Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-325
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 12 June 2017

Hydrology and Earth System Sciences
Open Access
Discussions

Does Nonstationarity in Rainfall Requires Nonstationary Intensity-Duration-Frequency Curves?

Poulomi Ganguli[1], Paulin Coulibaly[1]

[1]Department of Civil Engineering, McMaster Water Resources and Hydrologic Modelling Group, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4L7, Canada

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

Abstract. In Canada, increased risk of flooding due to heavy rainfall has risen in recent decades; most notable example include July 2013 storm in Greater Toronto region. We investigate nonstationarity and trends in the short-duration precipitation extremes in selected urbanized locations in Southern Ontario, Canada, and evaluate the potential of nonstationary Intensity-Duration-Frequency (IDF) curves, which form an input to civil infrastructural design. Despite apparent signals of nonstationarity in precipitation extremes in all locations, the stationary versus nonstationary models do not exhibit any significant differences in the design storm intensity. The signatures of nonstationarity in rainfall extremes do not necessarily imply the use of nonstationary IDFs for design considerations. When comparing the proposed IDFs with current design standards, for return periods (10-year or less) typical for urban drainage design, current design standards require an update up to 11%, whereas for longer recurrence intervals (50 - 100-year), ideal for critical civil infrastructural design, updates ranging between ~ 2 to 30% are suggested. We further emphasize that above findings need re-evaluation in light of climate change projections since intensity and frequency of extreme precipitation are expected to intensify due to global warming.

1 Introduction

Short-duration extreme rainfall events can have devastating consequences, damage to crops and infrastructures, leading to severe societal and economic losses in Canada (Canadian Climate Forum, 2013; Toronto Region Conservation Authority, 2013). In a warming climate, extreme precipitation events are expected to intensify due to moistening of the atmosphere (Donat et al., 2016; Fischer and Knutti, 2016; Pendergrass et al., 2015; Prein et al., 2016). The extent of urbanization also contributes to extreme regional precipitation through urban heat island effect and aerosol concentration (Mölders and Olson, 2004; Niyogi et al., 2007; Mohsin and Gough, 2012; Wang et al., 2015). For densely populated Southern Ontario, Canada, observations and multiple climate models suggest an increasing trend in regional surface temperature and extreme precipitation in recent decades (Stone et al., 2000; Paixao et al., 2011; Mailhot et al., 2012; De Carolis, 2012; Burn and Taleghani, 2013; Shephard et al., 2014; Deng et al., 2016). A recent study shows an increase in local surface temperature of 3.06 ± 0.18 °C/century in Greater

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-325
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 12 June 2017

Hydrology and Earth System Sciences Discussions

Toronto Area (GTA) since the 1960s (Berkeley Earth, 2017). In July 2013, a single storm event resulted in 126 mm of rainfall in GTA causing total insured losses of around $940 million and claimed to be the third most expensive weather-related event in Canada (CDD, 2015; TRCA, 2013).

Extreme rainfall statistics are often mathematically expressed using the concept of exceedance probability or *T*-year return period [*i.e.*, $T = 1/(1 - Fp(P))$, where $Fp(P)$ is the cumulative probability of the underlying distribution], and graphically as a decision relevant metrics in the form of Intensity-Duration-Frequency (IDF) curves (or relations) (ASCE, 2006; CSA, 2010; EC, 2012). These curves are based on a comprehensive statistical analysis of historical rainfall records and widely used for the design and operation of storm-water and sewerage systems, and other engineered hydraulic structures (Coulibaly and Shi, 2005; Durrans and Brown, 2001; Lima et al., 2016; Madsen et al., 2009; Rana et al., 2013; Sandink et al., 2016; Yilmaz et al., 2014a). At given return period and the storm duration, the average design-storm intensity (hereafter referred as DSI) is determined from the at-site IDF relationships. The IDF curves are based on fitting theoretical probability distribution to short-duration (hourly and sub-hourly) annual maximum precipitation (AMP). For Canada, information for preparation of IDFs and nation-wide IDF curves (EC (Environment Canada), 2012) are archived at Environment Canada (EC) Engineering database (http://climate.weather.gc.ca/prods_servs/engineering_e.html), which are produced based on short-duration available rainfall records from tipping-bucket rain gauges (TBRG). The methodology to derive existing IDF curves has certain drawbacks, *first*, the current IDF curves in Canada are based on the assumption of stationarity, which implies statistical properties of hydroclimatic time series will remain same over the period of time. However, impact of urbanization and human-induced climate changes (Field, 2012; Milly et al., 2008; Villarini et al., 2009) raises the question whether the stationarity assumption to derive IDF curves is still reliable for urban infrastructural planning (Sarhadi and Soulis, 2017; Cheng and AghaKouchak, 2014; Jakob, 2013; Yilmaz et al., 2014a; Yilmaz and Perera, 2013). The nonstationary behavior of rainfall extremes is already being reflected in the increase in frequency or magnitude of such events recently (Dixon and Mote, 2003; Guo et al., 2006; Mölders and Olson, 2004), resulting in a shift of its distribution [*Figure* SPM 0.3 in Intergovernmental Panel on Climate Change Special Report on Extremes, IPCC SREX Report: *Field*, 2012; Fig S1: IPCC AR5 working Group Report, (Stocker et al., 2013)]. This result is subsequently validated for winter temperature extremes over northern hemisphere (Kodra and Ganguly, 2014), and regional short duration precipitation extremes in India and Australia (Mondal and Mujumdar, 2015; Westra and Sisson, 2011). Two of the recent studies (Deng et al., 2016; Mailhot et al., 2012) analyzed large ensemble of CMIP3 Global Climate Model (GCM) runs and a sub-set of regional climate models that are part of North American Regional Climate Change Assessment Program

(NARCCAP) in terms of impact-relevant metrics over Canada. Both studies confirm relative increase in intensity and magnitude of rainfall extremes, especially over Southern Ontario. This issue has come to attention in the Guidelines for Canadian Water Resources Practitioner (CSA, 2010), that urges the need for updated IDF calculations: "…*climate change will likely result in an increase in the intensity and frequency of extreme precipitation events in most regions in the future. As a result, IDF values will optimally need to be updated more frequently than in the past* ….".

Secondly, for the development of IDF curves, a particular family of distribution function from the extreme value theory (*i.e.*, Gumbel distribution or extreme value type I, hereafter referred as EVI) is currently a recommended practice. EV1 has certain drawbacks, such that it is a non-heavy tailed distribution and displays symmetry in the right and left tails (Markose and Alentorn, 2005; Pinheiro and Ferrari, 2016). However, the short-duration AMP intensities often exhibit fat-tail behavior and have left asymmetries (skewed to the left relative to standard normal distribution). In fact, a few studies in the past have shown that EV1 fits poorly to the historical rainfall extremes (Burn and Taleghani, 2013; Coulibaly et al., 2015).

Motivated with these research gaps, here we address several important questions pertained to short-duration precipitation extremes in Southern Ontario, Canada, to improve pro-active management of storm-induced urban flooding. *First*, is there any signature of statistically significant nonstationary trends (slowly or varying changes), change points or regime shifts (occurrence of any abrupt changes in mean/variance of the distribution) in short-duration AMP in densely and moderately populated urbanized locations across Southern Ontario? *Second*, does signals of nonstationarity in the time series necessitates the use of nonstationary IDFs, barring economic consideration and mathematical complexity involved in the design? *Third*, how can we use this knowledge to assess the credibility of existing EC-generated IDFs in the backdrop of a changing climate? We do not attempt to provide a methodological comparison for EC-generated versus current approach but will focus on differences in estimated DSI values between the updated and EC-IDF. Further, to this end, we test the hypothesis that signatures of nonstationarity in rainfall extremes do not necessitate the use of nonstationary IDFs for design considerations. In general, urban drainage areas have substantial proportions of impervious or semi-impervious land cover, which significantly reduces response time to extreme precipitation and increases the peak flow, resulting into storm-induced floods (Miller et al., 2014). Hence, it is the short-duration precipitation extremes, which controls the design of urban infrastructure (Mishra et al., 2012). Therefore, we focus our analysis on AMP intensity. We select Southern Ontario as a test bed because of the majority of stations with more than 30-years of available rainfall record

(Adamowski and Bougadis, 2003; Deng et al., 2016; Shephard et al., 2014). *Secondly,* recent studies indicate the region is more vulnerable to climate change than any other part of Canada (Deng et al., 2016; Mailhot et al., 2012). Furthermore, southern Ontario is one of the prominent economic hubs with largest population concentration in Canada (Bourne and Simmons, 2003; Kerr, 1965; Partridge et al., 2007). In this context, we explore a robust statistical framework to evaluate possible nonstationary trends, analyze the frequency of urban precipitation extremes and assess the risk of severe rain-induced urban flooding in Southern Ontario (Table S1).

2 Study Area and Data

2.1 Study Area

Southern Ontario is situated on a Southwest-northeast transect, in the southernmost Canadian region, and separated from the United States by lakes Erie, Huron, and Ontario (Jien and Gough, 2013; Figure 1). The study includes nine densely and moderately populated urbanized and anthropogenically altered locations of the Windsor - Kingston corridor in Southern Ontario. The specific sites include (*in the order from southwest to northeast*): Windsor Airport, London International Airport, Stratford wastewater treatment plant (WWTP), Shand Dam in Fergus on the Grand River, Hamilton Airport, Toronto International Airport, Oshawa Water Pollution Control Plant (WPCP), Trenton Airport, and Kingston Pumping Station (Figure 1; Table S1). The Digital Elevation Model (DEM) of the study area was derived from Shuttle Radar Topography Mission (SRTM) 90-m Digital Elevation Database v4.1 (Jarvis et al., 2008), which indicates a shallow slope with maximum altitude of 670 m above mean sea level (MSL). The proximity to Great Lakes and topographic effect, especially in areas to the lee of Lakes Erie, Lake Ontario, and the Georgian Bay significantly modifies the climate in the region (Baldwin et al., 2011). Convective showers and thunderstorms primarily modulate the summer rainfall, but fall rainfall is dominated by reduced convective activity and increased lake effect precipitation (Lapen and Hayhoe, 2003). Further, the topographic features and associated westerly winds in the Niagara Escarpment and the Oak Ridge Moraine, play a significant role in modulating rainfall in Toronto region. On the other hand, Windsor metropolitan area, the southernmost urbanized location in the region, has the humid continental climate, which results in warm summer temperature (30°C or higher) with the greatest precipitation in the spring and summer seasons, and lowest in the fall and winter (Sanderson and Gorski, 1978). Moreover, because of the part of Windsor-Detroit international

transborder agglomeration, the extreme summer precipitation in the city of Windsor is primarily influenced by convection and urban heat island effect (Sanderson and Gorski, 1978; De Carolis, 2012).

2.2 Hydrometeorological Data

We identified the station locations (Figure 1b) based on the quality of long-range rainfall records (*e.g.*, more than 30 years or more) and 2011 Census information archived at Statistics Canada website (https://www12.statcan.gc.ca). The geographic areas of these locations are extracted from 2011 census digital boundary shape files (https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2011-eng.cfm). The Toronto metropolitan area is the most populous (over 5 million population) and known to be one of the fastest growing population base in Canada (http://torontosvitalsigns.ca/main-sections/demographics/), while Fergus is the least populated (population of around 19,000) (Table S1) city. The other cities have population ranges between ~ 500,000 (Hamilton) and 30,000 (Stratford) (Table S1). We obtained AMP observations at particular durations (15-, 30- minutes, 1-, 2-, 6-, 12- and 24-hours) with a few data gaps from Canada's National Climate Data Archive, maintained by the EC (http://climate.weather.gc.ca/prods_servs/documentation_index_e.html). The rainfall records collected from TBRG are thoroughly quality controlled (Shephard et al., 2014). These records have been previously analyzed for the assessment of national extreme rainfall trends (Burn and Taleghani, 2013; Shephard et al., 2014). We consider seven storm durations ranging from 15-, 30- minutes (the typical time of concentration for small urban catchments), and 1-, 2-, 6-, 12-, and 24- hours (the standard time of concentration for larger watersheds) following a previous study (Bougadis and Adamowski, 2006). Except for a few stations (for example, Toronto International Airport and Trenton Airport, for most of the sites the AMP observation is available either till the year 2007 or before (Table S1). Also, we found missing values in the AMP time series in all sites. We obtained daily and hourly rainfall records from the EC website and Toronto Region Conservation Authority (TRCA).

3 Methods

Figure 2 shows schematics of the overall analysis. In subsequent subsection we will discuss each of these steps in detail.

3.1 Imputation of Missing Extreme Precipitation Record

We infilled missing values and updated the AMP records by successively disaggregating daily rainfall values to hourly and sub-hourly time steps using multiplicative random cascade (MRC)-based disaggregation tool. The

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-325
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 12 June 2017

Cascade-based disaggregation model for continuous rainfall time series was suggested by (Olsson, 1995, 1998). 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. Due to freezing weather conditions during winter, most of the TBRGs' are inoperative from early November to late April of the following year. Therefore, when short-duration rainfall records were not available, the AMP values over moving windows of *n*- durations (*n* varies from 15-, 30- minutes and 1-,2-,6-,12- and 24-hours) are extracted from May to October (warm season) disaggregated rainfall volumes for remaining years. There are several reasons for selecting warm periods: *first*, extreme rainfall events mostly occur in the study area during the warm season (Cheng et al., 2010); *second*, the focus of our analysis is an investigation of extreme rainfall related flood risks and development of IDF curves using extreme rainfall statistics. We adjusted the occasional overestimation of extreme values at a higher order cascade step by a statistical post-processing method. We employed quantile matching (QM) approach (Li et al., 2010), which claims to outperform other simple bias correction methods and corrects not only the mean but also the variance of the distribution of interest (Gudmundsson et al., 2012; Teutschbein and Seibert, 2012). QM is based on equidistant cumulative probability distribution matching of observed and disaggregated AMP time series using three-parameter Generalized Extreme Value (GEV) distribution. Although like other statistical post-processing technique, QM relies on the stationarity assumption of the time series, in our case, we applied QM to entire time series of both observed and disaggregated AMP, which comes from the same station location (or similar spatial resolution) and a similar period. Therefore, we avoid potential consequences of inflation by quantile mapping (Maraun, 2013) in our analysis. We discuss the implementations of MRC, adjustment of extremes and associated model fits in more details in the Supplementary Information (SI1; Table S2; Figures S2-S8 and Tables S3-S4).

3.2 Detection of Nonstationarity

A series of statistical tests is employed to detect the presence of nonstationary trends and abrupt shifts in the short-duration AMP before frequency analysis. Figure 2 shows schematics of the overall analysis. Most of the trend and change-point detection algorithms assume observations are mutually independent. The presence of autocorrelation over/underestimates the statistical significance of trend and change-point detection algorithms (Serinaldi and Kilsby, 2016; von Storch and Navarra, 1999). We employed a Ljung-Box test with 20 lags to the short-duration AMP time series of each site to check if they show statistically significant autocorrelation (at 5% and 10% significance levels). For the time series with no serial autocorrelation, we test for trending behavior and nonstationarity. *First*, we check for a presence of nonstationarity in the time series by employing unit root-based

Augmented Dickey-Fuller (ADF; Dickey and Fuller, 1981) test. However, the test may have a low power against stationary near unit root processes (Dritsakis, 2004; Chowdhury and Mavrotas, 2006). Therefore, as a complementary to unit root test, KPSS test (Kwiatkowski et al., 1992) is employed to validate the results of the ADF test. Since both ADF and KPSS tests assume linear regression or normality of the distribution; alternatively, a log-transformation can convert a possible exponential trend present in the data into a linear trend. Therefore, following previous studies (Gimeno et al., 1999; Van Gelder et al., 2006), AMP time series is log-transformed before applying stationarity tests. The other test we employed is frequency-based Priestley and Subbarao test ['PSR'-test; (Priestley and Rao, 1969) for nonstationarity. Next, we detected the presence of smooth and abrupt changes in the time series. The continuous or monotonic trends in short-duration rainfall extremes are identified using non-parametric Mann-Kendall trend statistics with correction for ties (Hamed and Rao, 1998; Reddy and Ganguli, 2013) at 5 and 10% significance levels. In general, the abrupt change (or change point) in the time series occur at a single point in the record and bifurcate the time series into two halves, either with different means, variances, or both dissimilar means and variance together at each part. The change-point in location (or mean) is identified using non-parametric Pettit's (Pettitt, 1979) and Mann-Whitney tests (Ross et al., 2011). The shift in scale (or variance) is detected using non-parametric Mood's Test (Ross et al., 2011; See Figure 2 for details). We applied nonparametric tests due to their robustness to non-normality, which usually appears in the hydroclimatic time series. For the time series with significant autocorrelation, we employed a trend-free pre-whitening procedure (TPFW; SI 2) as described in (Yue et al., 2002, 2003) and later modified by (Petrow and Merz, 2009). Then, we applied trend and change point detection algorithms to the pre-whitened AMP extremes.

3.3 Extreme Value Analysis of Sub-daily and Daily Precipitation Extremes

Next, we perform frequency analysis of AMP intensity using GEV distribution. GEV distribution is a combination of Gumbel, Fréchet and Weibull distributions and is fitted to block or AM time series (Cheng and AghaKouchak, 2014; Katz et al., 2002; Katz and Brown, 1992). The GEV distribution contains three parameters, the location, the scale and the shape of the distribution, which describes the center of the distribution, the deviation around the mean and the shape or the tail of the distribution. The shape parameter zero indicates Gumbel, which is described by an unbounded light tailed distribution and the tail decreases rapidly following an exponential decay. On the other hand, the positive shape parameter denotes Fréchet and is a heavy-tailed distribution, and the tail drops relatively slowly following a polynomial decay (Towler et al., 2010). The negative shape parameter represents a Weibull distribution, which is a bounded distribution. Here we compare the performance of both stationary and nonstationary form of GEV distribution. For stationary model, we estimate parameters by maximizing the log-

likelihood function of the distribution (*i.e.*, Maximum Likelihood or ML-based) and Bayesian inference (BI) coupled with Differential Evaluation Markov Chain (DE-MC) Monte Carlo (MC) simulation as suggested by (Cheng et al., 2014; Cheng and AghaKouchak, 2014). For nonstationary model, the shape parameter is assumed as constant throughout. However, we incorporate time- varying covariates into GEV location (GEV I), and both in location and scale (GEV II) respectively, to describe trends as a function of time (in years) (See SI 3 for details). Then we estimate parameters of the nonstationary GEV distribution by integrating BI combined with DE-MC simulation. For AMP intensity, we derive the time variant parameter(s) from the 50^{th} (the median or the mid-point of the distribution) percentiles of the DE-MC sampled parameter(s). We obtain the associated 95% confidence intervals (the bounds) from the 5^{th} and 95^{th} percentiles of the simulated samples. We perform the calculations following (Cheng and AghaKouchak, 2014) using an MATLAB-based software package, Nonstationary Extreme Value Analysis (NEVA, Version 2.0). The Akaike information criterion (*AIC*) with small sample correction (AIC_c) is used to identify the best model, which claims to avoid overfitting the data as compared to traditional AIC (Burnham and Anderson, 2004; Hurvich and Tsai, 1995). Besides this, we assess the performance of the models using probability-probability (*PP*) plots. Following a previous study (El Adlouni et al., 2007), we select the model with fewer parameters as the best model when two models have comparable performances. For example, we chose GEV I as the best fit for 15-minute and 12-hour durations rainfall extremes at Hamilton and Trenton Airport respectively (Table S17 and Table S22). The derived model parameters are then utilized to obtain DSI using the concept of a *T*-year return period. We discuss the methods to estimate DSI and *T*-year return periods using stationary and nonstationary methods in detail in section SI 5. To test (statistically) significant difference in the estimated DSI from the best-selected stationary versus nonstationary models, we calculate standard z-statistics for selected return periods (Madsen et al., 2009; Mikkelsen et al., 2005). We applied the 2-sided option with 5% and 10% significance levels to assess the statistical significance of the test statistics (See SI 6 for details). Finally, we compared the DSI obtained from nonstationary and stationary models with existing EC-generated DSI estimates.

4 Results and Discussions

The extreme rainfall statistics show high standard deviation with positive skew behavior (Tables S5 and S6). The skewness is a measure of the asymmetry in the AMP distribution. Positive values of skewness indicate a shift towards an increase in the intensity of extreme events. For example, extremes at London International Airport, Trenton Airport, Stratford WWTP and Windsor Airport (Table S5 - S6 and Figure 3). The majority of stations show positive excess kurtosis (Tables S5 and S6), which indicates extreme values are more frequent in the time series. We find the presence of statistically significant autocorrelation in the AMP time series of Toronto

International Airport, Hamilton Airport, and Fergus Shand Dam (Table S7.1, Table S8.1, and Table S15). We apply nonparametric TPFW to precipitation extremes with a significant autocorrelation (Table S7.1, Table S8.1, and Table S15). However, two successive TPFWs fail to correct the effect of autocorrelation in 12- and 24-hour duration extremes in Shand Dam. Hence we exclude those two time series from frequency analysis (Table S15). The ADF-test for nonstationarity is statistically significant in all durations, as indicated by the higher *p*-values. As a complementary to ADF test, we also employed KPSS and PSR tests (Figure 2; Tables S7 – S15) to check significant nonstationarity. Figure 3 shows the distribution of spatial trends in short-duration rainfall extremes. We find co-occurrences of trends, change points and nonstationarities in extremes at multiple locations (Figure 3). In general, the three sites in the extreme Northeast, the Oshawa WPCP, Trenton Airport and Kingston P. Station show evidence of statistically significant upshifts and nonstationarities in the time series, whereas the rest of the sites in the Southwest exhibit downshifts and statistically significant nonstationarities (Figure 3). For 2-hour and beyond durations, London International Airport shows a presence of statistically significant downward trends with change points. An increase (decrease) in mean precipitation imply an increase in heavy precipitation and vice-versa. Further, it could also alter the shape of the right-hand tail, changing overall asymmetry in the distribution (Fig. S1), and hence affecting the nature of extremes (Stocker et al., 2013). Furthermore, the presence of opposite signs of trends within a proximity of sites are prominent in all durations, for example, except for 1-hour duration, extremes in all durations at Toronto International Airport and Oshawa WPCP show downwards and upward shifts respectively. Our findings confirm the other study (Burn and Taleghani, 2013), where authors report a lack of spatial structures and presence of different trends within a close vicinity of stations. We find even if trends in individual sites may not deem significant, the magnitude of trends (as measured by slope per decade, Tables S7 – S15) is never zero in any of the sites.

A weak trend can also have a significant impact on the results of probability analysis (Porporato and Ridolfi, 1998). Hence even if precipitation extremes often exhibit statistically insignificant trends in few durations, we assess the performance of both nonstationary and stationary models in all sites (Tables S16 – S24). Barring a few exceptions, the shape parameters in most of the models range between -0.30 and +0.3, which is an acceptable range of GEV shape parameter as shown in an earlier study (Martins et al., 2000). Our results corroborate well with recent research (Papalexiou et al., 2013; Wilson and Toumi, 2005), which showed that distribution with fat tails (with GEV shape parameter, $\xi < 0$) fits better for the precipitation extremes. As a measure of uncertainty, we also report the 95% confidence interval of design rainfall quantiles at 100-year return period as a ratio between the upper and the lower bounds, which ranges between the factor of 1.2-to-1 and 4.0-to-1 in all cases. The performance of time-varying

GEV models (Figure S9) closely follows the spatial pattern as indicated in the trend map (Figure 3). For example, Trenton Airport, which showed significant upward trends with change points and nonstationarity, is better modeled by the nonstationary GEV distributions for most of the durations. Likewise, except a few cases, we find that GEV II fits best if the time series exhibit (significant) evidence of a change point in variance, for example, 15-min and 1-hr extremes in Trenton Airport (Table S22) and Kingston P. station (Table S20) respectively. However, in many cases, the performance of nonstationary models are often comparable and even superseded by their stationary counterparts (Figure S9, Tables S16 - S24). In fact, the scatter of data points in the *PP*-plots (Figures S10 – S13) suggests a close resembles between stationary and nonstationary models across all durations. Figure 4 shows the relation between DSI and durations (ranges between 15-min and 24-hr) for 100-year return periods estimated by stationary versus nonstationary GEV distributions. While the interquartile range of the boxplot shows the uncertainty in estimated rainfall quantiles from the Bayesian analysis, the black circles denote the DSI obtained from the ML-based estimators (Figure 4). The median of DSI (as marked by the horizontal line inside the box), simulated by the stationary DE-MC GEV models are found to be approximately similar to the one obtain by the stationary ML-based models, indicating a close agreement between two methods. However, the spread of the boxes simulated by the nonstationary model is found to be relatively narrower as compared to the one simulated by the stationary model for most of the sites (Figure 4), indicating lesser uncertainty in the estimated quantiles. For return periods of less than 100 years, such as for 10- and 5-year, the DSI from stationary versus nonstationary models, show subtle differences (Figures S14 – S15).

Figure 5 displays the differences in DSI obtained using the best performing nonstationary model relative to the best selected stationary models using z-statistics for different durations and return periods. The standardized z-statistics show positive (negative) values indicating an increase (decrease) in DSI values assuming nonstationarity in return period estimates against its stationarity counterparts. Comparison between *T*-year event estimates from both models indicates statistically indistinguishable differences in rainfall intensity. We find for all return periods and durations, z-statistics ranges between -1 and +1 for all nine sites. Nonetheless, extreme precipitation intensity shows either positive or negative (statistically insignificant) changes in signs. The difference between DSI shows a decrease, especially between 15-min and 2-hour for Hamilton, Windsor, London International Airport, and Shand Dam for 50- and 100-year return periods (Figure 5, Tables S25-S26). These findings are in agreement with Figure 3, in which sites show a decrease in trends in extreme rainfall. In contrast, Windsor Airport shows an increase in DSI value at 12-hour duration (Figure 5; Tables S25-S26). However, the performance statistics in Windsor Airport indicate a comparable performance of the stationary versus nonstationary GEV models at 12- and 24-hour duration

rainfall extremes (Figure S9; Table S19). At 10-year return period, which is typical for most urban drainage planning, the differences are close to zero (Figure 5, Table S27) for most of the durations. For return period less than 10-year, we notice a mix performance (Figure 5, Table S28). These findings pose an important question: does the presence of nonstationarities in rainfall extremes require the design of nonstationary IDF curves? We speculate that although it is crucial to recognize nonstationarity in precipitation extremes, the stationary form of IDFs can still represent the extreme rainfall statistics for the present-day climate over Southern Ontario region. Our results are consistent with a recent study performed in the city of Melbourne, Southeast Australia (Yilmaz et al., 2014a; Yilmaz and Perera, 2013) that found despite the presence of (statistically) significant trends in rainfall extremes; nonstationary GEV models did not show any additional advantages over the stationary models. Similarly, another study (Singh et al., 2016) over India found noticeable differences in extreme rainfall quantiles of nonstationary versus the stationary method of analysis in urbanizing (rural to urban transition zone) areas as compared to the prime urbanized and rural areas. However, in their study, authors have not assessed the statistical significance of the difference in estimated precipitation quantiles using both methods. As supported by the previous literature (Singh et al., 2016), we attribute that the little or no changes in extreme rainfall statistics in the urbanized setting is due to the stabilization of urban development leading to no substantial variations in the land use pattern. Hence, no significant changes in synoptic scale circulations, which in turn affect space-time pattern in rainfall extremes (Moglen and Schwartz, 2006). More general, highlighting advantages and limitations of nonstationary versus the stationary methods of analyses (Koutsoyiannis and Montanari, 2015; Montanari and Koutsoyiannis, 2014; Serinaldi and Kilsby, 2015) are beyond the scope of the current study. Nevertheless, we hypothesize that above findings require a careful re-evaluation in light of climate projections, in which frequency and magnitude of extreme rainfall are expected to intensify in future (Mailhot et al., 2012; Deng et al., 2016; Fischer and Knutti, 2016; Prein et al., 2016).

Figure 6 compares the *T*-year event estimates of updated nonstationary versus EC-generated IDFs for different return periods. The median and associated lower and upper bounds of the ratio of regional nonstationary versus EC-generated *T*-year event estimates can interpret as analogous to most likely, minimum and maximum plausible scenarios. For T = 10-year return period, the ratio of updated versus old estimates of DSI are in the order of ~ 1.1 – 1.3 for Oshawa WPCP, Kingston P. station, Stratford WWTP, Windsor and Trenton Airport. The increase in the estimated ratio being more pronounced at 50- and 100-year return periods and are in the order of ~ 1.1 - 1.5 (Figure 6, Table S29.1 – Table S37.2). While for Toronto International Airport and Hamilton Airport, we find no increase in the short-duration rainfall extremes of less than 1- hour, the increase is more pronounced for longer durations

and larger return periods (12 and 24-hour, and 50- and 100-year return period, Figure 6, bottom panel). For longer recurrence intervals, while the heat maps of minimum bounds and the most likely scenario show a lesser number of stations and durations to reach a ratio of 1.5, the maximum bounds suggest a sharp increase in the ratio across all durations and locations. Further, for return periods more than 50-year, the increase in the ratio is more prominent in the upper bound of the stationary models (Figure S16) as compared to the nonstationary models. The resulting increase in *T*-year event estimates is because of the relatively wider confidence interval of estimated DSIs in stationary models than that of the nonstationary models (Figures 6; S14 – S15). In general, for larger return periods, our analysis reveals, the increase in the ratio of updated versus EC-generated rainfall intensity is more prominent in sites with statistically significant signatures of nonstationarity, upward trends, and change points. For example, the updated DSIs of Oshawa WPCP (Table S31.1) and Trenton Airport (Table S35.1) shows an increase in the ratio for most of the durations and return periods as compared to the EC-generated DSI values. On the other hand, the hourly precipitation extremes in London International Airport, in general, show a decrease in the ratio (Table S34.1), which is predominantly due to the presence of significant downward trends with change points in the time series.

Based on the study results and in anticipation of stakeholders' participation in adaptive management, we present updated nonstationary IDFs for the nine urbanized and semi-urbanized locations across Southern Ontario (Figure 7). It also makes the first attempt to compare the results of updated versus EC-generated IDFs, which are the part of contemporary Design Standards and widely used by the stakeholders and practitioners. Overall, the updated IDFs closely follow the pattern of trends analogous to EC-generated IDFs, except for the 100-year return period. At longer durations and higher return periods, stations in metropolitan areas (such as Toronto International Airport, Hamilton Airport, Oshawa WPCP and Windsor Airport) show large differences in DSIs, whereas moderately populated Stratford WWTP, shows substantial differences in less than 1-hour extremes. For T = 10-year or less, we find a decrease in the range of ~ 2 – 40% in the *T*-year event estimates (Table S29.1 – S37.2). Meanwhile, for T = 10-year return period, we find the increase is in the order of ~1.2 – 11% in several stations (*i.e.*, Oshawa WPCP, Windsor Airport, Kingston P. station, Trenton Airport and Stratford WWTP). For T = 50-year and more, the increase ranges between ~ 1.7 – 30%. We find the largest increase is for the 12-hour rainfall extreme in Toronto International Airport (~ 20 – 31.5%; Table S29.1), followed by 1-hour extreme at Stratford WWTP (~ 23 – 31%; Table S36.1). In summary, our findings confirm that updates in the order of ~ 2 – 30% are require based on locations and return periods to mitigate the risk of precipitation induced urban flooding irrespective of the choice of methods used in the IDF estimation (Table S29.1 – S37.2). The results are consistent with (Simonovic and Peck,

2009), in which authors recommends an average update of about 21%, with a difference, ranges between ~ 11 – 35% for the updated versus EC-IDF in London Metropolitan area. However, they assume stationarity in the method of IDF estimation.

The increase could also indicate a tendency towards an increase in mean precipitation and (or) a shift in the distribution, affecting its tail behavior. However, a few caveat remains, for example, a critical question could be: does an increase in DSI is potentially linked towards more frequent and more intense precipitation extremes or is it an artifact of the new dataset in the update process? It is worthwhile to note that results shown here are manifestations of present-day climate using ground-based hydrometeorological observations and the specific insights are nonetheless subject to the quality of available rainfall records. It remains an open-ended question to what extent we credibly develop IDFs in a changing climate (Coulibaly et al., 2015) since there is no uniformly accepted method of generating IDF information and related projection uncertainty in light of climate change. Further, several studies (Deng et al., 2016; Kunkel, 2003) has indicated an increase in frequency and magnitude of short-duration rainfall extremes in Southern Ontario due to global warming. Research towards this direction is currently underway for regional preparedness and to develop comprehensive adaptation strategies.

4 Conclusions

This paper has sought to assess signatures of nonstationarity in densely and moderately populated urbanized locations in Southern Ontario, which is one of the major economic hubs in Canada. We update short-duration rainfall extremes with latest available ground-based observations and present a comprehensive analysis to evaluate nonstationary versus the stationary method of IDF estimation. This analysis yields two principal findings. First, despite signatures of (statistically) significant nonstationarity and trends in extremes in most of the sites, the changes in design storm intensity remain statistically indistinguishable using stationary versus nonstationary methods. Second, comparison of at-site T-year event estimates of updated versus EC-generated IDFs shows for T = 10-year, the return period commonly used for urban drainage design, current design standards require updates up to 11% to mitigate the risk of urban flooding. Meanwhile, for longer recurrence interval (T = 50-year or more), typical for critical civil infrastructural design, comparison of updated versus EC-generated IDF curves shows a difference that ranges between 2% and 30% based on locations. Preliminary investigations based on regional and global climate model simulations in the study area confirm a considerable uncertainty in the projection of short-duration and high-intensity extreme rain events (Coulibaly et al., 2015). While short-duration precipitation extremes are typically controlled by synoptic scale moisture convergence (Ruiz-Villanueva et al., 2012; Westra

and Sisson, 2011), the daily extremes are often modulated by large-scale circulation patterns and local orographic factors (Carvalho et al., 2002; Gershunov and Barnett, 1998; Trenberth, 1999). Further, the role of natural variability and multidecadal modes of sea-surface temperature (SST) in modulating Canadian extreme rainfall intensity have already been shown in the past (Gan et al., 2007; Shabbar et al., 1997). The future research should include two aspects. First, investigation of physical drivers (such as temperature, decadal and multidecadal modes of SST) in influencing short-duration rainfall extremes. Second, the inclusion of these covariates in nonstationary IDF development (Mondal and Mujumdar, 2015; Yilmaz et al., 2014b). Given that these findings are for the current period climate, it is also recommended to explore hypothesis-driven studies for the future time periods using high-resolution climate model simulations.

Author Contribution

P. Ganguli and P. Coulibaly designed the experiment. P. Ganguli carried out the experiment. P. Ganguli prepared the manuscript with contributions from P. Coulibaly.

Acknowledgments

The annual maximum rainfall data used in this study is downloaded from Environment Canada website: ftp://ftp.tor.ec.gc.ca/Pub/Engineering_Climate_Dataset/IDF/. Hourly and daily rainfall data are obtained from Toronto and Region Conservation Authority (TRCA; https://trca.ca/) and Environment Canada Historical Climate Data Archive (http://climate.weather.gc.ca/). The authors would like to acknowledge financial support from the Natural Science and Engineering Research Council (NSERC) of Canada and the NSERC Canadian FloodNet. The first author of the manuscript would like to thank Dr. Jonas Olsson of Swedish Meteorological and Hydrological Institute (SMHI), Norrköping, Sweden for sharing MATLAB-based random cascade disaggregation tools and implementation details through email. While change point and nonstationarity analyses were conducted in the statistical software R version 3.30 with add-on packages “trend”, “fractal” and “cpm”, the remaining analyses were performed in a MATLAB platform (MATLAB R2016a). The nonstationary GEV analyses were performed using MATLAB-based NEVA toolbox, available at the University of California, Irvine website: http://amir.eng.uci.edu/neva.php (as accessed on May 2016).

References

Adamowski, K., and J. Bougadis: Detection of trends in annual extreme rainfall, *Hydrol. Process.*, *17*(18), 3547–3560, doi:10.1002/hyp.1353, 2003.

American Society of Civil Engineers (ASCE): *Standard Guidelines for the Design of Urban Stormwater Systems, Standard Guidelines for Installation of Urban Stormwater Systems, and Standard Guidelines for the Operation and Maintenance of Urban Stormwater Systems*, ASCE/EWRI 45-05, 46-05, and 47-05., American Society of Civil Engineers, Reston, VA, 2006.

Baldwin, D. J., J. R. Desloges, and L. E. Band: 2 Physical Geography of Ontario, *Ecol. Manag. Terr. Landsc. Patterns Process. For. Landsc. Ont.*, 12, 2011.

Adamowski, K. and Bougadis, J.: Detection of trends in annual extreme rainfall, Hydrol. Process., 17(18), 3547–3560, doi:10.1002/hyp.1353, 2003.

ASCE: Standard Guidelines for the Design of Urban Stormwater Systems, Standard Guidelines for Installation of Urban Stormwater Systems, and Standard Guidelines for the Operation and Maintenance of Urban Stormwater Systems, ASCE/EWRI 45-05, 46-05, and 47-05., American Society of Civil Engineers, Reston, VA. (Accessed 9 December 2016), 2006.

Bougadis, J. and Adamowski, K.: Scaling model of a rainfall intensity-duration-frequency relationship, Hydrol. Process., 20(17), 3747–3757, 2006.

Bourne, L. S. and Simmons, J.: New Fault Lines? Recent Trends in the Canadian Urban System and Their Implications for Planning and Public Policy, Can. J. Urban Res., 12(1), 22–47, 2003.

Burn, D. H. and Taleghani, A.: Estimates of changes in design rainfall values for Canada, Hydrol. Process., 27(11), 1590–1599, 2013.

Burnham, K. P. and Anderson, D. R.: Multimodel inference understanding AIC and BIC in model selection, Sociol. Methods Res., 33(2), 261–304, 2004.

Canadian Climate Forum (CCF) (2013), *Extreme Weather*, 1(1), 1- 4, Ottawa. Available at: http://www.climateforum.ca/, accessed on, December 2016.

Carvalho, L. M., Jones, C. and Liebmann, B.: Extreme precipitation events in southeastern South America and large-scale convective patterns in the South Atlantic convergence zone, J. Clim., 15(17), 2377–2394, 2002.

Canadian Disaster Database (CDD): Public Safety Canada, Available at: https://www.publicsafety.gc.ca/cnt/rsrcs/cndn-dsstr-dtbs/index-en.aspx, accessed on: December 2016.

Cheng, C. S., Li, G., Li, Q. and Auld, H.: A Synoptic Weather Typing Approach to Simulate Daily Rainfall and Extremes in Ontario, Canada: Potential for Climate Change Projections, J. Appl. Meteorol. Climatol., 49(5), 845–866, doi:10.1175/2010JAMC2016.1, 2010.

Cheng, L. and AghaKouchak, A.: Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate, Sci. Rep., 4, Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4235283/ (Accessed 7 September 2016), 2014.

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.

Chowdhury, A. and Mavrotas, G.: FDI and Growth: What Causes What?, World Econ., 29(1), 9–19, doi:10.1111/j.1467-9701.2006.00755.x, 2006.

Coulibaly, P. and Shi, X.: Identification of the effect of climate change on future design standards of drainage infrastructure in Ontario, Rep. Prep. McMaster Univ. Funding Minist. Transp. Ont., 82, Available from: http://www.cspi.ca/sites/default/files/download/Final_MTO_Report_June2005rv.pdf (Accessed 9 December 2016), 2005.

Coulibaly, P., Burn, D., Switzman, H., Henderson, J. and Fausto, E.: A comparison of future IDF curves for Southern Ontario., Technical Report, McMaster University, Hamilton., 2015.

Canadian Standards Association (CSA): *Technical Guide – Development, Interpretation and Use of Rainfall Intensity-duration-frequency (IDF) Information: Guideline for Canadian Water Resources Practitioners*. Technical Guide Plus 4013-12, ISBN 978-1-55491-866-9, Mississauga, Ontario, Canada, 2012.

Deng, Z., Qiu, X., Liu, J., Madras, N., Wang, X. and Zhu, H.: Trend in frequency of extreme precipitation events over Ontario from ensembles of multiple GCMs, Clim. Dyn., 46(9–10), 2909–2921, doi:10.1007/s00382-015-2740-9, 2016.

Dickey, D. A. and Fuller, W. A.: Likelihood ratio statistics for autoregressive time series with a unit root, Econom. J. Econom. Soc., 1057–1072, 1981.

Dixon, P. G. and Mote, T. L.: Patterns and Causes of Atlanta's Urban Heat Island–Initiated Precipitation, J. Appl. Meteorol., 42(9), 1273–1284, doi:10.1175/1520-0450(2003)042<1273:PACOAU>2.0.CO;2, 2003.

Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A. and Maher, N.: More extreme precipitation in the world's dry and wet regions, Nat. Clim. Change, 6(5), 508–513, doi:10.1038/nclimate2941, 2016.

Dritsakis, N.: Tourism as a Long-Run Economic Growth Factor: An Empirical Investigation for Greece Using Causality Analysis, Tour. Econ., 10(3), 305–316, doi:10.5367/0000000041895094, 2004.

Durrans, S. and Brown, P.: Estimation and Internet-Based Dissemination of Extreme Rainfall Information, Transp. Res. Rec. J. Transp. Res. Board, 1743, 41–48, doi:10.3141/1743-06, 2001.

EC (Environment Canada): Documentation on Environment Canada Rainfall Intensity-DurationFrequency (IDF) Tables and Graphs. V2.20., 2012.

El Adlouni, S., Ouarda, T. B. M. J., Zhang, X., Roy, R. and Bobée, B.: Generalized maximum likelihood estimators for the nonstationary generalized extreme value model, Water Resour. Res., 43(3), W03410, doi:10.1029/2005WR004545, 2007.

Field, C. B.: *Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change*, Cambridge University Press, 2012

Fischer, E. M. and Knutti, R.: Observed heavy precipitation increase confirms theory and early models, Nat. Clim. Change, 6(11), 986–991, 2016.

Gan, T. Y., Gobena, A. K. and Wang, Q.: Precipitation of southwestern Canada: Wavelet, scaling, multifractal analysis, and teleconnection to climate anomalies, J. Geophys. Res. Atmospheres, 112(D10), D10110, doi:10.1029/2006JD007157, 2007.

Gershunov, A. and Barnett, T. P.: ENSO influence on intraseasonal extreme rainfall and temperature frequencies in the contiguous United States: Observations and model results, J. Clim., 11(7), 1575–1586, 1998.

Gimeno, R., Manchado, B. and Mínguez, R.: Stationarity tests for financial time series, Phys. Stat. Mech. Its Appl., 269(1), 72–78, 1999.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E. and Engen-Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations–a comparison of methods, Hydrol. Earth Syst. Sci., 16(9), 3383–3390, 2012.

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.

Guo, X., Fu, D. and Wang, J.: Mesoscale convective precipitation system modified by urbanization in Beijing City, Atmospheric Res., 82(1–2), 112–126, doi:10.1016/j.atmosres.2005.12.007, 2006.

Hamed, K. H. and Rao, A. R.: A modified Mann-Kendall trend test for autocorrelated data, J. Hydrol., 204(1), 182–196, 1998.

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

Jakob, D.: Nonstationarity in extremes and engineering design, in Extremes in a Changing Climate, pp. 363–417, Springer. 2013.

Jarvis, A., Reuter, H. I., Nelson, A., Guevara, E. and others: Hole-filled SRTM for the globe Version 4, Available CGIAR-CSI SRTM 90m Database Httpsrtm Csi Cgiar Org. Available from: http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1 (Accessed 13 December 2016), 2008.

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

Katz, R. W. and Brown, B. G.: Extreme events in a changing climate: variability is more important than averages, Clim. Change, 21(3), 289–302, 1992.

Katz, R. W., Parlange, M. B. and Naveau, P.: Statistics of extremes in hydrology, Adv. Water Resour., 25(8), 1287–1304, 2002.

Kerr, D.: Some Aspects of the Geography of Finance in Canada, Can. Geogr. Géographe Can., 9(4), 175–192, doi:10.1111/j.1541-0064.1965.tb00825.x, 1965.

Kodra, E. and Ganguly, A. R.: Asymmetry of projected increases in extreme temperature distributions, Sci. Rep., 4, 5884, doi:10.1038/srep05884, 2014.

Koutsoyiannis, D. and Montanari, A.: Negligent killing of scientific concepts: the stationarity case, Hydrol. Sci. J., 60(7–8), 1174–1183, 2015.

Kunkel, K. E.: North American Trends in Extreme Precipitation, Nat. Hazards, 29(2), 291–305, doi:10.1023/A:1023694115864, 2003.

Kwiatkowski, D., Phillips, P. C., Schmidt, P. and Shin, Y.: Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?, J. Econom., 54(1–3), 159–178, 1992.

Lapen, D. R. and Hayhoe, H. N.: Spatial Analysis of Seasonal and Annual Temperature and Precipitation Normals in Southern Ontario, Canada, J. Gt. Lakes Res., 29(4), 529–544, doi:10.1016/S0380-1330(03)70457-2, 2003.

Li, H., Sheffield, J. and Wood, E. F.: Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching, J. Geophys. Res. Atmospheres, 115(D10), D10101, doi:10.1029/2009JD012882, 2010.

Lima, C. H., Kwon, H.-H. and Kim, J.-Y.: A Bayesian beta distribution model for estimating rainfall IDF curves in a changing climate, J. Hydrol., 540, 744–756, 2016.

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.

Mailhot, A., Beauregard, I., Talbot, G., Caya, D. and Biner, S.: Future changes in intense precipitation over Canada assessed from multi-model NARCCAP ensemble simulations, Int. J. Climatol., 32(8), 1151–1163, 2012.

Maraun, D.: Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue, J. Clim., 26(6), 2137–2143, 2013.

Markose, S. and Alentorn, A.: The Generalized Extreme Value (GEV) Distribution, Implied Tail Index and Option Pricing, *J. Deriv.*, *18*(3), 35–60, 2005.

Martins, E. S., Stedinger, J. R. and others: Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data, Water Resour. Res., 36(3), 737–744, 2000.

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, doi:10.1016/j.atmosres.2004.10.016, 2005.

Miller, J. D., Kim, H., Kjeldsen, T. R., Packman, J., Grebby, S. and Dearden, R.: Assessing the impact of urbanization on storm runoff in a peri-urban catchment using historical change in impervious cover, J. Hydrol., 515, 59–70, doi:10.1016/j.jhydrol.2014.04.011, 2014.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P. and Stouffer, R. J.: Stationarity Is Dead: Whither Water Management?, Science, 319(5863), 573–574, doi:10.1126/science.1151915, 2008.

Mishra, V., Dominguez, F. and Lettenmaier, D. P.: Urban precipitation extremes: How reliable are regional climate models?, Geophys. Res. Lett., 39(3), L03407, doi:10.1029/2011GL050658, 2012.

Moglen, G. E. and Schwartz, D. E.: Methods for adjusting US geological survey rural regression peak discharges in an urban setting, U.S. Geological Survey Scientific Investigation Report 2006 - 5270, 1-55, 2006.

Mölders, N. and Olson, M. A.: Impact of Urban Effects on Precipitation in High Latitudes, J. Hydrometeorol., 5(3), 409–429, 2004.

Mondal, A. and Mujumdar, P. P.: Modeling non-stationarity in intensity, duration and frequency of extreme rainfall over India, J. Hydrol., 521, 217–231, 2015.

Montanari, A. and Koutsoyiannis, D.: Modeling and mitigating natural hazards: Stationarity is immortal!, Water Resour. Res., 50(12), 9748–9756, 2014.

Olsson, J.: Limits and characteristics of the multifractal behaviour of a high-resolution rainfall time series, Nonlinear Process. Geophys., 2(1), 23–29, 1995.

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

Papalexiou, S. M., Koutsoyiannis, D. and Makropoulos, C.: How extreme is extreme? An assessment of daily rainfall distribution tails, Hydrol Earth Syst Sci, 17(2), 851–862, 2013.

Partridge, M., Olfert, M. R. and Alasia, A.: Canadian cities as regional engines of growth: agglomeration and amenities, Can. J. Econ. Can. Déconomique, 40(1), 39–68, 2007.

Pendergrass, A. G., Lehner, F., Sanderson, B. M. and Xu, Y.: Does extreme precipitation intensity depend on the emissions scenario?, Geophys. Res. Lett., 42(20), 8767–8774, 2015.

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.

Pettitt, A. N.: A non-parametric approach to the change-point problem, Appl. Stat., 126–135, 1979.

Pinheiro, E. C. and Ferrari, S. L. P.: A comparative review of generalizations of the Gumbel extreme value distribution with an application to wind speed data, J. Stat. Comput. Simul., 86(11), 2241–2261, 2016.

Porporato, A. and Ridolfi, L.: Influence of weak trends on exceedance probability, Stoch. Hydrol. Hydraul., 12(1), 1–14, 1998.

Prein, A. F., Rasmussen, R. M., Ikeda, K., Liu, C., Clark, M. P. and Holland, G. J.: The future intensification of hourly precipitation extremes, Nat. Clim. Change, advance online publication, 2016.

Priestley, M. B. and Rao, T. S.: A test for non-stationarity of time-series, J. R. Stat. Soc. Ser. B Methodol., 140–149, 1969.

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.

Reddy, M. J. and Ganguli, P.: Spatio-temporal analysis and derivation of copula-based intensity–area–frequency curves for droughts in western Rajasthan (India), Stoch. Environ. Res. Risk Assess., 27(8), 1975–1989, 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.

Ruiz-Villanueva, V., Borga, M., Zoccatelli, D., Marchi, L., Gaume, E. and Ehret, U.: Extreme flood response to short-duration convective rainfall in South-West Germany, Hydrol. Earth Syst. Sci., 16(5), 1543–1559, 2012.

Sarhadi, A. and Soulis, E. D.: Time-varying extreme rainfall intensity-duration-frequency curves in a changing climate, Geophys. Res. Lett., 44(5), 2016GL072201, doi:10.1002/2016GL072201, 2017.

Sanderson, M. and Gorski, R.: The effect of metropolitan Detroit-Windsor on precipitation, J. Appl. Meteorol., 17(4), 423–427, 1978.

Sandink, D., Simonovic, S. P., Schardong, A. and Srivastav, R.: A decision support system for updating and incorporating climate change impacts into rainfall intensity-duration-frequency curves: Review of the stakeholder involvement process, Environ. Model. Softw., 84, 193–209, 2016.

Serinaldi, F. and Kilsby, C. G.: Stationarity is undead: Uncertainty dominates the distribution of extremes, Adv. Water Resour., 77, 17–36, 2015.

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, doi:10.1007/s00477-015-1041-5, 2016.

Shabbar, A., Bonsal, B. and Khandekar, M.: Canadian Precipitation Patterns Associated with the Southern Oscillation, J. Clim., 10(12), 3016–3027, 1997.

Shephard, M. W., Mekis, E., Morris, R. J., Feng, Y., Zhang, X., Kilcup, K. and Fleetwood, R.: Trends in Canadian Short-Duration Extreme Rainfall: Including an Intensity–Duration–Frequency Perspective, Atmosphere-Ocean, 52(5), 398–417, 2014.

Simonovic, S. P. and Peck, A.: Updated rainfall intensity duration frequency curves for the City of London under the changing climate, Department of Civil and Environmental Engineering, The University of Western Ontario. (Accessed 13 January 2017), 2009.

Singh, J., Vittal, H., Karmakar, S., Ghosh, S. and Niyogi, D.: Urbanization causes nonstationarity in Indian Summer Monsoon Rainfall extremes, Geophys. Res. Lett., 43(21), 2016.

Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P. M.: Climate change 2013: the physical science basis. Intergovernmental panel on climate change, working group I Contribution to the IPCC fifth assessment report (AR5), N. Y., 2013.

von Storch, H. and Navarra, A.: Analysis of climate variability: applications of statistical techniques, Springer, 1999.

Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, J. Hydrol., 456, 12–29, 2012.

Toronto Region Conservation Authority (TRCA): Resilient City: Preparing for Extreme Weather Events, 2013.

Towler, E., Rajagopalan, B., Gilleland, E., Summers, R. S., Yates, D. and Katz, R. W.: Modeling hydrologic and water quality extremes in a changing climate: A statistical approach based on extreme value theory, Water Resour. Res., 46(11), W11504, doi:10.1029/2009WR008876, 2010.

Trenberth, K. E.: Atmospheric Moisture Recycling: Role of Advection and Local Evaporation, J. Clim., 12(5), 1368–1381, 1999.

Van Gelder, P., Wang, W. and Vrijling, J. K.: Statistical estimation methods for extreme hydrological events, in Extreme hydrological events: New concepts for security, pp. 199–252, Springer, 2006.

Villarini, G., Serinaldi, F., Smith, J. A. and Krajewski, W. F.: On the stationarity of annual flood peaks in the continental United States during the 20th century, Water Resour. Res., 45(8), W08417, doi:10.1029/2008WR007645, 2009.

Westra, S. and Sisson, S. A.: Detection of non-stationarity in precipitation extremes using a max-stable process model, J. Hydrol., 406(1–2), 119–128, doi:10.1016/j.jhydrol.2011.06.014, 2011.

Wilson, P. S. and Toumi, R.: A fundamental probability distribution for heavy rainfall, Geophys. Res. Lett., 32(14), L14812, doi:10.1029/2005GL022465, 2005.

Yilmaz, A. G. and Perera, B. J. C.: Extreme rainfall nonstationarity investigation and intensity–frequency–duration relationship, J. Hydrol. Eng., 19(6), 1160–1172, 2013.

Yilmaz, A. G., Hossain, I. and Perera, B. J. C.: Effect of climate change and variability on extreme rainfall intensity–frequency–duration relationships: a case study of Melbourne, Hydrol. Earth Syst. Sci., 18(10), 4065–4076, 2014a.

Yilmaz, A. G., Hossain, I. and Perera, B. J. C.: Effect of climate change and variability on extreme rainfall intensity–frequency–duration relationships: a case study of Melbourne, Hydrol. Earth Syst. Sci., 18(10), 4065–4076, 2014b.

Yue, S., Pilon, P., Phinney, B. and Cavadias, G.: The influence of autocorrelation on the ability to detect trend in hydrological series, Hydrol. Process., 16(9), 1807–1829, 2002.

Yue, S., Pilon, P. and Phinney, B. O. B.: Canadian streamflow trend detection: impacts of serial and cross-correlation, Hydrol. Sci. J., 48(1), 51–63, 2003.

Figure Captions

Figure 1. (**a**) Selected urbanized sites in Southern Ontario. The Southern Ontario (41° - 44°N, 84° - 76°W) is the southernmost region of Canada and is situated on a southwest-northeast transect, bounded by lakes Huron, Erie, and Ontario. The nine locations on the map are (*from southwest to northeast corner*): Windsor Airport, London International Airport, Stratford Wastewater Treatment Plant (WWTP), Fergus Shand Dam, Hamilton Airport, Toronto International Airport, Oshawa Water Pollution Control Plant (WPCP), Trenton Airport, and Kingston Pumping Station. Topography map indicates maximum slope of 670 m above mean sea level. (**b**) The population map shows six the sites: Windsor Airport, London International Airport, Hamilton Airport, Toronto International Airport, Oshawa WPCP, and Kingston P. Station to be located either in or the vicinity of densely populated

urbanized area. The remaining three sites are located in the moderately populated area. The rainfall records in all locations vary between the minimum of 46 and maximum of 66 years.

Figure 2. Schematics of the process flow (Blue - input step, orange - process step, and green – decision steps). All three tests – Mann-Kendall, Pettitt's and Mann-Whitney, check for shifts in the mean. While Mann-Kendall tests for monotonic trends, the other two tests, Pettitt's and Mann-Whitney check for change point or regime shift in the time series.

Figure 3. Spatial distribution of trends, change points and nonstationarities in hourly and sub-hourly rainfall extremes in nine urbanized locations, Southern Ontario (**a – g**). The population estimates in and the vicinity of urbanized sites varies between 5 Million and 19,000 with highest in Toronto Metropolitan Area and the lowest in Fergus Shand dam area. The up and down triangles in white indicate (statistically insignificant) up and downward shifts; the up and down triangles in cyan and orange indicate shifts with change points only; the up and down triangles in the dark blue and red show presence of (statistically significant) trends including change point(s). Sites with significant nonstationarity are marked with an 'x' sign. The trends in short-duration AMP extremes are detected using nonparametric Mann-Kendall trend test with correction for ties. The change point in precipitation extremes are identified either as a shift in the mean or the variance in AMP time series. We apply nonparametric Pettit (Pettitt, 1979) and Mann-Whitney (Ross et al., 2011) tests to identify change point in the mean and Mood's test (Ross et al., 2011) to detect change point in the variance respectively (See section 2 for details). All tests are performed at 5 and 10% significance levels.

Figure 4. Uncertainty in DSI for 100-year return periods for stationary (blue) versus nonstationary (red) models with durations ranging between 15- min and 24-hr for nine sites (a - i): Toronto International Airport, Hamilton Airport, Oshawa WPCP, Windsor Airport, Kingston P. station, London International Airport, Trenton Airport, Stratford WWTP, and Fergus Shand Dam. The boxplots indicate the uncertainty in estimated DSI from Bayesian inference, whereas the DSI obtained from maximum likelihood approach is shown as a black dot in the stationary simulation.

Figure 5. Z-statistics of at-site *T*-year event estimates for T = 2-year to T = 100-year return periods (**a – d**) with durations between 15-min and 24-hr in nine urbanized locations, Southern Ontario. The Z-statistic represents statistical significance of differences in DSI obtained from the best selected nonstationary versus the stationary model. The Z-statistics is statistically significant when $|Z| > 1.96$ and $|Z| > 1.64$ at 5 and 10% significance levels. The shades in blue and red denote decrease and increase in Z-statistics with the strength of shading represents the

magnitude of the test statistics. The cyan shading represents the site with significant autocorrelation, which we exclude from the analysis.

Figure 6. Central tendency (median, **b**) and the bounds (5 and 95% range, **a** and **c**) of the updated nonstationary versus EC-generated *T*-year event estimates for DSI at selected return periods with durations between 15-min and 24-hr. The minimum, median and maximum *T*-year event estimates of nonstationary models are obtained from time-variant GEV parameter(s) by computing the 5^{th}, 50^{th} and 95^{th} percentiles of DE-MC sampled parameters. The DSI and associated 95% confidence limits of EC-generated IDF is obtained from the national archive of Engineering Climate Datasets maintained by the Environment Canada (http://climate.weather.gc.ca/). The strength of shading represents the magnitude of the ratio between updated versus EC-generated DSI with a deeper shade indicates an increase in the ratio. The cyan shading indicates the site with significant autocorrelation.

Figure 7. Estimated nonstationary versus EC-generated IDFs for return periods T = 2, 5, 10, 25, 50 and 100-year for the selected urbanized locations in Southern Ontario, Canada. The nonstationary IDFs are shown using solid lines, while EC-generated IDFs are shown using dotted lines.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-325
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 12 June 2017

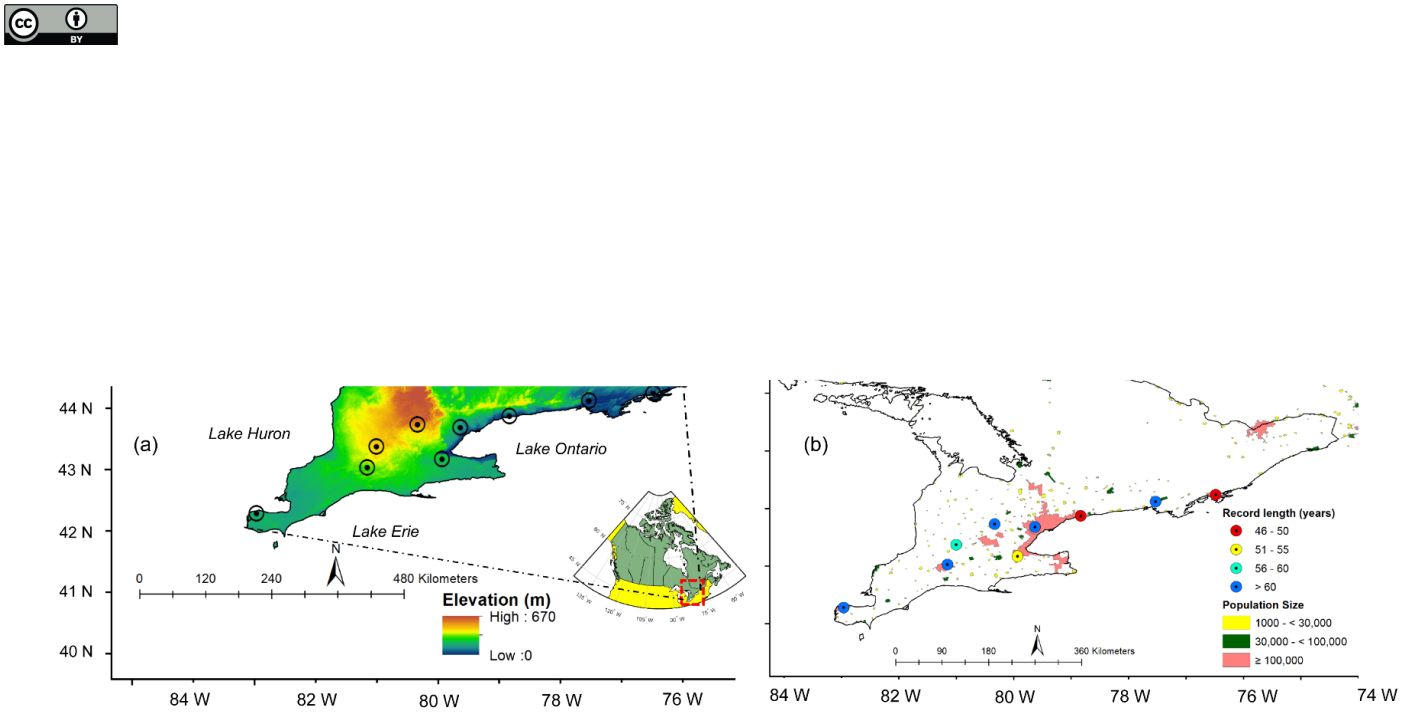

Figure 1. (**a**) Selected urbanized sites in Southern Ontario. The Southern Ontario (41° - 44°N, 84° - 76°W) is the southernmost region of Canada and is situated on a southwest-northeast transect, bounded by lakes Huron, Erie, and Ontario. The nine locations on the map are (*from southwest to northeast corner*): Windsor Airport, London International Airport, Stratford Wastewater Treatment Plant (WWTP), Fergus Shand Dam, Hamilton Airport, Toronto International Airport, Oshawa Water Pollution Control Plant (WPCP), Trenton Airport, and Kingston Pumping Station. Topography map indicates maximum slope of 670 m above mean sea level. (**b**) The population map shows six the sites: Windsor Airport, London International Airport, Hamilton Airport, Toronto International Airport, Oshawa WPCP, and Kingston P. Station to be located either in or the vicinity of densely populated urbanized area. The remaining three sites are located in the moderately populated area. The rainfall records in all locations vary between the minimum of 46 and maximum of 66 years.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-325
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 12 June 2017

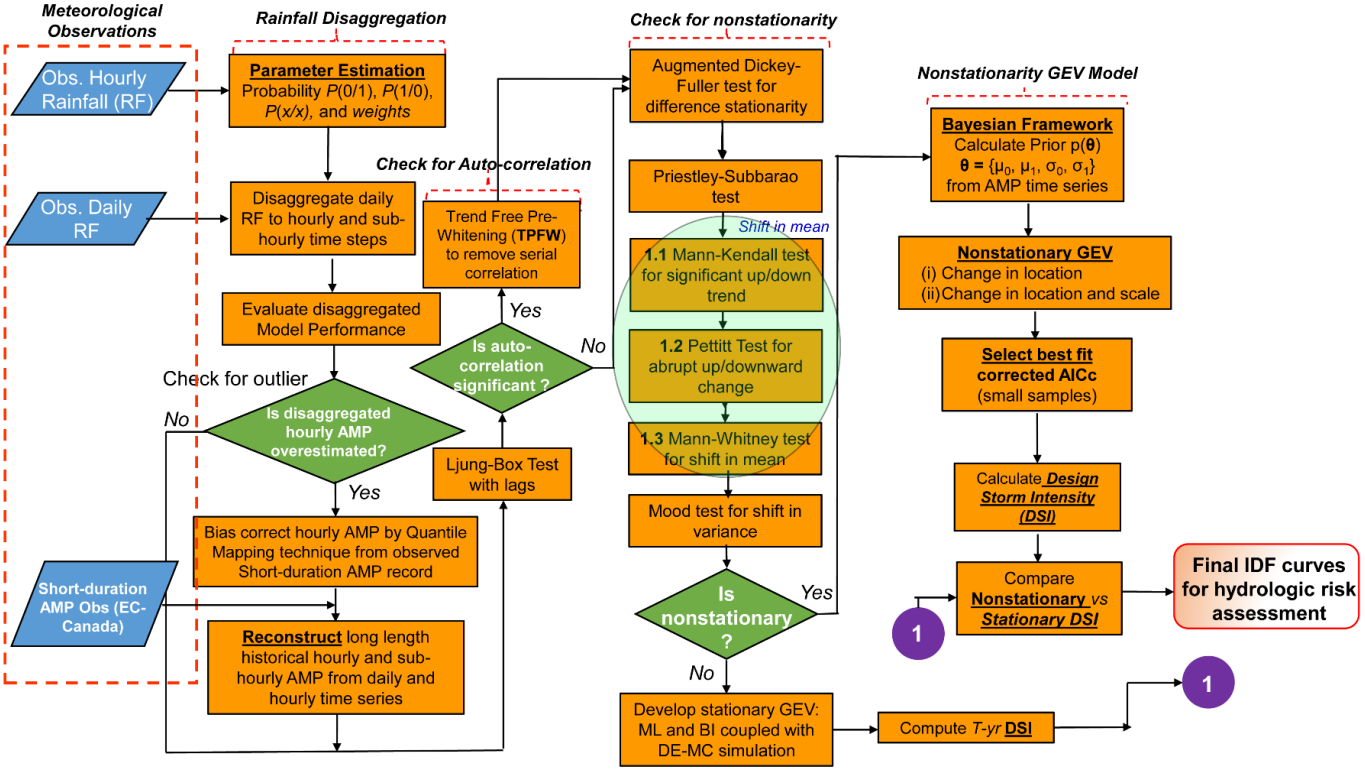

Figure 2. Schematics of the process flow (Blue - input step, orange - process step, and green – decision steps). All three tests – Mann-Kendall, Pettitt's and Mann-Whitney, check for shifts in the mean. While Mann-Kendall tests for monotonic trends, the other two tests, Pettitt's and Mann-Whitney check for change point or regime shift in the time series.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-325
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 12 June 2017

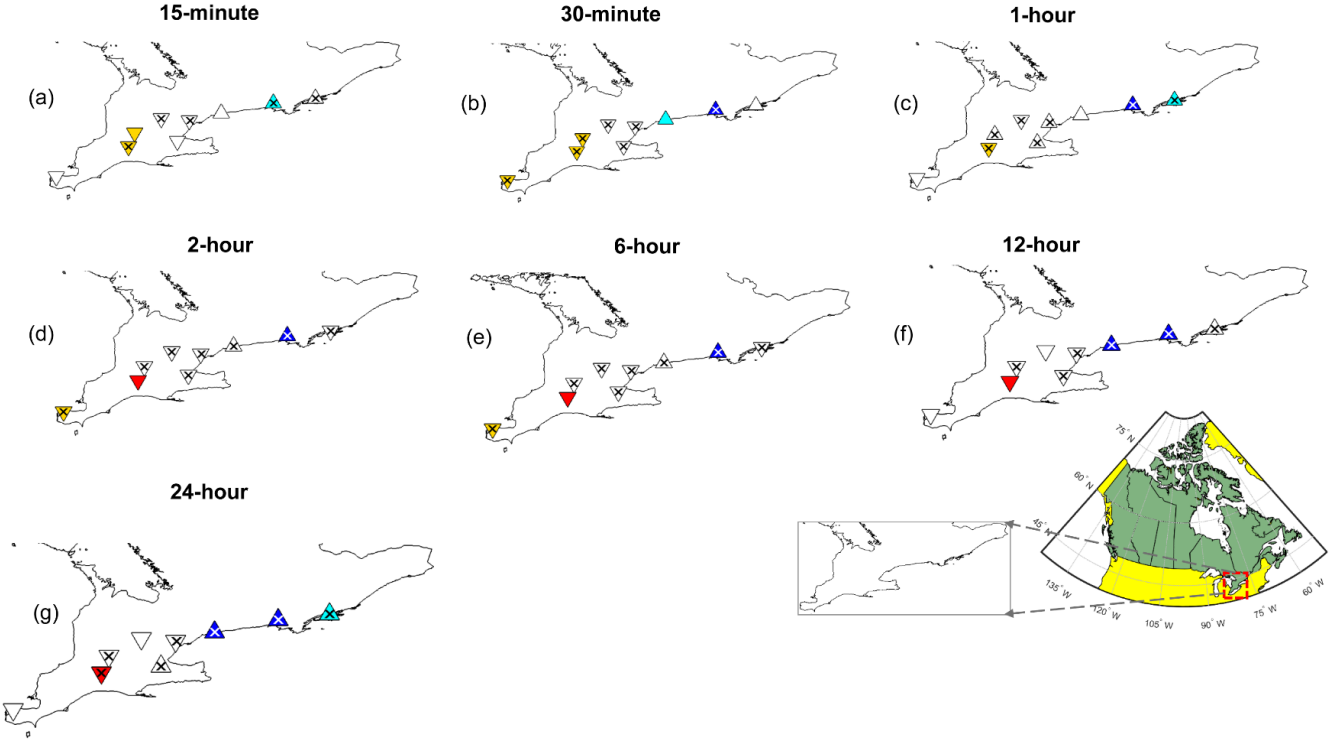

Figure 3. Spatial distribution of trends, change points and nonstationarities in hourly and sub-hourly rainfall extremes in nine urbanized locations, Southern Ontario (**a – g**). The population estimates in and the vicinity of urbanized sites varies between 5 Million and 19,000 with highest in Toronto Metropolitan Area and the lowest in Fergus Shand dam area. The up and down triangles in white indicate (statistically insignificant) up and downward shifts; the up and down triangles in cyan and orange indicate shifts with change points only; the up and down triangles in the dark blue and red show presence of (statistically significant) trends including change point(s). Sites with significant nonstationarity are marked with an 'x' sign. The trends in short-duration AMP extremes are detected using nonparametric Mann-Kendall trend test with correction for ties. The change point in precipitation extremes are identified either as a shift in the mean or the variance in AMP time series. We apply nonparametric Pettit (Pettitt, 1979) and Mann-Whitney (Ross et al., 2011) tests to identify change point in the mean and Mood's test (Ross et al., 2011) to detect change point in the variance respectively (See section 2 for details). All tests are performed at 5 and 10% significance levels.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-325
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 12 June 2017

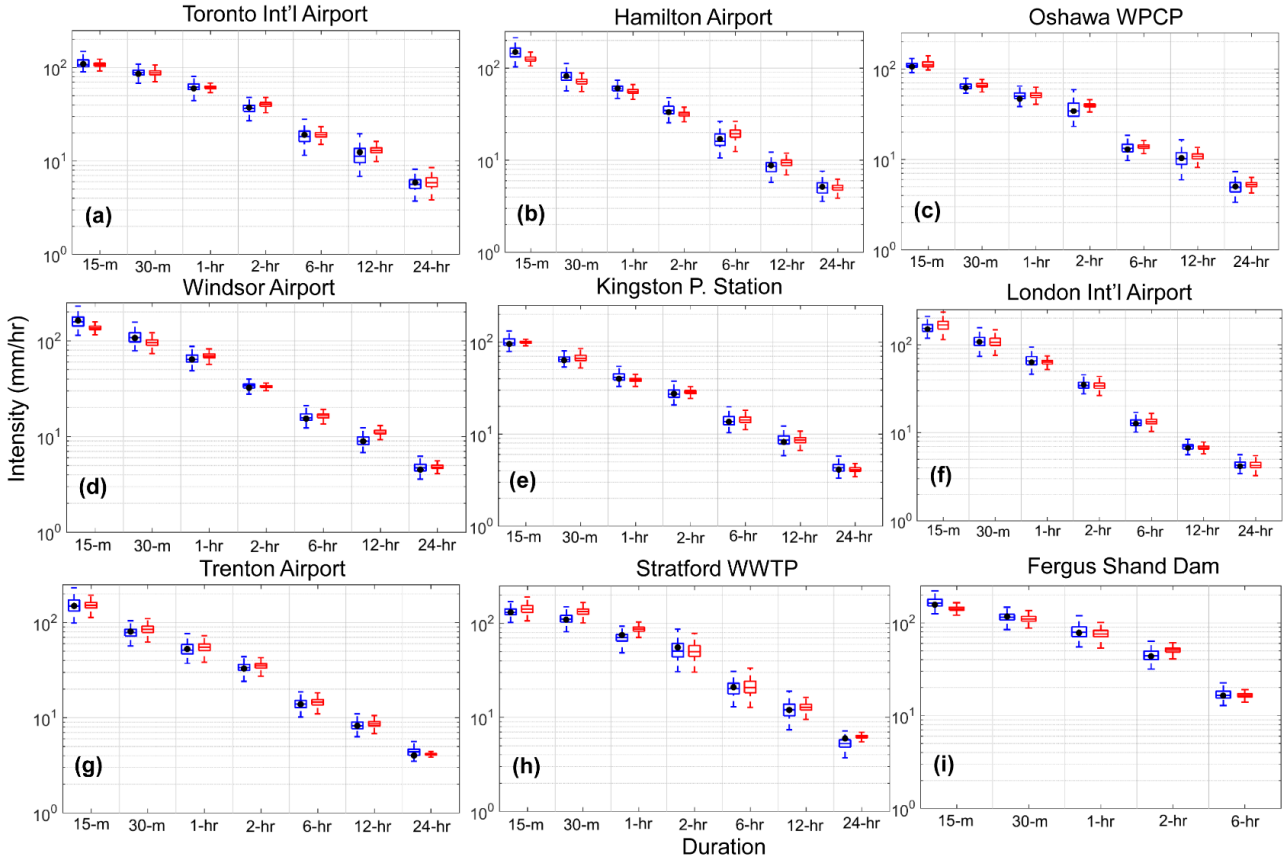

Figure 4. Uncertainty in DSI for 100-year return periods for stationary (blue) versus nonstationary (red) models with durations ranging between 15- min and 24-hr for nine sites (**a - i**): Toronto International Airport, Hamilton Airport, Oshawa WPCP, Windsor Airport, Kingston P. station, London International Airport, Trenton Airport, Stratford WWTP, and Fergus Shand Dam. The boxplots indicate the uncertainty in estimated DSI from Bayesian inference, whereas the DSI obtained from maximum likelihood approach is shown as a black dot in the stationary simulation.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-325
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 12 June 2017

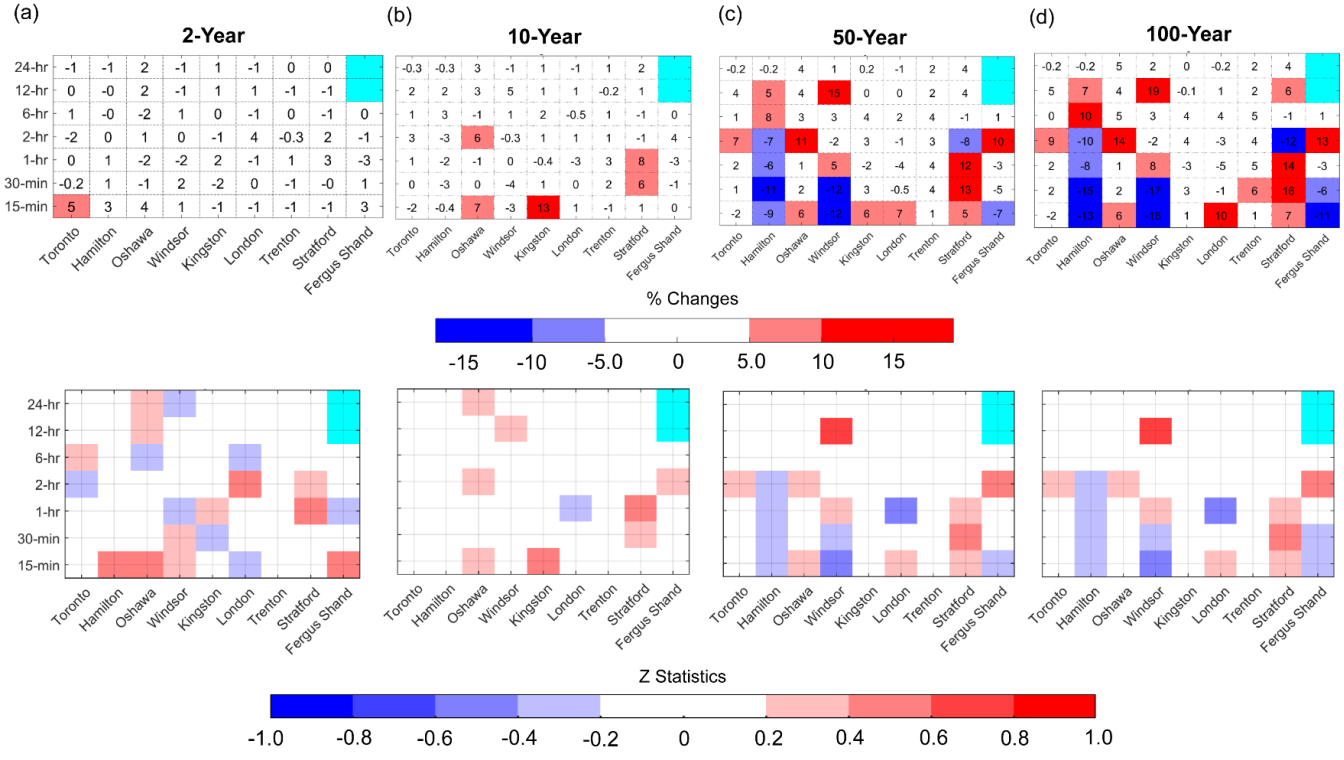

Figure 5. Z-statistics of at-site *T*-year event estimates for *T* = 2-year to *T* = 100-year return periods (**a – d**) with durations between 15-min and 24-hr in nine urbanized locations, Southern Ontario. The Z-statistic represents statistical significance of differences in DSI obtained from the best selected nonstationary versus the stationary model. The Z-statistics is statistically significant when |Z| > 1.96 and |Z| > 1.64 at 5 and 10% significance levels. The shades in blue and red denote decrease and increase in Z-statistics with the strength of shading represents the magnitude of the test statistics. The cyan shading represents the site with significant autocorrelation, which we exclude from the analysis.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-325
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 12 June 2017

Hydrology and Earth System Sciences
Open Access
Discussions

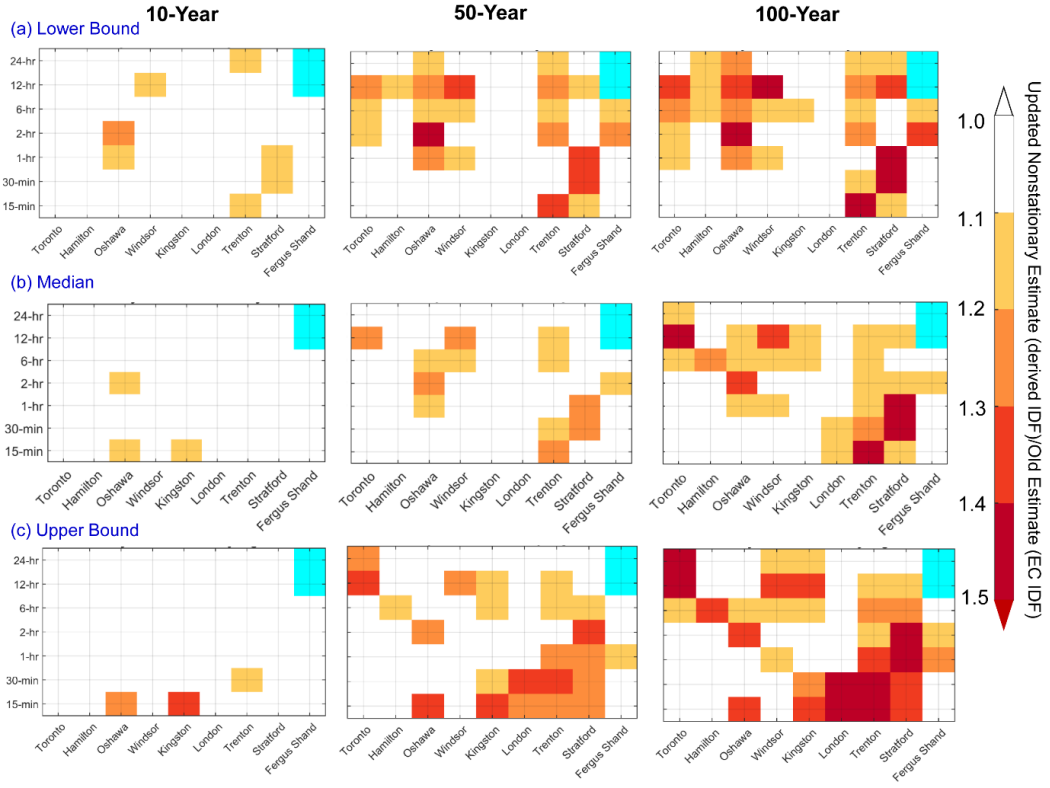

Figure 6. Central tendency (median, **b**) and the bounds (5 and 95% range, **a** and **c**) of the updated nonstationary versus EC-generated *T*-year event estimates for DSI at selected return periods with durations between 15-min and 24-hr. The minimum, median and maximum *T*-year event estimates of nonstationary models are obtained from time-variant GEV parameter(s) by computing the 5th, 50th and 95th percentiles of DE-MC sampled parameters. The DSI and associated 95% confidence limits of EC-generated IDF is obtained from the national archive of Engineering Climate Datasets maintained by the Environment Canada (http://climate.weather.gc.ca/). The strength of shading represents the magnitude of the ratio between updated versus EC-generated DSI with a deeper shade indicates an increase in the ratio. The cyan shading indicates the site with significant autocorrelation.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-325
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 12 June 2017

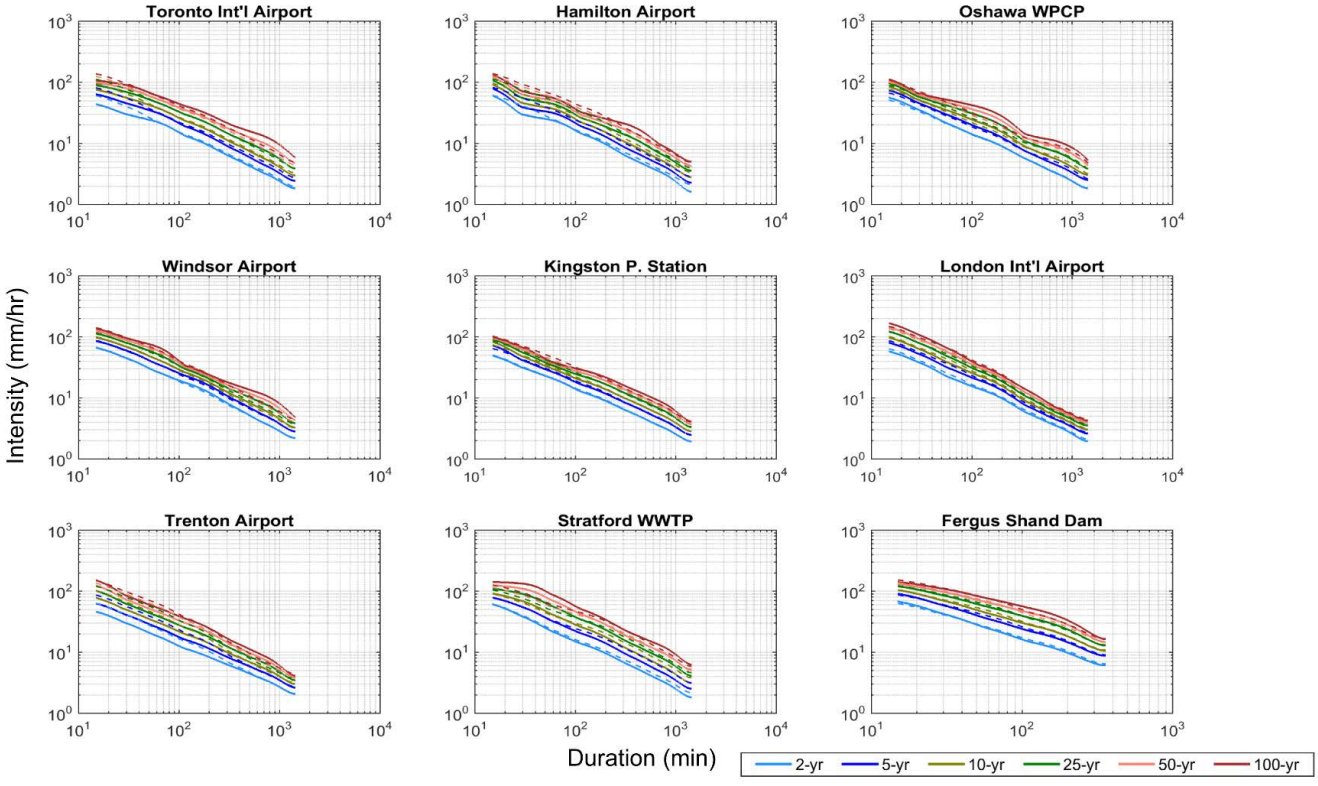

Figure 7. Estimated nonstationary versus EC-generated IDFs for return periods T = 2, 5, 10, 25, 50 and 100-year for the selected urbanized locations in Southern Ontario, Canada. The nonstationary IDFs are shown using solid lines, while EC-generated IDFs are shown using dotted lines.