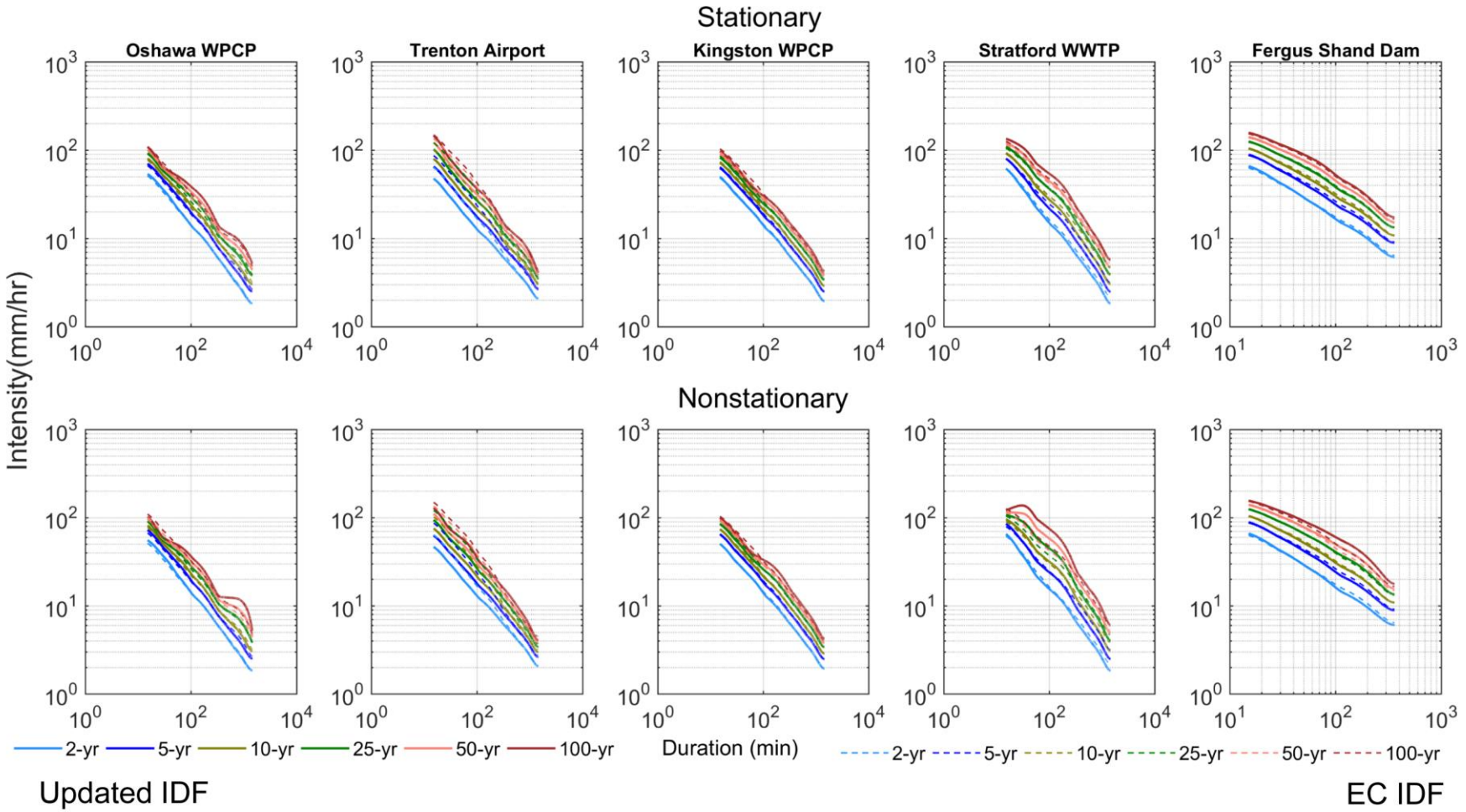

Figure S15. Estimated nonstationary versus EC-generated IDFs for return periods T = 2, 5, 10, 25, 50 and 100-year for the urbanized and moderately locations in Southern Ontario. The updated and EC IDFs are shown using solid and dotted lines respectively.

SI 5 Statistical Significance of Nonstationary versus Stationary DSI

The (statistically) significant differences in DSI is computed using the statistical test for the difference between two means or the standardized z-statistics, which is given by (Madsen et al., 2009; Mikkelsen et al., 2005)

$$z = \frac{\hat{z}_T^{NonSta} - \hat{z}_T^{Sta}}{\sqrt{0.5\left(var\{\hat{z}_T^{NonSta}\} + var\{\hat{z}_T^{Sta}\}\right)}} \tag{5.1}$$

Where $\hat{z}_T^{NonSta}$ is the *T*-year DSI obtained from the best selected nonstationary model, $\hat{z}_T^{Sta}$ describes the same but with the best stationary model. The denominator indicates predictive uncertainty; $var\{\hat{z}_T^{NonSta}\}$ and $var\{\hat{z}_T^{Sta}\}$ are the estimated variance obtained from the *T*-year event estimates and corresponding 95% credible interval. The z-statistic can be interpreted as statistically equivalent to quantiles of standard normal distribution, *i.e.*, z = ± 1.64 correspond to 10% significance levels. The null hypothesis of the test assumes the *T*-year event estimate obtained using the best fitted nonstationary model is significantly different from its best fitted stationary counterpart.

Table S27. Z-statistics* between best selected nonstationary and stationary model for 2-year return period

Duration	Toronto	Hamilton	Oshawa	Windsor	Kingston	London	Trenton	Stratford	Fergus Shand
15-min	-0.038	-0.362	0.251	-0.354	0.297	-0.032	-0.265	0.521	0.000
30-min	-0.050	-0.126	0.244	0.011	0.068	-0.158	0.030	0.118	-0.037
1-hour	-0.331	0.151	0.127	-0.288	0.158	-0.009	0.161	-0.168	0.016
2-hour	-0.170	0.041	-0.012	-0.153	-0.128	-0.041	0.246	0.371	-0.324
6-hour	0.389	-0.147	0.195	0.151	0.021	-0.088	0.183	0.000	-0.154
12-hour	0.214	-0.073	-0.031	0.088	-0.266	-0.038	0.032	-0.032	-
24-hour	0.000	-0.108	0.044	-0.116	-0.116	-0.049	-0.126	0.000	-

*The standardized Z-statistic is positive (negative) with an increasing (decreasing) trend, and statistically significant at 10% significance level when $|z| > \pm 1.64$.

Table S28. Z-statistics between best selected nonstationary and stationary model for 10-year return period

Duration	Toronto	Hamilton	Oshawa	Windsor	Kingston	London	Trenton	Stratford	Fergus Shand
15-min	0.373	-0.227	0.061	0.065	0.171	0.324	-0.281	0.196	-0.029
30-min	-0.201	0.001	-0.147	0.016	0.038	-0.295	-0.120	0.373	0.050
1-hour	-0.346	-0.007	-0.018	-0.068	0.075	0.151	0.224	0.348	0.046
2-hour	-0.324	0.309	-0.105	0.014	0.091	0.081	-0.070	0.397	0.129
6-hour	0.103	0.014	0.074	0.174	0.041	-0.277	0.212	0.140	0.017
12-hour	0.526	0.089	0.213	-0.043	-0.123	0.156	-0.264	0.201	-
24-hour	-0.039	0.061	0.000	-0.016	-0.071	-0.107	-0.104	0.047	-

Table S29. Z-statistics between best selected nonstationary and stationary model for 50-year return period

Duration	Toronto	Hamilton	Oshawa	Windsor	Kingston	London	Trenton	Stratford	Fergus Shand
15-min	0.569	-0.144	-0.137	0.287	0.081	0.636	-0.287	-0.214	-0.032
30-min	-0.244	0.080	-0.230	0.077	0.035	-0.234	-0.160	0.499	0.050
1-hour	-0.497	-0.153	-0.054	0.120	0.039	0.273	0.195	0.564	0.059
2-hour	-0.498	0.465	-0.126	0.045	0.207	0.162	-0.205	0.349	0.372
6-hour	0.016	0.107	-0.049	0.142	0.077	-0.290	0.204	0.198	0.050
12-hour	0.635	0.138	0.328	-0.064	-0.037	0.273	-0.308	0.345	-
24-hour	-0.049	0.103	0.005	0.000	-0.024	-0.108	-0.052	0.111	-

Table S30. Z-statistics between best selected nonstationary and stationary model for 100-year return period

Duration	Toronto	Hamilton	Oshawa	Windsor	Kingston	London	Trenton	Stratford	Fergus Shand
15-min	0.615	-0.120	-0.199	0.346	0.054	0.748	-0.284	-0.311	-0.032
30-min	-0.249	0.098	-0.239	0.092	0.035	-0.208	-0.166	0.529	0.046
1-hour	-0.531	-0.192	-0.059	0.175	0.031	0.309	0.182	0.613	0.062
2-hour	-0.534	0.510	-0.127	0.051	0.236	0.185	-0.231	0.327	0.441
6-hour	-0.002	0.130	-0.083	0.130	0.087	-0.280	0.200	0.209	0.058
12-hour	0.652	0.148	0.355	-0.067	-0.010	0.302	-0.311	0.390	-
24-hour	-0.052	0.110	0.007	0.000	-0.018	-0.104	-0.041	0.127	-

References

Anderson, D. R., Burnham, K. P. and White, G. C.: AIC model selection in overdispersed capture-recapture data, Ecology, 75(6), 1780–1793, 1994.

Bozdogan, H.: Akaike's information criterion and recent developments in information complexity, J. Math. Psychol., 44(1), 62–91, 2000.

Burnham, K. P. and Anderson, D. R.: Model selection and multimodel inference: a practical information-theoretic approach, Springer Science & Business Media., 2003.

Caroni, C. and Panagoulia, D.: Non-stationary modelling of extreme temperatures in a mountainous area of Greece, REVSTAT–Statistical J., 14(2), 217–228, 2016.

Cheng, L., AghaKouchak, A., Gilleland, E. and Katz, R. W.: Non-stationary extreme value analysis in a changing climate, Clim. Change, 127(2), 353–369, 2014.

Cheng, L. and AghaKouchak, A.: Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate, Sci. Rep., 4, doi: 10.1038/srep07093, 2014.

Coles, S. G. and Tawn, J. A.: A Bayesian Analysis of Extreme Rainfall Data, J. R. Stat. Soc. Ser. C Appl. Stat., 45(4), 463–478, 1996.

Dawson, C. W., Abrahart, R. J. and See, L. M.: HydroTest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts, Environ. Model. Softw., 22(7), 1034–1052, 2007.

Gelman, A., Shirley, K. and others: Inference from simulations and monitoring convergence, Handb. Markov Chain Monte Carlo, 163–174, 2011.

Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B.: Bayesian data analysis, Chapman & Hall/CRC Boca Raton, FL, USA. (2014).

Gilleland, E. and Katz, R. W.: Extremes 2.0: an extreme value analysis package in r, J. Stat. Software, 2016.

Gu, X., Zhang, Q., Singh, V. P., Xiao, M. and Cheng, J.: Nonstationarity-based evaluation of flood risk in the Pearl River basin: changing patterns, causes and implications, Hydrol. Sci. J., 62(2), 246–258, doi:10.1080/02626667.2016.1183774, 2017.

Güntner, A., Olsson, J., Calver, A. and Gannon, B.: Cascade-based disaggregation of continuous rainfall time series: the influence of climate, Hydrol. Earth Syst. Sci. Discuss., 5(2), 145–164, 2001.

Hirsch, R. M., Slack, J. R. and Smith, R. A.: Techniques of trend analysis for monthly water quality data, Water Resour. Res., 18(1), 107–121, 1982.

Hu, S.: Akaike information criterion, Cent. Res. Sci. Comput., 93, Available from: http://www4.ncsu.edu/~shu3/Presentation/AIC_2012.pdf, Accessed on June, 2017, 2007.

Hurvich, C. M. and Tsai, C.-L.: Model selection for extended quasi-likelihood models in small samples, Biometrics, 1077–1084, 1995.

Jebari, S., Berndtsson, R., Olsson, J. and Bahri, A.: Soil erosion estimation based on rainfall disaggregation, J. Hydrol., 436, 102–110, 2012.

Karmakar, S. and Simonovic, S.: Flood Frequency Analysis Using Copula with Mixed Marginal Distributions, Water Resour. Res. Rep. Available from: http://ir.lib.uwo.ca/wrrr/19, 2007.

Kass, R. E. and Raftery, A. E.: Bayes factors, J. Am. Stat. Assoc., 90(430), 773–795, 1995.

Madsen, H., Arnbjerg-Nielsen, K. and Mikkelsen, P. S.: Update of regional intensity–duration–frequency curves in Denmark: tendency towards increased storm intensities, Atmospheric Res., 92(3), 343–349, 2009.

Mandelbrot, B. B.: Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier, in Multifractals and $1/f$ Noise, J. Fluid Mechanics, 62(2), 331-358, 1974.

Mikkelsen, P. S., Madsen, H., Arnbjerg-Nielsen, K., Rosbjerg, D. and Harremoës, P.: Selection of regional historical rainfall time series as input to urban drainage simulations at ungauged locations, Atmospheric Res., 77(1–4), 4–17, 2005.

Olsson, J.: Evaluation of a scaling cascade model for temporal rain-fall disaggregation, Hydrol. Earth Syst. Sci. Discuss., 2(1), 19–30, 1998.

Olsson, J.: RANDOM CASCADE MODEL: documentation, Technical Report, Swedish Meteorological and Hydrological Institute, 2012.

Panagoulia, D., Economou, P. and Caroni, C.: Stationary and nonstationary generalized extreme value modelling of extreme precipitation over a mountainous area under climate change, Environmetrics, 25(1), 29–43, 2014.

Petrow, T. and Merz, B.: Trends in flood magnitude, frequency and seasonality in Germany in the period 1951–2002, J. Hydrol., 371(1), 129–141, 2009.
Rana, A., Bengtsson, L., Olsson, J. and Jothiprakash, V.: Development of IDF-curves for tropical india by random cascade modeling, Hydrol. Earth Syst. Sci. Discuss., 10(4), 4709–4738, 2013.
Renard, B., Sun, X. and Lang, M.: Bayesian methods for non-stationary extreme value analysis, in Extremes in a Changing Climate, pp. 39–95, Springer, 2013.
Ross, G. J.: Detecting changes in high frequency data streams, with applications, Ph.D. thesis, Imperial College London, 2013.
Ross, G. J., Tasoulis, D. K. and Adams, N. M.: Nonparametric Monitoring of Data Streams for Changes in Location and Scale, Technometrics, 53(4), 379–389, 2011.
Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, J. Am. Stat. Assoc., 63(324), 1379–1389, 1968.
Serinaldi, F. and Kilsby, C. G.: The importance of prewhitening in change point analysis under persistence, Stoch. Environ. Res. Risk Assess., 30(2), 763–777, 2016.
Ter Braak, C. J.: A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces, Stat. Comput., 16(3), 239–249, 2006.
Ter Braak, C. J. and Vrugt, J. A.: Differential evolution Markov chain with snooker updater and fewer chains, Stat. Comput., 18(4), 435–446, 2008.
Wang, X., Huang, G., Liu, J., Li, Z. and Zhao, S.: Ensemble projections of regional climatic changes over Ontario, Canada, J. Clim., 28(18), 7327–7346, 2015.
Willems, P.: Impacts of climate change on rainfall extremes and urban drainage systems, IWA Publishing, 2012.
Xie, H., Li, D. and Xiong, L.: Exploring the ability of the Pettitt method for detecting change point by Monte Carlo simulation, Stoch. Environ. Res. Risk Assess., 28(7), 1643–1655, 2014.
Yaglom, A. M.: The influence of fluctuations in energy dissipation on the shape of turbulence characteristics in the inertial interval, in Soviet Physics Doklady, vol. 11, p. 26, 1966.
Yue, S.: A bivariate gamma distribution for use in multivariate flood frequency analysis, Hydrol. Process., 15(6), 1033–1045, 2001.