Auxiliary Material for

Increasing Storm Tides in New York Harbor, 1844-2013

S.A. Talke[1], P.Orton[2], D.A. Jay[1]

(1): Civil & Environmental Engineering Department, Portland State University, P.O. Box 751 CEE, Portland, OR 97207-0751, s.a.talke@pdx.edu

(2) Stevens Institute of Technology, Castle Point on Hudson, Hoboken, NJ 07030

Overview

In this supplement we provide further information about New York harbor tide data (Section S.1), provide details of our quality assurance (Section S.2), and provide supporting information and graphs on our GEV analysis, correlation analysis, climate indices, and long-term changes to New York tides (Section S.3). Notes concerning our annual maximum storm tide (AMST) time series are found in Section S.4.

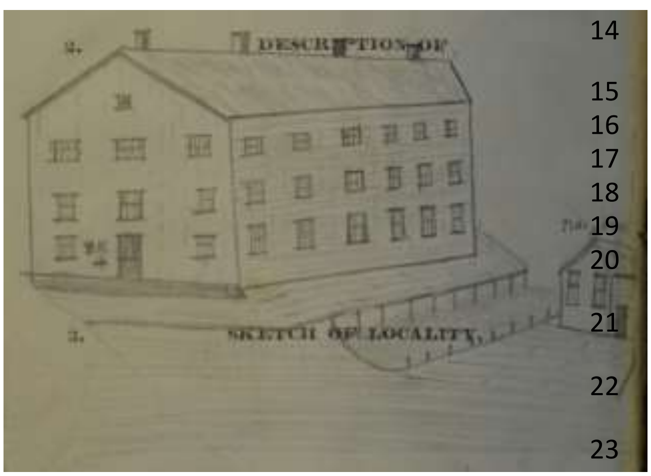


Figure S.1: The Brooklyn Tide Gauge House and Benchmark at the Hamilton Ferry Dock in 1861. The 'Tide House' is the small building at right, 'BM+' is the benchmark. Photograph by P. Lau at the US National Archives in College Park, MD.

S.1 Data

Before 1853, water levels were tabulated by a tide observer in the hour before or after high water, typically four times a day. An automatic, self-recording tide gauge (SRTG) was installed at Governors Island in December 1852, and yielded a continuous pencil trace of the tides on a scroll of paper known as a 'marigram'. Hand tabulated data were measured at the Brooklyn tide gauge into the 1860s, and form a valuable check against early automatic tide gauge data. Data were 'reduced' from marigrams into tabulated tide data using daily comparisons between the SRTG and a staff gauge, which (when they still exist) provide an estimate of error. Interestingly, the technology for measuring tides and reducing data changed little between the 1850s and the

1960s, and analog traces of the tide were recorded by NOAA until the 1980s/1990s. Hence, the tabulated precision— +/- 0.05 ft., or +/- 0.015m – is identical in 19th and 20th century data. Early SRTG data before 1859, however, are typically more error prone, with the largest errors occurring during rising and falling tides. See Talke et al., 2013 (and below) for more information. Sandy Hook data between September 1882 and 1885 was also noted by contemporary observers to be of lesser quality. Interestingly, the next most error-prone era seems to have been the late 1960s and 1970s, based on the authors' experience with west-coast tide data (also Agnew, 1986 and personal communication, P. Bromirski); however, more research is necessary to assess this qualitative observation.

Three types of data were used to construct our data set of Annual Maximum Storm Tides (AMSTs): tabulated annual extremes from Schureman (1934), hourly tide data (available from NOAA or recovered from the national archives), and daily tabulations of High Water (HW) and Low Water (LW). See Talke & Jay, 2013, for a discussion of data-types, historical data reduction, and locations. The main manuscript describes the data recovered and digitized for New York Harbor (NYH). A data set of annual maximum storm surge (AMSS) is also extracted from time periods in which hourly data was available: 1860-1885, 1889-1921, and 1927-present. The AMSS is constructed by subtracting the annual mean sea-level and the predicted tide from hourly data.

Examples of historical tabulated data are given below in Figs. S2, S3, and S4. Figure S2 shows the hourly tabulated data from Fort Hamilton in Aug. 1893, during the period of the hurricane described in Scilleppi& Donnelly (2007). Figure S3 shows an example of tabulated high/low data which was reduced from automatic gage data. Figure S4 shows an example of pre-1853, hand-tabulated high/low data from a box-gauge at Governors Island. The pre-1853 tide staff was inverted (i.e., z-axis was positive downwards), such that measurements for HW are less than LW. The weather observations on pre-1853 data were helpful (as in the indicated image) in identifying gale force winds and elevated water levels. Pictures were taken of available data in the US National archives in College Park, MD, and these data were manually digitized by students (See Acknowledgements).

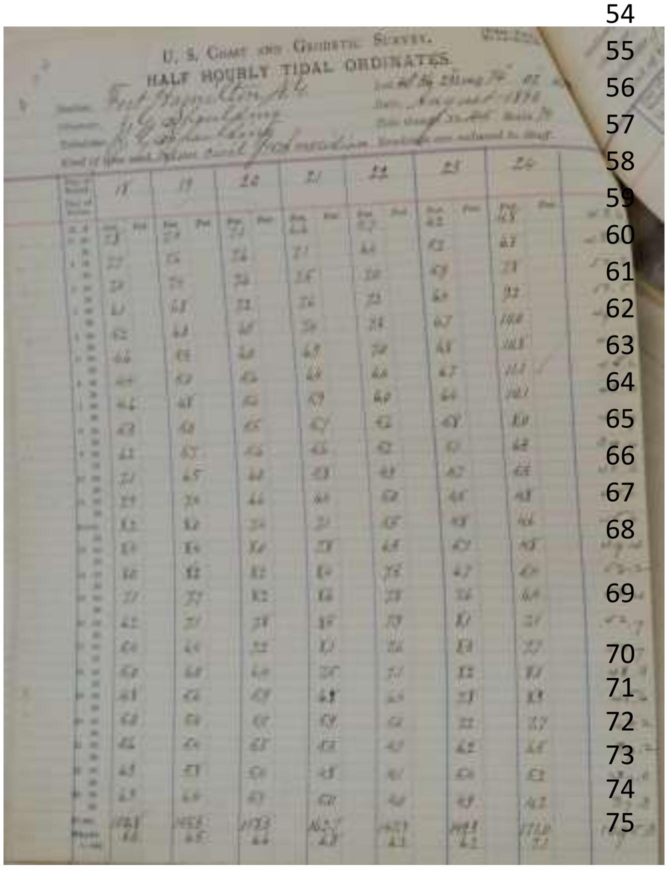


Fig.S2: Example of hourly tabulated data from Fort Hamilton, NY from Aug. 1893, photographed at the US National Archives in College Park, MD by S.A. Talke. Note the hurricane storm tide which occurred on Aug. 24, 1893 at 6:00am local civil time.

As described in the main text, data were quality assured by (a) manual inspection, and comparison with tabulated sums (see Fig. S2); (b) differencing sequential data to identify and eliminate unphysical spikes; (c) comparison of data with other gauge-data and newspaper archives to establish plausibility and (d) harmonic analysis to determine whether data produced reasonable constituent values (see e.g., Leffler & Jay, 2009). We will look at points (c) and (d) more closely here.

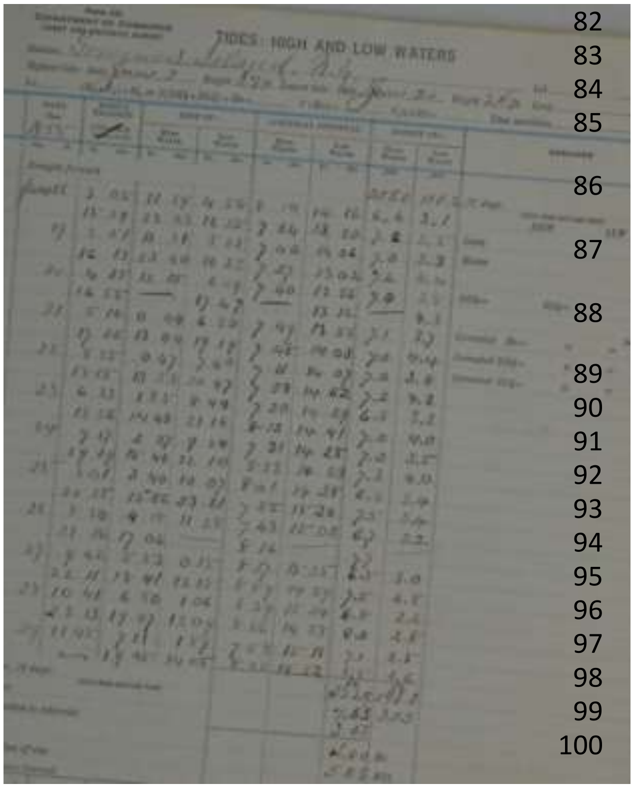


Fig. S3 (Left): Tabulated High/Low data from Governors Island, June 1855. Photograph by P. Lau at the U.S. National Archives, College Park MD.

Fig. S4 (below). Example of hand-measured tide gauge data from Governors Island on the 21st of November, 1851. Note the tide gauge has an inverted scale, such that HW is the minimum measurement. The note states "very squally with rain; the highest tide since I have been on the Island". The barometer reading (obscured) was 29.5inches mercury, and the air temperature 50 degrees Fahrenheit. A different observer was listed in early 1850; hence, the listed storm surge was the largest event of 1851, and perhaps 1850. Photograph by P. Lau.

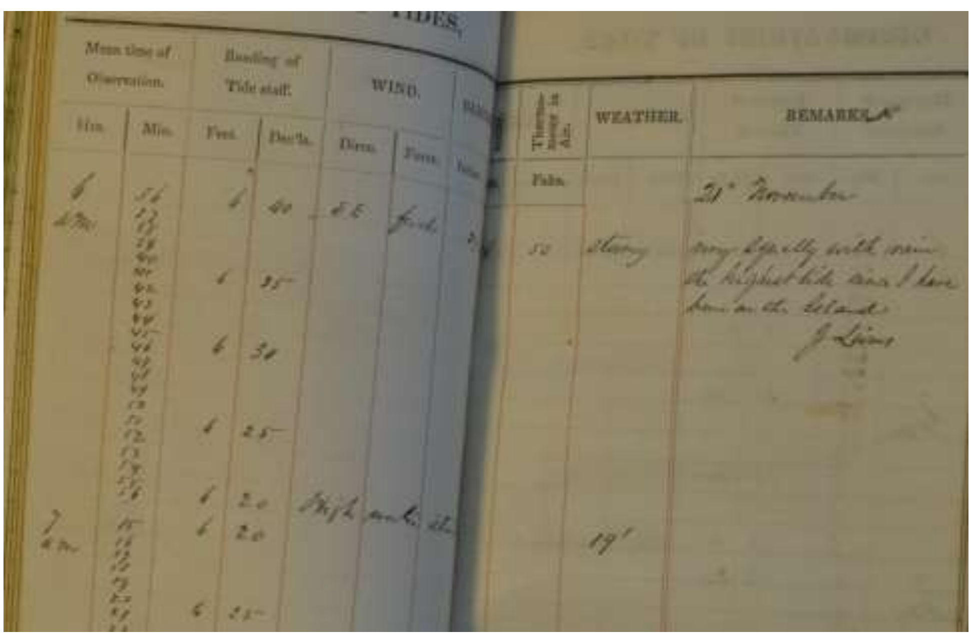


Figure S5 shows an example of early Governor's Island gauge checks from 1853; between 1853-1858, errors of +/-0.25ft (0.08m) were occasionally observed, and were attributed to problems with the stilling well. We note that these errors are largest during a rising or falling tide, with a less severe effect on extrema (see Agnew, 1986); hence, precision errors in tabulated AMST are likely much smaller. Later data, as also observed by Talke & Jay (2013), seems to become more accurate, although further archival research would be required to exhaustively check the nearly 160 years of automatic gauge data. Because of the early issues of the Governors Island automatic gauge (installed Dec. 1852), we used hand-tabulated Brooklyn data (1856-1862) whenever possible for the pre-1860 time period.

Figure S6 shows a plot of one year of hourly tide data from Fort Hamilton and six months from Sandy Hook from 1893. The hurricane storm tide and storm surge from the Aug. 24th, 1893 hurricane tabulated in Fig. S2 is shown in Fig. S7. Results show that a storm surge of ~1.2m occurred at a HW of 0.4m, producing a storm tide of 1.6m that was not even the largest water level of the year. The AMST for 1893 occurred on the night of April 20th, as measured at the Battery Port-Docks (not shown), Fort Hamilton, and Sandy Hook tide gauges. While the Sandy Hook tide series ends in June 1893, two gauges (The Battery and Fort Hamilton) therefore suggest that the 1893 hurricane was relatively muted in NYH. The Battery extreme data is from Schureman (1934). The hourly data on April 20th from Fort Hamilton was not available; hence, the only extrema depicted is from Sandy Hook. No Governors Island data was measured in 1893, and hourly Battery data is unavailable.

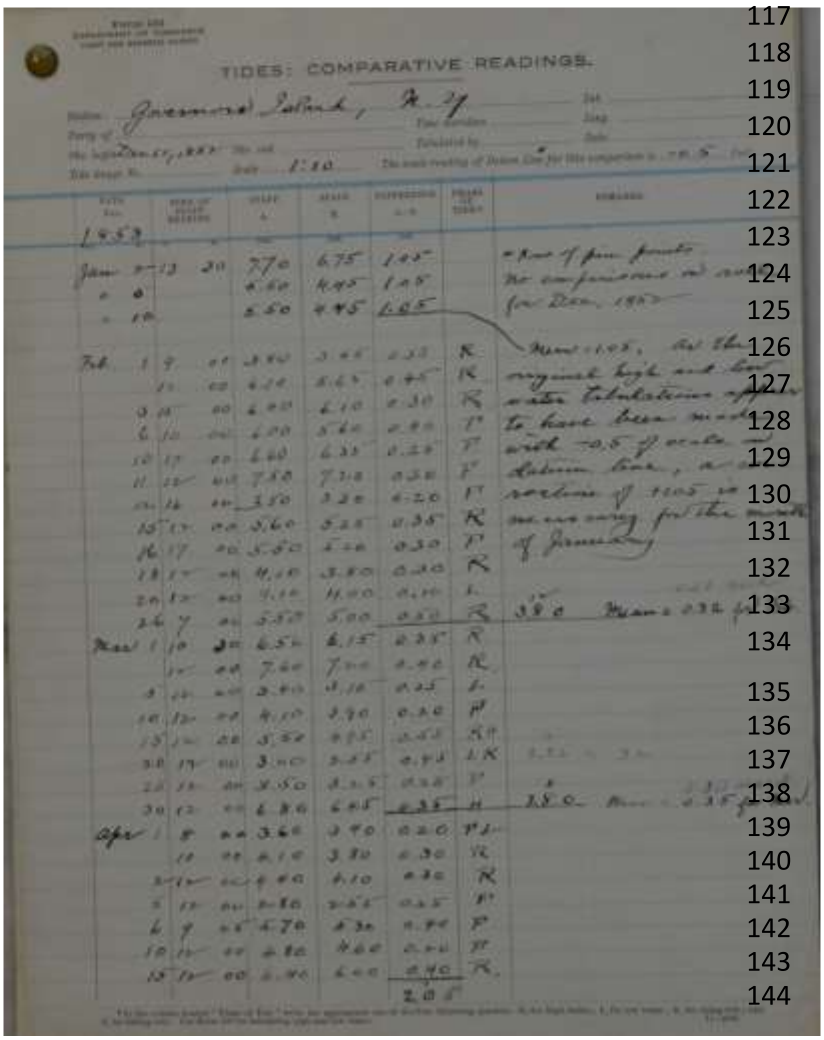


Fig. S5: Gage checks (Comparison sheet) for Governors Island from June 1853. During the early years of the automatic gauge (1853-1858), variability between the water level inside the gauge stilling well and the water level measured on a tide staff often varied by +/- 0.1 to +/- 0.25ft. Photograph by P. Lau.

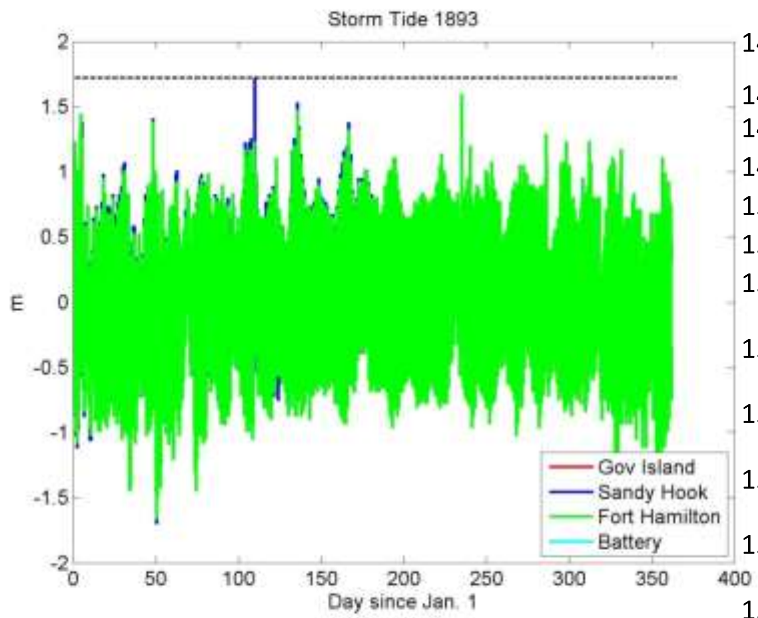


Fig. S6: Example of one year of hourly data, from 1893. The mean sea-level has been removed. The AMST is depicted by the horizontal dashed line.

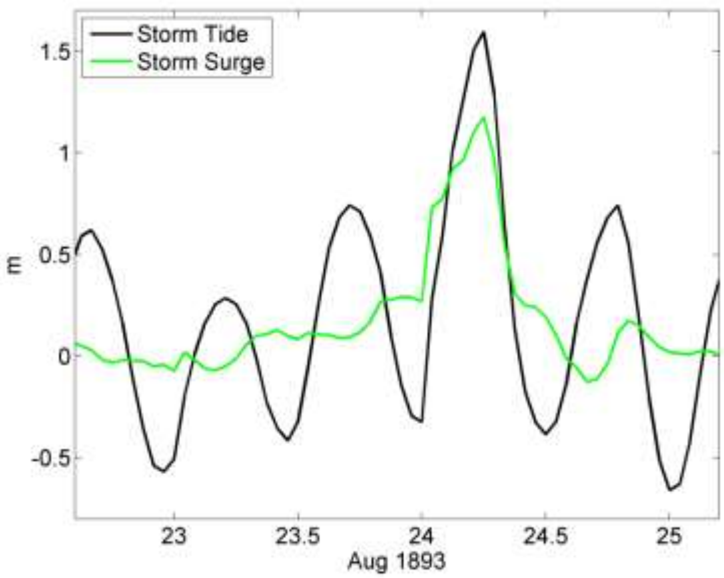


Fig. S7: The hurricane storm tide and storm surge for the last hurricane to directly strike New York City before hurricane Sandy in 2012, measured at Fort Hamilton, NY.

S.1.2: Battery Annual Extremes, 1886-1932

The annual extremes at Battery Park from 1886 to 1932 are listed in Schureman (1934) relative to the mean-sea-level datum from Sandy Hook tide data, 1876-1881. These extrema data overlap the modern hourly data available for 1920-1921 and 1927-2013, which allows us to convert the Schureman (1934) tabulations to the modern Battery station datum. The Schureman (1934) datum is a median of -1.52m below the modern Battery station datum (range 1.5-1.53m, standard deviation 0.016m). The variance occurs because hourly maxima are slightly different than the actual tabulated maxima (see discussion in Section S.2). A similar result is found using annual means. We add a correction of +1.52m to the Schureman data to obtain the extreme values relative to the modern Battery station datum, recognizing that a slight bias of order 0.01m may occur. Interestingly, the value 1.52m is 5 feet; hence, the Battery station datum appears to have been set at 5 feet above Sandy Hook MSL measured from 1886-1881, though this may be coincidental.

We next convert the Battery maxima to an ‘annual maximum storm tide’ (AMST) by subtracting out the annual sea-level between 1886 and 1926, using tabulated sea-level for New York from the Proudman Service for Mean Sea-level (psmsl.org), which is obtained from NOAA (National Oceanographic and Atmospheric Administration). The MSL for 1889-1892 was obtained from Sandy Hook hourly data, after adjusting the 1893 Sandy Hook data to the Fort Hamilton datum. In the units of PSMSL, the 1889-1892 MSL was: 6783mm, 6754mm, 6765mm, and 6752mm, respectively. For the 1886-1888 period, we subtracted out the average of 1874-1878 and 1893-1897 MSL. Since we do not have access to the original Battery Port-Docks data, we check the 1886-1920 Battery Park AMST for plausibility against the Sandy Hook and Fort Hamilton data (section S.2). Note that while the Fort Hamilton extremes in Schureman (1934) for 1893-1920 have been validated by digitized data, The Battery Port-Docks data (1886-1920) have not been recovered. To reduce possible gauge errors in the Schureman (1934) data (e.g., due to undiagnosed dock subsidence), we average the 1893-1926 extreme values from the Fort Hamilton and The Battery stations.

S.2 Quality Assurance and Bias Corrections

In this section we describe the consistency checks we made to quality assure our digitized data, as well as assess and correct for bias within our AMST time series. The overall AMST estimated from different gauges is shown from 1860-2013 in Fig. S8. To obtain the closest possible geographical congruence between measurements, data from Battery Park, Governors Island, or Brooklyn were used in the AMST time series whenever possible. These measurements were all within 2km of the modern gauge. However, as described in the main manuscript, gauge recording gaps (e.g., in 2001 and 2004 for The Battery) necessitated using the Sandy Hook gauge for 13 AMST values (see Fig. 2 in main text). Outliers in Battery/Gov. Island data in Fig. S8 denote periods in which incomplete data necessitated using other gauge data. The Fort Hamilton/Battery data from 1893-1926 were also averaged together, as described below.

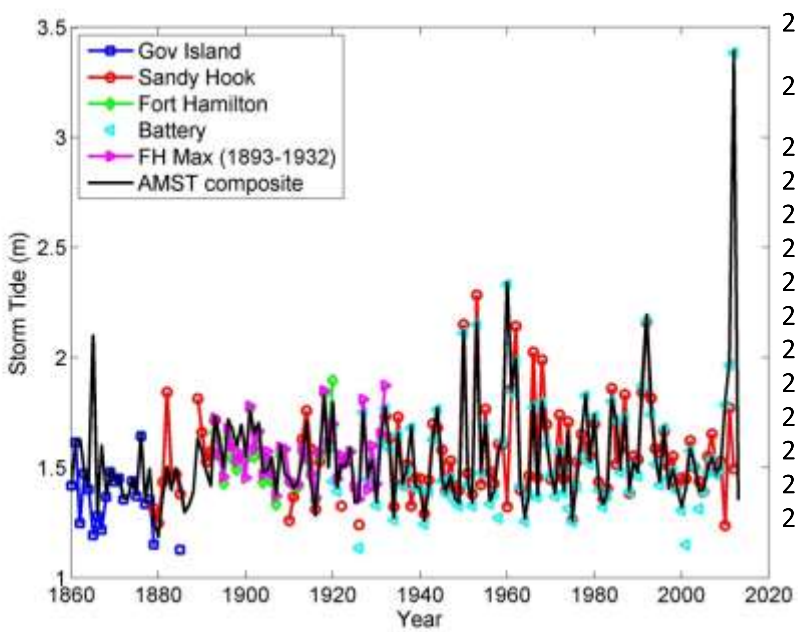


Fig. S8: Composite AMST (black) and its relationship to data from tide gauges, 1860-2013. The ‘FH max’ time series is derived from Schureman (1934) from 1893-1932; likewise, the composite time series is derived from Battery extremes listed in Schureman (1934) from 1886-1926.

The difference between measured AMST at Sandy Hook and The Battery are shown in Fig. S9. The pre-1920 differences appear to be slightly larger than post-1932 data. This is confirmed by statistics: Overall, the Sandy Hook storm tides are larger than Battery storm tides from 1886-1920 (12 events; +0.07 +/- 0.13m larger), but only 0.052 +/- 0.058m larger between 1932-2012 (69 events). The mode for the 1932-2012 and 1886-1920 periods was -0.125m and 0.052m, respectively. The larger offset, variance, and mode for the earlier data set may stem from the smaller sample size, but could also result from storm variability, increased hydraulic connectivity (decreased friction) between Sandy Hook and the Battery, or an undiagnosed but small bias between the two AMST data sets. Both the Sandy Hook and Battery tide gauge location changed slightly several times, which could also lead to small local variations. Still, the overall AMST offset is only 0.02m more at Sandy Hook during the 1886-1920 period than the 1932-2012 period. In addition, no statistically significant trend is observed between the AMST magnitude (i.e., storm strength) and the offset; larger storms do not yield a greater difference between gauges. Hence, we bias correct all the Sandy Hook AMST by the median offset over the entire 1886-2012 data set (81 values; also 0.052m).

The pre-1920 and 1920-1926 AMST are based on the tabulated Dock-A and 'Barge Office' extremes at Battery Park, respectively (Schureman, 1934), for which period we have only limited data from 1918-1921. Therefore, we compare this time period against the Sandy Hook and Fort Hamilton gauge data carefully to assess their validity. Fig. S10 shows AMST from the Battery, Fort Hamilton and Sandy Hook hourly data as well as AMST estimated from Schureman (1934), over the 1893-1932 data set. Over this 40y time period, the mean difference between the Fort Hamilton and the Battery AMST is less than 0.01m, with a standard deviation of 0.07m. This suggests that Battery and Fort Hamilton data are on average interchangeable. Since we have confirmed Fort Hamilton tabulations with hourly data (Fig. S10), we therefore average data between the Battery sites and Fort Hamilton; while small local differences may occur for individual events, this strategy serves to minimize any undiagnosed problems with the Battery record (e.g., small datum shifts, dock subsidence, etc.).

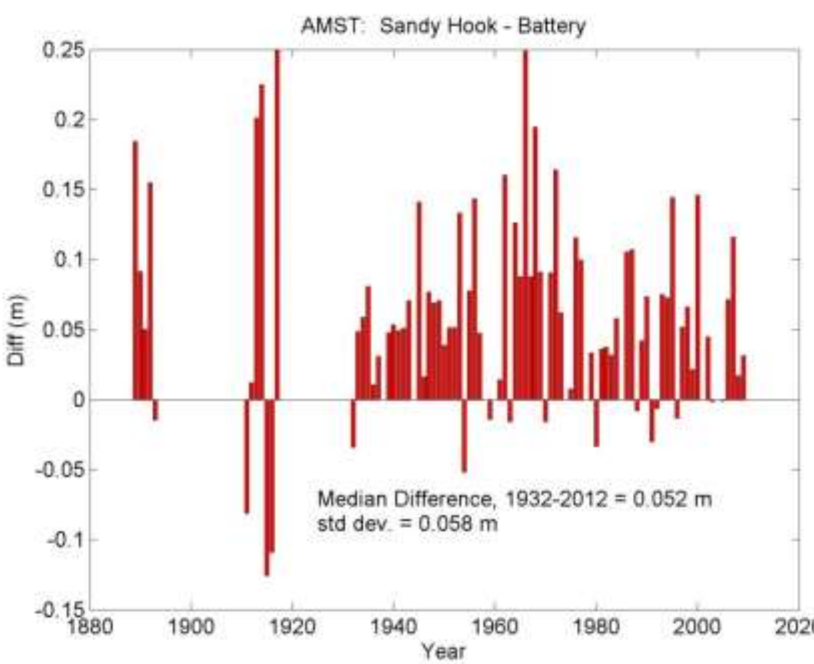


Figure S9: Difference between AMST measured at Sandy Hook and at Battery Park, 1886-2012.

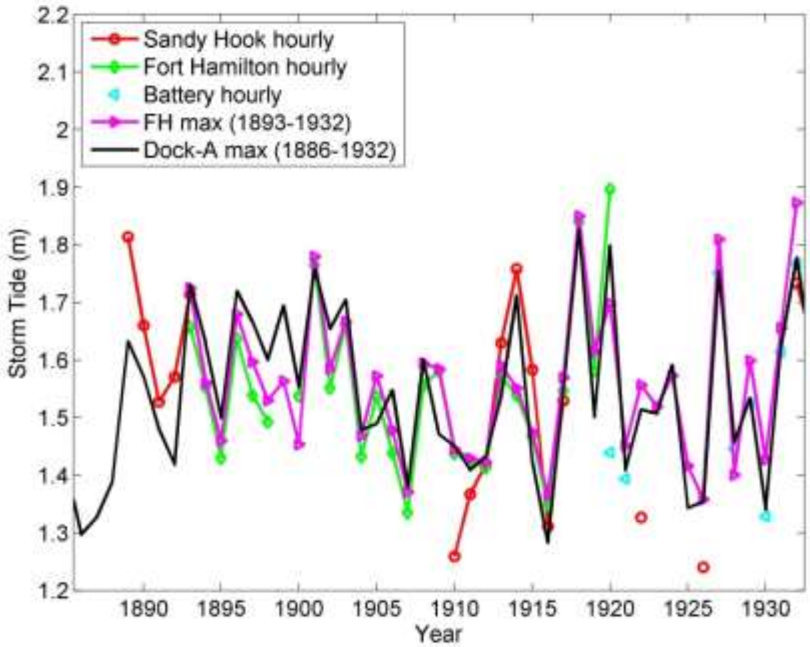


Figure S10: Close-up of Fig. S8 from 1886-1932. Dock-A data is from near the present-day location of the Battery tide gauge.

We next compare the quartiles and quartile difference which are found by using Fort Hamilton and the Battery separately from 1893-1932, rather than averaging to obtain a composite estimate (Fig. S11). Using Fort Hamilton data produces a negligible difference in the median and the 25% quartile statistics, but does reduce the upper quartile estimate by up to 0.05m, compared to the Battery. Because the 25% and 50% estimate is approximately equal with both data sets, we infer that local differences in height during large events is a likely cause. However, altered tidal range at the two stations may lead to a changed risk of AMST. We note that tidal constituents were not stationary during this time period, but rather increased at Fort Hamilton (Marmer, 1935) and decreased at The Battery/Governors Island (see section S4). Close inspection of Fig. S8 supports this interpretation: Battery AMST values are slightly larger than Fort Hamilton before 1913 (16 out of first 20 yrs.), and slightly smaller than Fort Hamilton from 1913-1932 (16 out of 20 yrs.). Differential land subsidence between the gauges or dock instability at the Fort Hamilton gauge could also produce the observation, as could changes to gauge and gauge location in 1920 (Battery), 1921 (Fort Hamilton) and 1926 (Battery), or the small upwards datum shift in 1898 at the Dock-A gauge (Koop, 1915, p. 83).

Another observation apparent in Fig. S10 is that the Fort Hamilton extremes (magenta line) from Schureman (1934) are up to several cm larger than AMST calculated from hourly data. This can occur when a maximum falls between the hourly tabulated data. We investigate this for The Battery data in Fig. S12, which shows the difference between the extremes and the hourly maximum measured annually between 1966 and 2012. Over this time period, the median difference was 0.008m, with a standard deviation of approximately 0.02m. We therefore conclude that any statistical bias from using extrema data vs. hourly data is small.

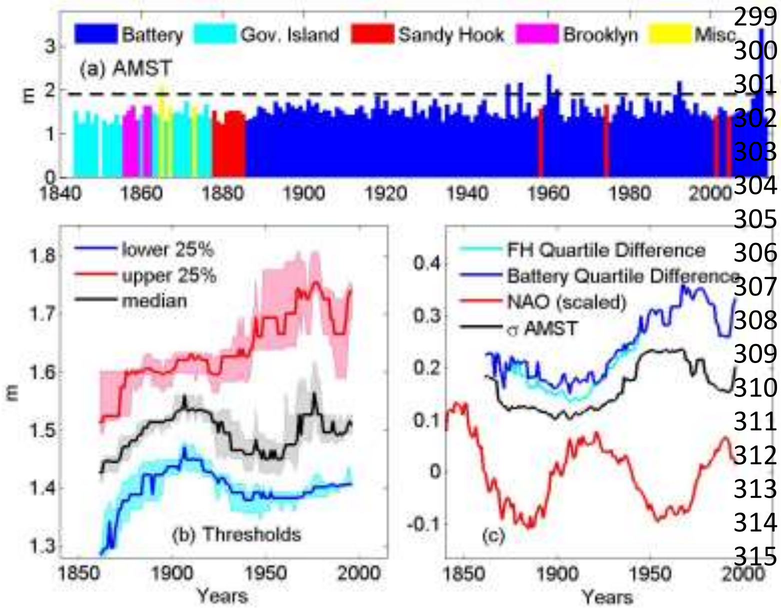


Fig. S11: (a) AMST from New York Harbor area; (b) The 25%, 50%, and 75% quartile AMST over sequential 36y periods; the Fort Hamilton AMST quartiles are the solid line, while the Battery AMST quartiles are the dotted line. c) The difference between the upper and lower quartile AMST vs. the annual NAO index, scaled by ¼ for presentation purposes. A 36yr median has been applied to the NAO index. The shaded areas in (b) denotes the +/- 5% quintile. A 36 year running standard deviation is shown in (c)

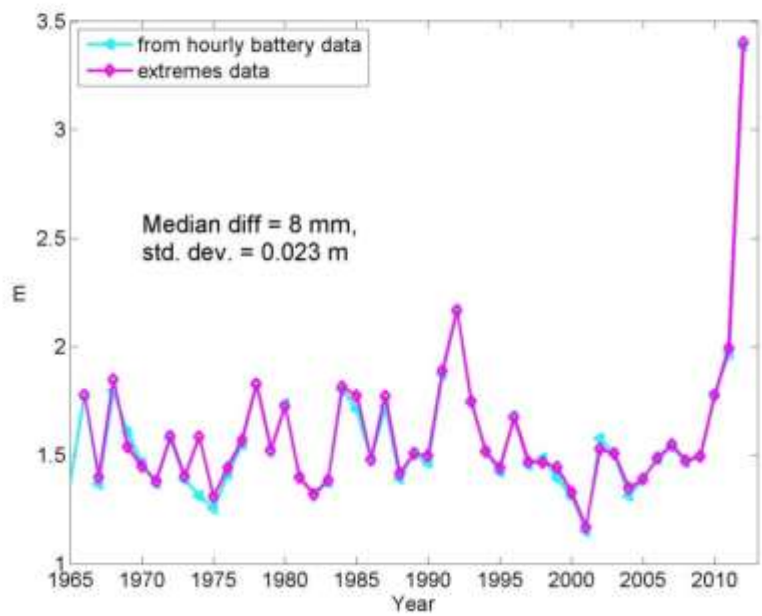


Fig. S12: Difference between tabulated annual extremes and the AMST estimated using hourly tide data at the Battery, 1966-2012. Sea-level has been removed from both data series.

Pre-1860 Quality Assurance

We examine the pre-1860 high/low data by generating tidal hind-casts using hourly data from Governors Island and comparing these against measurements. Figure S.13 shows that tidal hind-casts reproduce the 1853-1860 high/low data fairly well, indicating that these data are generally reliable. The hand-tabulated data from 1845 to 1849 is larger than the predicted tidal range and HW, but underestimates the spring-neap and diurnal variability (Fig. S14 and S15). Compared to 1853-1855 data, the average HW for the 1840s was 0.12m larger (0.76 vs 0.64m), while the standard deviation was 0.065m less (0.16m vs 0.235m). Data from 1850-1852 deviate even

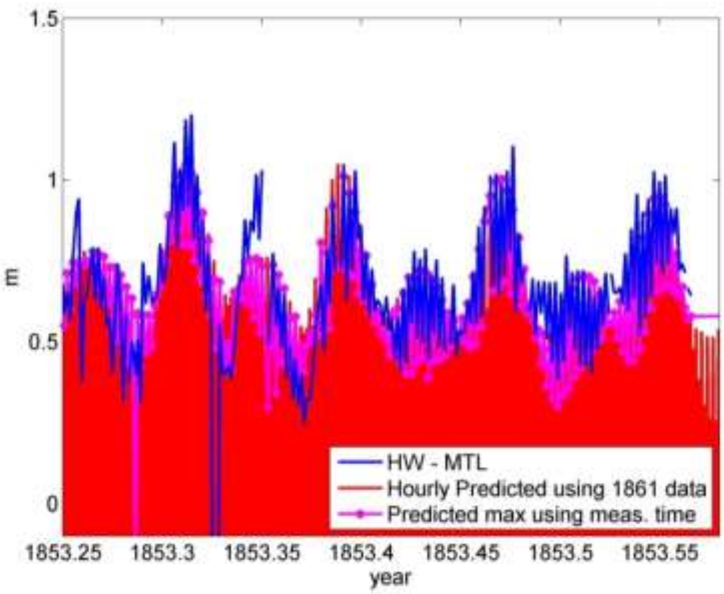


Fig. S13: Measured HW at Governors Island in 1853 after removing the Mean Tide Level (blue), compared against hourly predictions (red) and the predicted value at the tabulated time of HW (magenta). A large difference between the peak red and magenta values may indicate timing errors (i.e., the measured HW time was offset from the predicted HW time).

more from expectations (see Fig. S16). A careful look at Figs. S14-S16 indicates that timing errors also occurred. The most likely cause of tidal irregularities pre-1853 are measurement errors; however, the effect of deepening of Hells Gate beginning in 1851 (causing greater connectivity to Long Island Sound; Marmer, 1935) cannot be discounted.

Despite the gauge issues pre-1853, peaks are clearly visible in the tabulated data and can be related to individual storm events (e.g., compare note on Fig. S.3 with Fig. S16). There is some value therefore in tabulating these values, with the recognition that these approximations could be better constrained (in some cases) by further archival research (e.g., by recovering additional Sandy Hook data from the 1840s or by qualitative accounts). The reduced variance observed in pre-1853 data suggest that the stilling-well used in the box-gauge was over-damping the measurements. However, we have as yet no good explanation for the elevated tidal range; it does not appear to be a nodal cycle issue (again, it is possible that deepening of Hells Gate beginning in 1851 altered NY Tides). Since the tidal values are both over-damped and offset upwards, additional information about how the measurements were made is necessary to make bias corrections.

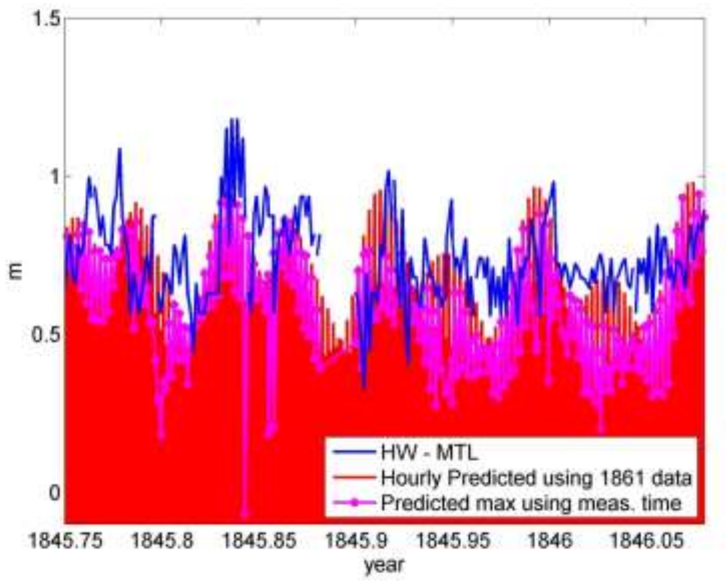


Fig. S14: Measured HW at Governors Island in 1845 after subtracting out the Mean Tide Level (blue), compared against hourly predictions (red) and the predicted value at the tabulated time of HW (magenta). A large difference between the peak red and peak magenta values may indicate timing errors (i.e., the measured HW time was offset significantly from the predicted HW time). Note that the spring-neap cycle variance is not well represented in the historical data.

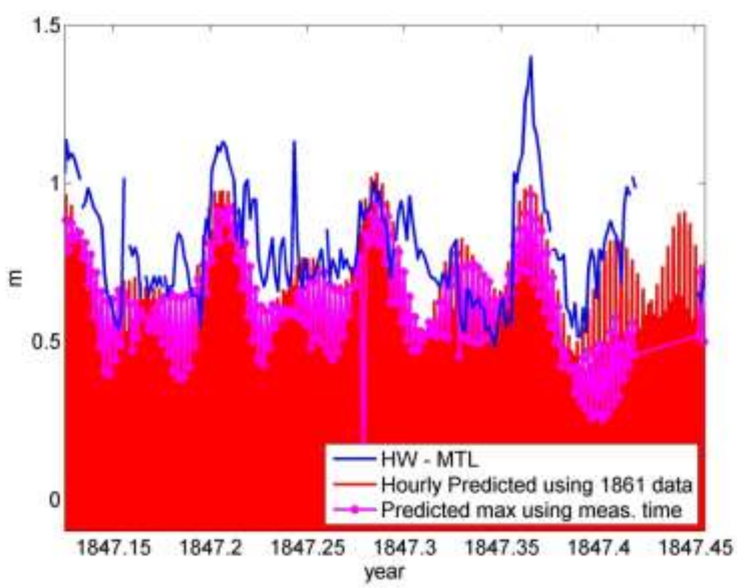


Fig. S15: Measured HW at Governors Island in 1847 after subtracting out the Mean Tide Level (blue), compared against hourly predictions (red) and the predicted value at the tabulated time of HW (magenta). A storm tide is observed around t = 1847.35. A large difference between the peak red and peak magenta values may indicate timing errors, e.g., around t=1847.4.

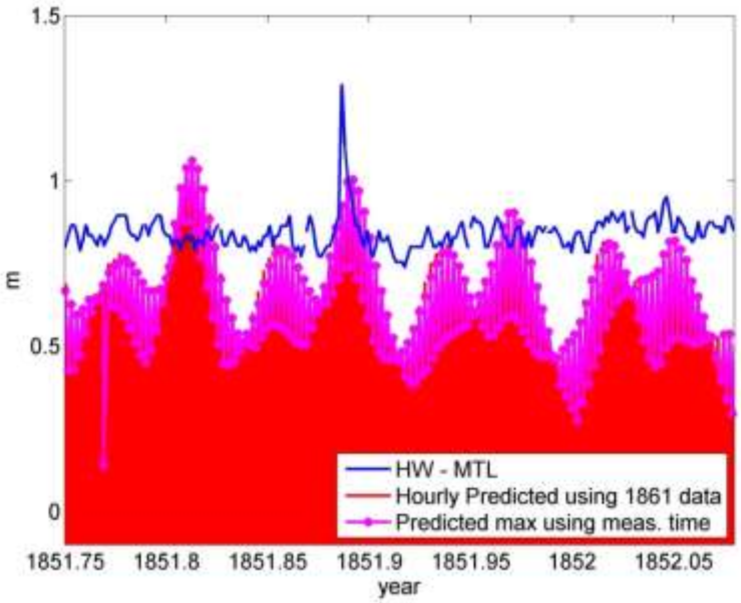


Fig.. S16: Measured HW at Governors Island in 1851 after subtracting out the Mean Tide Level (blue), compared against hourly predictions (red) and the predicted value at the tabulated time of HW (magenta). Note that the spring-neap cycle variance is poorly represented in the historical data, although the peak storm-tide on Nov. 21st, 1851 is clearly measured.

S3: Notes on analysis

The annual maximum storm tide (AMST) is defined by the maximum water level measured for a year, minus the annual mean sea level (MSL). Our definition of a storm tide is therefore the sum of the astronomical tide and the meteorological surge, similar to the definition by NOAA. The MSL is estimated either by (a) hourly spaced data; (b) the tabulated NOAA sea-level available at PSMSL back to 1856 and (c) the Mean Tide Level (MTL) before 1856. The MTL, which is calculated from the annual average of High/Low extrema, was determined by the US Coast and Geodetic survey to be $MTL + 0.08 = MSL$ at the Governors Island tide gauge. The annual maximum storm surge (AMSS) is defined by subtracting the annual mean and the predicted tide from hourly tide data, available for the following years: 1860-1885, 1889-1921, and 1927-present.

Climate indices

The North Atlantic Oscillation (NAO) index was put together as follows: from 1899-present, we used the PCA analysis from Hurrell (2013); from 1864-1899, we use the station data from Hurrell (2013); and pre 1864, we use the Jones et al. (1997) reconstruction.

For comparison, we plot in Fig. S17 the de-trended quartile difference (QD) and the de-trended standard deviation (σ_{AMST}) against three climate indices: The NAO, the Artic Oscillation (AO), and the multivariate El Niño/Southern Oscillation index (MEI). All data shown are a 36y running average. The AO index is a composite drawn from the following websites: http://jisao.washington.edu/data/aots/aojfm18992002.ascii (1899-2002) and http://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/monthly.ao.index.b50.current.ascii (2003-2013). Similarly, the MEI is a composite of http://www.esrl.noaa.gov/psd/enso/mei.ext/table.ext.html (1871-2005) and //www.esrl.noaa.gov/psd/enso/mei/table.html (2006-2013). The AO index between 2003 and 2013 is scaled by 1.6, to take into account the median ratio between the two indices.

The results in Fig. S17 show that the AO index and the NAO index are virtually indistinguishable over their period of record, and are therefore both anti-correlated with AMST variability. No obvious relationship is observed between the MEI index and AMST characteristics.

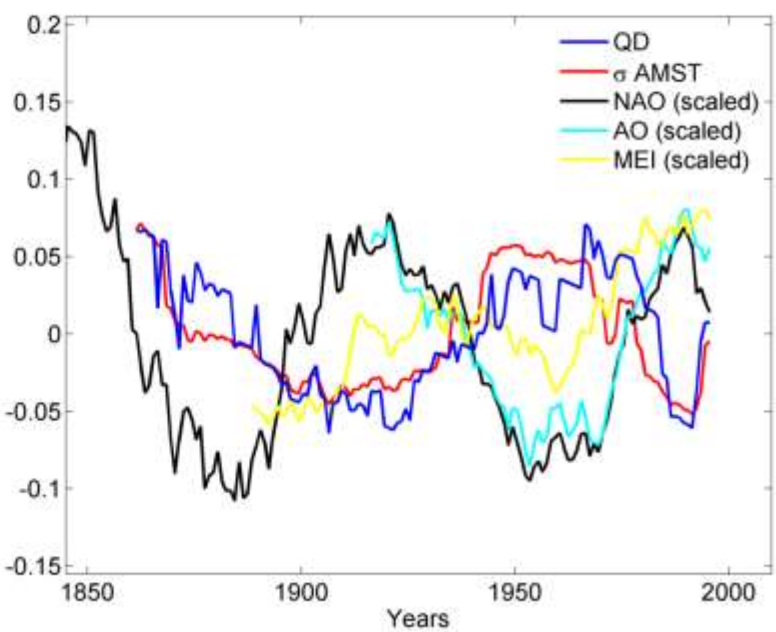


Fig. S17: Comparison of the NAO, AO, and MEI indices and their relationship to the de-trended AMST quartile difference (QD) and the de-trended standard deviation of AMST. All indices have been divided by 4 for plotting purposes.

Generalized Extreme Value Analysis

The Generalized Extreme Value (GEV) distribution used in the manuscript has a cumulative density function defined by the following equation, where μ is the location parameter, k is the shape parameter, and σ>0 is the scale parameter:

$$y = f(x|k,\mu,\sigma) = \frac{1}{\sigma} exp\left(-\left(1+k\frac{(x-\mu)}{\sigma}\right)^{\frac{-1}{k}}\right)\left(1+k\frac{(x-\mu)}{\sigma}\right)^{-1-\frac{1}{k}}, \quad \text{(S1)}$$

for $1+k\frac{(x-\mu)}{\sigma} > 0$

This function is typically applied to evenly spaced annual time series (e.g., Sweet et al., 2013), and the method of maximum likelihood estimation (MLE) is used to estimate the parameters μ, k

and σ. Confidence intervals were generated by resampling data using a bootstrapping technique and recalculating return period statistics. The standard deviation of results is used to plot the 2σ envelope of 5y and 10y AMST levels in Fig. 4 (main text). We compare the estimate of the return period based on Eq. S1 to an unbiased return period estimated calculated as follows:

$$T_R = \frac{1}{1-P},$$

Where $P = \frac{m}{n+1}$, n= # events, and m is a ranked order of integers from 1 to n, corresponding to the smallest (m =1) and largest (m = n) events.

Four examples of the GEV fit and its comparison to the unbiased return period are given in Figure S18 below. As can be seen, the GEV fits the data well, particularly at the 5 and 10 year return intervals which are the focus of the manuscript. At larger return intervals, the 95% confidence intervals become quite large, indicating that these estimates are less certain.

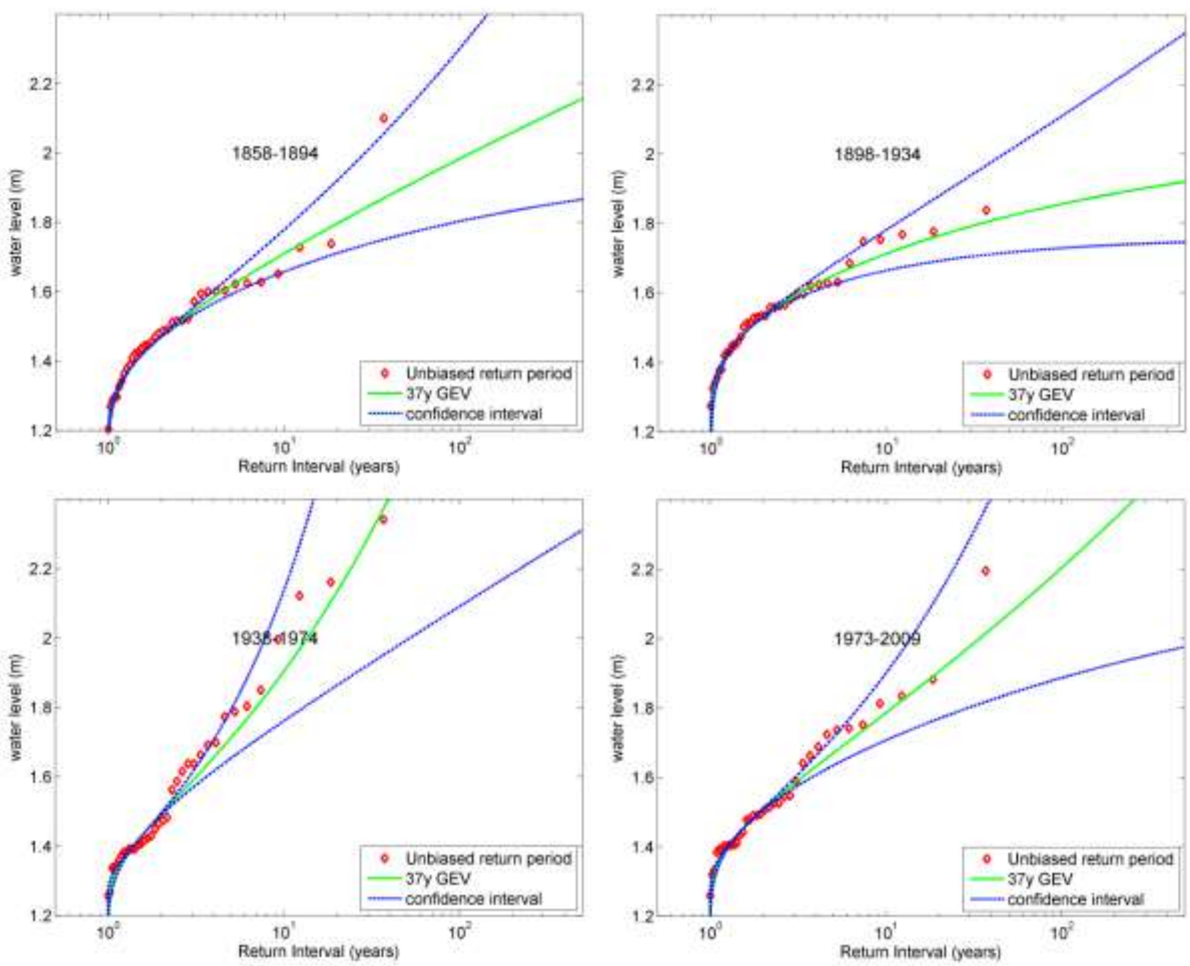


Figure S18: A comparison of the GEV analysis and an unbiased return interval analysis for the 37 year data sets evaluated in the manuscript.

AMSS analysis

Figure S19 below shows the change in the annual maximum storm surge (AMSS) as a function of time. Since the mid-19th century, the magnitude of the 10 year storm surge has increased from 1.27m to 1.62m. This increase is consistent with the 0.28m increase in the magnitude of the 10 year storm tide over a similar time period. Due to gaps in hourly data, the analysis in the 19th and early 20th century is not continuous.

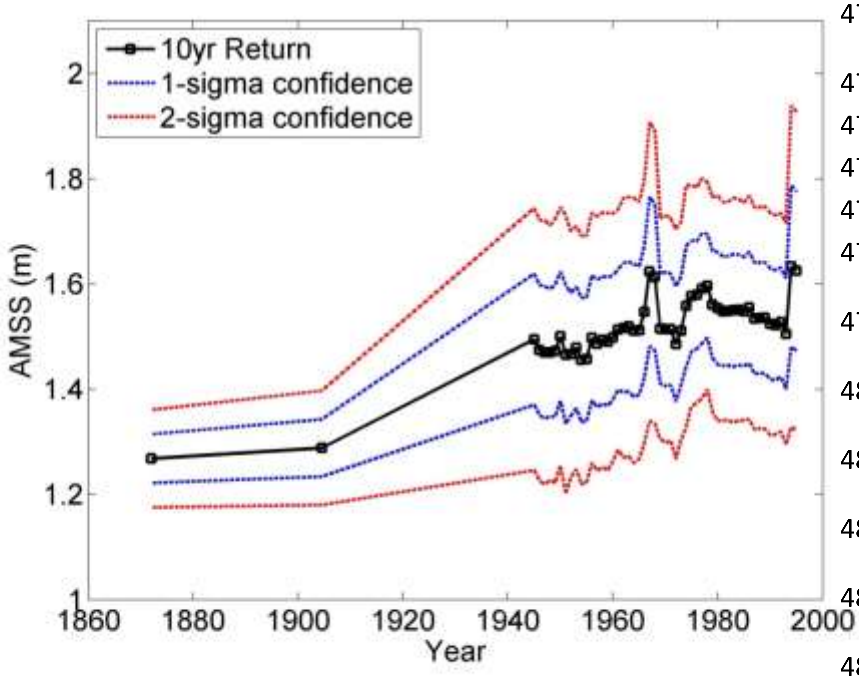


Figure S19: Variability of the 10 year return period for the annual maximum storm surge (AMSS).

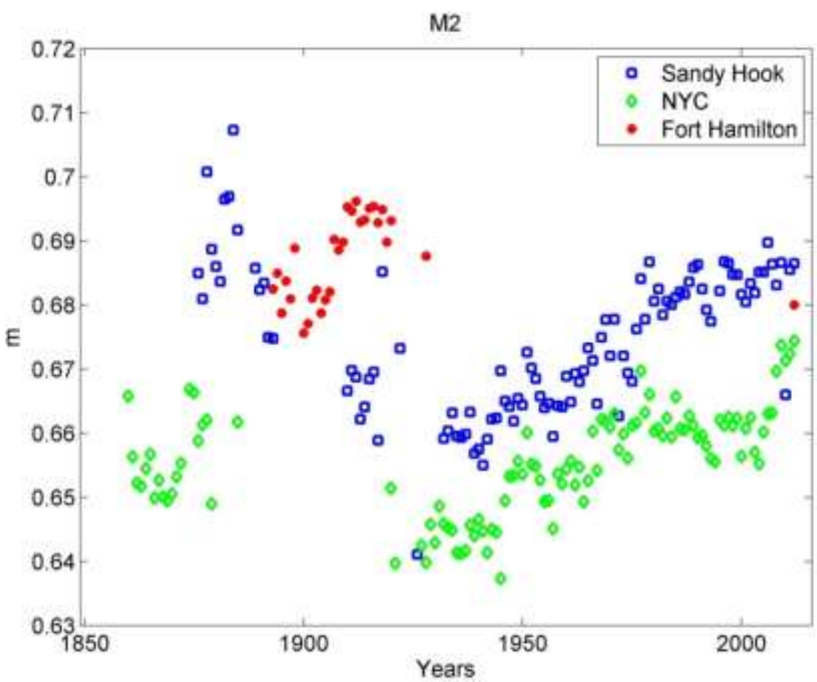
Correlation analysis

We estimate the correlation between the North Atlantic Oscillation (NAO) and the quartile statistics using a bootstrapping method. In our approach, we first divide the AMST data series into 4 independent regions that are separated by more than 35 years (recall that each point has been filtered over 36 data points). We next select NAO/Quartile pairs from each of these four groups using a random number generator, and regress the points against each other. The procedure is repeated until the mean statistics converge (usually after > 1000 experiments). From these results we extract the mean and median R and p-value reported in the text.

Tidal behavior and trends in New York Harbor

To explain why AMST data between Sandy Hook and The Battery are nearly equivalent, we next estimate tidal constituents in New York harbor using harmonic analysis (The Matlab program r_t_tide of Leffler & Jay, 2009). Results show that nodally corrected tidal constituents are nearly constant within New York harbor (Fig. S20). In the 19th century, the M2 constituent decreased by an average of 0.03m between Sandy Hook and Governors Island (from 0.688m to 0.656m), whereas between 1993 and 2012 the M2 constituent decreased by ~0.015m (from 0.678 to 0.663m) between Sandy Hook and The Battery (Fig. S14). The M2 constituent at Fort Hamilton was slightly larger than at Sandy Hook in the early 20th century, and is now slightly less (the modern estimate is from NOAA air-gap height data from the Verrazano Narrows bridge). Note that the M2 constituent varies from approximately 0.63 to 0.69m at The Battery over an 18.6 y cycle (not shown). Hence, tidal range variability over the 18.6y cycle is more than the secular changes observed in Fig. S16. The large nodal variability made it important to analyze AMST return period probabilities using an integer number of nodal periods.

Because the primary component of the tide (M2) in New York is nearly constant, any given storm surge between Sandy Hook and The Battery will occur at a nearly equal tidal phase and amplitude. Moreover, storm surge and tide-waves are both shallow water waves and are likely affected similarly by friction and convergence in estuaries. By analogy with the tides, storm surge attenuation is likely to be small, and is therefore not that surprising that AMST values are nearly equal between Sandy Hook and Battery Park (see Fig. S8).



.

Fig. S20: Long-term variation of the M2 tidal constituent near NYC (Governors Island, Battery Park) since 1860, Sandy Hook, NJ since 1876, and Fort Hamilton (Verrazano Narrows) since 1893. A nodal correction has been applied.

The M2 constituent is not constant over time, but is now (slightly) larger at NYC/Governors Island than in the 19th century, after reaching an early 20th century minimum. Sandy Hook tides exhibit the same minimum in the early 20th century; the M2 estimate has increased since 1930, and is now approximately equal to the 19th century values (though large variability is apparent in the 19th century). Between the Sandy Hook and Manhattan Island areas, the M2 difference decreased from 0.03m to 0.015m. These results suggest that friction is dissipating both tides and storm surge slightly less than in the past. However, trends are not spatially homogenous, as also observed in Marmer (1935): Fort Hamilton tidal constituents and AMST values seem to increase between the 1890s and the 1920s (see Figs. S9, S16).

S4: Specific Notes on the AMST time series

The following adjustments were made to construct the AMST tide series, in addition to the adjustments made to the Schureman(1934) data discussed in section S2.

1844-1859: For data before 1860, mean sea-level was estimated by the formula: MTL + 0.08 = MSL, where 0.08ft is the historical adjustment applied by the United States Coast and Geodetic Survey to convert the average of HW and LW (the Mean Tide Level, or MTL) to mean sea-level (MSL).

1860-1885, 1927-2013: An upwards adjustment of 0.008m was applied to all AMST values based on hourly tide data, to take into account the median difference between hourly maxima and extremes (see Fig. S12).

1873, 1878-1885, 1958, 1974, 2001, and 2004: For years in which Sandy Hook tide data was used, a bias correction of -0.052m was applied to the Sandy Hook AMST values to take into account the median difference between NYC and Sandy Hook storm tide magnitudes.

1893-1926: For these years, the AMST values for Fort Hamilton and The Battery were averaged.

Notes for specific years:

1851: Observer states that the Nov. 21 HW was the largest since he started measurements (likely in 1850). The gauge was not working properly (see Fig. S15), so it is unclear how accurate this AMST value is. Since the datum appears to be shifting in the first half of the year, MTL was calculated based on the September-December time frame.
1852: The 1852 AMST value is based on automatic tide gauge data from Dec. 1852; though no hand tabulated data from 1852 exceeded this value, the box gauge from Jan. to Nov. was most-likely not working correctly. Therefore, this peak value is also somewhat uncertain.
1854: Adjustments were necessary on high/low data due to datum shifts that occurred on Mar. 22 and May 23 after adjustments to the gauge chain. The average of gauge checks was used to reduce the entire year to a common datum.
1858: The New York Times from 10/26/1858 quotes a dockworker who states that water levels were higher than in at least 14 years. This report is consistent with our AMST data set.
1861: The AMST value for the 'Expedition' hurricane in Nov. 1861 was obtained from the Brooklyn tide gauge, since the measurements during the storm peak are missing from the Governors Island tide gauge. Presumably, the pencil 'ran off the marigram chart'—this was a

common occurrence during storms, unless the tide observer specifically shortened the chain, thereby moving the tidal trace lower on the paper roll.

1862: This AMST value was estimated based on note from Brooklyn observer. Since the HW occurred around 4:00 am, the value was not recorded on the Brooklyn gauge. The automatic gauge at Governors Island was also not recording during this period, probably due to ice. HW was reported to be higher than the 11-2-1861 AMST; since it is unclear by how much (if any), we set the AMST value approximately equal to the 1861 value.

1865: The 1865 AMST was based on a Mar. 22, 1871 NY Times article, which states that the Nov. 21, 1865 high water was "5 ft. above the usual (ordinary) high tide'. It is presumed that the New York Times value comes from the Brooklyn tide gauge or another tide-gauge which we have not recovered. Assuming that the 'usual high tide' is mean high water, and adding 5 feet, we obtain an AMST of 2.1m. This large value is plausible if a storm surge of 1.38m occurred at the predicted HW of +0.72m. A recording gap of 5 hrs occurred on the Governors Island gauge during the peak of the storm. Since we estimate a decreasing surge (Storm-tide – predicted water level) of 0.8m at 3hrs after HW, and an increasing surge of 0.5m 2hrs before HW, we conclude that a 1.4m surge is possible. Further archival research is needed to confirm this estimate.

1867: Gauge was not working in late January due to ice. The AMST was therefore estimated based on a Jan 22nd, 1867 NY Times article, which stated that the high tide was one foot higher than any other in the past several years. We therefore add 1 foot to the AMST value for 1866 to estimate the 1867 value.

1868: The HW value for March 2, 1868 was missing, and probably 'ran off the chart'. We linearly interpolate over the storm surge (measured-predicted tide) to estimate the surge for the 2 missing hourly values, and add the predicted tide to estimate the AMST value.

1871: The Mar. 22, 1871 NY Times article mentions flooding at HW that is 3 feet above ordinary high water. However, no anomalous storm-tide was observed in the Governors Island data.

1871: The HW value for the Nov. 14, 1871 storm tide is missing, and probably 'ran off the chart'. The values +/- 1 hr are available, and suggest that the storm surge increased from 0.76 to 1.26 m between these two data points. The predicted tide increased by 0.03m from *t*=-1hr to t=0hr at HW. We linearly interpolate to estimate the storm surge at HW and add this value to the predicted tide to obtain the AMST.

1873: No Governors Island data was recovered for 1873, and the largest peak measured between July-Oct at Sandy Hook was a biased corrected 1.38m on Oct. 10th. The Oct. 20th storm-tide was not measured, but no evidence of flooding was found in newspaper archives. However, the storm on April 12th caused flooding in cellars along the water front (New York Times, April 13 1873). Similar flooding occurred in 1858 (AMST =1.63m), 1867 (estimated 1.6m), 1870(1.44m), 1871 (1.74m), 1861 (1.62m), and 1862 (estimated 1.63m), and 1865 (estimated 2.1m). Besides 1870, we have found no other archival evidence for flooding for AMST<1.6m. Reported flooding was not as bad as 1865 or 1871. Therefore, we assume the Apr. 1873 storm tide was approximately 1.6m, to be checked against data if/when more is recovered.

1878: This maximum was based on a value given in Sandy Hook ancillary data (peak was missing from both Gov. Island and Sandy Hook hourly data).

1881 The AMST for this year on Mar. 3rd was adjusted up by 0.1m due to two missing data points at peak surge. The surge linearly decreased by 0.01m per hour, while the tide rose 0.11m and then fell, according to predictions. A large rainstorm and high-tide caused flooding in New York City On Mar. 19th, but only produced a surge of 1.29m at Sandy Hook.

1882 Three different gauges were used at Sandy Hook in 1882, due to the Sept. 1882 hurricane that destroyed the gauge: Jan-Aug, mid-Sept to Oct., and Nov-Dec.. Each time series was de-meaned separately. Not enough information was found to estimate the hurricane storm-tide. However, no wide-spread evidence of catastrophic flooding or damage was found in newspaper archives, so we tentatively presume that the storm-tide for the hurricane was not especially large.
1883 The AMST for this year was adjusted up by 0.1m due to two missing data points at peak water level on January 10th. By interpolation we estimated a linear increase in surge of 0.14m/hour, while the predicted tide fell 0.04m from HW in the first hour and then more thereafter.
1884: The 19th century 'computer' who reduced the Sandy Hook 1884 data from the marigram sheet estimated the peak value on February 27th. The data +/- 1 hour are fine, suggesting the pencil ran off the marigram sheet.
1885: The Sandy Hook data set for this year was adjusted in monthly increments by up to 0.3m to match the monthly sea-level from Governors Island measurements, which were made between April and October. It is known from archival notes that the tide observer between 1883-1885 made few gauge checks, with the result that it was difficult to relate the marigram trace to the correct datum during the 'reduction' process.
1895: The date and value of the maximum extreme tabulated by Schureman (1934) does not agree with hourly tabulated data. The Schureman(1934) extreme value is 9.7ft above Sandy Hook MSL, on 1/26/1895. The tabulated data suggests a peak of 9.6 ft. on 12/12/1895. We use the Schureman (1934) value, since it is probably based on HW.
1900: The tabulated hourly extreme for Fort Hamilton on Dec. 4th is 0.4 ft. larger than the value tabulated by Schureman (1934) at both Fort Hamilton and The Battery. The hourly data may have been transcribed incorrectly, e.g., '10.9' ft. was entered instead of '10.4' ft. We use the Schureman(1934) data in our composite time series.
1920: The hourly tabulated extreme for Fort Hamilton on Feb. 5th is 0.8 ft. larger than the extreme for Fort Hamilton used in Schureman (1934), and would yield an AMST of 1.896m. Unclear what is causing the discrepancy; However, the hourly data were already revised downwards from a tabulated '12.6' ft. to a tabulated '12.4' ft., suggesting some uncertainty in the data reduction. Given the data issue, we use the tabulated extreme from The Battery.

Acknowledgements

The authors would like to acknowledge Pat Lau for his help in photographing the data used in this manuscript, as well as the following students for helping to digitize the data: Michelle Yuen, Muhanned Al-Murib, Jake Coppola, Mike Waxter, Claudia Estrada, Conrad Hilley, Zhara Kavianpour, Lauriane Wales, and Kamal Ahmed.

References

Agnew, D.A., 1986, Detailed analysis of tide gauge history: A case analysis, *Marine Geodesy,* 10, Issue 3-4. **DOI:** 10.1080/01490418609388024

Hurrell, J. & National Center for Atmospheric Research Staff (Eds). Last modified 08 Oct 2013. The Climate Data Guide: Hurrell North Atlantic Oscillation (NAO) Index (PC-based) Retrieved from https://climatedataguide.ucar.edu/climate-data/hurrell-north-atlantic-oscillation-nao-index-pc-based.

Jones P.D., T. Jonsson and D. Wheeler (1997), Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and South-West Iceland. *Int. J. Climatol.* 17, 1433-1450.

Koop, F.W., 1915. *Precise Leveling in New York City*, City of New York, Board of Estimate and Apportionment, Office of the Chief Engineer. 275pp.

Leffler, K. & D. A. Jay, (2009), enhancing tidal harmonic analysis: Robust (hybrid L1/L2) solutions, *Continental Shelf Research* 28, 10.1016/j.csr.2008.04.011.

Marmer, H.A., 1935. Tides and Currents in New York Harbor. *U. S. C.& G. S. Special Publication #111*,. Rev. ed. Washington, D. C. 198 pp.

Schureman, P., 1934. Tides and Currents in Hudson River, *Coast and Geodetic Survey*, *Special Publication No. 180*, United States Government Printing Office, 106p.

Scileppi, E., & J. P. Donnelly (2007), Sedimentary evidence of hurricane strikes in western Long Island, New York, *Geochemistry, Geophysics, Geosyst*ems,,8, Q06011, doi:10.1029/2006GC001463.

Sweet, W, Zervas,C., Gill,S. & Park,J (2013). Hurricane Sandy Inundation Probabilities Today and Tomorrow, *Bull. Amer. Meteor. Soc.* **94 (9),** S17-S20.