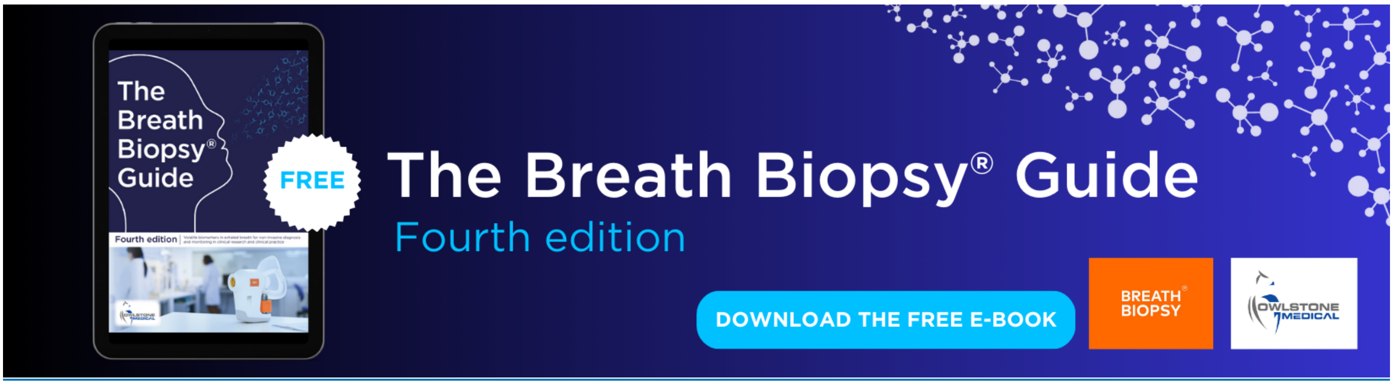
ENVIRONMENTAL RESEARCH
LETTERS

LETTER • **OPEN ACCESS**

Structural change to the persistence of the urban heat island

To cite this article: Debbie J Dupuis and Luca Trapin 2020 *Environ. Res. Lett.* **15** 104076



Environmental Research Letters

OPEN ACCESS

RECEIVED
29 May 2020

REVISED
1 August 2020

ACCEPTED FOR PUBLICATION
21 August 2020

PUBLISHED
6 October 2020

Structural change to the persistence of the urban heat island

Debbie J Dupuis[1] and **Luca Trapin**[2]

[1] Department of Decision Sciences, HEC Montréal, 3000, chemin de la Côte-Sainte-Catherine, Montréal (Québec) H3T 2A7, Canada
[2] Department of Statistics, University of Bologna, Via delle Belle Arti, 41, 40126 Bologna, Italy

E-mail: debbie.dupuis@hec.ca

Keywords: urbanization, extreme temperature, clustering

Supplementary material for this article is available online

Abstract

The term urban heat island (UHI) is used to describe the effect of urban temperatures rising several degrees above concurrent temperatures in surrounding suburban or rural areas. This is typically assessed through records of daily extreme temperatures. However, on a hot day the temperature can exceed an extreme threshold for several consecutive hours, forming a cluster of extremes. We use the statistical theory of extreme values combined with a model that allows structural breaks to show that there has been a significant upward shift in the length of clusters in New York City. No such shift is found at a Connecticut location where the usual UHI assessment indicates that the two sites are comparable. Our study is the first to highlight this danger of the UHI. Prolonged exposure to extreme temperatures has deleterious effects on both health and the environment.

In 2018, 55% of the world's population resided in urban areas and North America was the most urbanized region with 82% of its population living in towns and cities [1]. While the urbanization of the United Stated has progressed throughout its entire history, concerns for the effects of urban development on inner-city temperatures are more recent [2]. The term 'urban heat island' (UHI) is used to describe the effect of urban temperatures rising several degrees above concurrent temperatures in surrounding suburban or rural areas. Temperature differentials are mostly related to the size, building density and materials, and land-use distribution of the city, as well as the climate and weather to which it is subjected [3]. In the recent literature, there are numerous quantifications of UHI in the United States' largest metropolitan area, New York City [4–7]. There are two distinct phenomena: the urban canopy layer UHI, and the urban boundary layer UHI [3]. We study the former, a principally nocturnal phenomenon, and simply refer to it as the UHI. Two characteristics of UHI for New York City are well established: (i) there is considerable intra-city variability, and (ii) while nocturnal UHI can be quite strong, there are only small differences in urban and non-urban temperatures at mid-day peaks during the summer months [4]. For example, in the summer months, maximum daily temperatures at LaGuardia International Airport (LGA) in Queens, NY, are often more comparable to those in relatively non-urban Windsor Locks, CT (200 km northeast) than to those at JFK International Airport (16 km southeast). Differences in daily maximum temperatures in July (figure 1(A)) and August (figure S1) for the period 1973-2018 show no UHI for LGA w.r.t. Bradley International Airport (BDL) in Windsor Locks, and show BDL warmer than JFK. The differences between BDL and JFK, and the lack of such between BDL and LGA, show the complexity of the UHI phenomenon as the level of urbanization at LGA and JFK is much higher than at BDL. There has been an almost 100% increase in Settlements area (defined as developed land, including transportation infrastructure and human settlements of any size not qualified under forest, cropland, grassland or wetlands) from 1986 to 2009 in New York City [8], and population densities for 2016 are 21,284 inhabitants per square mile in Queens and 1,389 inhabitants per square mile in Windsor Locks, respectively [9]. Can this disparity really have no effect on extreme temperatures, or has the situation so far been incompletely assessed? We look beyond maximum temperature levels and examine the extreme temperature temporal dynamics at LGA, JFK, and BDL. Hourly temperatures during heat waves in 1977, 2006, and 2013, respectively, show changing dynamics at LGA (figure S2). While the maximum temperature reached can be quite similar during the three heat waves, the paths of hourly temperatures are quite different, with an increasing number of hours spent over 30.6 °C, the 90th quantile of hourly temperatures for LGA in July

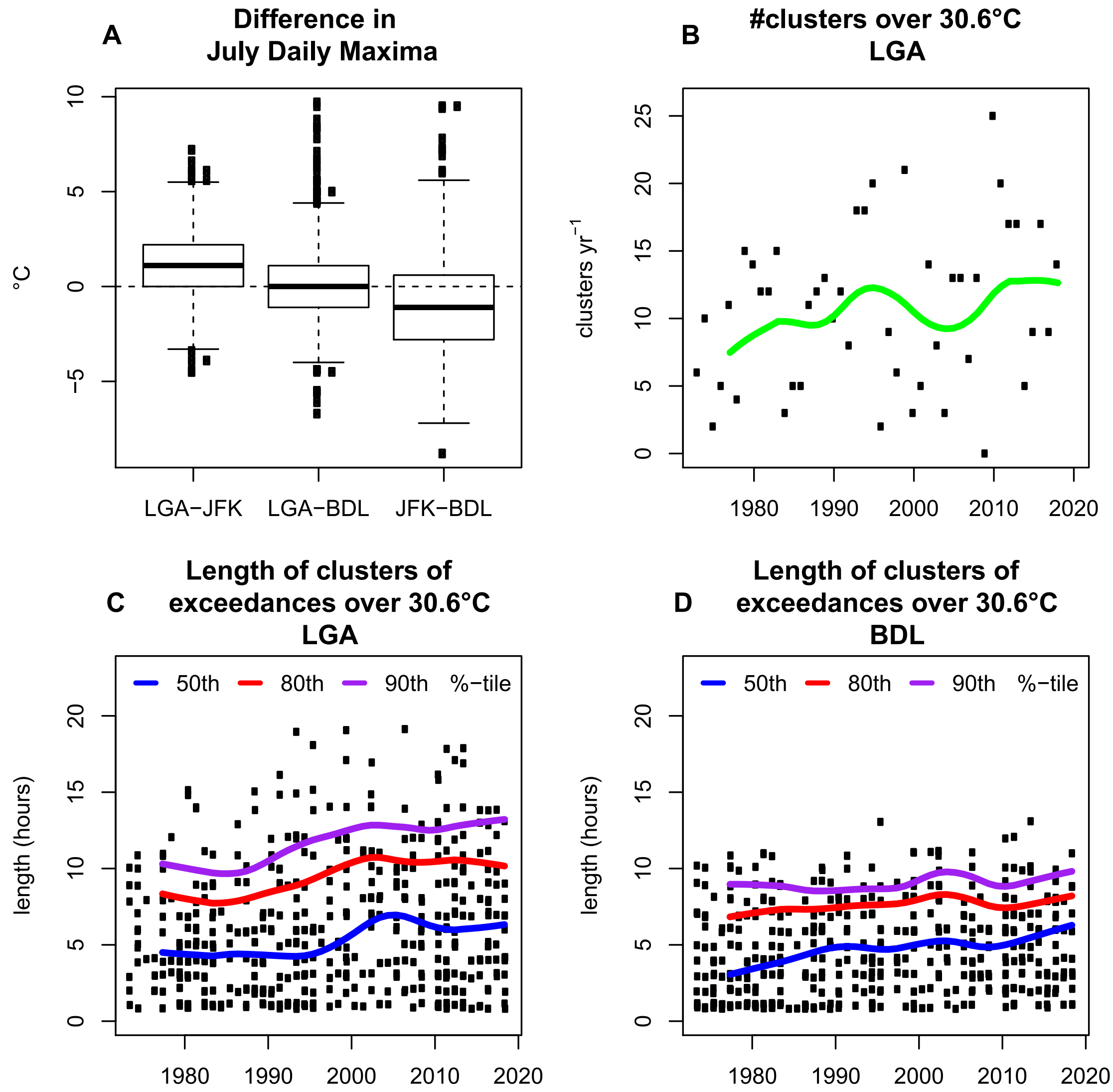


Figure 1. Hourly temperature characteristics. (**A**) Pairwise difference in daily maximum temperatures in July for the period 1973-2018. (**B**) Number of clusters per year at LGA. Line is smoothed annual mean based on 5-yr moving window. (**C**) Length (in hours) of clusters of temperatures exceeding 30.6°C at LGA in the month of July. Lines are smoothed annual percentiles. (**D**) Length (in hours) of clusters of temperatures exceeding 30.6°C at BDL in the month of July. Lines are smoothed annual percentiles.

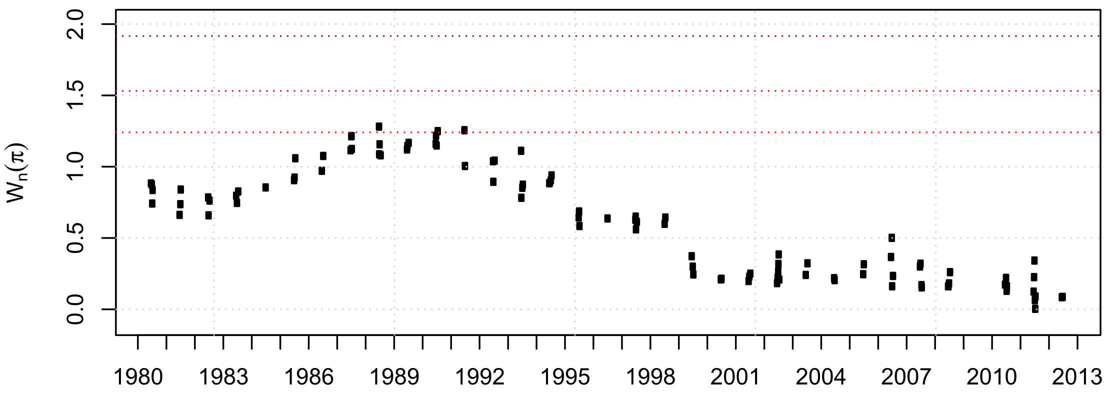


Figure 2. Test statistics for July at LGA. Value of the test statistic $W_n(\pi)$ for each $\pi \in [0.15, 0.85]$ for length of clusters of exceedances over 30.6° C for the month of July at LGA. D-GPD threshold is $c = 2$. $\pi \in [0.15, 0.85]$ means that dates from 1980 to 2012 are tested. Dotted lines show the bootstrap ($B = 500$) critical values for $\alpha = 0.01$, 0.05, and 0.10, respectively. The statistic W_n attains its maximum value on 1988-07-10, showing the presence of a structural break in the length of the clusters on that date.

during the 1973-2018 period, at LGA in more recent years. Has extreme temperature persistence changed? Consecutive exceedances of the time series of temperatures over 30.6°C form clusters, and we are interested in the length of such clusters over such high temperature thresholds. The distribution of the length of clusters of hourly temperature exceedances over 30.6°C in July at LGA seems to have shifted over 1973-2018 (figure 1(C)). In contrast, the analogous distribution at BDL shows its median smoothly doubling, but very little change in the upper-tail (figure 1(D)). The number of clusters at LGA does not appear to have increased over time (figure 1(B)). The possible shift in the length of clusters could lead to serious public health and sustainability issues, and we investigate further.

The length of the clusters is modeled using a discrete Generalized Pareto distribution (D-GPD). Testing the parameter stability verifies the presence of a structural break (SB). The test statistic for a SB in the scale parameter σ of the D-GPD fitted to clusters longer than $c = 2$ hours exceeds the bootstrap ($B = 500$) critical value ($\alpha = 0.10$) on three occasions and reaches its maximum value of 1.28 on 1988/07/10 (figure 2). The p-value for the bootstrap test is 0.096. The partial-sample estimates of the D-GPD structural break model parameters on 1988/07/10 are $\widehat{\xi} = -0.46$, $\widehat{\sigma}_1 = 7.9$, and $\widehat{\sigma}_2 = 10$, indicating longer clusters of exceedances since 1988. More precisely, the estimated mean length of clusters over 30.6°C has increased by 20%, from 6.9 hours before 1988/07/10 to 8.3 hours after 1988/07/10. Similarly, the estimated probability of a cluster longer than 12 hours has doubled from 10.5% to 21.5% before and after the SB, respectively. The D-GPD model is a good fit before and after the SB (figure S3). Testing for clusters longer than $c = 3$ and $c = 4$ hours yields similar results. We also test the length of clusters over the more extreme 92nd, 94th, and 95th quantiles of hourly temperatures for July at LGA during the 1973-2018 period as a robustness check. All the tests coherently identify the 1988-1991 period as the estimated time of structural change. Values of the test statistic in figure 2 are typical of those for all our tests for July at LGA except that the 1988-1991 peak period does not exceed the bootstrap critical values in all tests (table 1). A simulation study in the Supplementary Material (stacks.iop.org/ERL/15/104076/mmedia) shows that the test has good power when $n = 150$, but detecting breaks in smaller samples is more challenging.

Similar analyses are performed for August at LGA (table 1) and show the good fit of the D-GPD SB in scale model before and after the SB (figure S4). The estimated dates of the SB are 1991/08/01 and 2001/08/01, based on tests for 30.6°C and 30.77°C, respectively, indicating that the SB in the length of clusters was later for August than for July at LGA. There are fewer significant test statistics for August than July at LGA. Temperatures are also not as extreme in August as they are in July at LGA. Warmer temperatures could have an amplified effect on tail dynamics just like they augment UHI intensity during heat waves in New York City [10]. The relative increases from pre- to post-structural break are, however, more important for August than for July, although clusters are shorter. For $u = 30.6$°C and $c = 1$, the partial sample estimates are $\widehat{\xi} = -0.30$, $\widehat{\sigma}_1 = 5.05$, and $\widehat{\sigma}_2 = 6.91$, indicating longer clusters of exceedances since 1991. More precisely, the estimated mean length of clusters over 30.6°C has increased by 32%, from 4.4 hours before 1991/08/01 to 5.8 hours after 1991/08/01. Similarly, the estimated probability of a cluster longer than 12 hours has increased fivefold from 1.6% to 8.7% before and after the SB, respectively.

A model with SB in scale only is inadequate for July at JFK. Both the scale and the shape are changing through time, and a model with common shape gives a very poor fit at several values of potential structural break π. The model with a SB in scale and shape provides a good fit to the length of clusters of the exceedances over 28.9°C, the 90th quantile of hourly temperatures in July at JFK, for all $c = 1$ to 4 and a significant break is found in each case (table 1). The test statistic reaches its maximum on 1995/07/14 ($c = 1, 2, 3$) and 1993/07/03 ($c = 4$), and exceeds the bootstrap critical value ($\alpha = 0.10$) on many dates over the 1988 to 2001 period. With $c = 1$, the partial-sample estimates on 1995/07/14 are $\widehat{\sigma}_1 = 7.9$, $\widehat{\xi}_1 = -0.40$, $\widehat{\sigma}_2 = 12.4$, and $\widehat{\xi}_2 = -0.61$, indicating longer clusters after 1995. More precisely, the estimated mean length of clusters over 28.9°C has increased over 30%, from 6.2 hours before 1995/07/14 to 8.2 hours after 1995/07/14. Similarly, the estimated probability of a cluster longer than 12 hours has more than doubled from 10% to 23.2% before and after the break, respectively. The latter increases are similar to those suffered at LGA in July over its 90th quantile of 30.6°C. At JFK, 30.6°C represents the 95th quantile in July. Estimated mean length of clusters over 30.6°C has increased over 40%, from 4.9 hours before 1998/07/20 to 7 hours after 1998/07/20. Similarly, the estimated probability of a cluster longer than 12 hours has increased more than twofold from 5.2% to 13.6% before and after the break, respectively. Relative increases for length of clusters over 30.6°C are thus slightly greater at JFK than LGA, but lengths remain shorter. Similar analyses are also performed for August at JFK. The D-GPD model with structural break in scale is a good fit in all partial samples, but estimated scale parameters differ very little. We find no evidence of a SB (table 1). Finding a break for July at JFK, but not for August at JFK which is somewhat cooler, is again consistent with warmer temperatures amplifying the UHI intensity during heat waves in New York City, and also consistent with the July vs August results at LGA.

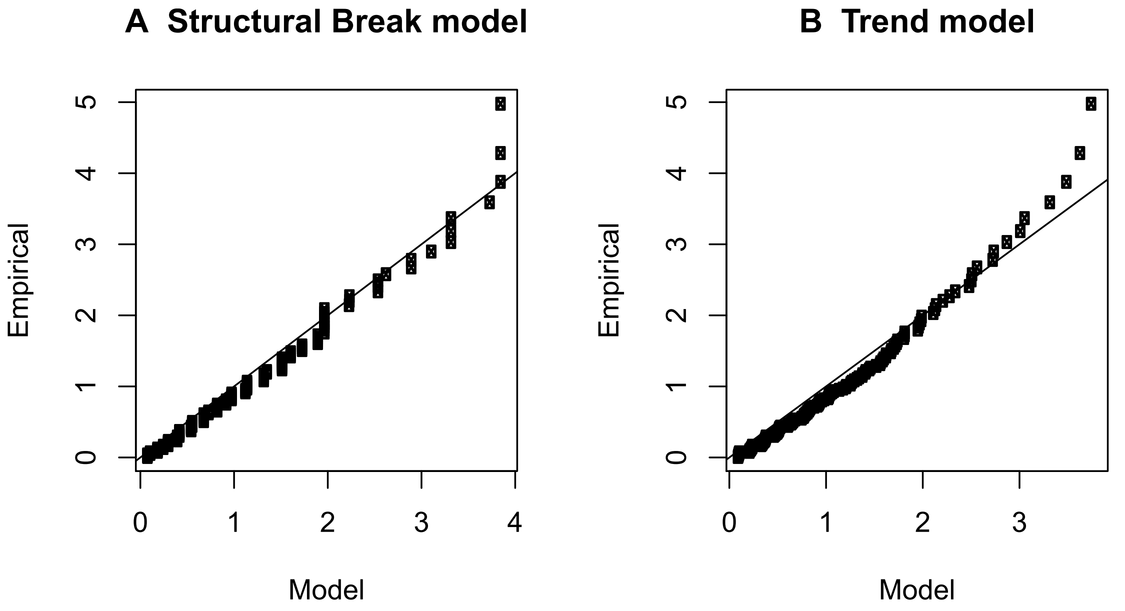


Figure 3. Structural break vs trend. (**A**) Quantile-quantile plot for the SB model fitted to lengths of clusters of exceedances over 30.6° C for the month of July at LGA using GMM. (**B**) Quantile-quantile plot for the trend model fitted to lengths of clusters of exceedances over 30.6° C for the month of July at LGA using GMM. The value of the test statistic T_n is 5.15, heavily favoring the SB model. Fit of the points in **A** is also better than in **B**. Q-Q plots are on exponential scale.

Finally, tests are carried out for July and August at BDL. The D-GPD model with structural break in scale only is a good fit in all partial samples. Estimates of σ_1 and σ_2 are always quite close, in some cases we even find $\widehat{\sigma}_1 > \widehat{\sigma}_2$, and the test statistic for a SB is always small. There is no evidence of a significant change in the length of clusters over high temperatures at BDL over the 1973-2018 period (table 1). So while temperatures at BDL can peak at values similar to those at LGA and JFK (figure 1(A) and figure S1), there is no evidence that BDL has suffered a SB in the distribution of the length of clusters of exceedances over high temperatures over the last 45 years. Long exposure times to extreme temperatures were, and remain, unlikely at BDL. The estimated probability of a cluster longer than 12 hours is $\approx$ 1% for over 30.5° C in July at BDL and less than 10^{-10}% for over 30.5° C in August at BDL.

Tests for comparing the suitability of a D-GPD model with a linear trend in the scale parameter (model $\mathcal{M}_1$) to a structural break in the scale parameter (model $\mathcal{M}_2$) reveal strong evidence that $\mathcal{M}_2$ is better than $\mathcal{M}_1$ in almost all cases and never favors $\mathcal{M}_1$ (table S4). The heavily favored model $\mathcal{M}_2$ fits the data very well (figure 3).

Many researchers have looked at UHI in New York City, but all studies focus on marginal behavior and are empirical in nature, e.g. comparing all, maximum, or average temperatures at different locations over different periods with no formal statistical testing [4–7]. We look beyond marginal maximum temperature levels, and study the extreme hourly temperature temporal dynamics at LGA, JFK, and BDL, finding a significant upward *shift* in the length of clusters of exceedances in July and August at LGA, in July only at JFK, and no similar shift at BDL over the last 45 years.

Finding different structural breaks at LGA and JFK is consistent with related results in the literature. For example, when analyzing the spatial variability in the thermal structure of the boundary layer over New York City in 2016, it is found that while the coastal JFK site is influenced by the sea breeze during the summer afternoon periods, the sea breeze is unable to penetrate the inland LGA site which is influenced by the northerly land breezes [12]. Our different structural breaks are also consistent with the complex thermal structure and high intra-city variability found during heat waves [5]. An estimated structural break in 1988 at the inland LGA site suggests an immediate impact of the 100% increase in Settlements area from 1986 to 2009 in New York City [8]. Sudden changes in other temperature characteristics at LGA have been detected during our estimated time period, e.g. tails of daily maximum summer temperatures at LGA were found to have a structural break in the early nineties [11].

While we find no increases to the persistence of extreme temperatures at BDL in the summer, we find that the average time with temperatures above 30.6° C has increased from 6.9 hours (to 1988/07) to 8.3 hours (after 1988/07) at LGA in July; increased from 4.4 hours (to 1991/08) to 5.8 hours (after 1991/08) at LGA in August; and increased from 4.9 hours (to 1998/07) to 7.0 hours (after 1998/07) at JFK in July. It is common to investigate the effects of high temperatures on mortality [13] and morbidity [14] using exposure measured by mean temperature calculated as the average of the highest and lowest hourly

Table 1. Bootstrap tests for structural break. D-GPD fitted to length of clusters of exceedances over u set to the 90th, 92nd, 94th, and 95th quantile of hourly temperatures for the given month and location during the 1973-2018 period. Sample size n and p-value (p-val) for tests at D-GPD threshold c are reported. The August 2006 North American heat wave [20] yields extraordinary observations at LGA, e.g. a 60-hour cluster over 29.4°C, and these observations are removed so that any inference does not rest solely on them.

	u	30.6°C		31.1°C		31.7°C		32.13°C	
	c	n	p-val	n	p-val	n	p-val	n	p-val
	1	158	0.138	142	0.106	131	0.114	129	0.262
LGA	2	144	0.096	132	0.066	122	0.098	117	0.136
July	3	134	0.062	119	0.132	107	0.078	105	0.118
	4	122	0.042	108	0.048	94	0.220	89	0.302
	u	29.4°C		30°C		30.6°C		30.77°C	
	c	n	p-val	n	p-val	n	p-val	n	p-val
	1	163	0.204	149	0.106	135	0.072	133	0.074
LGA	2	146	0.314	129	0.128	109	0.290	107	0.266
August	3	140	0.206	110	0.178	97	0.228	95	0.178
	4	120	0.346	98	0.194	88	0.230	84	0.106
	u	28.9°C		29.4°C		30°C		30.6°C	
	c	n	p-val	n	p-val	n	p-val	n	p-val
	1	175	0.066	157	0.032	136	0.032	131	0.032
JFK	2	154	0.016	138	0.064	124	0.112	114	0.118
July	3	138	0.040	122	0.126	115	0.192	93	0.234
	4	127	0.070	112	0.168	97	0.190	83	0.218
	u	28.3°C		28.8°C		29.33°C		29.4°C	
	c	n	p-val	n	p-val	n	p-val	n	p-val
	1	175	0.590	169	0.728	165	0.626	147	0.964
JFK	2	157	0.556	147	0.746	140	0.746	130	0.962
August	3	141	0.744	132	0.846	126	0.910	117	0.970
	u	29.8°C		30.5°C		31.1°C		31.6°C	
	c	n	p-val	n	p-val	n	p-val	n	p-val
	1	179	0.886	176	0.746	146	0.486	133	0.470
BDL	2	167	0.788	158	0.940	123	0.140	112	0.674
July	3	150	0.876	143	0.856	107	0.270	104	0.578
	u	28.7°C		29.4°C		30.0°C		30.5°C	
	c	n	p-val	n	p-val	n	p-val	n	p-val
	1	179	0.744	167	0.686	139	0.776	138	0.480
BDL	2	168	0.766	153	0.692	124	0.794	122	0.842
August	3	154	0.874	130	0.772	113	0.602	107	0.890

measurements recorded within each day. Our results suggest that actual exposure may not be tied to such a mean temperature, may be location dependent and may have shifted significantly over time in a location dependent manner. Any sudden increase in the continuous exposure to very high temperatures negatively affects other public health issues like physical inactivity [15], worker safety [16], and mental health [17], with subsequent economic impacts [16]. Finally, air conditioning was responsible for 9 percent of large buildings' source energy use in New York City in 2015, and that number was expected to grow [18]. Given the demonstrated longer hourly exposure to very high temperatures, recent legislation to improve energy efficiency for certain buildings [19] could be even more beneficial in New York City than other cities with similar maximum daily temperature statistics. Increased energy efficiency will be essential as anthropogenic heat release due to larger cooling loads could otherwise further contribute to UHI.

Acknowledgment

The first author acknowledges the support of the Natural Sciences and Engineering Research Council of Canada RGPIN-2016-04114 and the HEC Montréal Foundation. The authors would like to thank two anonymous referees for comments that improved the presentation of the work.

Author contributions

DJD: conceptualization, software, formal analysis, visualization, writing. LT: methodology, software, writing.

Competing interests

Authors declare no competing interests.

Data and materials availability

The data that support the findings of this study were obtained without charge from the National Oceanic and Atmospheric Administration and are available upon reasonable request from the authors. R code is available from the authors.

ORCID iD

Debbie J Dupuis https://orcid.org/0000-0003-4277-4605

References

[1] United Nations 2019 *World Urbanization Prospects: The 2018 Revision* (ST/ESA/SER.A/420) (New York: United Nations) (https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf)
[2] Oke T 1973 *Atmospheric Environment (1967)* **7** 769
[3] Oke T 1982 *Quarterly J. R. Meteorol. Society* **108** 455
[4] Gaffin S *et al* 2008 *Theor. Appl. Climatol.* **94** 1
[5] Ramamurthy P, González G, Ortiz L, Arend M and Moshary F 2017 *Environ. Res. Lett.* **12** 054011
[6] Ortiz L *et al* 2018 *J. Appl. Meteorol. Climatol.* **57** 837
[7] González J *et al* 2019 *Ann. N.Y. Acad. Sci.* **1439** 30
[8] Bagan H and Yamagata Y 2014 *Environ. Res. Lett.* **9** 064015
[9] (https://www.opendatanetwork.com/entity/0500000US36081/Queens_County_NY/geographic.population.density?year=2016)
[10] Tewari M *et al* 2019 *Environ. Res. Lett.* **14** 034002
[11] Dupuis D 2012 *J. Am. Stat. Assoc.* **107** 24
[12] Melecio-Vázquez D, Ramamurthy P, Arend M and González-Cruz J 2018 *Bound.-Layer Meteorol.* **169** 151
[13] Gasparrini A and Armstrong B 2011 *Epidemiology* **22** 68
[14] Lin S *et al* 2009 *Epidemiology* **20** 738
[15] Heaney A, Carrión D, Burkart K, Lesk C and Jack D 2019 *Environmental Health Perspectives* **127**
[16] Kjellstrom T *et al* 2016 *Ann. Rev. Public Health* **37** 97
[17] Evans G 2019 *Ann. Rev. Psychol.* **70** 449

[18] The City of New York Urban Green Council 2017 *New York City's Energy and Water Use 2014 and 2015 Report* (http://www.nyc.gov/html/gbee/downloads/pdf/UGC-Benchmarking-Report-101617-FINAL.pdf)

[19] Council N Y C 2019 (https://council.nyc.gov/data/green/)

[20] Pérez-Peña R 2006 *The New York Times* www.nytimes.com/2006/08/08/nyregion/08mbrfs-005.html?ex=1155960000&en=8bd47f117bee4573&ei=5070