Is precipitation in northern New England becoming more extreme?

A statistical analysis of extreme rainfall in Massachusetts, New Hampshire and Maine and updated estimates of the 100-year storm.

Ellen M. Douglas* and Chelsea A. Fairbank

Environmental, Earth and Ocean Science, University of Massachusetts Boston, 100 Morrissey Blvd, Boston, MA 02125.

Phone: 617-287-7437; email: ellen.douglas@umb.edu

*corresponding author

Journal of Hydrologic Engineering

Submitted: April 19, 2010

Accepted for publication: July 16, 2010

Abstract

The objective of this study was to investigate the presence of trends in extreme precipitation. (denoted MAXP and defined as the annual maximum daily precipitation depth) time series for coastal northern New England and to assess changes in the magnitude of the so-called "100-year storm". MAXP depths from 48 stations with long, continuous records in Maine, New Hampshire and Massachusetts were analyzed. At those same stations, the number of daily precipitation depths ≥ 2 inches (denoted as GT2in) was also quantified for each year. The seasonally-averaged MAXP was found to be fairly uniform throughout the year, but the frequency of MAXP is highest during August through October, the typical hurricane season in New England. The presence of trends in MAXP and GT2in was evaluated over four time frames (1954-2005, 1954-2008, 1970-2005 and 1970-2008) using two statistical methods (linear regression and the Mann-Kendall trend test) and at two scales ("at-site" and regional). The trend analysis over the time period 1954-2005 indicated that MAXP was amazingly stationary; however, a trend in GT2in was found at some stations. More trends in both MAXP and GT2in were present in the time period 1954-2008. The majority of stations in southern NH and eastern MA showed evidence of trends in MAXP (but not GT2in) for the time period 1970-2008. The fact that the number of trends in MAXP increased despite the shorter record length suggests a strong increase in the magnitude of extreme precipitation in northern coastal New England in the last few decades. The stationarity of the 1954-2005 record was confirmed by the regional trend analysis as was the presence of stronger trends in coastal stations when the record was extended through 2008. Most stations that had trends in MAXP also had trends in GT2in.

The Generalized Extreme Value (GEV) distribution was used to estimate 100-yr precipitation depth quantiles for the 1954-2005 record which were then compared to TP-40 100-year, 24-hour precipitation depths. Estimates for stations along coastal MA, NH and ME all exceeded 7 inches and exceeded TP-40 by one inch or more. Stations in northeastern MA,

southeastern NH and southern ME exceeded 8 inches and also exceeded TP-40 estimates by more than 2 inches. These findings indicate that TP-40 under-represents coastal storm depths. This study as well as recent record-breaking events in northern New England strongly suggests the need for updating of design storm estimates. Furthermore, extreme precipitation events of longer than 1-day duration have caused large-scale flooding in the region over the last decade. The magnitude of longer duration storms (particularly 2-day storms) may also be increasing, calling for engineered infrastructure that can accommodate increases in both storm magnitude and duration.

Introduction

Changes in the risk and magnitude of extreme rainfall events rank among the most studied impacts, and indicators (symptoms) of climatic variations (Schnur, 2002). Both natural and human systems are generally able to cope with the historical range of climatic variability in their location. For instance, aquatic and riparian ecosystems are adapted to, and often depend on, the seasonal cycle of high and low flows in a river system. And, a snow storm in New England or the northern Midwest U.S. will cause only temporary inconveniences (delays of one day or less), whereas the same storm in the southern U.S can lead to major disruptions in transportation and public services and large losses in economic productivity. Hence, climate change is most likely to have its largest impact on both natural and human systems through changes in the characteristics (magnitude, frequency, duration) of extreme events (USCCSP, 2008). Extreme rainfall often translates into extreme flooding and consequently great material losses, collapse of lifeline infrastructure, and the breakdown of public health services among others. Across the contiguous U. S., the southeastern and Gulf coast states have been the most prone to weather related disasters, having suffered more than 15 “billion dollar” climate and weather related disasters between 1980 and 2005. Hurricane Katrina was the costliest, wreaking more than $100 billion (normalized to 2002 dollars) in damages (Lott and Ross, 2006). During this same time period, New England states suffered relatively few disasters (only 4 to 6). However, from October 2005 through April 2007, coastal northern New England (eastern Massachusetts, southern New Hampshire and southern Maine) experienced three extreme weather related events. On October 8 and 9, 2005, southwestern New Hampshire experienced damaging flooding as a result of a storm that produced over 7 inches (180 mm) of rain in a 30-hour period. The heavy, intense rainfall resulted in runoff and severe flooding, especially in regions of steep topography that are vulnerable to flash flooding (Olson, 2006). Five counties were later declared a federal disaster area. The second event (a.k.a., the Mother’s Day storm) occurred May 12 through 16, 2006, resulting from a strong low pressure centered over the Great Lakes region at 500mb and a nearly-stationary occluded surface front situated off of the southern New England coast. During this record-breaking event, rainfall totals exceeded 13.8 inches (350 mm) over the five-day period in a relatively concentrated area along the northern Massachusetts, New Hampshire, and southern Maine coastlines (Brown and Keim, 2006). New Hampshire State Climatologist David Brown stated that the storm was truly of historic proportions, a "better than a one-hundred year" storm in some areas along the Seacoast. Within 72 hours, some communities received anywhere between 8 and 14 inches (200 and 360 mm) of rain, the majority of which fell within the first 24 hours (http://www.fosters.com/apps/pbcs.dll/article?AID=/20061230/FOSTERS0104/112190227/-1/fosters0104). Within 4 months of this event, state and federal assistance to Massachusetts alone exceeded $70 million. Parts of Massachusetts and New Hampshire were declared a federal disaster area. New England experienced the third extreme storm event April 15 through 17, 2007. Known as the Patriot’s Day storm, it was one of the largest springtime storms to hit New England in memory (http://www.fema.gov/about/regions/regioni/patriotsdaynoreaster.shtm). The storm dropped over 6 to 8 inches (160 to 200 mm) of rainfall in southern Maine and New Hampshire (http://www.erh.noaa.gov/gyx/patriots_day_storm_2007.htm) and heavy snowfall to northern areas and caused flooding of many rivers throughout the region. The storm also packed hurricane force winds which caused storm surge and flooding in coastal areas. Once again, parts of central New Hampshire and southern Maine were declared federal disaster areas. The occurrence of these events over such a short time period begs the question: was New England simply overdue or are these events an indication of what is to come?

Climate variability and change in the Northeastern U.S.

It is now incontrovertible that the climate is changing across North America. Precipitation events, heat waves and droughts have been occurring more frequently and with much greater intensity in the last few decades than has been seen in the past (USCCSP, 2008). The Northeast Climate Impacts Assessment (NECIA; Frumhoff et al., 2007) found a similar regional pattern of change in climatic indicators. They report that temperatures across the Northeastern U.S. have been rising at a rate of about 0.5 °F per decade since 1970 and that winter temperatures have risen at a much faster rate of 1.3 °F per decade over the same time period. They noted other observations consistent with this warming, including a higher percentage of winter precipitation falling as rain, smaller snowpack extent, and earlier ice-out and snow melt. Hodgkins et al. (2003) reported that the winter/spring center of volume (WSCV) dates have become significantly earlier at all 11 river gaging stations analyzed in areas of New England where snowmelt runoff has the most effect on spring river flows. Most of this change has occurred in the last 30 years with dates advancing by 1–2 weeks. The center of volume date is the date by which half of the total volume of water for a given period of time flows past a river gaging station, and is a measure of the timing of the bulk of flow within the time period. They found that higher January precipitation is related to earlier WSCV dates and concluded that changes in the WSCV dates over the last 30 years are consistent with observed changes in New England last-frost dates, lilac bloom dates, lake ice-out dates, and spring air temperatures. Collins (2009) found that 25 of 28 long-term annual flood series for New England watersheds with dominantly natural streamflow showed upward trends, with 10 of these trends being statistically significant. These observed trends are superimposed on the natural, decadal-scale variability due to hemispheric-scale phenomena such as the North Atlantic Oscillation, the Pacific Decadal oscillation and regional sea surface temperatures (Bradbury et al., 2002; 2003). Most recently, Anderson et al. (2010) used a regional model to assess the affects of anthropogenically-induced climate change on summertime hydroclimatology in the Northeast U.S. They report that the combination of decreased rainfall and increased evaporation will lead to reduced soil moisture over much of the Northeast and increased humidity and temperatures will yield 350-400% increase in the number of days with heat stress at the level of “extreme caution”.

Estimating the risk of extreme hydrologic events

TP-40 (NWS, 1961) estimates of extreme precipitation for the eastern U.S. were based on a statistical analysis of available hourly and daily precipitation observations through 1958. Annual maximum precipitation (MAXP) values from each year were selected and the time series of MAXP were fit to a Fisher-Tippet type I statistical distribution, now known as the Extreme Value (EV) type I or Gumbel distribution (Stedinger et al., 1993). The mean value of each annual maximum (AM) time series was compared to the mean value of the corresponding partial-duration (PD) series. While the AM time series includes only the largest value from each year, the PD series includes all values in a year that exceed a specified threshold. Because the PD included all the highest values, regardless of year, it accounts for the fact that the second highest value in a year sometimes exceeds the maximum value in other years. The PD mean values exceeded the AM means by approximately 11%; the relationship between the AM and PD means was used to adjust the results of the AM frequency analysis.

A number of recent papers highlight the linkages between climate processes and the frequency of hydroclimatic events. Barros and Evans (1997) examined the climatic coincidence of super floods (floods with calculated return periods > 100 years for basins of at least 4000 km^2) on the Upper Mississippi River and found that super floods are not only coincident with extreme rainfall, but also with the negative phase of the Southern Oscillation Index (SOI). Keim et al. (2003) reported that flood estimates doubled between El Nino and La Nina years. Sankarasubramanian and Lall (2003) developed a method for estimating floods conditioned on large-scale climate indices and found dramatic departures from standard “unconditional” flood estimates. Yiou and Nogaj (2004) found that heavy precipitation in the northeastern U.S. is related to a positive North Atlantic Oscillation (NAO) index. Durkee et al. (2008) note that the frequency of extreme events over the eastern U.S. is related to the strength and phase of the NAO. A few recent studies have focused on the interactions of climate and hydrologic processes in New England (Hodgkins et al., 2005; Keim et al., 2005; Bradbury et al., 2002, 2003; Moore et al., 1997; Hartley and Dingman, 1993); however, only one (Wilks and Cember, 1993) has sought to update the TP-40 design storm estimates. They used both AM and PD methods for estimating precipitation extremes and there estimates were substantially higher than the TP-40 estimates. This study seeks to assess what a TP40-like analysis would yield today. The objectives of this study were to 1) investigate the presence of trends in both the magnitude and frequency of extreme precipitation events in coastal northern New England and 2) to assess changes in the so-called “100-year storm” estimates using the available historical record.

Methodology

Data source

Daily precipitation data collected at rain gages throughout coastal northern New England, Massachusetts (MA), New Hampshire (NH) and Maine (ME), were downloaded from the National Climatic Data Center (http://www.ncdc.noaa.gov/oa/climate/climatedata.html), a climate data archiving and retrieval system operated by the National Oceanographic and Atmospheric Administration (NOAA) Satellite and Information Service. Initially, 57 stations that had a record length of at least 50 years were selected: 24 in MA, 19 in NH, and 14 in ME. However, due to a large number of missing data at some of these stations, the number of stations was reduced to 48 stations that had relatively continuous daily precipitation depths over the time period of 1954 through 2008. Table 1 lists the 48 stations, COOP IDs, geographic coordinates, elevation and record length. The average and median record lengths (subtracting out missing years) was 87.4 and 81.0 years, respectively with a standard deviation of 22.5 years. About a quarter of these stations had record lengths greater than 100 years (MA: Amherst, Blue Hills, Haverhill, Lawrence, Middleboro, Plymouth-Kingston, Taunton; NH: Durham, Hanover, Plymouth, Nashua; ME: Farmington, Gardiner, Lewiston, Madison). Station elevations ranged broadly from 18 ft (5.5 m; Newburyport, MA) to 6,262 ft (1908.7 m; Mt. Washington, NH) with average and median elevations of 617.5 ft (188.2 m) and 348.1 ft (106.1 m), respectively. An initial data analyses used the time frame 1954 through 2005 and later precipitation data from 2006 through 2008 was added and the results compared. In this way, the influence of the recent extreme events in 2006 and 2007 on our results could be assessed.

Two stations were moved during the time frame of this data analysis: Portsmouth, NH (276980) was moved to Greenland, NH (276273) in 1973; Northbridge, MA (195514) was moved to Northbridge 2 (195524) in 1964. The Portsmouth-Greenland (Ports-Grnld) timeseries were combined and the Northbridge-Northbridge 2 (Northbridge) timeseries were combined. A double mass curve analysis was performed on the annual rainfall depths for each combined dataset to estimate changes that may have occurred in the precipitation characteristics due to the move. The Northbridge annual rainfall depths were accumulated from 1927 through 2005 and plotted against accumulated rainfall depths at Franklin, MA, the closest station. The double mass curve showed a small reduction in the slope after 1964 (linear regression slope of 1.05 before and 0.94 after move). Prior to the move (1927 to 1964), the annual Northbridge rainfall depth was, on average, 5 % higher than Franklin; after the move (1965 – 2006), the Northbridge rainfall depths was, on average, 5% lower than Franklin, so it appears that rainfall depths are 10% lower at the new location. The standard deviation of the differences between Northbridge and Franklin were similar (9% before and 8% after the move). A similar reduction in rainfall depths was observed in the Ports-Grnld time series. The Ports-Grnld annual depths were accumulated and plotted against accumulated annual depths for Durham, NH, the closest station. The double mass curve analysis of the combined Portsmouth/Greenland stations indicated a larger reduction in slope after 1973 (linear regression slope of 0.98 before and 0.87 after move). Prior to the move (1955 – 1973), the Portsmouth annual rainfall depths were, on average, 1.4% lower than at Durham. After the move (1974 – 2005), the Greenland annual rainfall depths were, on average 11.5%, lower than Durham annual rainfall depths, suggesting that the annual rainfall depths at the Greenland location were approximately 10% lower than at Portsmouth. The standard deviation of the differences between Ports-Grnld and Durham were also similar (9% before and 10% after the move). Although the changes in observed rainfall at the new locations are substantial (approximately 10% lower in both cases), the combined stations were retained for this analysis because the lower rainfall catches after the moves would make the analysis of trends and quantile estimates conservative. Interestingly, in both cases, the break in slope in the double mass curves occurred several years after the location change, hence it is possible that these observed reductions in rainfall were due to factors other than the changes in location.

Data Validation

In order to evaluate changes in extreme precipitation events, continuous and complete time series of MAXP for each station were required. In many cases, entire months were missing from the record at a station, therefore a simple unsupervised selection of annual maxima from the data records would not be sufficient to ensure that the selected maximum value was actually the annual maximum value for a given year. To validate the selection of annual maxima, the month and Julian day in which the MAXP occurred each year was tabulated and the total number of times an annual maximum occurred in each month during 1954 through 2005 was counted. Figure 1 shows the distribution of annual maximum by month categorized by state (columns) and for the combination of all stations (line). If these events were uniformly distributed, then approximately 8.3% of the events would occur within each month. It is clear from Figure 1 that this is not the case. The months of December through July have combined percentages less than 8.3%, while August through November have percentages higher than 8.3%. Approximately 50% of extreme rainfall events occurred during the peak hurricane season (August through November) and only 14% of extreme precipitation events occurred during January through March. Sixty-six percent of the extreme events occurred in the six month period of June through November. The average value of the MAXP depths at each station was also computed. This statistical information, as well as the timing of annual maxima at stations with complete data, was used to help us validate the annual maxima in years that had mising data. Table 2 summarizes the validation criteria (from highest to lowest priority) used for validation of MAXP depths for a station with missing months. The validation approach was as follows. Typically, the selected annual maximum values for years that had 3 or more months of missing data were deleted unless the selected maximum value was greater than the station average maximum annual precipitation depth or unless the annual maximum at one or more surrounding stations with complete data occurred at the same time. If the selected maximum in a year with missing months occurred in June through August, this was likely to be the true annual maximum and the value was retained. If it occurred in any other month, the value was discarded. The retained annual maxima were compared to the annual maxima at nearby stations. After the validation procedure was completed, there were 20 missing yearly values between 1954 and 2005, representing less than 1% of all data (20 missing values divided by 2476 observed values = 0.008). These missing values were filled with the station average values.

Trend Analysis

The presence of trends in the MAXP was evaluated at individual stations ("at-site" analysis) and for groups of stations based on common statistical variability (regional analysis).

These analyses were performed over four different time periods: 1954-2005, 1954-2008, 1970-2005 and 1970-2008. As previously noted, the two different end years (2005 and 2008) were selected in order to assess the impact of recent extreme events on observed trends because high values at the end of the time series can have a large influence on the linear slope. Two different begin times (1954 and 1970) were used to assess changes in precipitation characteristics over the time period when other studies in New England have reported an acceleration due to climate change (Frumhoff et al., 2007; Collins, 2009). Two different methods for identifying trends were used: the test of slope of an ordinary least squares regression line (a parametric statistical method) and the Mann-Kendall trend test, a non-parametric, rank-based method. For the Mann-Kendall trend test, the data are ranked according to time and then each data point is successively treated as a reference data point and is compared to all data points that follow in time. The test statistic, Kendall's S, (Kendall, 1962) is calculated as

$$S = \sum_{i-1}^{n-1} \sum_{j=i+1}^{n} sign(x_i - x_j) \tag{1}$$

where x is the data point (in this case, MAXP depth) at times i and j and *sign( )* is equal to +1 if x_i is greater than x_j and -1 if x_i is less than x_j. For independent, identically distributed random variables, $E(S) = 0$ and

$$Var(S) = \frac{n(n-1)(2n+5)}{18} = \sigma^2 \tag{2}$$

For more details on either test of slope method, see Helsel and Hirsch (2002). These methods were selected because they have been used extensively in the literature and in practice, making the results of this study comparable to other studies.

From the viewpoint of stochastic hydrology, the time series generated at each individual station can be viewed as one realization of an infinite number of possible realizations generated by a larger regional process, in this case, the process that generates extreme precipitation.

Because of this, variations in the statistics computed at individual stations within a region can be viewed as sampling variability around a regional mean value. This, coupled with the fact that spatial correlation between stations inflates the variance of the test statistics, makes it difficult to draw conclusions from the results of trend tests at individual stations. Hence, stations were grouped into "regions" so that regional trend methods could be applied. Five factor groups (regions) were delineated using a factor analysis of the MAXP time series. Stations were assigned a group (region) based on which of the five factor loadings had the largest value. Four of the factor groups were spatially coherent (see Figure 2), although the physical reasons for these factors are not yet fully understood. The 20 stations in Factors 1 and 4 are near to the coast, and the 10 stations in Factor 3 are within or near the Connecticut River Valley region. The 11 stations in Factor 2 were distributed at higher elevations across northern New Hampshire and central Maine. The 4 stations in Factor 5 did not appear to be spatially related, however, three of them (First CT Lake,NH; Caribou and Presque Isle, ME) are located in the extreme northern parts of the study area.

For the "regional" linear regression test of slope, a regional average time series was computed by averaging the MAXP depths for each year within a region. For example, the regional average MAXP for 1954 in Factor Group 1 (FG1) was computed as the average of 1954 MAXP values observed at Concord, Durham, Ports-Grnld, Newburyport, Ipswich, Middleton, Haverhill, Nashua, Milford and Lawrence. The regional average MAXP timeseries for FG1 consisted of the yearly averages of these stations. This assumes that the MAXP observed in a given year at each station within a factor group represents the sampling variability around the regional mean MAXP for that year and the regional average is an estimate of that regional mean value. A linear regression analysis on the regional average timeseries was then performed. For the non-parametric test, the regional trend test developed by Douglas et al. (2000) was used, which is a modification of the Mann-Kendall test that incorporates the influence of spatial correlation. Kendall's S (1) is computed at each station and a regional average Kendall's $\bar{S}_m$ is computed as

$$\bar{S}_m = \frac{1}{m}\sum_{k=1}^{m} S_k \qquad (3)$$

where S_k is the value at an individual station and m is the number of stations within a region. Douglas et al. (2000) showed that $E(\bar{S}_m) = 0$ and

$$Var(\bar{S}_m) = \frac{\sigma^2}{m}\left[1 + (m-1)\bar{\rho}_{xx}\right] \qquad (4)$$

where $\bar{\rho}_{xx}$ is the average cross-correlation estimate within a region. For a sufficiently large number of stations within a region, the test statistic $Z_m = \bar{S}_m \big/ \sqrt{Var(\bar{S}_m)}$ follows a standard normal distribution. However, the number of stations was small (less than 15) in each region, and therefore Z_m was assumed to follow a *t*-distribution with degrees of freedom m-2. The presence of trends is indicative of a change over the time period of the analysis: a positive slope indicates an upward trend (the magnitude is increasing with time), negative slopes indicate a downward trend (the magnitude is decreasing with time). That being said, Cohn and Lins (2005) argued that the statistical significance computed by trend tests that do not account for long-term persistence is likely to be meaningless, hence the presence of trends is identified based on p-values thresholds and no argument about the statistical significance of the results is made. Two important thresholds were selected: p01 which denotes trends with p-value < 0.01 and p05 which denotes trends with p-values < 0.05. These thresholds have been used in the past to denote statistical significance at 99% and 95%, respectively. In this paper, p-values less than these thresholds indicate the strength of the change in the magnitude or frequency of precipitation, depending on the time series tested as will be explained later. This is appropriate because the presence of trends, regardless of the underlying cause or statistical significance, is meaningful over the timeframe of hydrologic design and planning.

Results and Discussion

Characteristics of extreme precipitation

The average of all MAXP depths for the different time periods used in this analysis are: 1954-2005: 2.74 inches; 1954-2008: 2.77 inches; 1970-2005: 2.76 inches; 1970-2008: 2.79 inches. The standard deviations for the same time periods were 1.22, 1.22, 1.16 and 1.17 inches, respectively. The largest MAXP depth was 12.25 inches, recorded in Belchertown, MA in 1955. In this same year, a maximum MAXP depth of 10.05 inches was recorded at Southbridge, MA, 10.01 inches at Northbridge, MA, 9.66 inches at Franklin, 8.83 inches at Hadwick, 8.07 inches at Blue Hill, MA, 7.88 at Brockton, MA and 7.06 inches at Boston Logan Airport. Hurricane Diane passed by New England on August 19 and 20, 1995 and indeed this is reflected in the record MAXP depths observed at these stations, all of which occurred in August 1955. Hence, Hurricane Diane appears to have produced the record MAXP depths in Massachusetts and in New England for more than 50 years. The record MAXP depths for New Hampshire and Maine occurred in 1996: 11.74 inches in Portland, ME and 10.82 inches at Mt. Washington, NH. These records were set in October 1996, as was the record MAXP depth in Ports-Grnld (8.41 inches). These record rainfall depths were due to an intense nor’easter during October 20-21, 1996 that dropped unprecedented amounts of rainfall in some locations of coastal Maine, Massachusetts and New Hampshire and storm totals just shy of the all time largest storm in New England (Keim, 1998). With respect to the recent extreme events that prompted this analysis, Keene and Lakeport, NH and Amherst and Birch Hill Dam, MA measured their record MAXP (8.64, .5.43, 7.56 and 6.32 inches, respectively) in October 2005, as a result of the storm described at the beginning of this paper. Belchertown, MA recorded 6.28 inches, which would have been a record had it not been for Hurricane Diane in 1955. Despite the widespread damage that resulted from the Mother’s Day storm in May 2006 and the Patriot’s Day storm in April 2007, only

Concord, NH measured a record MAXP on 5.12 inches in May 2006 and none of the stations analyzed measured a record MAXP in 2007. This attests to the fact that extreme floods in New England often result from a culmination of factors, such as successive days of heavy rain (as in May 2006) or heavy rain falling on a ripe snow pack, which caused record flooding in Maine and Massachusetts in April 1987.

As previously noted, four time periods were analyzed, 1954-2005, 1954-2008, 1970-2005 and 1970-2008. The latter two time periods were investigated because trends have been found in other hydrologic time series since 1970. Figure 3 shows boxplots of the stations averages over these time period and compares the distribution of these averages. The post-1970 median values are slightly higher (2.80 inches through 2005 and 2.81 inches through 2008) than the pre-1970 medians (2.75 inches through 2005 and 2.76 inches through 2008). The 1954-2005 maximum station average (4.44 inches) was considered an outlier, while the 1954-2008 maximum stations average (4.38 inches) was not. The 1970-2005 and 1970-2008 maximum station averages (4.61 and 4.52 inches) were both considered outliers. These outliners were from Mt. Washington, which is substantially higher in elevation than any other station analyzed. The boxplots and statistics do not suggest a dramatic difference in average rainfall characteristics between time periods. Although Figure 1 shows that the number of MAXP varies by month, Figure 4 shows that the seasonally-averaged MAXP is relatively uniform throughout the year. Seasons were defined as follows: Winter, December 21 of previous year through March 20 current year (Julian day 357 of the current year through Julian day 80 of the following year); Spring: March 21 through June 20 (Julian day 81 through 172); Summer, June 21 through September 20 (Julian day 173 through 264); and Fall, September 21 through December 20 (Julian day 265 through 356). Maine and New Hampshire seasonal averages were found to be quite similar and 0.3 to 0.5 inches lower than MA seasonal averages in Fall, Winter and Spring, and 0.8 to 1 inches lower in Summer. Except for MA summer average of 3.37 inches, all seasonal averages ranged between 2 and 3 inches. Averaged over the entire region, seasonally MAXP averages ranged from 2.45 inches in Winter to 2.81 inches in Summer.

At-Site Trend Analysis

The number of p05 and p01 slopes for the linear regression and Mann-Kendall (M-K) tests of trend is presented in Table 3. The spatial distribution of the linear test of slope results are shown in Figure 5. Nine and three of the slopes were negative for 1954-2005 and 1970-2005 time periods, respectively. Six and two of the slopes were negative for the 1954-2008 and 1970-2008 time periods, respectively. However, all negative slopes had p-values > 0.05. All other things being equal, one would expect fewer identified trends in the time series beginning at 1970 because of the smaller sample size. However, this was clearly not the case. For the time periods ending in 2005, the number of p01 did not change but p05 slopes increased 4-fold, suggesting evidence of increasing extreme precipitation since 1970. For the time periods ending at 2008, the p01 slopes tripled while the p05 slopes doubled, indicating that the extreme precipitation events in 2006 through 2008 do indeed have a large influence on the magnitude of the slopes. The Mann-Kendall results showed more p01 and p05 trends after 1970 than over the entire time series, but overall the number of p01 and p05 slopes in the Mann-Kendall analysis were much greater than the linear regression analysis. This is most likely due to the fact that the increases were not monotonic over the entire time series, which is an assumption behind the M-K test. An evaluation of selected time series with p01 slopes indicated that the linear regression residuals were reasonably homeoscadastic and normally distributed. Hence, the linear regression results are more meaningful for the "at-site" analysis than the M-K results.

Figure 5 shows that spatially consistent patterns in the p05 and p01 linear regression slopes appear in the different time frames. For 1954-2005 (Figure 5a), one p05 slope was found in western NH (Fitzwilliam) and one p01 slope at a nearby stations in central MA (Birch Hill

Dam). Both stations had their record events occur in 2005. Extending the time series to 2008 (Figure 5b) shows an appearance of p05 trends in coastal MA and central ME but no change in p01 trends. For the time periods beginning in 1970 (Figure 5c and 5d), the number of p05 trends increased quite dramatically in both central MA/western NH and along the coast. For the latter time period, most of the stations in southern NH and central MA show p05 or p01 slopes and there is a cluster of stations with p05 slopes in northeastern MA. Only two of the stations in ME and none of the stations in northern NH showed trends.

At the majority of stations across northern New England, there is little evidence to suggest that annual maximum precipitation has increased over the 52 years between 1954 and 2005. However, a few stations closer to the coast showed trends when the time period is extended to 2008. The story becomes quite different when we narrow our focus to the time period after 1970. As noted previously, the fact that many more slopes are p05 and p01 after 1970, even though the sample size is about 1/3 smaller, suggests a strong increase in the magnitude of extreme precipitation events over the last three decades, especially in the eastern half of MA and the southern half of NH. The magnitude of these trends ranged between 0.3 and 0.6 inches per decade or 0.9 to 2 inches since 1970. Hence, it appears that precipitation is getting more extreme in the eastern MA and southern NH. However, as noted previously, differences in “at-site” trend results may be attributable to sampling variability, therefore a regional analysis was performed to confirm or deny these results.

Regional Trend Analysis

Although the physical reasons for the factor groups are not fully understood, it appears that Factors 1 and 4 are related to proximity to the ocean and Factor 2 is related to elevation and latitude. Stations in Factor 3 are located within or near the Connecticut River valley. Stations in Factor 5 do not appear to be spatially related, however three of the four stations are in the extreme northern portion of the study domain. The slope of the regression line for each of the regionally averaged time series (described in Section 3.3) was computed and the p-value of the slope was quantified. The modified Mann-Kendall trend test (modified M-K, Douglas et al., 2000) was used for the regional trend analysis. Table 4 shows the results of these regional tests of trend along with the number of stations and the average correlation coefficient for each factor. Because Factors 1 and 4 both looked to be related to the proximity to the ocean, stations in these two groups were combined and linear slope and modified M-K tests were performed on the combined group. All slopes had p-values > 0.05 for the time periods beginning in 1954. Both factor groups 1 and 4 showed p05 trends during the time period 1970-2008; the combined group also show p05. These results are consistent with the increase in p05 "at-site" slopes shown for 1970-2008 in figure 5c and 5d. Factor 3 showed the strongest trend, with a p01 linear trend after 1970. Again, this is consistent with Figure 5.

The modified M-K trend tests overall showed no trends during 1954-2005, but Factor groups 1 through 3 showed p05 trends when the record was extended to 2008. There were no p01 trends shown in the individual groups, but when groups 1 and 4 were combined and analyzed as one region, the trend was p01 during both time periods that ended in 2008. This is mostly likely due to the larger number of sites (m) and the slightly reduced spatial correlation in the combined group (see Table 4). While averaging the time series within a region lowered the variance for the resulting time series, the modified M-K inflates the variance based on the magnitude of the correlation coefficient. Hence for the regional analysis, the results of the modified M-K results are more appropriate. The modified M-K suggests an increase in extreme precipitation for stations in coastal and central MA/NH, similar to the results of the "at site" analysis. The magnitude of the p01 slopes were approximately 0.2 to 0.3 inches per decade or approximately 0.8 to 1.2 inches over the time period.

There are 9 stations in northern New England that have more than 100 years of record: Lewiston and Gardiner, ME; Plymouth, Hanover and Keene, NH; and Taunton, Lawrence, Blue Hills and Amherst, MA. Figure 6 shows a time series plot of these stations along with the average time series. To ensure that the aforementioned results were not related to sample size, a linear test of slope was performed for 1893-2005 on seven of the long-term stations; Blue Hills and Keene time series were not included because they had more than 10% of the daily data missing. Interestingly, none of the individual times series showed a p01 slope, and only Lewiston and Gardiner showed a p05 slope. The p-value for the aveage time series was greater than 0.05. This analysis supports the assertion that annual maximum precipitation in northern New England was relatively stationary up through 2005 (at least over the last 100 years), and accentuates the fact that trends observed since 1970 are indeed unusual and may be indicative of change.

Changes in the frequency of extreme precipitation events.

From the previous section, it appears that the magnitude of extreme precipitation (as measured by the annual maximum daily precipitation) is increasing in southern NH and eastern MA since 1970. To assess whether or not the frequency of extreme events was increasing as well, at-site and regional trend tests were performed on the number of daily precipitation events that equaled or exceeded 2 inches (denoted GT2in) over the same four time frames (1954-2005, 1970-2005, 1954-2008 and 1970-2008). To do this, annual time series of GT2in for each station were created and "at-site" test of linear slope and Mann-Kendall's test of trend were performed. For the regional analysis, only the regional modified M-K test of trend was perforemd. Results are shown in Table 6. Similar to the MAXP trend results, there are more p05 and p01 in the GT2in records through 2008 than in the GT2in records through 2005. The number of trends in GT2in was similar to that number of trends in MAXP also. However, there were fewer p05 and p01 trends in GT2in during the period after 1970. For those stations that had p01 or p05 slopes from 1954, many (but not all) of the linear slopes from 1970 were similar to the slope of MAXP trends, but the p-values were > 0.05, probably due to the shorter record length. Hence, the frequency of extreme rainfall events appears to be increasing in northern New England, but the rate of increase does not appear to be as strong as the rate of increase in MAXP. Table 6 lists the stations with p01 and p05 trends in MAXP from 1970 (left side) and p01 and p05 trends in GT2in from 1954 (right side). Stations that have trends in both magnitude and frequency are highlighted; one can see that the majority of stations that have an increase in magnitude also have an increase in the frequency of extreme events. All highlighted stations are located in southern NH and central to eastern MA.

Estimation of Extreme Precipitation Quantiles

The next step in this study was to compute estimates of the so-called “100-year design storm” following a method similar TP-40. This analysis was limited to the time period 1954 through 2005 because this time frame was shown to be reasonably stationary. For large samples, the cumulative distribution function for the maximum values of many probability distributions converges to one of three extreme value distributions (EV type I, II, or III) described by Gumbel (1958). The Generalized Extreme Value (GEV) distribution is a general mathematical form which incorporates Gumbel’s type I, II and III distributions for maxima (Stedinger et al., 1993). The Gumbel (EV type I) was used by the NWS (1961) to estimate the 100-year, 24-hour storm quantile. A three-step process was used to select a probability distribution function (PDF) that would best reproduce the observed time series of MAXP. First, sample L-moments were computed and plotted on theoretical L-moment diagrams to evaluate which PDFs could make good candidates. L-moment diagrams (not shown) indicated that several distributions, namely the Gumbel (GUM), the GEV, the log-Pearson type 3 (LP3), the Generalized Pareto (GPA) and the log normal (LN) distributions were potential candidates. The ln-transformed L-moments appeared to fit better than the raw L-moments. For the next step, L-moment based parameters for each of the candidate PDFs were estimated for each station (following Hosking and Wallis, 1997) and then modeled maximum MAXP values were compare to the observed MAXP values at each station. Using the Weibul plotting position formula (p_e = m/N+1 where p_e is the exceedance probability, m is the rank from largest to smallest and N is the sample size), the maximum value in a time series of 52 years (1954-2005) would have a $p_e = 0.019$ and a cumulative probability, $p = 1 – p_e = 0.981$. The quantile for p = 0.981 was estimated using each of the fitted pdfs at each station, and then plotted against the observed maximum values, as shown in Figure 7. Two things became clear. First, while the GEV generated the closest values, all quantiles based on PDFs fit to raw data (GEV, GUM, GPA) underestimated maximum values greater than about 7 inches (see Fig. 7a). In fact, the GUM (used by NWS, 1961) diverged after about 5 inches (the consequences of this will become clear later). Second, the PDFs fit to ln-transformed data (lnGUM, lnGEV, lnGPA, LP3) visually appeared to generate better results across the range of observed annual maxima. A comparison of model error statistics (Table 7) shows that this is true for the GUM but not for the GEV. The LP3 results were similar to lnGEV and lnGUM.

The third step in the selection process was to follow the method presented in Vogel et al. (1993) for assessing the best PDF. For this method, the quantiles associated with average return periods (T_r) of 100 and 1000 years were computed from each distribution of interest and then the number of observed MAXP that equaled or exceeded these quantiles were counted. The expected probability adjustment formulas presented in Vogel et al. (1993) were used, which account for the fact that for finite record lengths, the T_r-year quantile will not be equaled or exceeded with a probability equal to $1/T_r$.

$$p(T_r = 100) = 0.01\left(1 + {}^{26}\!/_{N^{1.16}}\right) \quad \text{and} \quad p(T_r = 1000) = 0.001\left(1 + {}^{280}\!/_{N^{1.55}}\right) \qquad (5)$$

For a record length of 52 years, this resulted in $p(T_r=100)$ = 0.0127 and $p(T_r=1000)$ = 0.0016. Again, following Vogel et al. (1993), a 95% confidence interval for the expected number of exceedances, *C*, was estimated assuming that *C* follows a binomial distribution with mean $E[C] = np$ and variance $Var[C] = np(1 - p)$, where *n* is the number of independent trials and *p* is the exceedance probability associated with each event (p = $1/T_r$). In this case, *n* = 48 stations times 52 years = 2496. Hence for $T_r=100$, $E(C)$ = 25 and $Var(C)$ = 24.7. The 95% confidence interval was computed as $E(C) \pm 1.96\sqrt{Var(C)}$ = [15, 35]. For $T_r=1000$, $E(C)$ = 2.5 and $Var(C)$ = 2.49 and the 95% confidence interval was computed as [-1, 6]. Results were rounded to the nearest integer. To account for the spatial correlation within this dataset would have inflated the $Var(C)$ and increased the width of the confidence intervals to the point that discrimination between distributions would not have been possible. Therefore, the assumption of independence was retained for this analysis. Table 8 presents the results of this analysis for the following distributions: Gumbel (GUM), Generalized Extreme Value (GEV), Generalized Pareto (GPA), lnGUM, lnGEV and the log-Pearson type III (LP3). The distributions that generated values of C within the theoretically expected range were GEV, lnGEV and LP3. Interestingly, although the lnGUM gave the best error statistics, the expected counts for the 100-year exceedences were slightly lower than the theoretical range. Both the GEV and LP3 distributions have been shown to be a useful for predicting extreme events (ie., Castellarin et al., 2007; Thompson et al., 2007; Douglas and Vogel, 2006; Hosking and Wallis, 1997; Vogel et al., 1993; USGS, 1981). Because the purpose of this study was to follow a method similar to TP-40 (NWS, 1961), the GEV was chosen because it is in the same family of PDFs as the Gumbel (Stedinger et al, 1993) used in TP-40. The difference between GEV and Gumbel is that Gumbel assumes that $\kappa = 0$, whereas the GEV does not.

Table 9 presents the 100-year precipitation depth quantiles predicted by the GEV distribution. Also, included are GEV estimates that were adjusted upward using the 11% difference between the PD and AM time series used in NWS (1961). Both series of estimates are compared to the TP-40 estimates for each station, which were obtained by interpolation of the 100-year, 24-hour precipitation depth contour map presented in NWS (1961). Estimates higher than TP-40 are highlighted in yellow. There are fewer unadjusted GEV estimates that exceed the TP-40, but those that do are predominantly stations along the coast, indicating that coastal storms may have been were underrepresented in the TP-40 analysis. This affect can be seen more clearly in Figure 8. Filled circles show the magnitude of the adjusted GEV 100-yr storm estimates and the open symbols show the differences between the GEV and TP-40 estimates. Larger filled circles (representing GEV estimates of 7 to 8 inches and > 8 inches, respectively) are located mostly at coastal stations. Large open squares indicate estimates that are more than 2 inches greater than TP-40 and large open circles indicate differences between 1 and 2 inches greater than TP-40. These differences are located mostly at coastal stations in ME, NH and northeastern MA, with a few also in central MA. The stations in southeastern MA exceed the TP-40 estimates by less than 1 inch. Our 100-yr storm depths near the coast are consistent with estimates by Wilks and Cember (1993). For instance, in Boston, the GEV adjusted and the Wilks and Cember 100-year storm estimates are both 8 inches, which exceeds the TP-40 estimate of 6.75 inches. Part of the difference between our estimates and TP-40 is due to the fact that we estimated TP-40 depths from the published contour map. But the biggest contributor to the differences between the adjusted GEV and TP-40 noted here are most likely due to the fact more stations and longer time series are available now than at the time of the TP-40 analysis.

So where do we go from here?

Since the beginning of the analysis reported in this paper, northern New England has continued to be deluged by extreme, sometimes record breaking, events. July 2008 was the wettest July on record in southeastern New Hampshire; July 2009 was among the wettest months and was reported to be the cloudiest month on record at the Blue Hills Observatory (BHO) in Milton, MA; December 2008 brought an ice storm that devastated a broad area of eastern MA, southern NH and southern ME and disrupted electric power to more than 700,000 customers in the region (some were out of power for more than two weeks). Finally, March 2010 brought three intense and persistent rainstorms within the same number of weeks, making it not only the wettest March on record, but at some locations, the wettest month on record! The BHO reported a March 2010 total rainfall of 18.81 inches, which exceeded the March average by over 14 inches and the previous monthly record depth (August 1955) by 0.03 inches. All of this rain came on the heels of an intense nor'easter at the end of February 2010 that packed hurricane-force winds and heavy rains and left more than 300,000 customers without power in NH alone. This sequence of events resulted in a seemingly endless string of floods in coastal rivers, some of which exceeded record flood stage by several feet! Clearly, the upward trends in extreme precipitation reported in this paper have real and destructive consequences to property, communities and infrastructure. To address these consequences, we must update the design storm estimates that are currently being used for engineering design. This has been understood for quite a while and indeed, the National Oceanographic and Atmospheric Administration (NOAA) is working towards that end with NOAA Atlas 14 (Bonnin et al., 2006a, 2006b). Unfortunately, the northeast US (including New England) has not yet funded the Atlas 14 study (Geoffrey Bonnin, NOAA, personal communication, April 18, 2010).

We end this paper with a call to begin thinking about design methods that go beyond the analysis of 1-day annual maximum events as illustrated in this study. Many, if not most, of the damaging events over the last decade or so have been due to storm events with durations of two or more days. And it appears that the magnitude (and perhaps the frequency) of such longer duration events is increasing as well. Figure 9 shows the time series of 2-day and 3-day annual maximum precipitation depths overlain on the 1-day annual maximum analyzed in this study for Ports-Grnld NH (near the lead author's home town). Some interesting observations are apparent in this figure. First, the record event in October 1996 was quite extraordinary; the 1-day total exceeded most of the annual maximum 2- and 3-day totals over the period of record. Second, for the Mother's Day storm (May 2006), which until 2010 was the largest storm of this decade, the 1-day precipitation depth was large but not terribly unusual. It was the 2nd and 3rd days of consecutive rain that turned this storm into a destructive event for the region. The same is true for the largest of the March 2010 storms (Mar 13-15, 2010, not shown). The third interesting observation of Figure 9 is that the slope of the trend line for the 2-day annual maximum precipitation depth (0.034 inches per year) is almost double that of the 1-day annual maximum depth (0.018 inches per year). Furthermore, the p-value of the 2-day slope is 0.03 (meeting the p05 trend threshold) whereas the 1-day trend p-value is 0.10 (which would not qualify as a trend in this study). Both the slope and p-value of the 2-day trend line are influenced by the magnitude and timing of the 1996 and 2006 storms, but if this trend continues into the future, it suggests that we need to consider expanding our design strategies to accommodate storms of larger magnitude AND longer duration. A detailed analysis of 2-day and 3-day events across the regions needs to be performed.

Summary and Conclusions

The objective of this study was to investigate the presence of trends in extreme precipitation (denoted MAXP and defined as the annual maximum daily precipitation depth) time series for coastal northern New England and to assess changes in the magnitude of the so-called "100-year storm". MAXP depths from 48 stations in Maine, New Hampshire and

Massachusetts were analyzed. At those same stations, the number of daily precipitation depths ≥ 2 inches was also quantified (denoted as GT2in) for each year. The seasonally-averaged MAXP is fairly uniform throughout the year, but the frequency of MAXP is highest during August through November, the typical hurricane season in New England. The presence of trends in MAXP and GT2in was evaluated over four time frames (1954-2005, 1954-2008, 1970-2005 and 1970-2008) using two statistical methods (linear regression and the Mann-Kendall trend test) and at two scales (“at-site” and regional). Five spatially coherent regions were delineated using a factor analysis of the extreme precipitation time series, although the physical reasons for these factors is not yet fully understood. The presence of trends was determined using p-value thresholds based on the results of the trend tests: p01, denoting a p-value < 0.10 and p05 denoting a p-value < 0.05.

The trend analysis over the time period 1954-2005 indicated that MAXP was amazingly stationary; however, the frequency of GT2in did increase at some stations. More trends in both MAXP and GT2in were present in the time period 1954-2008. The majority of stations in southern NH and eastern MA showed evidence of trends in MAXP (but not GT2in) for the time period 1970-2008. The fact that the number of trends increased despite the shorter record length suggests a strong increase in MAXP in northern coastal New England in the last few decades. This is consistent with other hydroclimatic studies in the region. The stationarity of the 1954-2005 record was confirmed by the regional trend analysis as was the presence of stronger trends in coastal stations when the record was extended through 2008. An analysis of stations with records beginning in 1893 confirmed the stationarity of annual maximum precipitation depths through 2005. Most stations that had trends in MAXP also had trends in GT2in.

The Generalized Extreme Value (GEV) distribution was used to estimate 100-yr precipitation depth quantiles for the stationary 1954-2005 record. These quantiles were adjusted upwards by 11% (following NWS, 1961) and then compared to TP-40 100-year, 24-hour precipitation depths. Estimates for stations along coastal MA, NH and ME all exceeded 7 inches and most exceeded TP-40 estimates by one inch. Stations in northeastern MA, southeastern NH and southern ME exceeded 8 inches and exceeded TP-40 estimates by more than 2 inches, indicating that TP-40 under-represented coastal storm depths. This analysis as well as the continuation of record-breaking events in New England strongly suggests the need for updating of design storm estimates. Furthermore, it appears that the magnitude of longer duration storms (particularly 2-day storms) may also be increasing, hence it would be prudent to consider ways that engineered infrastructure can be built to accommodate increases in both storm magnitude and duration.

Acknowledgements

This research was partially supported by an award from the Joseph P. Healey Grant program at the University of Massachusetts, Boston (UMass Boston), awarded to Dr. Douglas in summer 2007. Ms. Fairbank graduated with a BS in Environmental, Earth and Ocean Sciences (EEOS) and with honors in major in May 2008. Part of this research served as her honors thesis. We would like to acknowledge Steven Kichefski, Barry Fradkin and Doina Oglavie, Master's students in EEOS at UMass Boston, who assisted with the regional Kendall's trend tests.

References

Anderson, B. T.,. K. Hayhoe, and X-Z. Liang, 2010. Anthropogenic-induced changes in twenty-first century summertime hydroclimatology of the Northeastern US, Climatic Change, 99 (3-4): 403-423.

Barros, A. P. and J. L. Evans, Designing for climate variability, *Journal of Professional Issues in Engineering Education and Practice*, 123 (2): 62-65, 1997.

Bonnin, G. M., Martin, D., Lin, B., Parzybok, T., Yekta, M., Riley, D., 2006a: NOAA Atlas 14 Precipitation-Frequency Atlas of the United States Volume 1 Version 4.0: Semiarid Southwest

(Arizona, Southeast California, Nevada, New Mexico, Utah), U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, Maryland.

Bonnin, G. M., Martin, D., Lin, B., Parzybok, T., Yekta, M., Riley, D., 2006b: NOAA Atlas 14 Precipitation-Frequency Atlas of the United States Volume 2 Version 3.0: Delaware, District of Columbia, Illinois, Indiana, Kentucky, Maryland, New Jersey, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, Virginia, West Virginia, U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, Maryland.

Bradbury J. A., B. D. Keim, and C. P. Wake. 2003. The influence of regional storm tracking and teleconnections on winter precipitation in the northeastern United States, *Annals of the Assoc. of American Geographers,* 93 (3): 544-556.

Bradbury J. A., B. D. Keim, and C. P. Wake. 2002. U.S. east coast trough indices at 500hPa and New England winter climate variabilitiy, Journal of Climate, 15 (23): 3509-3517.

Brown, D. P. and B. D. Keim, 2006. Record rainfall associated with the 12-16 May 2006 New England storm, American Geophysical Union, Fall Meeting 2006, abstract H21B-1376.

Castellarin, A., R.M. Vogel and N.C. Matalas, Multivariate probabilistic regional envelopes of extreme floods, *Journal of Hydrology*, doi:10.1016/j.jhydrol.2007.01.007, 2007.

Cohn, T. A. and H. F. Lins, 2005. Nature's style: naturally trendy. Geophysical Research Letters, 32 (23): article L23402.

Collins, M.J., Evidence for changing flood risk in New England since the late 20th century, *Journal of the American Water Resources Association*, 45 (2): 279-290, 2009

Douglas, E.M., and R.M. Vogel, The probabilistic behavior of the flood of record in the United States, *Journal of Hydrologic Engineering*, Vol. 11, No. 5, 482-488, 2006.

Douglas, E. M.. R. M. Vogel and C. N. Kroll, 2000. Trends in floods and low flows in the United States: impact of spatial correlation, *Journal of Hydrology*, 240: 90-105.

Durkee, J. D., J. D. Frye, C. M. Fuhrmanny, M. C. Lacke, H. G Jeong and T. L. Mote, 2008. Effects of the North Atlantic Oscillation on precipitation-type frequency and distribution in the eastern United States, Theoretical and Applied Climatology, 94 (1-2): 51-65.

Frumhoff, P.C., J.J. McCarthy, J.M. Melillo, S.C. Moser, and D.J. Wuebbles. Confronting Climate Change in the U.S. Northeast: Science, Impacts, and Solutions. Synthesis report of the Northeast Climate Impacts Assessment (NECIA). Cambridge, MA: Union of Concerned Scientists (UCS) 2007.

Gumbel, E. J. (1958) Statistics of Extremes, Columbia University Press, New York.

Hartley, S. and S. L. Dingman, 1993. Effects of climatic variability on winter-spring runoff in New England river basins, *Physcial Geography*, 14 (4): 379-393.

Helsel, D. R. and R. M. Hirsch, 2002. Statistical methods in water resources, Techniques of Water-Resources Investigations of the United States Geological Survey Book 4, Hydrologic Analysis and Interpretation, Chapter A3. Publication available at: http://water.usgs.gov/pubs/twri/twri4a3/

Hodgkins, G.A., R.W. Dudley, and T.G. Huntington, Changes in the timing of high river flows in New England over the 20th Century, *Journal of Hydrology*, 278(1-4): 244-252, 2003.

Hodgkins, G. A., R. W. Dudley, and T. G. Huntington. 2005. Changes in the number and timing of days of ice-affected flow on northern New England rivers, 1930-2000, *Climatic Change*, 71 (3): 319-340.

Hosking, J. R. M., and J. R. Wallis (1997). Regional frequency analysis: an approach based on L-moments. Cambridge University Press, Cambridge, U.K.

Keim, A. S., S. W. Franks and G. Kuczera, 2003. Multi-decadal variability of flood risk, *Geophysical Research Letters,* 30 (2): 1035, doi: 10.1029/2002GL015992.

Keim B. D., L. D. Meeker LD, and J. F. Slater. 2005. Manual synoptic climate classification for the East Coast of New England (USA) with an application to PM2.5 concentration, *Climate Research*, 28 (2): 143-154.

Keim, B. D., 1998. Record precipitation totals for the coastal New England rainstorm 20-21 October 1996, Bulletin of the American Meteorological Society, 79 (6): 1061-1067.

Kendall, M.G., 1962. Rank Correlation Methods. 3rd ed. Hafner Publishing Company, New York.

Lott, N. and T. Ross, 2006. Tracking and evaluating U.S. billion dollar weather disasters, 1980-2005, NOAA National Climatic Data Center, Asheville, North Carolina, available on-line at http://www1.ncdc.noaa.gov/pub/data/papers/200686ams1.2nlfree.pdf

Moore, MV, M. L. Pace, J. R. Mather, P. S. Murdoch, R. W. Howarth, C. L. Folt, C. Y. Chen, H. F. Hemond, P. A. Flebbe, and C. T. Driscoll CT. 1997. Potential effects of climate change on freshwater ecosystems of the New England/Mid-Atlantic Region, *Hydrological Processes*, 11 (8): 925-947).

National Weather Service (NWS). 1961. Rainfall Frequency Atlas of the United States for Durations from 30 minutes to 24 Hours and Return Periods from 1 to 100 Years, *Technical Paper 40*, National Weather Service.

Olson, S.A., 2006, Flood of October 8 and 9, 2005, on Cold River in Walpole, Langdon, and Alstead and on Warren Brook in Alstead, New Hampshire: U.S. Geological Survey Open-File Report 2006–1221, 46 p., online only.

Sankarasubramanian, A. and U. Lall, 2003. Flood quantiles in a changing climate: Seasonal forecasts and causal relations, Water Resources Research, 39 (5): Art. No. 1134.

Schnur, R. 2002. Climate Science: The investment forecast, Nature, 415 (6871): 483-484.

Stedinger, J. R., R. M. Vogel and E. Foufoula-Georgio, 1993. Chapter 18: Frequency analysis of extreme events, in Handbook of Hydrology, D. R. Maidment, ed., McGraw-Hill.

Thompson, E. M., L. G. Baise, and R. M. Vogel (2007), A global index earthquake approach to probabilistic assessment of extremes, *J. Geophys. Res.*, 112, B06314, doi:10.1029/2006JB004543.

USCCSP, 2008. Weather and climate extremes in a changing climate, U.S. Climate Change Science, Program, Synthesis and Analysis Product 3.3, 161 pages.

USGS, 1981. Guidelines for determining flood flow frequency, Bulletin 17B of the hydrology subcommittee, Interagency Advisory Committee on Water Data, U. S. Geological Survey, Reston, VA.

Vogel, R. M., W. O. Thomas, Jr. and T. A. McMahon, 1993. Flood flow frequency model selection in southwestern United States, *Journal of Water Resources Planning and Management*, 119 (3): 353-366.

Wilks, D.S. and R.P. Cember, 1993, Atlas of Precipitation Extremes for the Northeastern United States and Southeastern Canada, *Northeastern Regional Climate Center, Research Publication RR 93-5*, Cornell University, Ithaca, NY, 40 p

Yiou, P. and M. Nogaj, 2004. Extreme climatic events and weather regimes over the North Atlantic: When and where? Geophysical Research Letters, 31, L07202, doi: 10.1029/2003GL019119.

Figure captions

Figure 1**:** Percentage of annual maximum precipitation events that occurred within each month, categorized by state. Diamond-shaped symbols represent the percentage of all events that occurred within each month.

Figure 2: Location of 48 precipitation stations used in this study. Also shown are the groups (1 through 5) that each station belongs to based on factor analysis.

Figure 3**:** Boxplots of station average annual maximum precipitation depths (MAXP) for the different time periods used in this analysis.

Figure 4**:** Seasonally-averaged annual maximum precipitation depths for Maine, New Hampshire and Massachusetts and for the region overall.

Figure 5: Linear regression slopes with p-value thresholds for the four analysis timeframes: a) 1954-2005, b) 1954-2008, c) 1970-2005 and d) 1970-2008.

Figure 6**:** Annual maximum precipitation for long term stations in Northern New England that had less than 10% missing daily data over the period of record. Larger filled circles represent the annual averages of all long-term stations. Dashed line is the regression line for the long-term average time series.

Figure 7: Comparison of a) modeled and observed maximum precipitation values and b) modeled and observed ln-transformed precipitation values. Dashed line is 1:1.

Figure 8: Spatial distribution of the 100-year precipitation depth (in inches) from the adjusted GEV (GEVadj) shown in filled circles. Open symbols represent the differences between GEVadj and TP-40 (GEV – TP40) estimates.

Figure 9: Annual maximum 1-day, 2-day and 3-day precipitation depths for the Portsmouth-Greenland station in NH. The record event (October 1996) and recent extreme events (May 2006) are highlighted. Also shown are linear trend lines for the 1-day annual maximum (lower) and the 2-day annual maximum time series.

Table 1: Stations used in this analysis.

COOP ID	**Station Name**	**Latitude**	**Longitude**	**Elevation (m)**	**Begin year[1]**	**End year[2]**	**No. missing years**
New Hampshire							
271683	Concord (airport)	43.2000	-71.5000	105.5	1921	2008	0
272174	Durham	43.1500	-70.9500	24.4	1893	2008	2
272842	Errol	44.7833	-71.1333	390.1	1927	2008	1
272999	First CT Lake	45.0833	-71.2833	506.0	1918	2008	6
273024	Fitzwillilam (2 W)	42.7833	-72.1833	353.6	1931	2008	4
273850	Hanover	43.7000	-72.2833	183.8	1885	2008	4
274399	Keene	42.9500	-72.3167	155.4	1893	2008	2
274480	Lakeport (2)	43.5500	-71.4667	152.4	1927	2008	3
275412	Milford	42.8167	-71.6500	91.4	1945	2007	4
275639	Mt Washington	44.2667	-71.3000	1908.7	1948	2008	0
275712	Nashua (2 Nnw)	42.7833	-71.4833	39.6	1886	2008	14
275868	Newport	43.3833	-72.1833	240.8	1929	2008	2
276818	Pinkham Notch	44.2667	-71.2500	612.3	1930	2008	1
276944/5	Plymouth	43.7500	-71.6833	153.0	1887	2008	3
276980/273626	Portsmouth/Greenland[3]	43.0667	-70.7167	18.0	1955	2006	1
Massachusetts							
190120	Amherst	42.3833	-72.5333	45.7	1893	2008	4
190190	Ashburnham	42.6167	-71.8833	335.3	1907	2008	40
190562	Belchertown	42.2833	-72.3500	170.7	1942	2008	1
190666	Birch Hill Dam	42.6333	-72.1167	263.3	1949	2008	2
190736	Blue Hill	42.2167	-71.1167	192.0	1893	2008	0
190770	Boston Logan Airport	42.3667	-71.0167	6.1	1936	2008	0
190860	Brockton	42.0500	-71.0000	24.4	1894	2008	20
192451	East Wareham	41.7667	-70.6667	6.1	1913	2008	14
192997	Franklin	42.0833	-71.4167	73.2	1927	2008	5
193401	Hardwick	42.3500	-72.2000	295.7	1927	2006	5
193505	Haverhill	42.7667	-71.0667	5.5	1900	2008	7
193876	Ipswich	42.6667	-70.8667	24.4	1926	2008	3
194105	Lawrence	42.7000	-71.1667	18.3	1893	2008	2
194711	Middleboro	41.8833	-70.9167	18.3	1893	2008	7
194744	Middleton	42.6000	-71.0167	27.4	1926	2008	2
195285	Newburyport (4 Nnw)	42.8333	-70.9333	5.5	1893	2008	1
195514/24	Northbridge/Northbridge 2[4]	42.1167	-71.6833	96.0	1926	2008	3
196486	Plymouth-Kingston	41.9833	-70.7000	13.7	1893	2008	6
197627	Southbridge	42.0500	-72.0833	219.5	1927	2008	1
198367	Taunton	41.9000	-71.0667	6.1	1893	2007	0

Table 1 (cont'd)

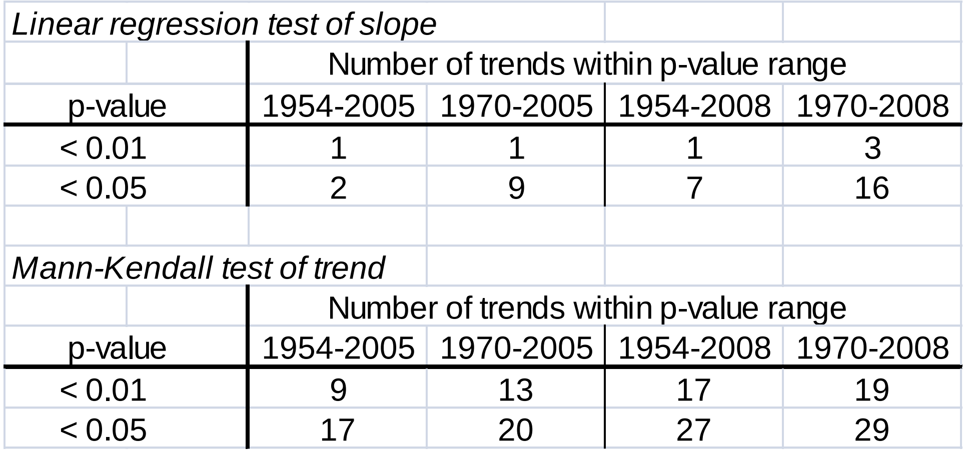
COOP ID	Station Name	Latitude	Longitude	Elevation (m)	Begin year[1]	End year[2]	No. missing years
Maine							
170275	Augusta (State Ap)	44.3167	-69.8000	106.7	1949	2008	1
170480	Belfast	44.4000	-69.0000	9.1	1946	2008	0
170814	Brassua Dam	45.6667	-69.8167	323.1	1930	2008	2
171175	Caribou (Muni Ap)	46.8667	-68.0333	190.2	1939	2008	0
172765	Farmington	44.6833	-70.1500	128.0	1893	2008	1
173046	Gardiner	44.2167	-69.7833	42.7	1887	2008	3
174086	Jackman	45.6333	-70.2667	362.7	1894	2008	27
174566	Lewiston	44.1000	-70.2167	54.9	1893	2008	0
174927	Madison	44.8000	-69.8833	79.2	1894	2008	9
175261	Middle Dam	44.7833	-70.9167	445.0	1927	2008	0
175460	Moosehead	45.5833	-69.7167	313.3	1931	2008	2
176905	Portland (Intl Jetport)	43.6500	-70.3000	13.7	1920	2008	0
176937	Presque Isle	46.6500	-68.0000	182.6	1983	2008	2

Notes:
[1] Begin year starts at the first January 1 in the dataset.
[2] End year for this study was 2008, unless the dataset ended before that.
[3] Portsmouth station was moved to Greenland in 1973.
[4] Northbridge station was moved to Northbridge 2 in 1964.

Table 2: Selection criteria for determining the annual maximum precipitation in stations with missing data.

Factors Listed In Order From Greatest Significance to Least Significance	**1)** Shares Significant Rainfall Event With Other Surrounding Stations.	**2)** Does Not Have Months Missing That Are Months of Maximum Annual Precipitation Events For Surrounding Stations.	**3)** Has Higher Than Average Recorded Maximum Annual Precipitation.	**4)** Has One or Less "More Significant" Months Missing/Has Three or Less "Less Significant" Month Missing.

Table 3: Results of “at-site” test of trend analysis for MAXP. Time series at 48 stations were tested in each case.



Linear regression test of slope

	Number of trends within p-value range			
p-value	1954-2005	1970-2005	1954-2008	1970-2008
< 0.01	1	1	1	3
< 0.05	2	9	7	16

Mann-Kendall test of trend

	Number of trends within p-value range			
p-value	1954-2005	1970-2005	1954-2008	1970-2008
< 0.01	9	13	17	19
< 0.05	17	20	27	29

Table 4: Results of “regional” test of trend analysis for MAXP.

Linear regression test of slope

Factor	m	ρ_{xx}	1954-2005	1970-2005	1954-2008	1970-2008
1	11	0.66	ns	ns	ns	p05
2	14	0.48	ns	ns	ns	ns
3	10	0.52	ns	**p01**	ns	**p01**
4	9	0.48	ns	ns	ns	p05
5	4	0.27	ns	ns	ns	ns
1+4	20	0.46	ns	ns	ns	p05

Modified Mann-Kendall test of trend

Factor	m	ρ_{xx}	1954-2005	1970-2005	1954-2008	1970-2008
1	11	0.66	ns	ns	p05	p05
2	14	0.48	ns	ns	p05	ns
3	10	0.52	ns	p05	p05	p05
4	9	0.48	ns	ns	ns	ns
5	4	0.27	ns	ns	ns	ns
1+4	20	0.46	ns	ns	**p01**	**p01**

ns denotes p-value > 0.05

p05 denotes p-value < 0.05

p01 denotes p-value < 0.01

Table 5: Test of trend analyses for GT2in.

At-site

Linear regression test of slope

	Number of trends within p-value range			
p-value	1954-2005	1970-2005	1954-2008	1970-2008
< 0.01	4	1	5	1
< 0.05	1	1	10	9

Mann-Kendall test of trend

	Number of trends within p-value range			
p-value	1954-2005	1970-2005	1954-2008	1970-2008
< 0.01	4	1	5	0
< 0.05	5	4	16	7

Regional

Modified Mann-Kendall test of trend

Factor	m	ρ_{xx}	1954-2005	1970-2005	1954-2008	1970-2008
1	11	0.66	ns	ns	**p01**	ns
2	14	0.48	ns	ns	ns	ns
3	10	0.52	ns	ns	**p01**	ns
4	9	0.48	ns	ns	ns	ns
5	4	0.27	ns	ns	ns	ns
1+4	20	0.46	ns	ns	ns	ns

ns denotes p-value > 0.05

p05 denotes p-value < 0.05

p01 denotes p-value < 0.01

Table 6: Stations with linear trends in magnitude (MAXP) (left) and frequency (GT2in) (right). P05 denotes p-values < 0.05; p01 denotes p-values < 0.01. Stations that show trends in both magnitude and frequency are highlighted in yellow.

Trends in Magnitude				*Trends in Frequency*			
p01		**p05**		**p01**		**p05**	
1970-2005	**1970-2008**	**1970-2005**	**1970-2008**	**1954-2005**	**1954-2008**	**1954-2005**	**1954-2008**
Birch Hill Dam	Nashua	Ipswich	Haverhill	Belchertown	Belchertown	Birch Hill Dam	Birch Hill Dam
	Lakeport	Newburyport	Ipswich	Franklin	Franklin		Ipswich
	Birch Hill Dam	Lakeport	Middleton	Haverhill	Haverhill		Middleboro
		Belchertown	Newburyport	Fitzwilliam	Concord		Newburyport
		Fitzwilliam	Ashburnham		Fitzwilliam		Plymouth-Kingston
		Hardwick	Belchertown				Taunton
		Northbridge	Fitzwilliam				Keene
		Blue Hill	Hardwick				Lakeport
			Newport				Nashua
			Northbridge				Farmington
			Southbridge				Portland
			Taunton				

p05 denotes p-value < 0.05
p01 denotes p-value < 0.01

Table 7: Comparison of error statistics for selected quantile functions. For ln-transformed functions (lnGUM, lnGEV, LP3), quantiles were transformed to real space before computing error statistics.

	Model error statistics (in inches)		
PDF	**Mean error**	**MAE**	**RMSE**
GUM	-1.47	1.48	2.00
lnGUM	-0.37	0.99	1.40
GEV	-0.90	1.01	1.46
lnGEV	-0.71	0.99	1.45
GPA	-1.13	1.16	1.60
LP3	-0.89	1.07	1.56

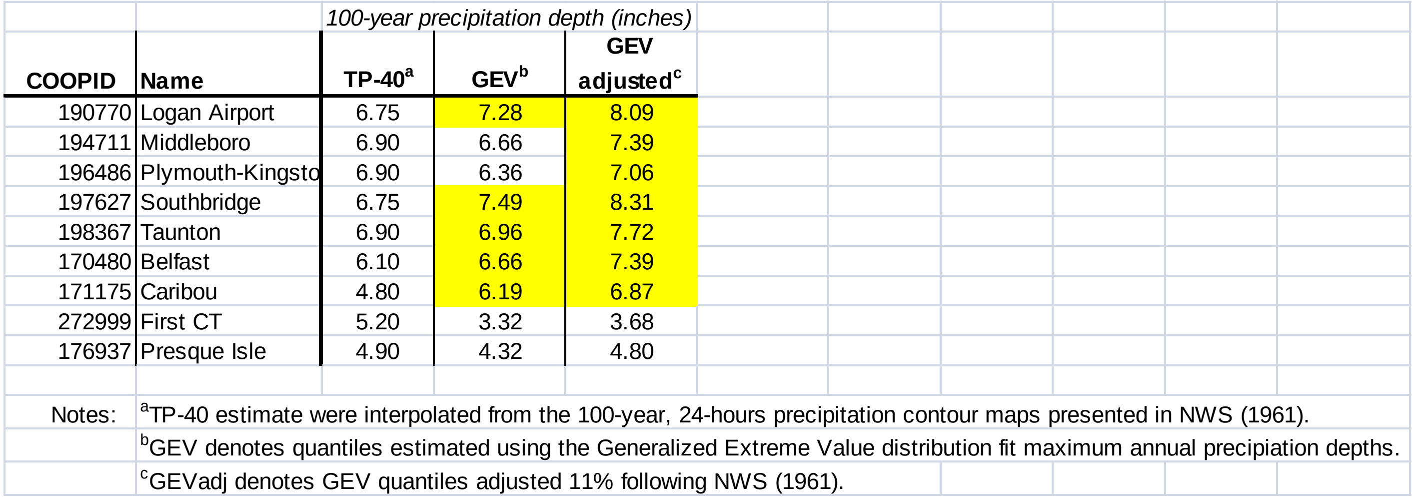
Table 8: Results of theoretical exceedences analysis. Numbers of exceedences (C) that fell within the 95% confidence interval (CI) are italicized.

	T_r (years)	
	100	**1000**
E(C)	**25**	**2.5**
95% CI	**[15, 35]**	**[-1, 6]**
GUM	52	13
GEV	*29*	*1*
GPA	41	10
lnGUM	14	*0*
lnGEV	*21*	*1*
LP3	*28*	*1*

Table 9: Comparison of 100-year precipitation depths computed with GEV with the TP-40 estimates. Adjusted and unadjusted GEV estimates are shown and highlighted if greater than TP-40 estimates.

		100-year precipitation depth (inches)		
COOPID	**Name**	**TP-40[a]**	**GEV[b]**	**GEV adjusted[c]**
271683	Concord	6.20	5.09	5.65
272174	Durham	6.40	6.91	7.68
193505	Haverhill	6.50	8.16	9.06
193876	Ipswich	6.50	6.99	7.76
194105	Lawrence	6.50	7.22	8.02
194744	Middleton	6.50	7.75	8.60
275412	Milford	6.25	5.36	5.95
275712	Nashua	6.40	6.12	6.79
195285	Newburyport	6.50	9.21	10.22
176905	Portland	6.40	10.50	11.65
276980	Portsmouth	6.40	8.12	9.01
170275	Augusta	6.10	5.97	6.63
170814	Brassua Dam	5.25	4.22	4.69
272842	Errol	5.60	4.04	4.49
172765	Farmington	5.80	5.69	6.32
173046	Gardiner	6.10	7.21	8.00
174086	Jackman	5.25	5.35	5.93
274480	Lakeport	6.20	6.01	6.67
174566	Lewiston	6.10	6.82	7.57
174927	Madison	5.80	4.42	4.90
175261	Middle Dam	5.70	4.45	4.94
175460	Moosehead	5.25	4.18	4.64
275639	Mt Washington	7.30	11.43	12.69
276818	Pinkham Notch	7.20	7.06	7.84
276944	Plymouth	6.00	4.62	5.13
190120	Amherst	6.50	5.87	6.52
190190	Ashburnham	6.40	6.07	6.73
190562	Belchertown	6.50	8.40	9.32
190666	Birch Hill Dam	6.40	5.93	6.58
273024	Fitzwilliam	6.25	5.79	6.43
273850	Hanover	5.80	4.67	5.19
193401	Hardwick	6.50	6.48	7.19
274399	Keene	6.20	6.19	6.87
275868	Newport	6.00	6.44	7.14
195514	Northbridge	6.75	8.01	8.89
190736	Blue Hill	6.75	8.36	9.28
190860	Brockton	6.80	7.24	8.03
192451	East Wareham	7.00	6.36	7.06

Table 9: cont'd

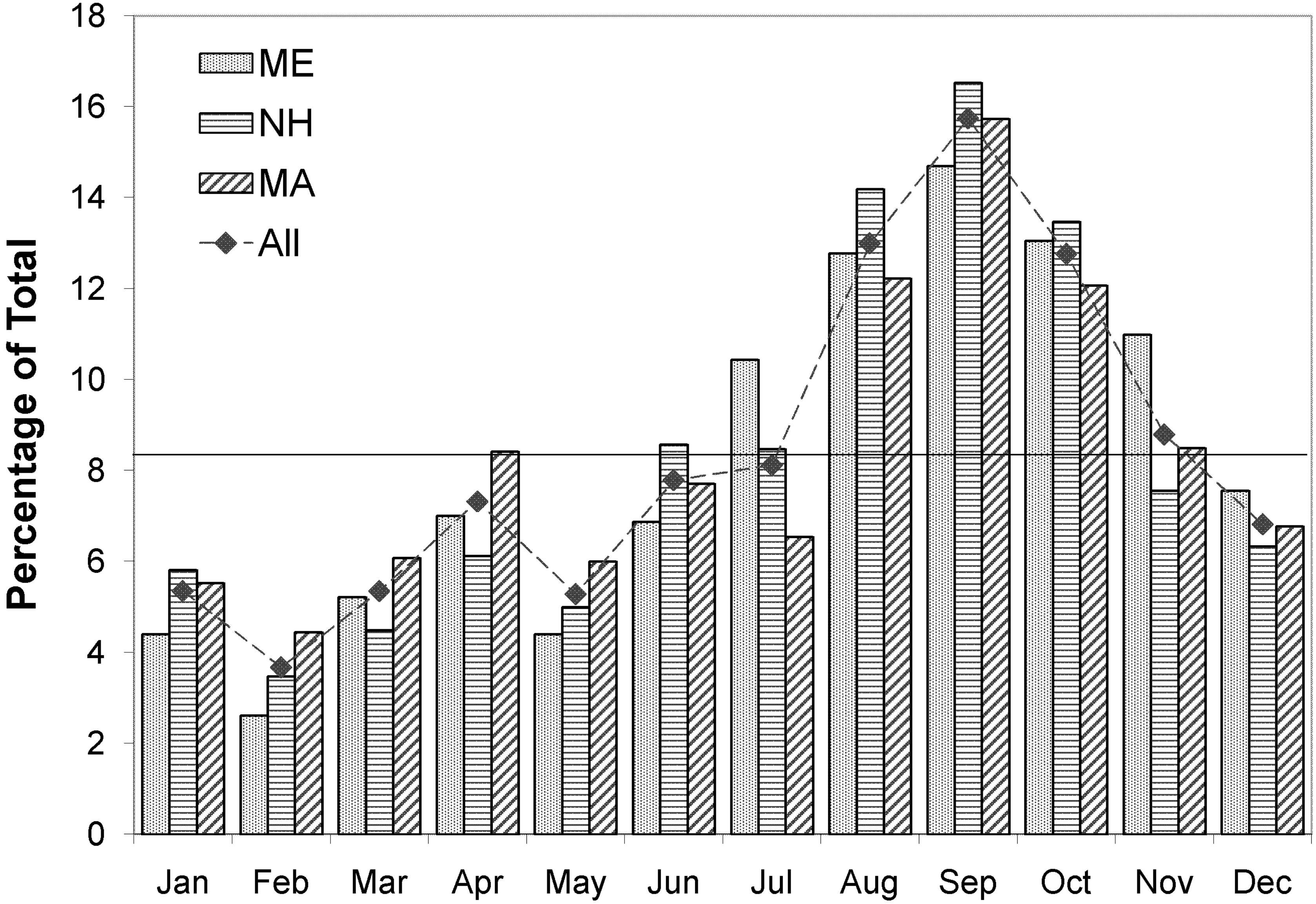


		100-year precipitation depth (inches)		
COOPID	**Name**	**TP-40**[a]	**GEV**[b]	**GEV adjusted**[c]
190770	Logan Airport	6.75	7.28	8.09
194711	Middleboro	6.90	6.66	7.39
196486	Plymouth-Kingsto	6.90	6.36	7.06
197627	Southbridge	6.75	7.49	8.31
198367	Taunton	6.90	6.96	7.72
170480	Belfast	6.10	6.66	7.39
171175	Caribou	4.80	6.19	6.87
272999	First CT	5.20	3.32	3.68
176937	Presque Isle	4.90	4.32	4.80

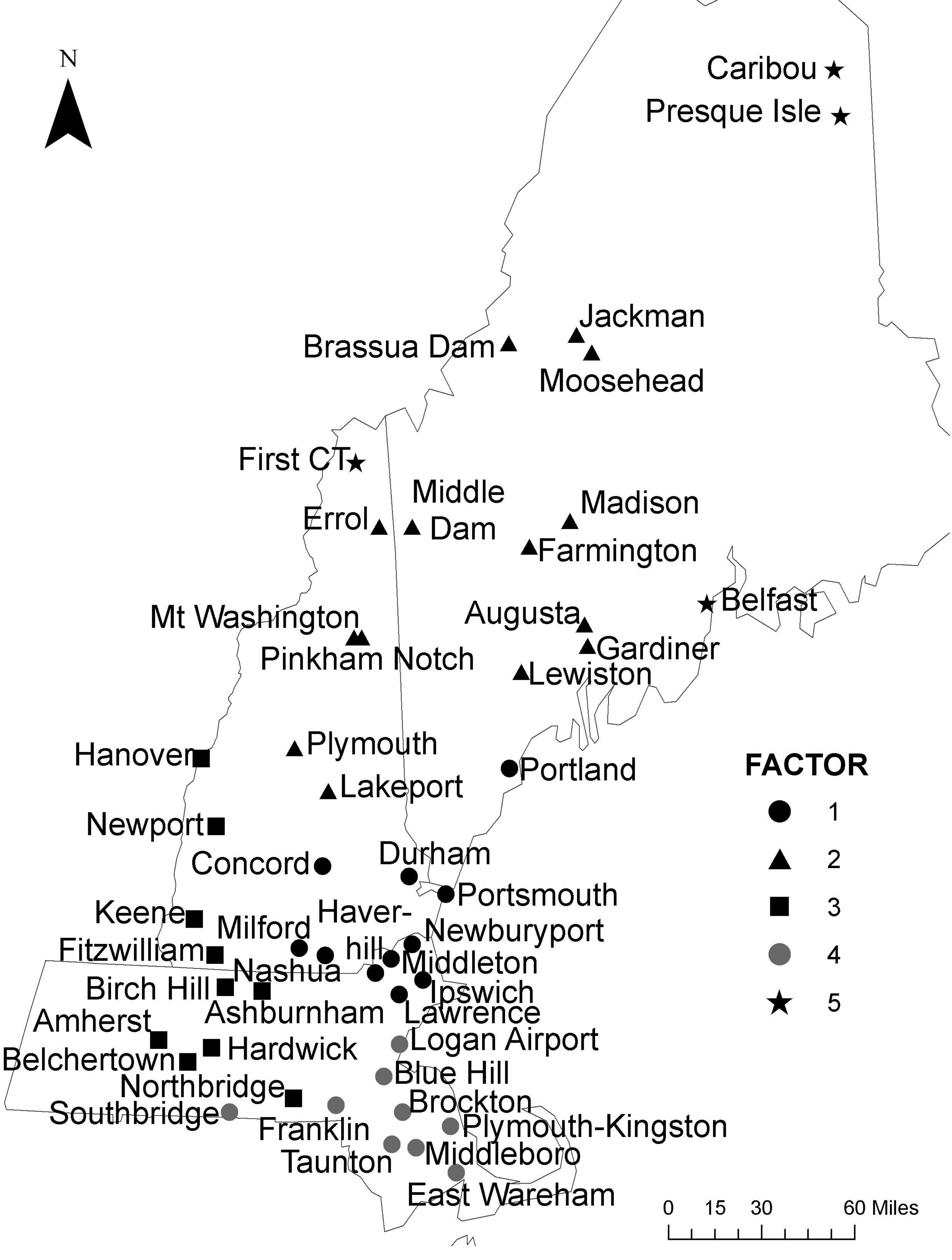
Notes: [a] TP-40 estimate were interpolated from the 100-year, 24-hours precipitation contour maps presented in NWS (1961).

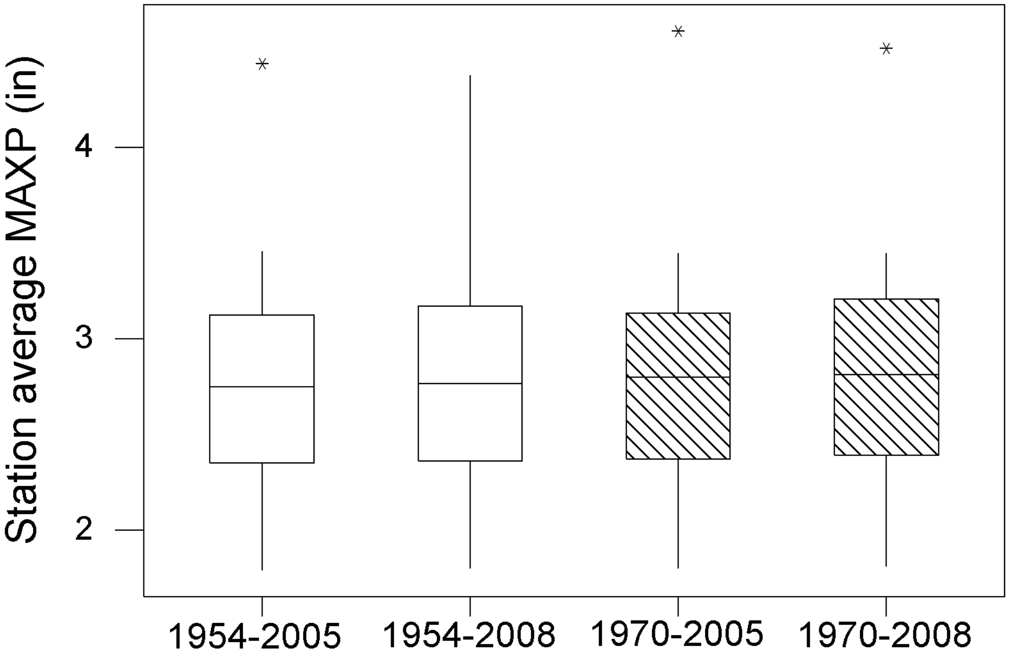
[b] GEV denotes quantiles estimated using the Generalized Extreme Value distribution fit maximum annual precipiation depths.

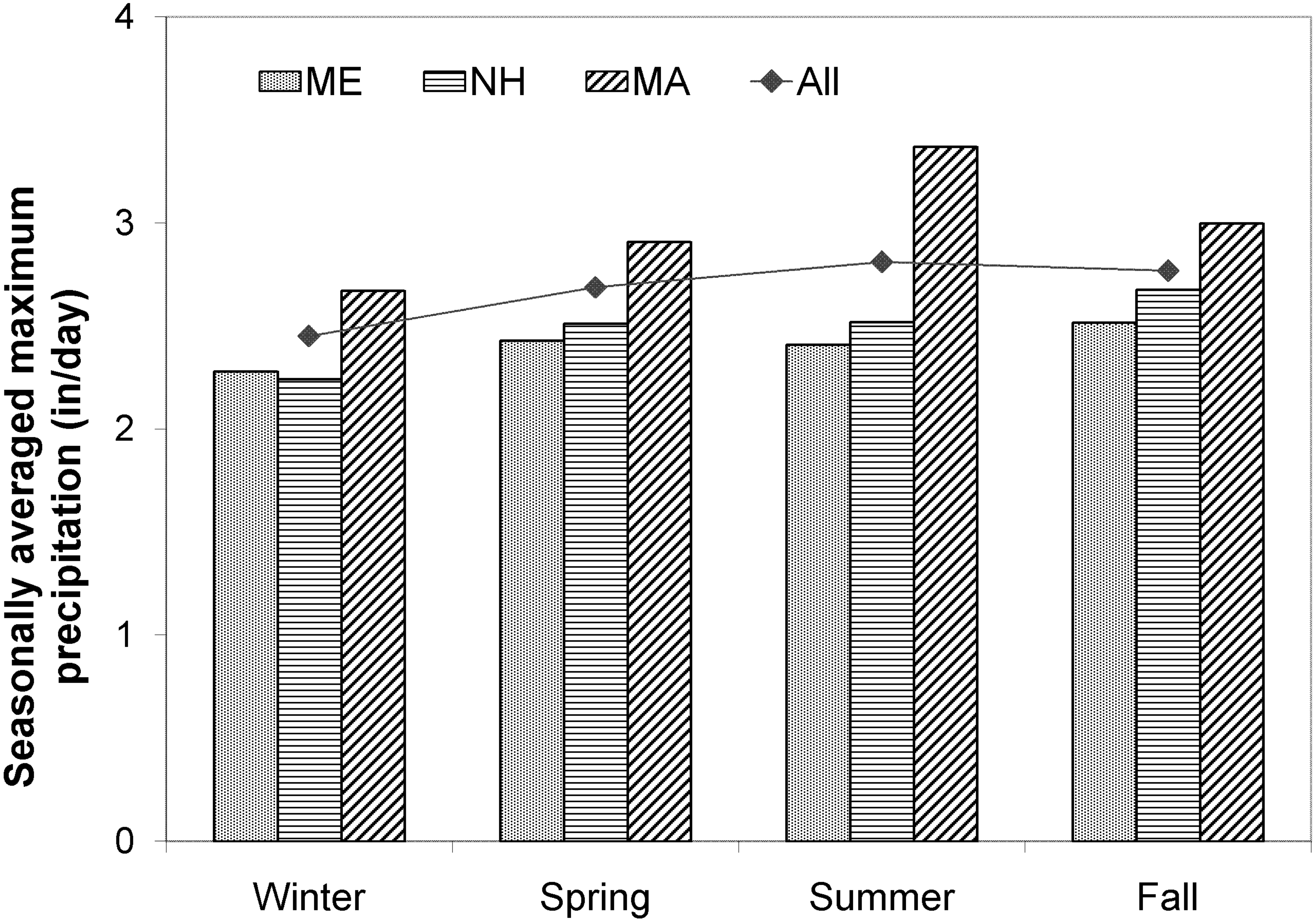
[c] GEVadj denotes GEV quantiles adjusted 11% following NWS (1961).



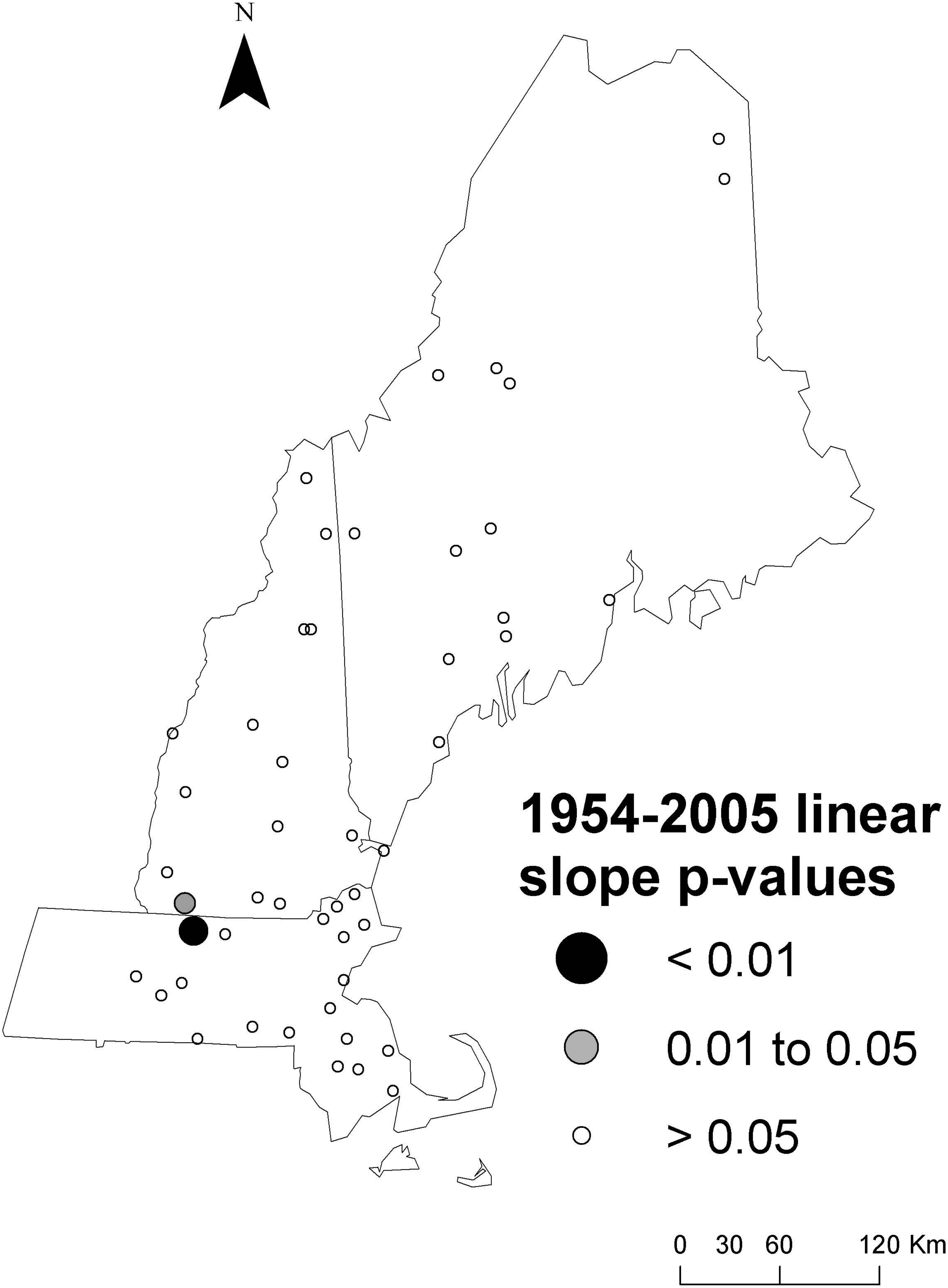
Percentage of Total
0
2
4
6
8
10
12
14
16
18
ME
NH
MA
All
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

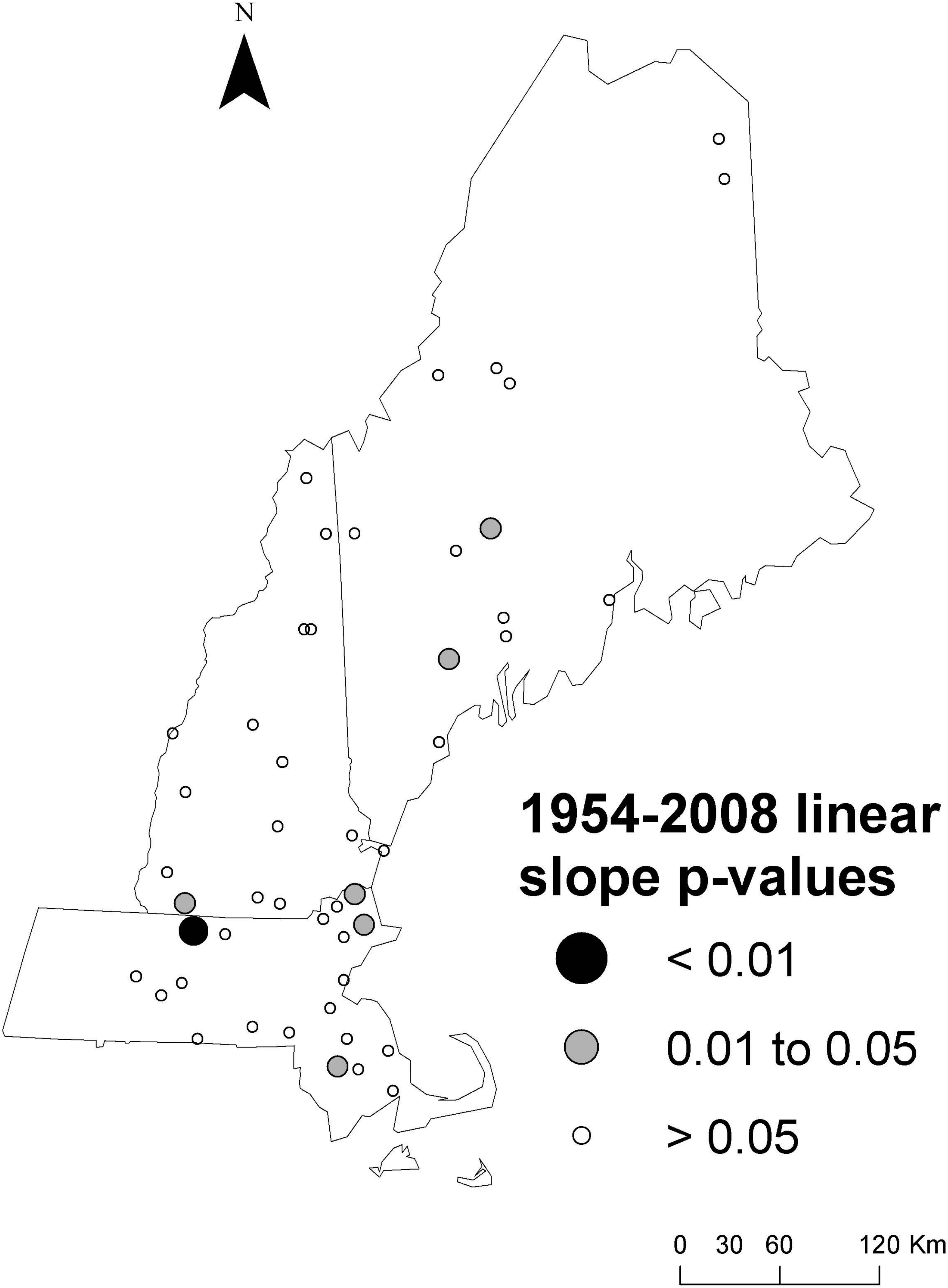

N
Caribou
Presque Isle
Jackman
Brassua Dam
Moosehead
First CT
Middle Dam
Madison
Errol
Farmington
Belfast
Augusta
Mt Washington
Gardiner
Pinkham Notch
Lewiston
Hanover
Plymouth
Portland
Lakeport
Newport
Concord
Durham
Portsmouth
Keene
Haverhill
Milford
Newburyport
Fitzwilliam
Middleton
Nashua
Birch Hill
Ipswich
Ashburnham
Lawrence
Amherst
Logan Airport
Hardwick
Belchertown
Blue Hill
Northbridge
Brockton
Southbridge
Franklin
Plymouth-Kingston
Taunton
Middleboro
East Wareham
FACTOR
● 1
▲ 2
■ 3
● 4
★ 5
0 15 30 60 Miles

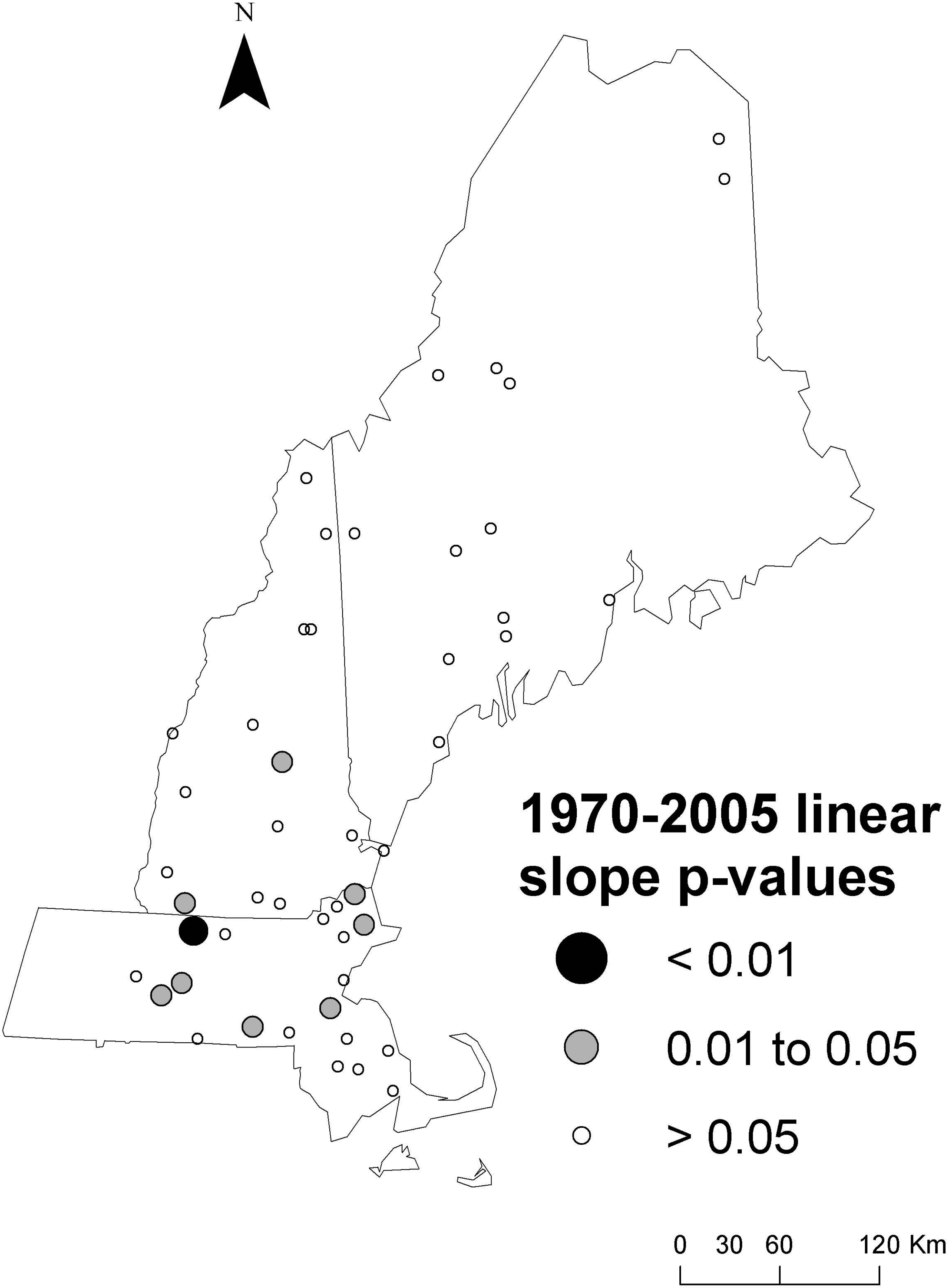

Station average MAXP (in)
2
3
4
1954-2005
1954-2008
1970-2005
1970-2008

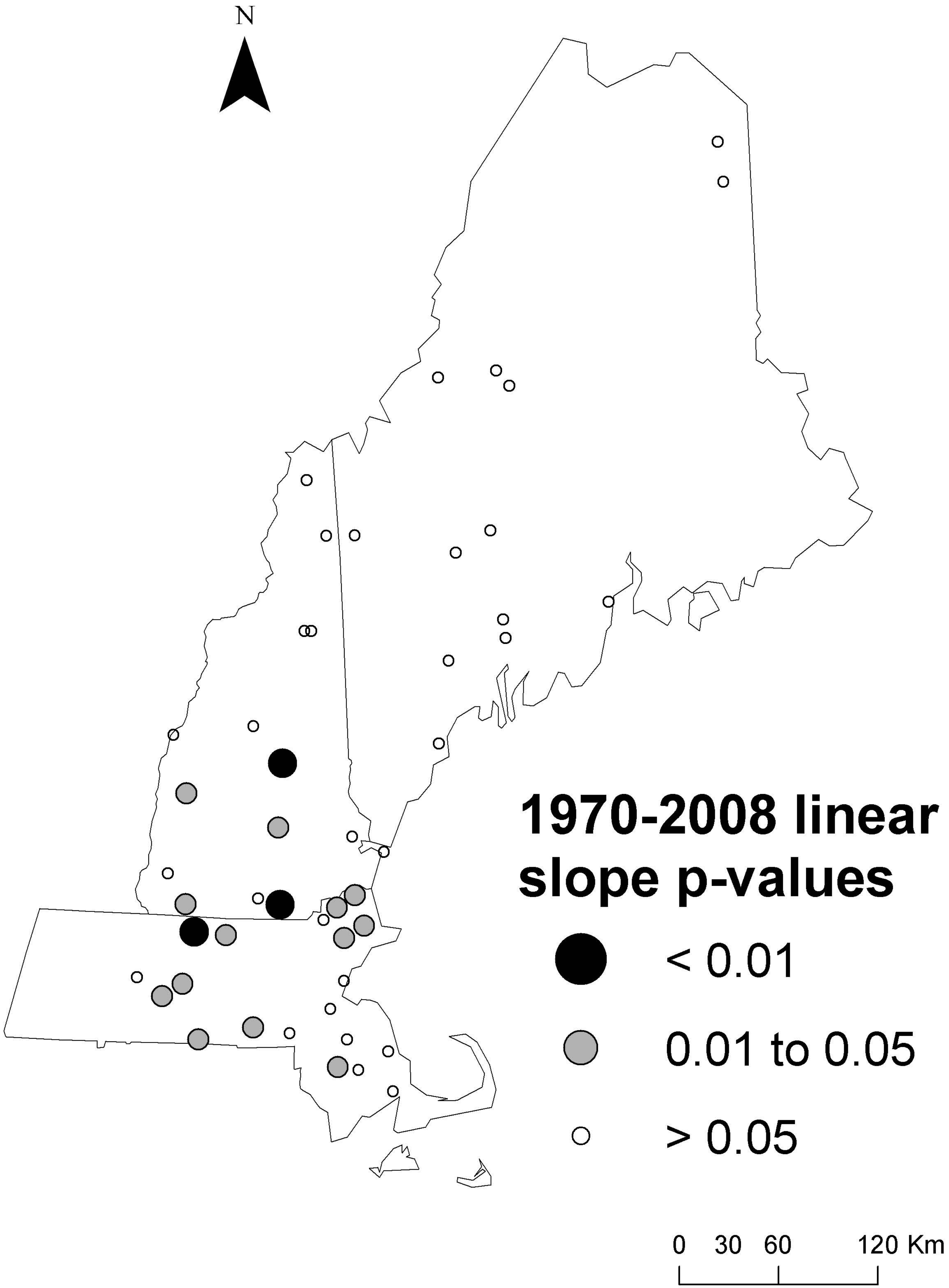


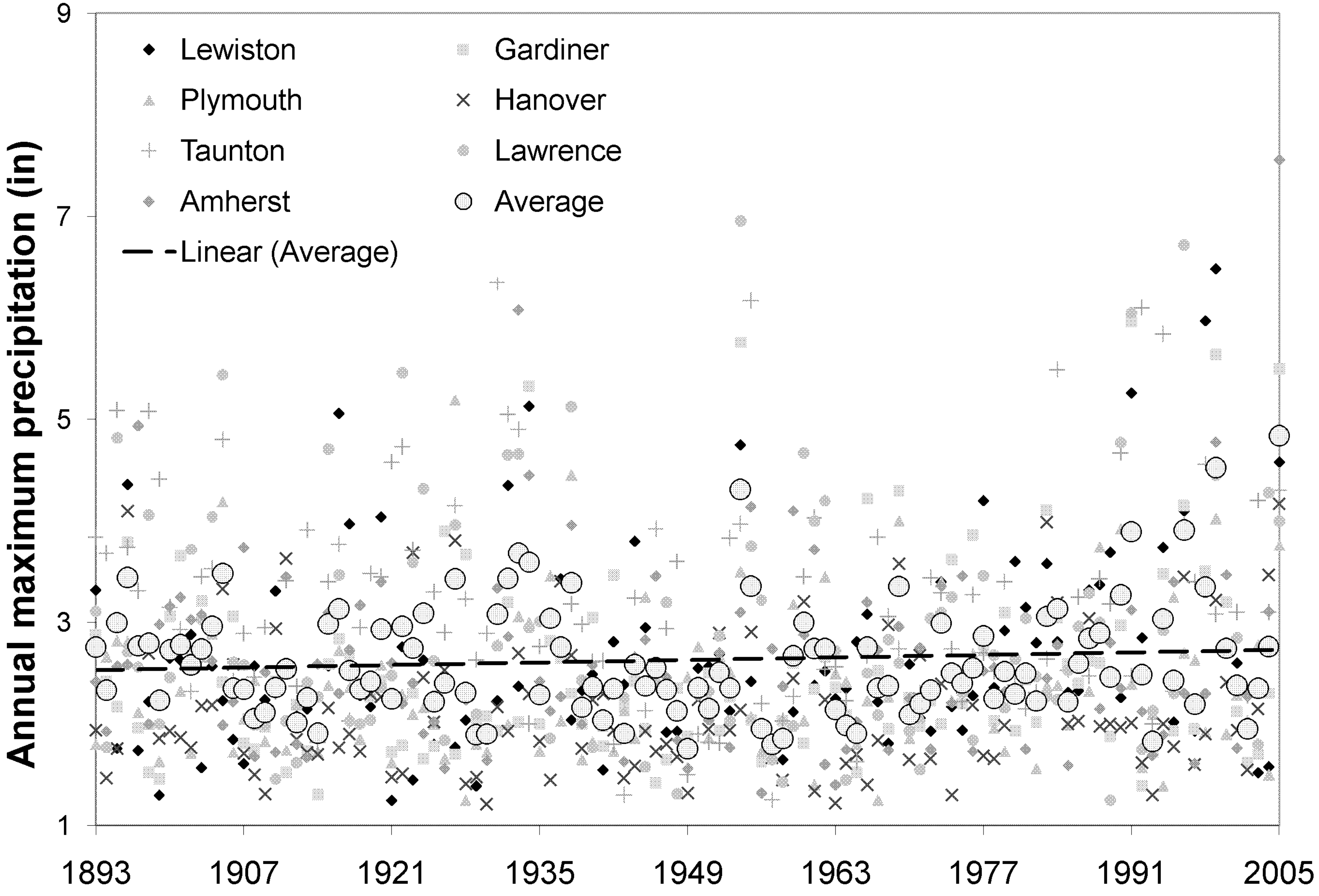
ME
NH
MA
All
Seasonally averaged maximum precipitation (in/day)
4
3
2
1
0
Winter
Spring
Summer
Fall

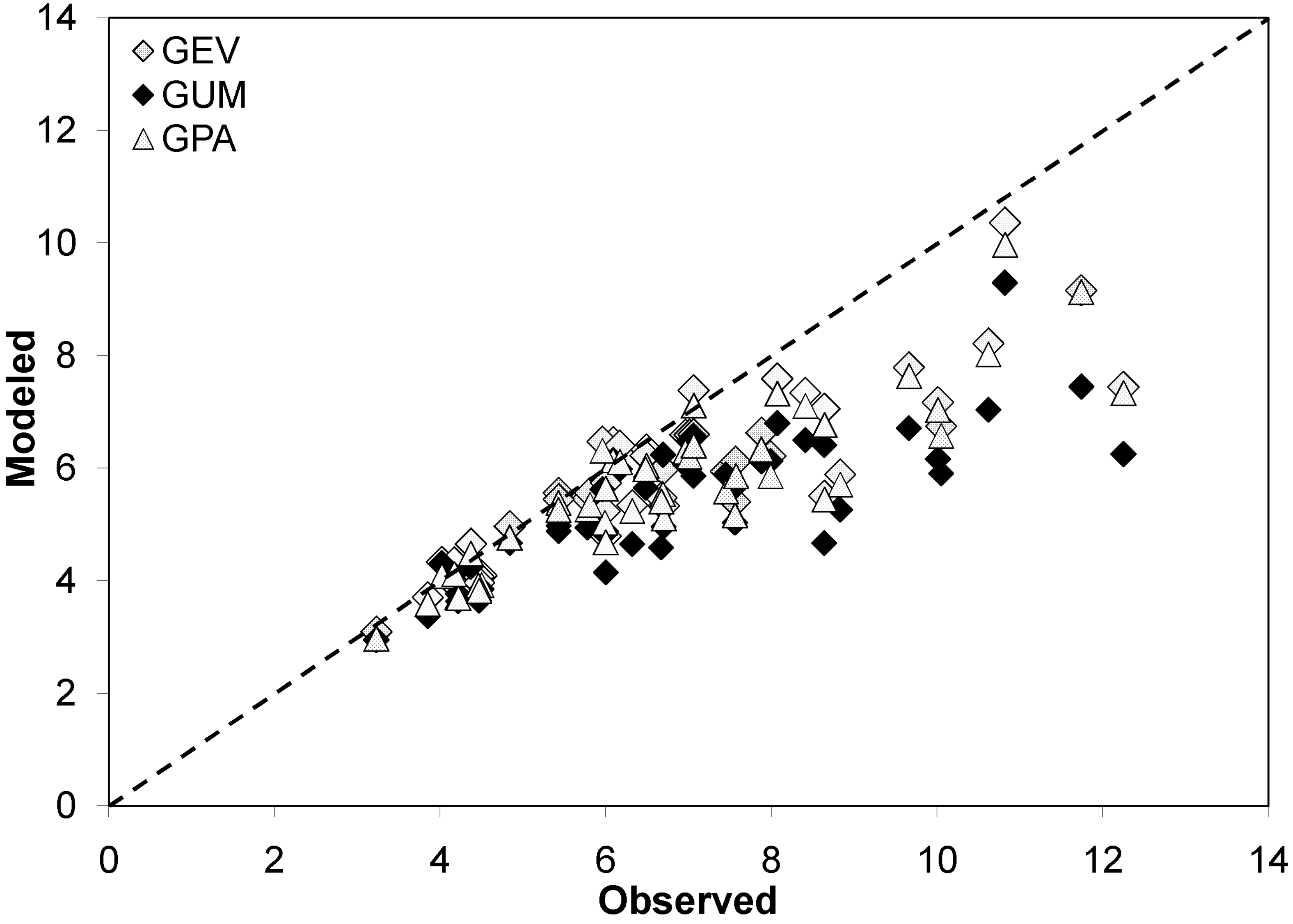

N
1954-2005 linear slope p-values
● < 0.01
● 0.01 to 0.05
○ > 0.05
0
30
60
120 Km


N
1954-2008 linear slope p-values
● < 0.01
● 0.01 to 0.05
○ > 0.05
0 30 60 120 Km

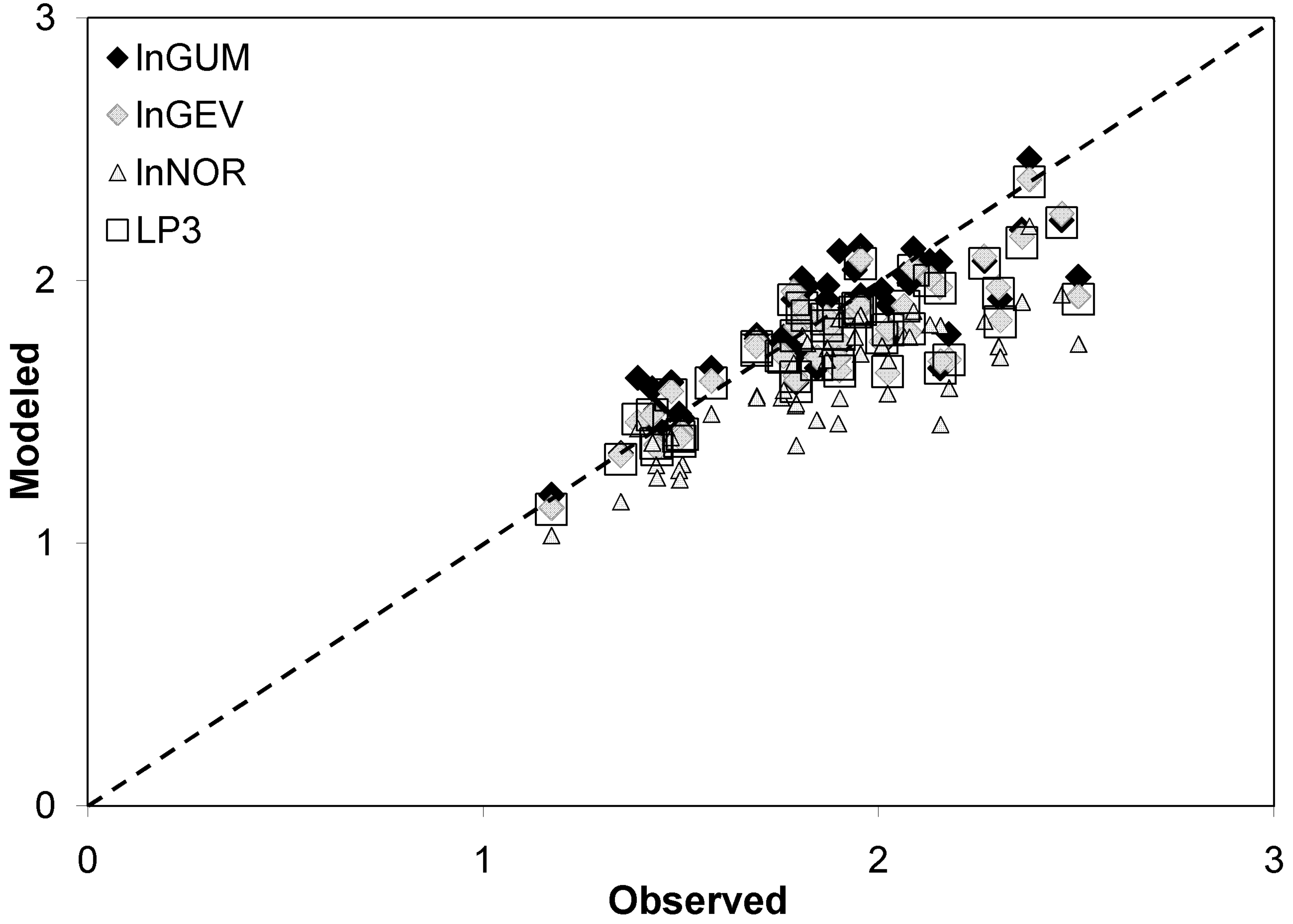

N
1970-2005 linear slope p-values
● < 0.01
● 0.01 to 0.05
○ > 0.05
0 30 60 120 Km

N
1970-2008 linear slope p-values
< 0.01
0.01 to 0.05
> 0.05
0
30
60
120 Km

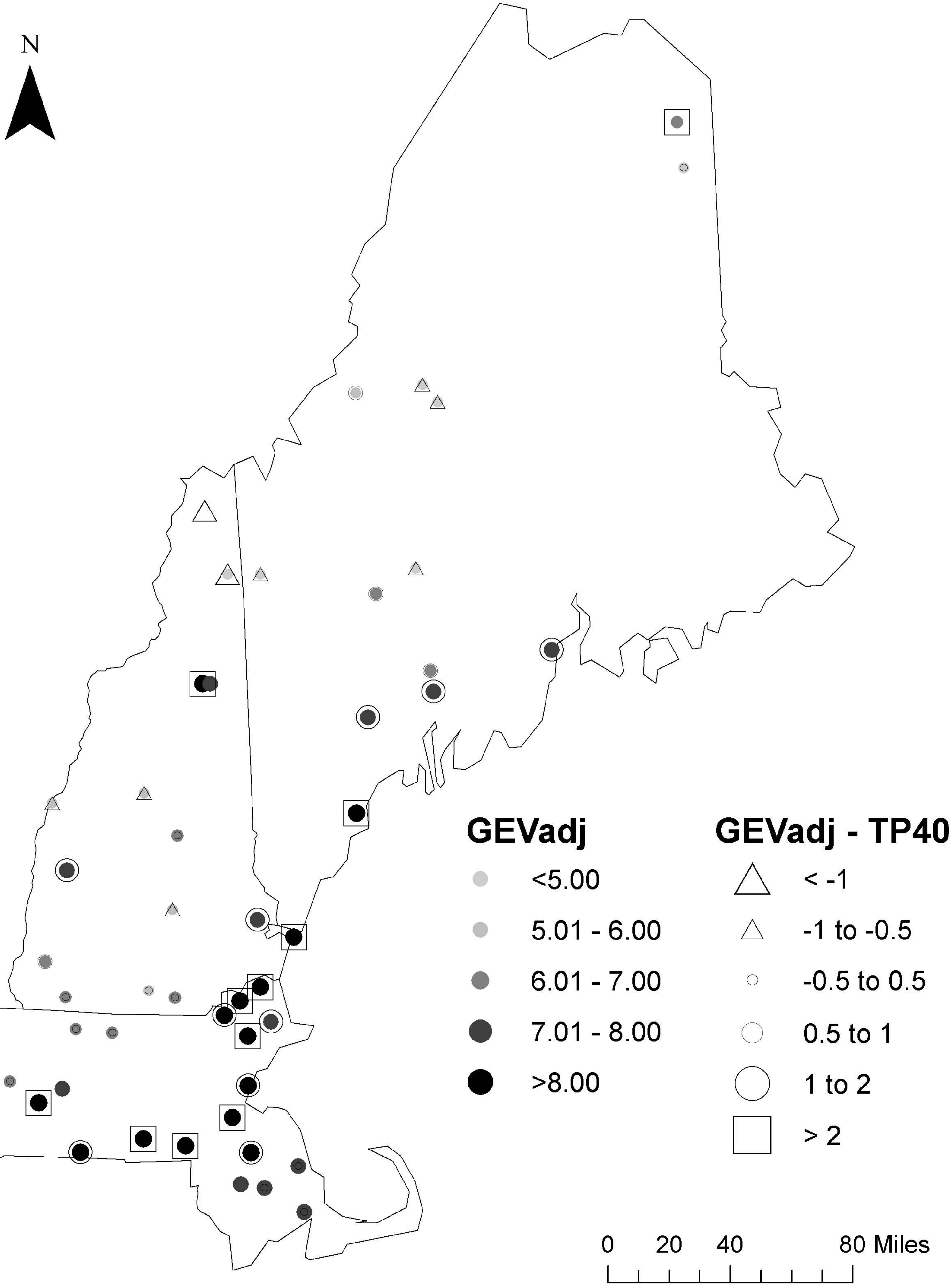

Annual maximum precipitation (in)
9
7
5
3
1
1893
1907
1921
1935
1949
1963
1977
1991
2005
Lewiston
Plymouth
Taunton
Amherst
Linear (Average)
Gardiner
Hanover
Lawrence
Average

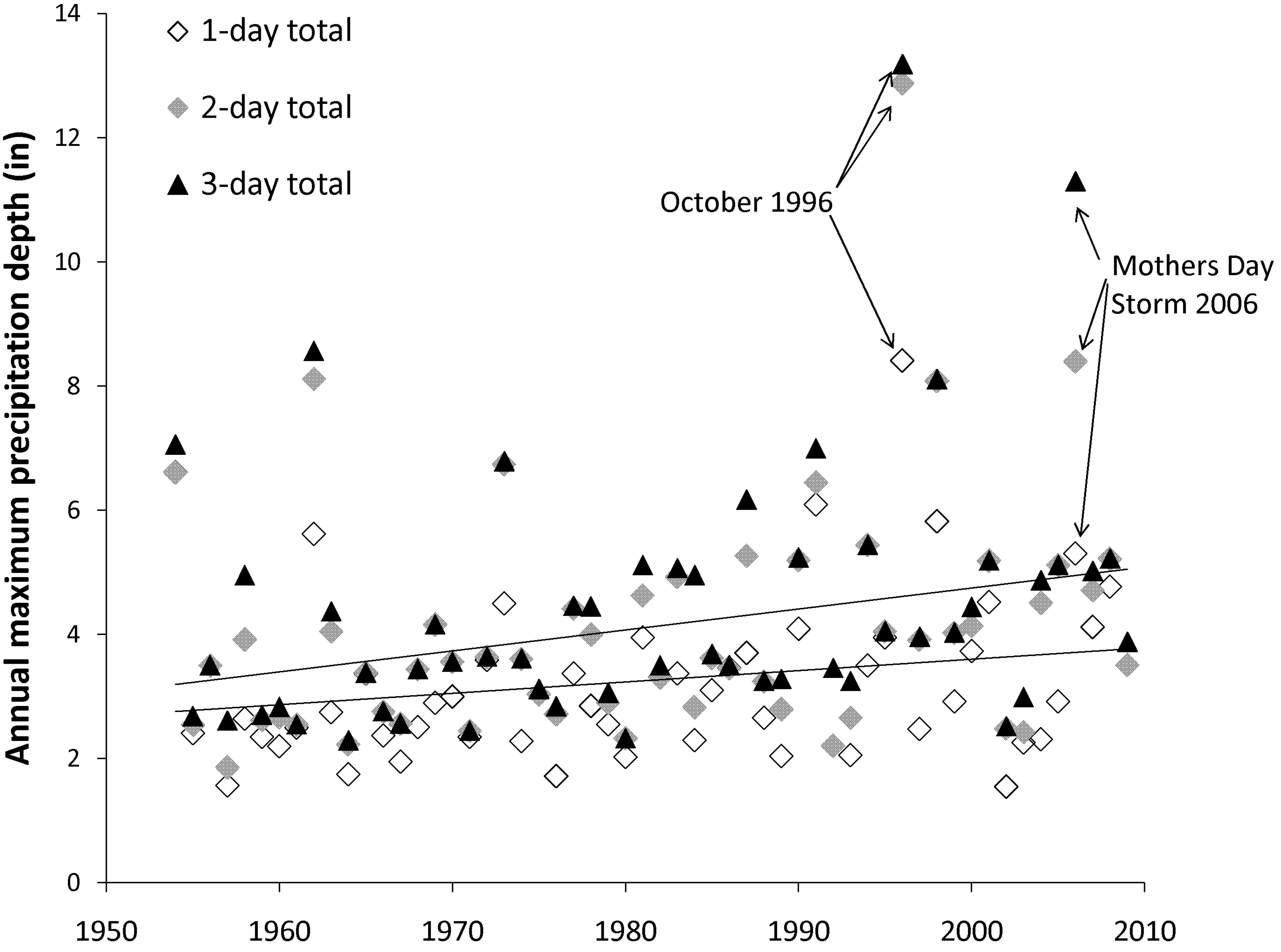


◇ GEV
◆ GUM
△ GPA
Modeled
Observed
0
2
4
6
8
10
12
14
0
2
4
6
8
10
12
14



3
2
1
0
0
1
2
3
Modeled
Observed
lnGUM
lnGEV
lnNOR
LP3


N
GEVadj
<5.00
5.01 - 6.00
6.01 - 7.00
7.01 - 8.00
>8.00
GEVadj - TP40
< -1
-1 to -0.5
-0.5 to 0.5
0.5 to 1
1 to 2
> 2
0
20
40
80 Miles


◇ 1-day total
◆ 2-day total
▲ 3-day total
October 1996
Mothers Day Storm 2006
Annual maximum precipitation depth (in)
0
2
4
6
8
10
12
14
1950
1960
1970
1980
1990
2000
2010