Spatial Analysis of High-Resolution Radar Rainfall and Citizen-Reported Flash Flood Data in Ultra-Urban New York City

Brianne Smith* and Stephanie Rodriguez

Department of Earth and Environmental Sciences, Brooklyn College, Brooklyn, NY 11210, USA;
brianne.smith43@brooklyn.cuny.edu
* Correspondence: brianne.smith43@brooklyn.cuny.edu

Abstract: New York City (NYC) is an ultra-urban region, with over 50% impervious cover and buried stream channels. Traditional flood studies rely on the presence of stream gages to detect flood stage and discharge, but these methods cannot be used in ultra-urban areas. Here we create a high-resolution radar rainfall dataset for NYC and utilize citizen and expert reports of flooding throughout the city to study flash flooding in NYC. Results indicate that interactions between the urban area and land–sea boundary have an important impact on the spatial variability of both heavy rainfall and flooding, sometimes in contrast to results obtained for other cities. Top days of daily and hourly rainfall exhibit a rainfall maximum over the city center and an extended region of higher rainfall downwind of the city. The mechanism for flooding appears to vary across the city, with high groundwater tables influencing more coastal areas and high rain rates or large rain volumes influencing more inland areas. There is also a strong relationship between sewer type and flood frequency, with fewer floods observed in combined sewer areas. Flooding is driven by maximum one-hour to one-day rainfall, which is often substantially less rain than observed for the city-wide daily maximum.

1. Introduction

Flash flooding is a costly and deadly problem throughout the United States. In 2015, flash flooding caused more fatalities and cost more dollars in damages than any other weather-related hazard [1]. The impacts of flash flooding have grown over time, even as impacts due to other weather hazards (tornadoes, hurricanes) have decreased [2]. These increasing effects are largely due to increases in flash flood frequency and severity, as well as the difficulty in predicting flash flooding.

Urban areas are particularly prone to flash flooding. Many studies have detailed the increase in flood volumes and discharge peaks associated with increasing impervious area (for example, [3–7]). A national study found that the flashiest watersheds in the country are urban [8]. The U.S. is urbanizing as the population continues to move into cities and suburbs. The 2010 census estimated that 80.7% of the population lived in urban areas, up from 79.0% in 2000 (https://www.census.gov/geo/reference/ua/urban-rural-2010.html). Increases in urbanization have been found to increase flash flooding volume [3] and frequency [9,10].

Additionally, impacts of global climate change may increase flash flooding. Climate change is expected to increase the intensity of rainfall extremes in most areas [11,12], which should increase the severity of flash flood events. In the northeastern U.S., extreme rainfall (defined as the 1% daily rainfall) increased by 71% between 1958 and 2012 [13]. In the east, urban areas may be experiencing an

even greater increase in extreme rainfall than rural areas, but this trend is less clear in the northeastern U.S. [14].

Finally, flash floods are difficult to predict. Urban flash floods are frequently caused by convective rainfall, which can be difficult to predict and difficult to place in space [15]. Flash flooding is often predicted using a rainfall guide: the volume of rainfall which will cause flooding in a catchment based on the volume of water which can be held in the soil at the time. This procedure is less accurate in urban areas where soil properties are less important, and rainfall intensity, not volume, drives flooding [16].

Unfortunately, in ultra-urban areas like New York City, flash floods are also difficult to detect. Traditional definitions of flash flooding rely on a short time between rainfall and flood peak or a fast flood water rise in a stream. Flash flood detection typically involves analyzing a stream gaging record to find sharp rises in streamflow after intense periods of rainfall. Researchers have used various techniques to distinguish flash floods, such as streamflow peaks-over-threshold [8,17] and "flashiness," a combination of flood stage, action stage, flooding rise time, and basin area [18]. Other studies have focused on identifying flash floods in un-gaged watersheds on a case-by-case basis with post-flood field investigations using eye-witness accounts and high-water marks [19], or have constructed a longer record using historical documents and tree ring data [20]. However, in ultra-urban areas with substantial impervious cover, streams have been buried and turned into sewer systems. In New York City, streams such as the Minetta Brook are now sewer interceptors and occasionally cause flooding in building basements [21]. The only stream gaging record in New York City is the Bronx River (US Geological Survey ID number 01302020), which has a short (11-year) history. This is a problem in city centers throughout the country. In this study, we create a high-resolution radar rainfall dataset for ultra-urban New York City. We use this dataset to investigate relationships between rainfall and flooding as reported by New York City residents. The study is the first to combine two such datasets to study the spatial occurrence of flash flooding in an ultra-urban area. The research questions which guide include:

1. What are the spatial characteristics of warm-season rainfall in New York City?
2. What are the spatial and temporal characteristics of citizen-reported flooding in New York City?
3. Can citizen-reported flooding be related to specific rainfall characteristics or other land use/sewershed characteristics?
4. What kind of rainfall causes flooding in New York City?

The paper is structured as follows: Section 2 provides a background of studies relevant to this work, Section 3 describes the data and methods used in this study, Section 4 describes results from this study, and Section 5 details the conclusions reached.

2. Background

In recent years, high-resolution radar rainfall data have allowed for advances in urban hydrology, especially for studies of flash flooding and stormwater management. Urban flooding occurs at extremely rapid rates: the time from rainfall to flooding can be as little as 15 min [22]. Study of these floods require rainfall data with as short a time resolution as possible in order to capture the temporal heterogeneities that can drive flooding and stormwater generation. Widely available, spatially explicit rainfall data have resolutions ranging from 1 h upwards. Furthermore, rainfall which causes urban flooding is extremely concentrated both in time and in space. The devastating Fort Collins flood of 1997 was caused by rainfall totals in excess of 10 inches near the flood site, but rainfall reached only 2 inches 4 km away [23]. The rainfall was so concentrated in space that it was missed by local weather spotters and had to be verified through residents' backyard rain gages [24]. High spatial resolution rainfall fields are required to capture the spatially concentrated extreme rainfall which drives urban flooding. The importance of developing high-resolution radar rainfall data for cities was explained in a recent study. The authors found that the commonly used Stage IV rainfall

data are too coarse for urban flood modeling and urban hydrology and recommended the use of the higher-resolution data developed here [25].

A handful of studies have attempted to understand the spatial patterns of observed extreme rainfall in New York City as it is impacted by urbanization. Bornstein and LeRoy investigated the impacts of the city on bifurcation of thunderstorms. They found that the urban heat island can cause convection while the city roughness can induce divergence; together these result in a rainfall minimum within the city and rainfall maxima surrounding and downwind of the city [26]. Yeung used high-resolution radar rainfall fields from the Fort Dix (New Jersey) radar, where weather surveillance radar (WRF)-modeled storm events were used to investigate the role of urban areas on convective storm tracks in the greater NYC region. The results showed an increased number of days exceeding 25 mm of rainfall over New York City (averaging nine per summer season [27]). Recently, Stage-IV rainfall data were used to classify rainfall in New York City down to 1-h events [28]. The results showed that rainfall extremes have substantially higher rainfall rates at a 1-h scale in the summer as opposed to the winter. Additionally, summertime extreme rainfall is more spatially concentrated and associated with localized frontal systems than extreme winter rainfall. Furthermore, Queens is most likely and Staten island least likely to experience high intensity, large areal extent 1-h summertime precipitation extremes [28].

As previously mentioned, city centers suffer from a lack of streams and stream gaging sites with which to study flash flooding. The combination of these uncertainties, and the fact that flash floods often happen in places not covered by stream gages, have led researchers to look for other ways to define flash flooding. Researchers have turned to flash flood reports, in which trained observers detail flash flood extent and damage, to identify flooding and have used these reports, along with more common stream gaging records, to create flash flood databases (in the U.S. [29] and in Europe [30]). Other researchers have called for the development of a Flash Flood Severity Index, based on damage, similar to the Enhanced Fujita scale for tornadoes [31]. Meanwhile, some researchers are turning to citizen science to record the presence of flooding or to verify hydrodynamic model output [32,33] and have created smartphone apps to report flooding [34]. In ultra-urban areas, which lack typical surface channel gages, citizen flood reports represent a great resource and pathway forward in studying flash floods.

3. Data and Methodology

3.1. Rainfall

Development of the bias-corrected radar rainfall field data for New York City utilized radar reflectivity scans from the Fort Dix (KDIX) Weather Surveillance Radar–1998 Doppler (WSR-88D) in Mount Holly, New Jersey, operated by the National Weather Service. The weather radar produces three-dimensional scans of the regional atmosphere every 5–6 min with reflectivity values as data. This reflectivity data can be transformed into rainfall rates using a Z-R relationship: $R = a \times Z^b$. In this equation, R represents rain rate in mm h^{-1}, Z represents radar reflectivity in mm^2 m^{-3}, and a and b are coefficients which depend on the type of rainfall and weather system, respectively. Radar reflectivity data require processing to obtain reasonable estimations of rainfall values at a consistent time scale. This processing must contend with radar problems such as the radar hitting the ground, hailstones, conversion of three-dimensional (3D) radar scans into two-dimensional (2D) rainfall maps, and range errors associated with variations with the vertical variation of reflectivity data. The Hydro-NEXRAD algorithms have been shown to successfully process radar reflectivity data to account for the Z-R relationships and above issues [35]. A 15-year period (2001–2015) of warm-season (April–September) radar rainfall fields was produced for the KDIX radar.

The Hydro-NEXRAD-processed radar rainfall fields were then bias-corrected using a daily multiplicative bias. Rainfall values calculated from weather radar can be inaccurate due to miscalibration of the radar, improper use of Z-R relationships, and other physical problems [36].

These inaccuracies create consistent differences between the actual amount of rainfall and the radar estimate. Bias correction has been shown to adjust for these inaccuracies in radar rainfall estimates [36,37]. Multiplicative bias correction requires use of ground observations to correct radar rainfall data. The form of the bias is: $B = \sum_i \frac{G_i}{R_i}$. Where B represents the bias, G represents rainfall depth for individual ground observations, and R represents rainfall depth for the radar pixel in which the ith ground observation was taken. The bias is calculated by dividing the sum of ground observation depths over a time period by the sum of radar observation depths over the same time period. It is necessary to use radar rainfall data from the same area, or radar pixel, that the ground observation is collected in.

Ground observations were taken from rain gages in the NOAA National Center for Environmental Prediction Meteorological Assimilation Data Ingest System (NCEP MADIS, https://madis.ncep.noaa.gov/) and the NOAA National Centers for Environmental Information US Historical Climatology Network (NCEI HCN, https://www.ncdc.noaa.gov/cdo-web/) which includes several long-term gages within New York City. Rain gages throughout the New York City greater metropolitan region were chosen for bias correction, and the gage locations are shown in Figure 1. An effort was made to exclude gages which fell within a blocked radar band which covers much of Staten Island. Rain gages were used to calculate biases at a daily time-step. Smaller time-steps can result in random errors in bias-calculation [37]. Biases were calculated for any day in which at least four rain gages registered rainfall.
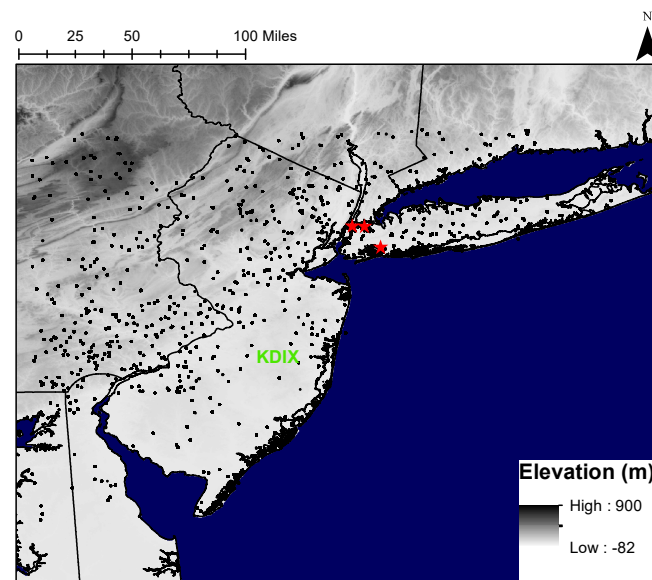

Figure 1. Study area map with elevation. The radar location in Mount Holly, NJ, is represented by the green "KDIX", HCN rain gages are represented by red stars (Central Park to the west, JFK to the south, and LaGuardia to the northeast), and all other rain gage locations are represented by black points.

3.2. Flood Reports

What New York City lacks in stream gages, it makes up for with a dense and vocal population. Residents can register flood reports by calling or accessing on-line the local 311 non-emergency city services and information number/website. NYC 311 records flood observations as well as noise complaints, requests for street trees, rodent sightings, graffiti reports, missed trash collections, and numerous other city occurrences. Data for the day and location of these complaints are available on-line through the NYC open data website: (https://opendata.cityofnewyork.us/). In this study we utilized street and highway flooding observations between 2004 and 2015. We excluded clogged catchbasin flooding reports, as these reports are largely influenced by local practice and season. A total of 15,418 flood observations were used for this study, as shown in Figure 2. The flood observations were normalized to control for factors that could make residents of certain areas of the city more likely to report flooding than in other areas of the city (including population density or perceived city

responsiveness). All 311 complaints, excluding complaints which were clearly geographically biased, were used to normalize the flood 311 reports. Geographically-biased 311 reports which were removed included any reports likely to occur in certain housing complexes or institutions (those related to the Department of Housing Preservation). Figure 2 shows floods reports as a measure of flood report per 311 report per km^2.
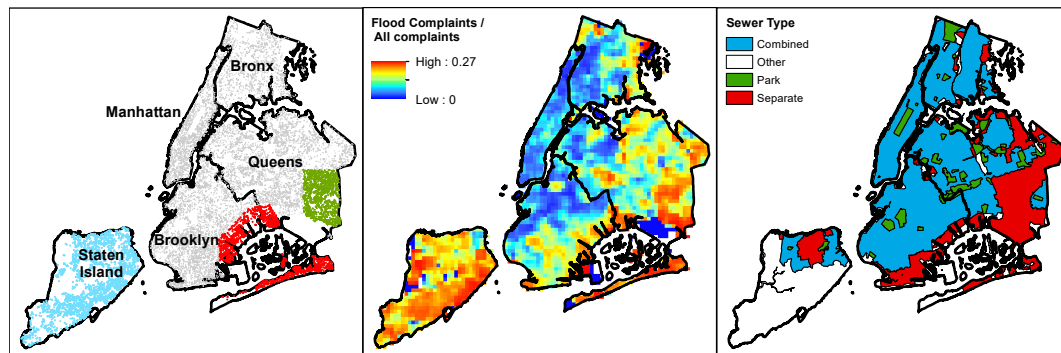

Figure 2. Flood observations through 311 for the period 2004–2015. **Left panel**: Each point represents one 311 flood observation. Blue points are in Staten Island, red points are near Jamaica Bay, and green points are in southeast Queens; **Center panel**: Flood observations through 311, normalized by all 311 observations. Units are in flood observations per any observation in 1 km^2; **Right panel**: New York City (NYC) sewer type.

4. Results and Discussion

4.1. Radar Rainfall Dataset

The spatial pattern of daily-averaged rainfall for the entire period (warm season 2001–2015) shows the presence of a blocked radar band across Staten Island (Figure 3). The blocked radar band can be seen as the blue line of lower rainfall depth extending from Staten Island northeast through New York State. Rainfall gages located in the blocked radar band were excluded from bias correction procedures. Results of bias correction are shown in Table 1. Average annual biases range from 0.75 to 1.5 and are consistent with the range of values found in other studies [37,38]. The annual percentage of radar periods with complete radar data ranges from 87% to 97%, with earlier periods missing more radar data. Finally, the number of rain gages used in the bias-correction analysis increases substantially, from 13.4 in 2001 to 120.9 in 2015. This increase represents the increasing presence of Automated Surface Observing System (ASOS) gages in the NYC region.
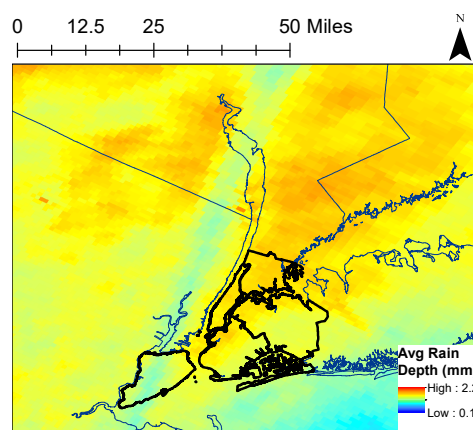

Figure 3. Daily averaged radar rainfall for all days in the warm seasons of 2001–2015. Navy lines are state outlines and black lines are NYC borough outlines.

Table 1. Radar rainfall daily bias correction results by year. Columns detail the average multiplicative daily bias, the % completeness of 15-min radar rainfall periods, and the average number of rainfall gages used for the calculation of a daily bias.

| Year	Average Bias	% Radar Record Completeness	Average Number of Gages
2001	0.76	86.6	13.4
2002	1.1	95.0	14.0
2003	1.1	96.3	14.4
2004	1.2	88.4	29.2
2005	0.99	97.1	40.5
2006	1.2	94.9	52.5
2007	1.1	91.6	51.8
2008	1.1	95.5	71.7
2009	1.3	96.5	88.6
2010	1.2	96.3	75.4
2011	1.2	96.9	103.7
2012	1.5	95.4	109.8
2013	1.5	94.4	124.6
2014	1.5	95.8	138.2
2015	1.4	96.1	120.9

The Pearson correlation coefficient was calculated to compare rain gage observations to the radar rainfall data for the 1-km^2 pixel in which the rain gage is located. The correlation for all rain gages on all days for the un-bias corrected data was 0.77. The correlation coefficient for all rain gages on all days for bias-corrected data was 0.87. Bias correction substantially increased the correlation between the radar rainfall dataset and the rain gage observations. Within New York City, the relationship between radar rainfall and rain gage observations is shown for the three NYC HCN gages in Figure 4. The Central Park rain gage, located in the middle of Manhattan, is NYC's longest standing rain gage. The JFK rain gage is located at John F. Kennedy Airport in southeastern Queens, and the LaGuardia rain gage is located at LaGuardia International Airport in northwestern Queens. Bias correction brought radar rainfall observations closer to radar observations, which can be observed through distance from the points to the red 1-1 line. This improvement is particularly noticeable for more extreme rainfall observations. At Central Park and JFK the un-bias corrected radar rainfall fields systematically under-estimated the most extreme rainfall days, and at LaGuardia the un-bias corrected radar fields systematically over-estimated the extreme rainfall days. Over all days, bias correction improved radar–rain gage correlations at Central Park from 0.78 to 0.94, at JFK from 0.83 to 0.93, and at LaGuardia from 0.85 to 0.95.

The pattern of rainfall in the New York City region (Figure 3) is heavily influenced by the complex land–sea boundaries in the area. The impact of water bodies on rainfall patterns appears to vary both by the size of the water body and its position relative to land masses. Weather sytsems in the New York City region typically move southwest to northeast [39]. There is substantially less rainfall south (upwind) of the city in the larger New York Harbor/Atlantic Ocean, but no noticeable reduction in rainfall to the north (downwind) of the city in smaller Long Island Sound. Additionally, the location of New York Harbor and Atlantic Ocean appear to influence the pattern of rainfall over the city; rainfall is lower in the southeastern portions of Brooklyn and Queens and higher in the Bronx. Regionally, there is more rainfall north of the city than within the city. Urban rainfall modification has been found to increase downwind rainfall in urban regions across the world [40]. Land–sea boundaries have been found to shift rainfall further inland [41] and displace urban rainfall maximums [42]. The land–sea boundary in NYC is particularly complex, and studies have shown its ability to create multiple sea-breeze fronts [43].
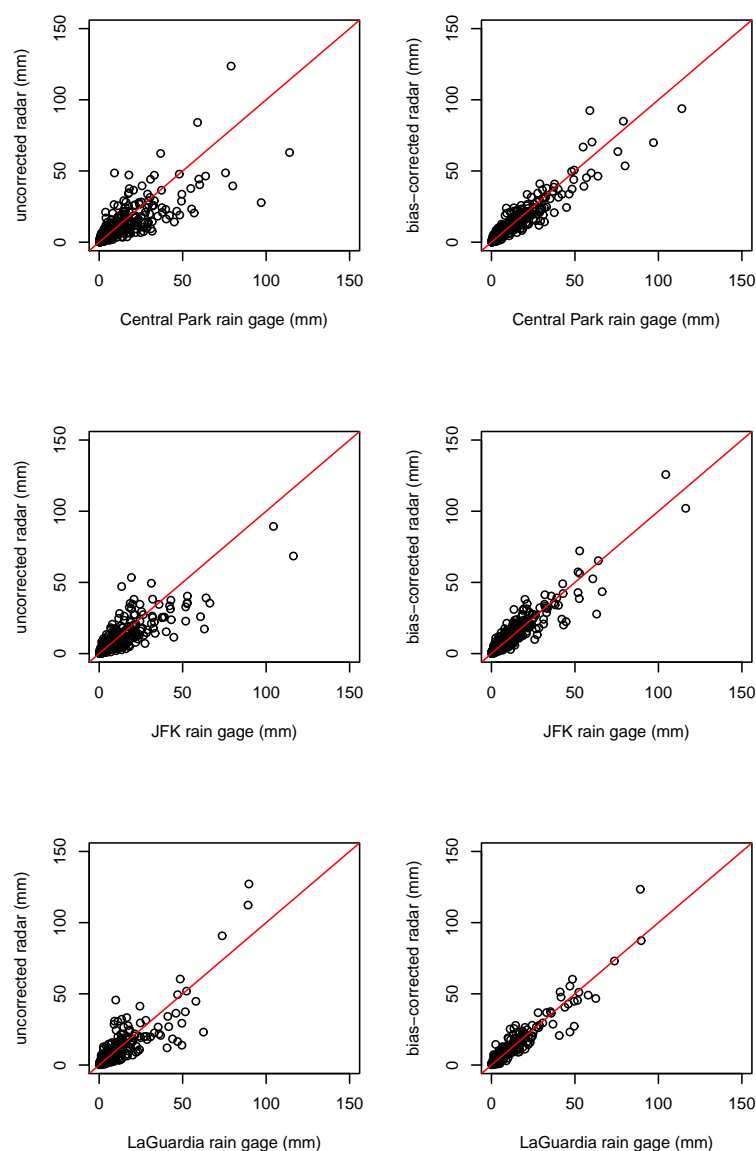
Figure 4. Comparison of daily rain gage rain depths and radar rainfall rain depths for un-bias corrected (**left column**) and bias-corrected (**right column**) data. The top panel represents observations over 15 years of warm seasons from the Central Park rain gage, the middle panel represents observations from the JFK rain gage, and the bottom panel represents observations from the LaGuardia rain gage. The red line represents a 1 to 1 line.

Heavy rainfall in New York City shows a slightly different pattern (Figure 5). The 50 rainiest days based on the averaged rainfall of New York City display a rainfall maximum of 55 mm per day over the geographical center of the city in southwest Queens and northern Brooklyn. Rainfall is again diminished over Staten Island due to the blocked radar band. Rainfall changes relatively rapidly to the south of the city with the 35- and 50-mm rainfall contours falling within 36 km. Rainfall contours are more extended northeast (downwind) of the city with 66 km between the 35 and 50 mm contours. It is difficult to distinguish whether this pattern of extended higher rainfall downwind of NYC is due to urban rainfall modification or to the presence of New York Harbor, the Atlantic Ocean, and the Long Island Sound. The heavy rainfall pattern does not show the separate downwind maximum of rainfall that has been found in other urban regions (e.g., St. Louis [44], Dallas [45], Indianapolis [46], Atlanta [47], and Baltimore [37]).
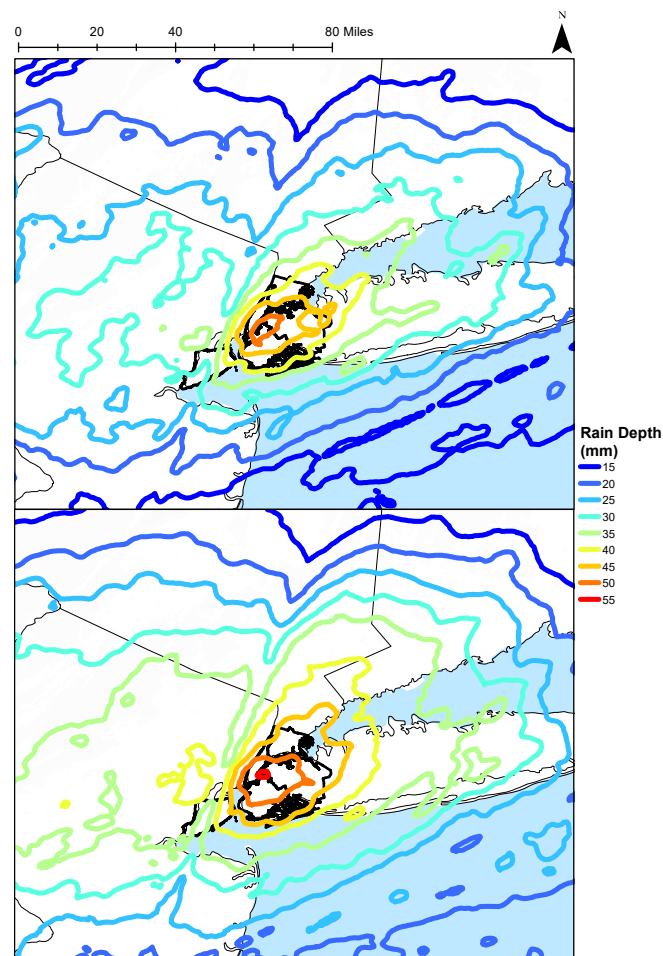
Figure 5. Rainfall depth for daily-averaged rainfall for the 50 rainiest days in NYC (**top**) and the days with the highest 3-h rainfall in NYC (**bottom**). Bold black lines represent NYC borough boundaries, white areas represent land masses, and blue areas represent water bodies.

Flash flooding is often caused by short-term, intense rainfall. Intense rainfall for a 1–6 h period can be more likely to cause urban flash flooding than intense rainfall for a longer period [22]. In order to account for this, the 50 maximum rainfall days were determined as days with: (a) the highest total rainfall in 24 h; (b) the highest total rainfall in 1 h; (c) the highest total rainfall in 30 min; or (d) the highest total rainfall in 3 h. The rainfall map for the 50 maximum 1-h rainfall days shows a similar pattern to the rainfall map for the 50 days with the highest 24-h rainfall (Figure 5). The days with the highest 1-h rainfall have a slightly lower spatial rainfall max of 50 mm, falling in the geographical city center. Daily rainfall for these highest 1-h rainfall days is more concentrated around the center of the city and also displays a more extreme difference in upwind and downwind rainfall than the highest 24-h rainfall days. For the 1-h rainfall days, the 35- and 50-mm contours are as close as 15 km apart south of the city. Northeast of the city, the 35- and 50-mm contours are as far as 57 km apart. High short-term rain rates are likely to be caused by thunderstorm systems, which may be more influenced by urban modification [22]. These differences in short-term rainfall patterns are important for current and future flood prediction within the city as well as drainage system planning.

4.2. Flood Observations

Normalized 311 flood reports display a strong spatial pattern over New York City (Figure 2). Certain areas of NYC are entirely lacking in flood reports and appear solid blue in the normalized flood complaint map. The results in these areas can be explained by a lack of residents. These areas

include the small islands surrounding Manhattan, JFK airport in southeast Queens, and the Gateway National Recreation Area at the western tip of the Rockaway Peninsula and along the western portion of Jamaica Bay. The Jamaica Bay area, southeast Queens, and Staten Island (see Figure 2, left panel) stand out as having a large number of flood reports, with up to 0.27 flood reports per 311 complaint in the area. The causes of this flooding are likely quite different. The Rockaway Peninsula and much of the land surrounding Jamaica Bay suffer from high groundwater tables and frequent storm surges. Much of the flooding in these areas is likely due to compound flooding in which the combination of high groundwater tables, storm surges in the groundwater, and rain combine to cause a flood [48]. Scattered flooding across southeastern Queens is more likely due to heavy rainfall, as it lies farther from local water bodies. Staten Island has struggled with rainfall drainage issues because it does not have the combined or separate sewer systems present in much of the rest of New York City. Recent projects, such at the Staten Island Blue Belt, have sought to correct these issues.

New York City has two types of sewer systems and several other systems for dealing with drainage. The majority of the city has combined sewers, in which sewage and runoff are combined and piped to wastewater treatment plants. A substantial part of the city is served by separate sewers, in which runoff is concentrated into storm sewers separately from sewage. Some portions of the city, and the majority of Staten Island, are served by a combination of other methods, including direct drainage and the Staten Island Blue Belt. Flood occurrence in the city appears to be most closely related to sewer type. Areas with combined sewers have less than half the average incidence of flooding (0.005 flood complaints/all complaints/km^2) than do areas with separate sewers (0.011 flood complaints/all complaints/km^2) and other drainage systems (0.018 flood complaints/all complaints/km^2). Jamaica Bay, southeast Queens, and Staten Island are all areas without substantial combined sewer coverage.

Seasonality of 311 reports presents interesting patterns between the flood-prone regions of NYC (Figure 6). The basic patterns of all 311 reports for any problem are higher in the summer (day 200, mid-July) and lower in the winter (day 15, mid-January). This baseline can be used to compare against the seasonal variation of flood reports. Flood reports through all the city are less common in the wintertime and more common in the late spring (day 160, early June) through late fall. This suggests that the predominant mechanisms for these flood events are not large extratropical storm systems. Nor'easters are most severe in New York City between September and April (days 244–91). The more likely mechanisms for these floods are convective storms, which are strongest in the late spring through summer, and tropical cyclones, which are strongest from mid-August to mid-October (days 227–288). Floods caused by convective storms are most likely to be flash floods triggered by intense rainfall, and floods caused by tropical cyclones may be flash floods from heavy rainfall or coastal floods caused by storm surge. Flooding in southeast Queens closely follows the pattern of the city as a whole with only slightly more flooding in June through October (days 160 through 300) indicating that it may be impacted even less by nor'easters. Flooding in the Jamaica Bay area exhibits a large maximum in mid-May (day 140) when groundwater levels are highest and compound flooding is most likely. Flooding in Jamaica Bay exhibits a pronounced low in the heat of summer (days 200–250) suggesting that convective storms may not be the dominant cause of flooding in Jamaica Bay. Staten Island shows the most unique seasonal flood pattern, with a high incidence of flood reports through the winter months and a low incidence in late summer and early fall. Other flood reports systems (see Significant Flood Events below) similarly indicate a large number of winter-time floods which only impact Staten Island/Richmond County. This seasonal variation suggests that extratropical cyclones may play a larger role in flooding Staten Island than the rest of the city. Overall, southeast Queens appears to flood from intense short-term rainfall, Jamaica Bay appears to flood when groundwater tables are high, and Staten Island appears to flood from substantial long-term rainfall. These differences suggest important differences in managing and predicting flooding throughout the city. The diurnal cycle of flood reports is similar throughout all regions of the city (Figure 7). Flood reports reach a peak around noon and are lowest in the early morning hours. This suggests that flood reports are

more common in hours when residents are out in public, and that seasonality results must be viewed through this lens.
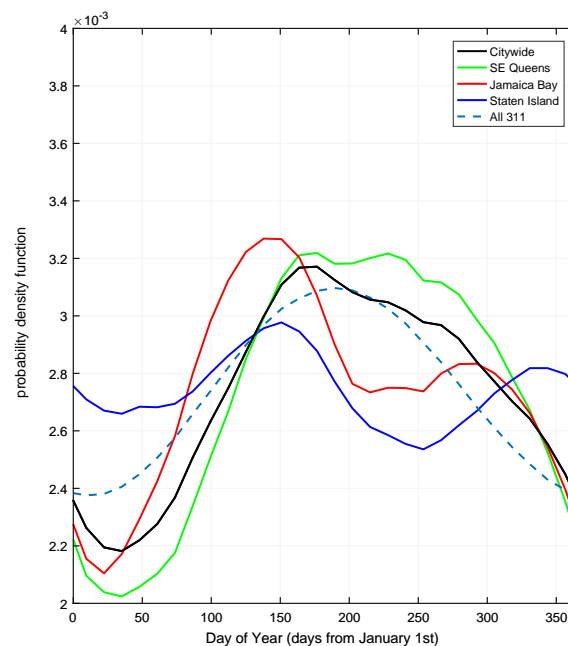
Figure 6. Probability density function for the seasonal variation in 311 reports. "Citywide" represents city-wide flood reports, "Southeast (SE) Queens Jamaica Bay and Staten Island" represent flood reports in these areas (see Figure 2), and "All 311" represents any 311 complaint.
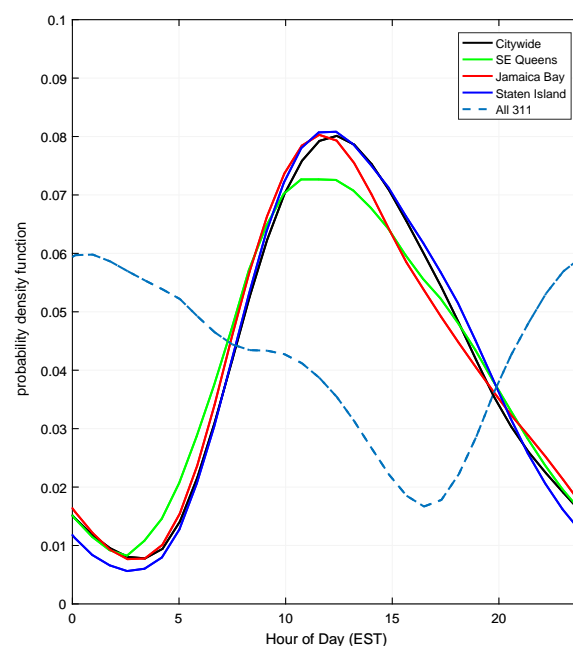

Figure 7. Probability density function for the diurnal variation in 311 reports. "Citywide" represents city-wide flood reports, "Southeast (SE) Queens Jamaica Bay and Staten Island" represent flood reports in these areas (see Figure 2), and "All 311" represents any 311 complaint. EST stands for Eastern Standard Time.

4.3. Rainfall/Flood Interactions

One benefit of citizen flood reports is that the data are spatial and can be compared to other spatial datasets. Several potential flood factors (rainfall, impervious area, etc.) were mapped and

normalized (0 to 1) with the same resolution and grid location (Figure 8). Next, 1000 points were randomly placed on these maps (excluding the blocked rain band in Staten Island) and used to find Spearman correlation coefficients between flood reports and potential flood factors (Table 2). Results indicate that the factor most highly correlated with flood reports is population density. Population density potentially impacts flood reports in two ways: (1) through an increased number of residents available to report flooding; and (2) through increased impervious surfaces with increased population density. Impervious surface percentage is similarly correlated to flood reports and ranks as the third most highly correlated factor. The second most highly correlated factor is rainfall from days with the highest 3-h rainfall. This suggests that 3 h of intense rainfall is most likely to cause flooding citywide. The highest 24-h rainfall and highest 1-h rainfall days are similarly correlated, while the highest 30-min rainfall days and entire warm-season rainfall are substantially less correlated. It appears that intense 3-h to 1-day rainstorms are most important for causing flooding in the city. Elevation, as both absolute elevation and flow accumulation (the amount of water which would accumulate from the surface), is less correlated than other factors. This is somewhat surprising, but may indicate that the city surface drainage is primarily controlled by sewers rather than elevation.
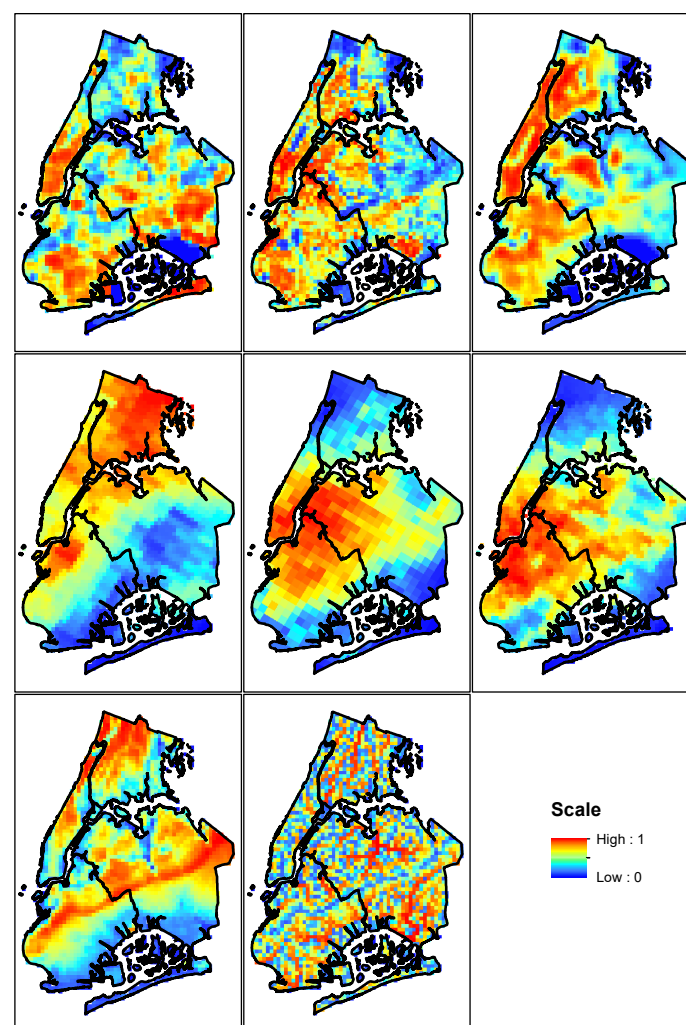
Figure 8. Normalized (0 to 1) maps for flooding (**top left**), impervious surface percentage (**top center**), population density (**top right**), daily-averaged rain for entire study period (**center left**), daily-averaged rain for the 50 rainiest days (**center center**), daily-averaged rainfall from days with top 50 one-hour rainfall (**center right**), absolute elevation (**bottom left**), and flow accumulation (**bottom center**). Color bar represents all maps.

Table 2. Flooding correlations between number of flood events and potential flood factors.

Characteristic	Correlation Coefficient		
	Citywide	Jamaica Bay	Southeast Queens
Daily-averaged rainfall for entire study period	−0.20	0.09	0.02
Daily-averaged rain for the 50 highest 24-h rainfall days	0.30	0.25	0.16
Daily-averaged rain for the 50 highest 3-h rainfall days	0.34	0.31	0.17
Daily-averaged rain for the 50 highest 1-h days	0.26	0.27	0.12
Daily-averaged rain for the 50 highest 30-min rainfall days	0.11	0.26	0.08
Impervious surface percentage	0.32	0.38	0.11
Absolute elevation	0.05	0.33	−0.14
Population density	0.42	0.57	0.10
Flow accumulation	0.17	0.25	0.18

Analyses were repeated for two high-flood areas, Jamaica Bay and southeast Queens (as depicted in Figure 2). Jamaica Bay flood reports have higher correlation coefficients for all factors than do citywide flood reports, but the ranking of factor correlation coefficients is generally similar. The major differences are that elevation and flow accumulation are more highly correlated to flood reports in Jamaica Bay, and that maximum 30-min rainfall days are more highly correlated. This suggests that Jamaica Bay is sensitive to short-term extreme rainfall and that surface drainage is a more important flooding factor in Jamaica Bay than city-wide. This difference may be explained by the separate sewers and surface drainage in Jamaica Bay (Figure 2). Surprisingly, elevation is positively correlated with flood reports in Jamaica Bay, suggesting that the flood reports are for more upland areas because elevation increases with distance from water in this area. The upland locations and increased importance of 30-min rainfall indicate that these flood reports may include many flash floods. Southeast Queens flood reports have low correlation coefficients with all factors. In southeast Queens the dominant factors are flow accumulation and 3-h rainfall, again suggesting the importance of overland flow in a separate sewered area and the importance of short-term rain rates. The correlations in southeast Queens may be impacted by the small area and relative uniformity of datasets within the area, particularly population and impervious surface percentage. Staten Island was not included in these analyses due to the lack of rainfall data.

4.4. Significant Flood Events

While 311 reports provide a good coverage of observed floodwater in the city, other resources are available to identify significant flood events. The New York City Office of Emergency Management publishes a list of flooding events and damages in its 2014 Hazard Mitigation Plan and its annual updates. Additionally the National Climactic Data Center (https://www.ncdc.noaa.gov/stormevents/) publishes information for significant floods as reported by National Weather Service employees. Neither of the datasets is comprehensive, and reports have tended to increase in more recent years. These reports do, however, provide a basis of 72 expert-selected flood-producing storm events to analyze. Flood events which were only reported in Staten Island were excluded from the dataset due lack of rainfall data. Rainfall for these flood-producing events is substantially less (Figure 9, up to 32 mm) than maximum rainfall amounts (Figure 5, up to 57 mm). Spatial rainfall patterns are similar for significant flood rainfall and maximum rainfall, but rainfall highs are less concentrated and extend further downwind of the city for these significant flood events. Lower total rainfall, more urban modification of rainfall, and a stronger impact of the Long Island Sound, suggest that convective storms may be a dominant factor in these flood-producing events.
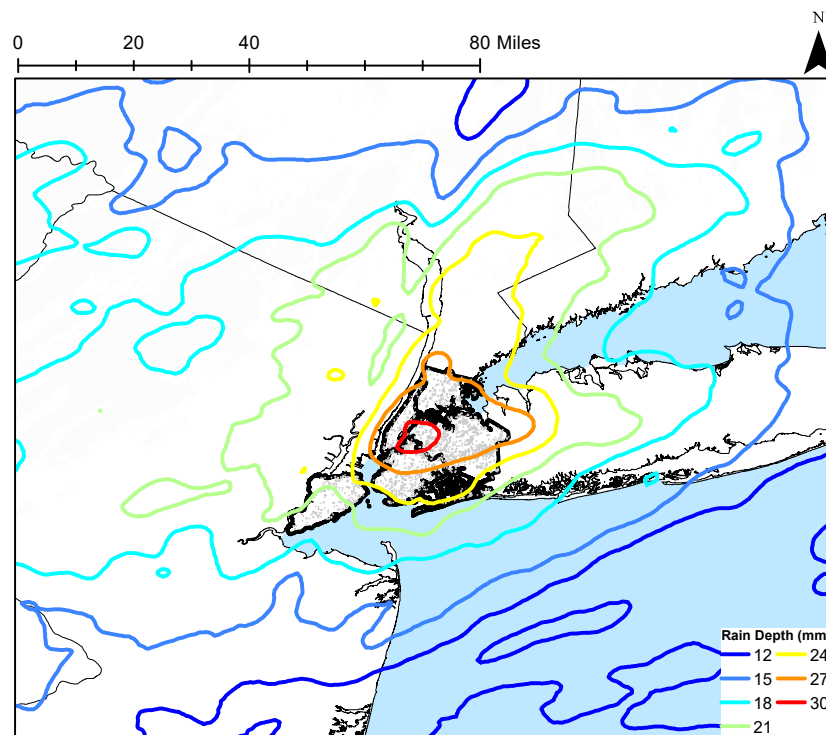
Figure 9. Daily-averaged rainfall for days of known flooding. Grey points represent locations of 311 flood reports on these days. Bold black lines represent NYC borough boundaries, white areas represent land masses, and blue areas represent water bodies.

Rain rates, averaged across four boroughs (Manhattan, Brooklyn, Queens, and the Bronx), indicate the variation in rainfall driving these floods (Figure 10). Daily-averaged rain rates range from 0.09 to 4.79 mm h^{-1} (2.2 to 115.0 mm day^{-1}) with a mean of 1.18 mm h^{-1} (28.4 mm day^{-1}). The minimum flood-producing daily rain on 2 July 2004 caused a localized flood in the Bronx where thunderstorms caused a higher borough-averaged rainfall of 6 mm day^{-1}. While localized, the flood did result in a flooded highway and two water rescues. The maximum flood-producing daily rainfall fell on 15 April 2007 caused by a spring-time nor'easter which stalled off the coast of New York City and produced the largest rainfall recorded at the Central Park rain gage (224 mm). This storm caused extensive street flooding throughout Manhattan, Brooklyn, and Queens. Short-term 30-min rainfall varied from 2.75 to 45.22 mm h^{-1} with an average value of 14.73 mm h^{-1}. The maximum 30-min flood-producing rainfall was caused by a severe thunderstorm on 8 August 2007 which resulted in several highway lane closures, damage to homes in Brooklyn, delays and closure of subway lines, and the first tornado to touch down in Brooklyn in at least 50 years. The minimum 30-min flood-producing rainfall was again the localized 2 July 2004 flood, with a higher 30-min rain rate (10.46 mm h^{-1}) in the Bronx indicating that nearly all the rain fell within 30 min. Among these flood event maximum rain rates, there is strong correlation in storm-specfic values between the 30-min and 3-h rainfall (Pearson correlation coefficients between 0.83 and 0.97), but there is less of a correlation between the shorter-term rain rates and daily rainfall (between 0.42 and 0.61). Results suggest that New York City is vulnerable to two types of pluvial floods: widespread floods caused by high daily-rainfall throughout the city, and localized floods caused by thunderstorm systems with high short-term rainfall in the area. The localized nature of some flood-producing rain events indicates the importance of using local rainfall rather than city-wide data.
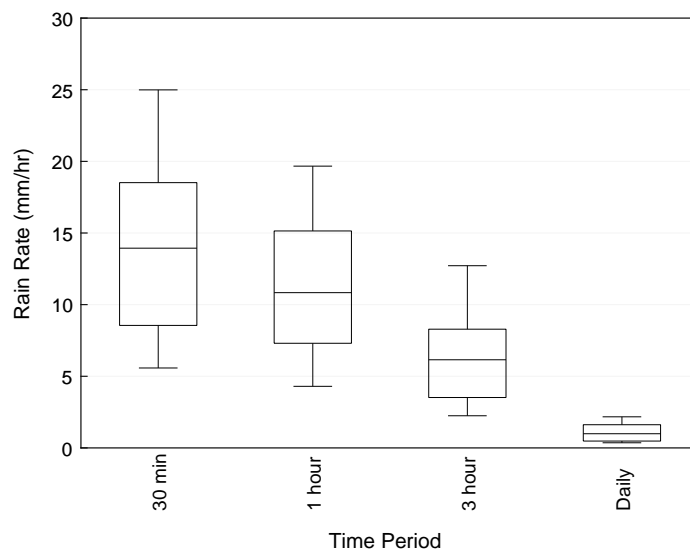
Figure 10. Boxplot for NYC-averaged max rainfall rates over 30 min, 1 h, 3 h, and 1 day for the 86 days of known flooding. Whiskers represent the 10th and 90th percentiles, the box represents the 25th and 75th percentiles, and the line represents the median.

5. Summary and Conclusions

A 15-year (2001–2015) high-resolution (1 km^2, 15 min) radar rainfall dataset was created for the warm season (April–September) in the New York City metropolitan area. Hydro-NEXRAD-processed radar-rainfall fields were bias-corrected on a daily basis to create the dataset. Spatial characteristics of the rainfall fields were analyzed to observe patterns across NYC. Flash flood data were taken from the NYC 311 report line, which residents can call to report flooding on streets and highways. These data were mapped, normalized, and compared to sewer type, rainfall, and several other spatial datasets across the city. Rainfall was analyzed for days of significant flooding, as identified by experts in the NOAA Storm Events Database and NYC Hazard Management Plans. In this way, this study was able to bridge the difficulties of studying flash flooding in an ultra-urban area with no surface stream channels. The following conclusions were formed:

1. Warm-season rainfall is impacted by both the presence of the urban area and complex land–sea boundaries. Urban modification of rainfall appears to be less pronounced than in other regions, but is more observable for large shorter-term (1 h, Figure 5) and flood-producing rainfall (Figure 8) than for large daily rainfall (Figure 5). Urban modification appears to increase short-term rainfall down-wind of the city, but the combination of urban modification and the nearby coast pushes the rainfall maximum over the center of the city.

2. Normalized flood reports identify areas of the city with more flood reports per all other reports (Figure 2), and many of the higher flooding areas are similar to areas discussed by local flood experts (Jamaica Bay, southeast Queens, and Staten Island). Spatial variation among normalized flood reports closely resembles a map of the sewer-types throughout the city. Additionally, areas without combined sewers (Jamaica Bay and southeast Queens) exhibit a stronger correlation between spatial flood reports and flow accumulation than does the rest of the city (Table 2), suggesting than surface flow is more important in these areas. Areas with combined sewers appear to have substantially less flooding than areas without, possibly due to the larger carrying-capacity of sewers built for stormwater and sewage and an extended drainage network.

3. Flooding type varies throughout the city. The Jamaica Bay area is near the coast and floods occur predominantly in the spring when groundwater tables are highest (Figure 6); these floods are likely to be compound floods where rainfall interacts with groundwater surge. Southeast Queens

floods throughout the summer and fall, suggesting that thunderstorms and tropical cyclones create the heavy rain to flood this non-coastal area. Staten Island has more flooding in the winter than other portions of the city. These floods are likely caused by longer-term, large-volume rainfall from extra-tropical cyclones. Staten Island drainage is predominantly surface drainage and may take longer to drain. Flood-type variation between upland and coastal areas indicates that the land–sea boundary impacts both flood-producing rainfall and flooding mechanisms within the city.

4. Flood reports appear to be driven by 1-h to 1-day rainfall durations. There are substantially lower correlations (Table 2) between flooding and maximum 30-min rainfall and seasonally-averaged daily rainfall in comparison to the correlations between 1-h, 3-h, and daily maximum rainfall.

5. Flood-causing rainfall is substantially less than maximum daily or hourly rainfall (maximum of 30 mm versus a maximum of 55 mm, Figures 5 and 7). City-averaged rain rates of as little as 2.2 mm in one day and 2.75 mm h^{-1} over 30 min can produce local flooding (with borough-wide rates of 6 mm in a day and 10.46 mm h^{-1} over 30 min). The correlation between city-wide rainfall in the sub-daily and daily time-periods is low, suggesting that floods in NYC are caused by either intense, localized short-term rain or long-term city-wide rain, but rarely both.

6. Flood reports, made by residents or by experts, are an important resource in an ultra-urban city with no surface channels and a dense population. Methods of utilizing these reports are necessarily different than methods for stream gaging flood identification. Report timing may be influenced by resident behavior rather than the diurnal cycle of flooding (Figure 7), but the geographic nature of the data allows for spatial analyses which are not possible with stream gage point measurements.

Author Contributions: B.S. conceived and designed the study; S.R. deveoped the spatial flood database. B.S. bias-corrected the rainfall data, analyzed all data, and wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. National Weather Service Office of Climate, Water and Weather Services. *Summary of Natural Hazard Statistics for 2015 in the United States*; NWS: Silver Spring, MD, USA, 2016.
2. Kunkel, K.E.; Pielke, R.A.; Changnon, S.A. Temporal fluctuations in weather and climate extremes that cause economic and human health impacts: A review. *Bull. Am. Meteorol. Soc.* **1999**, *80*, 1077–1098.
3. Leopold, L.B. *Hydrology for Urban Planning—A Guidebook on the Hydrologic Effects of Urban Land Use*; Circular 554; United States Geological Survey: Reston, VA, USA, 1968.
4. Anderson, D.G. *Effects of Urban Development on Floods in Northern Virginia*; USGS Paper; United States Geological Survey Water Supply: Reston, VA, USA, 1970; p. 22.
5. Robbins, J.C.; Pope, B.F. *Estimation of Flood-Frequency Characteristics of Small Urban Streams in North Carolina*; United States Geological Survey Water-Resources Investigations Report 964084; United States. Geological Survey: Reston, VA, USA, 1996; p. 21.
6. Shuster, W.D.; Bonta, J.; Thurston, H.; Warnemuende, E.; Smith, D.R. Impacts of impervious surfaces on watershed hydroloy: A review. *Urban Water J.* **2005**, *2*, 263–275.
7. Moglen, G.E.; Kim, S. Limiting Imperviousness. *J. Am. Plan. Assoc.* **2007**, *73*, 161–171.
8. Smith, B.K.; Smith, J.A. The Flashiest watersheds in the contiguous US. *J. Hydrometeorol.* **2015**, *16*, 2365–2381.
9. Hollis, G.E. The effect of urbanization on floods of different recurrence intervals. *Water Resour. Res.* **1975**, *11*, 431–435.
10. Villarini, G.; Smith, J.A.; Serinaldi, F.; Bales, J.; Bates, P.D.; Krajewski, W.F. Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. *Adv. Water Resour.* **2009**, *32*, 1255–1266.

11. Lenderink, G.; van Meijgaard, E. Linking increases in hourly precipitation extremes to atmospheric temperature and moisture changes. *Environ. Res. Lett.* **2010**, *5*, 025208.

12. Sillman, J.; Kharin, V.; Zwiers, F.; Zhang, X.; Bronaugh, D. Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *J. Geophys. Res. Atmos.* **2013**, *118*, 2473–2493.

13. Walsh, J.; Wuebbles, D.; Hayho, K.; Kossin, J.; Kunkel, K.; Stephens, G.; Thorne, P.; Vose, R.; Wehner, M.; Willis, J.; et al. *Climate Change Impacts in the United States: The Third National Climate Assessment*; Chapter Our Changing Climate; U.S. Global Climate Change Research Program: Washington, DC, USA, 2014; pp. 19–67.

14. Niyogi, D.; Lei, M.; Schmid, P.; Shepherd, M. Urbanization Impacts on the Summer Heavy Rainfall Climatology over the Eastern United States. *Earth Interact.* **2017**, in press.

15. Doswell, C.; Brooks, H.; Maddox, R. Flash flood dorecastin: An ingredients-based methodology. *Weather Forecast.* **1996**, *11*, 560–581.

16. Hapuarachchi, H.A.P.; Wang, Q.J.; Pagano, T.C. A review of advances in flash flood forecasting. *Hydrol. Process.* **2011**, *25*, 2771–2784.

17. Ledbetter, M. On avbasis for "Peaks over Threshold" modeling. *Stat. Probab. Lett.* **1991**, *12*, 357–362.

18. Saharia, M.; Kirstetter, P.E.; Vergara, H.; Gourley, J.J.; Hong, Y.; Giroud, M. Mapping Flash Flood Severity in the United States. *J. Hydrol.* **2017**, *18*, 397–411.

19. Gaume, E.; Livet, M.; Desbordes, M.; Villeneuve, J.P. Hydrological analysis of the river Aude, France, flash flood on 12 and 13 November 1999. *J. Hydrol.* **2004**, *286*, 135–154.

20. Ruiz-Villanuevaa, V.; Diez-Herrero, A.; Bodoque, J.; Canovas, J.B.; Stoffel, M. Characterisation of flash floods in small ungauged mountain basins of Central Spain using an integrated approach. *Catena* **2013**, *110*, 32–43.

21. Kadinsky, S. *Hidden Waters of New York City: A History and Guide to 101 Forgotten Lakes, Ponds, Creeks, and Streams in the Five Boroughs*; WW Norton: New York City, NY, USA, 2016.

22. Smith, B.K.; Smith, J.A.; Baeck, M.L.; Villarini, G.; Wright, D.B. Spectrum of storm event hydrologic response in urban watersheds. *Water Resour. Res.* **2013**, *49*, 2649–2663.

23. Petersen, W.; Carey, L.D.; Rutledge, S.; Knievel, J.C.; Doesken, N.J.; Johnson, R.H.; McKee, T.B.; Haar, T.V.; Weaver, J.F. Mesoscale and radar observations of the Fort Collins flash flood of 28 July 1997. *Bull. Am. Meteorol. Soc.* **1999**, *80*, 191–216.

24. Cifelli, R.; Doesken, N.; Kennedy, P.; Carey, L.D.; Rutledge, S.; Gimmestad, C.; Depue, T. The Community Collaborative Rain, Hail, and Snow Network: Informal Education for Scientists and Citizens. *Bull. Am. Meteorol. Soc.* **2005**, *86*, 1068–1077.

25. Wright, D.B.; Smith, J.A.; Villarini, G.; Baeck, M.L. Long-Term High-Resolution Radar Rainfall Fields for Urban Hydrology. *J. Am. Water Resour. Assoc.* **2013**, *50*, 713–734.

26. Bornstein, R.; Leroy, M. *Urban Barrier Effects on Convective and Frontal Thunderstorm*; preprint volume for Fourth Conference on Mesoscale Processes; American Meteorological Association: Boston, MA, USA, 1990.

27. Yeung, J.; Smith, J.; Baeck, M.; Villarini, G. Lagrangian Analysis of rainfall structure and evolution for organized thunderstorm systems in the urban corridor fo the Northeastern US. *J. Hydrometeorol.* **2015**, *16*, 1575–1595.

28. Hamidi, A.; Devineni, N.; Booth, J.; Hosten, A.; Khanbilvardi, R.K. Classifying urban rainfall extremes using weather radar data: An application to the greater New York area. *J. Hydrometeorol.* **2017**, *18*, 611–623.

29. Gourley, J.J.; Hong, Y.; Flamig, Z.L.; Arthur, A.; Clark, R.; Calianno, M.; Ruin, I.; Ortel, T.; Kirstetter, P.E.; Clark, E.; et al. A Unifired Flash Flood Database Aross the United States. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 799–805.

30. Gaume, E.; Bain, V.; Bernardara, P.; Newinger, O.; Barbuc, M.; Bateman, A.; Blaskovicova, L.; Bloschl, G.; Borga, M.; Dumitrescu, A.; et al. A compilation of data on European flash floods. *J. Hydrol.* **2009**, *367*, 70–78.

31. Schroeder, A.; Gourley, J.J.; Hardy, J.; Henderson, J.; Parhi, P.; Rahmani, V.; Reed, K.; Schumacher, R.; Smith, B.; Taraldsen, M. The development of a flash flood severity index. *J. Hydrol.* **2016**, *541*, 523–532.

32. Cheung, W.; Houston, D.; Schubert, J.E.; Basolo, V.; Feldman, D.; Matthew, R.; Sanders, B.F.; Karlin, B.; Goodrich, K.A.; Contreras, S.L.; et al. Integrating Redsident digital sketch maps with expert knowledge to assess spatial kmnowledge of lfood risk: A case study of participatory mapping in Newport Beach, California. *Appl. Geogr.* **2016**, *74*, 56–64.

33. Poser, K.; Dransch, D. Volunteered geographic information for disaster management with application to rapid flood damage estimation. *Geomatica* **2010**, *64*, 89–98.

34. Elmore, K.L.; Lakshmanan, V.; Kaney, B.T.; Farmer, V.; Reevers, H.D.; Rothfusz, L.P. Crowd-sourcing weather reports for research. *Bull. Am. Meteorol. Soc.* **2014**, *95*, 1335–1342.

35. Seo, B.C.; Krajewski, W.F.; Kruger, A.; Domaszczynski, P.; Smith, J.A.; Steiner, M. Radar-rainfall estimation algorithms of Hydro-NEXRAD. *J. Hydroinform.* **2010**, *13*, 277–291.

36. Krajewski, W.F.; Smith, J.A. Radar hydrology: Rainfall estimation. *Adv. Water Resour.* **2002**, *25*, 1387–1394.

37. Smith, J.A.; Baeck, M.L.; Villarini, G.; Welty, C.; Miller, A.J.; Krajewski, W.F. Analysis of a long-term, high-resolution radar rainfall data set for the Baltimore metropolitan area. *Water Resour. Res.* **2012**, *48*, 1–14.

38. Villarini, G.; Smith, J.A.; Baeck, M.L.; Smith, B.K.; Sturdevant-Rees, P. Hydrologic analyses of the 17–18 July 1996 flood in Chicago and the role of urbanization. *J. Hydrol. Eng.* **2011**, *18*, 250–259, doi:10.1061/(ASCE)HE.1943-5584.0000462.

39. Yeung, J.K.; Smith, J.A.; Villarini, G.; Ntelekos, A.A.; Baeck, M.L.; Krajewski, W.F. Analyses of the warm season rainfall climatology of the Northeastern US using regional climate model simulations and radar rainfall fields. *Adv. Water Resour.* **2011**, *34*, 184–204, doi:10.1016/j.advwatres.2010.10.005.

40. Shepherd, J.M. A review of the current investigations of urban-induced rainfall and recommendations for the future. *Earth Interact.* **2005**, *9*, 12.

41. Ryu, Y.H.; Smith, J.A.; Baeck, M.L.; Bou-Zeid, E. The influence of land-surface heterogeneities on heavy convective rainfall in the Baltimore-Washington metropolitan area. *Mon. Weather Rev.* **2016**, *144*, 553–573.

42. Changnon, S. More on the La Porte anomaly: A review. *Bull. Am. Meteorol. Soc.* **1980**, *61*, 702–711.

43. Novak, D.; Colle, B. Observations of multiple sea breeze boundaries during an unseasonably warm day in metropolitan New York City. *Bull. Am. Meteorol. Soc.* **2006**, *87*, 169–174.

44. Changon, S. Rainfall changes in summer caused by St. Louis. *Science* **1979**, *205*, 402–404.

45. Shepherd, J.M.; Pierce, H.; Negri, A.J. Rainfall modification by major urban areas: Observations from spaceborne rain radar on the TRMM satellite. *J. Appl. Meteorol.* **2002**, *41*, 689–701.

46. Niyogi, D.; Pyle, P.; Lei, M.; Arya, S.; Kishtawal, C.; Shepard, J.; Chen, F.; Wolfe, B. Urban modification of thunderstorms: An Observational Storm Climatology and Model Case Study for the Indianapolis Urban Region. *J. Appl. Meteorol. Climatol.* **2011**, *50*, 1129–1144.

47. Wright, D.B.; Smith, J.A.; Villarini, G.; Baeck, M.L. The hydroclimatology of flash flooding in Atlanta. *Water Resour. Res.* **2012**, *48*, 1–14.

48. Wahl, T.; Jain, S.; Bender, J.; Meyers, S.D.; Luther, M.E. Increasing Risk of Compound Flooding fro Storm Surge and Rainfall for Major US Cities. *Nat. Clim. Chang.* **2015**, *5*, 1093–1097.