A probabilistic framework for homeowner flood insurance cost-benefit analysis under rising seas

D. J. Rasmussen*

CEE 460, Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ and Woodrow Wilson School of Public & International Affairs, Princeton University, Princeton, NJ

**Corresponding author address:* Woodrow Wilson School of Public & International Affairs, Princeton University, Princeton, NJ
E-mail: dj.rasmussen@princeton.edu

1. Introduction

Local sea level rise (LSLR) over the coming decades is expected to substantially alter the flood insurance and coastal real estate market (e.g. Houser et al. 2015; Abel et al. 2011; Bin et al. 2011). It is no longer sufficient for buyers of homes in coastal regions to assume that sea levels and flood risk in their region will remain constant over the typical multi-decade ownership period, including the three decade duration of a typical home mortgage. A home is often the biggest financial investment that an individual will make in their lifetime. As such, buyers of homes in coastal regions could substantially benefit from being better informed of the financial risks stemming from evolving local natural hazards, such as LSLR.

While several climate change assessments provide projections of future sea level rise to inform financial decision making, these projections generally do not 1) provide probabilistic estimates of local physical climate variables, such as LSLR 2) provide time frames of when particular levels of sea level rise will occur and 3) estimate the probability of flood events over time.

This study presents a probabilistic risk assessment framework for determining the costs and benefits of purchasing flood insurance, and applies it to a $725,000 home in Linwood, New Jersey. The home is currently located in a minimal flood risk area, but the flood risk is expected to increase over the 21st century due to factors that include LSLR. A cost-benefit analysis is done to asses whether or not the homeowner should purchase insurance at present at a low rate from the National Flood Insurance Program (NFIP), or if they should not purchase insurance and instead pay for flooding damages throughout the home ownership period. This framework, and the associated computer code (attached in the appendix), could be applied to any structure in a region vulnerable to sea level rise to characterize future flood risk.

2. Data and Methods

Following the approach used by Tebaldi et al. (2012), this study's analysis begins with estimating the historical rate of LSLR at the tide gauge nearest to the home. Hourly tide gauge data are detrended using the local rate of LSLR to isolate historical storm events and astronomical tidal components. Extreme value theory is used to estimate annual return levels for flood heights in the current climate. Probabilistic projections of relative sea level rise (RSL; annual sea level rise relative to the year 2000) from Kopp et al. (2014) are used to generate future return levels for flood heights. Finally, both empirical depth-damage functions and the projected flood heights are used to calculate expected losses over the homeownership period and are compared to the costs of paying flood insurance premiums throughout the same time frame. An overview of the risk assessment framework used in this study is shown in Fig. 1.

a. Current trends and extreme value analysis

1) CURRENT TRENDS AND HISTORICAL STORM TIDE

Hourly tide gauge data observed at the nearest tide gauge station to the home in Linwood, New Jersey are obtained for a tide gauge at Atlantic City, New Jersey (Station ID: 8534720) from the University of Hawaii Sea Level Center (UHSLC) (retrieved `http://uhslc.soest.hawaii.edu/`, April 2016). The UHSLC maintains a database of sea level observations from a global network of tide gauges. Data from the UHSLC have undergone a thorough quality control process. This study's hourly values are adjusted so that they are relative to the historical mean high water (MHW) datum at the tide gauge[a]. Hourly data for the period 1911–2015 are aggregated to annual means. The trend in annual mean sea level at Atlantic City from 1911 to 2015 is shown in Fig. 2. The rate of LSLR is estimated to be 4.1 mm yr^{-1} and is roughly constant throughout the 105 year period. The remainder of the historical storm tide analysis uses hourly data for the period 1986–2015. This 30-year record is long enough to include multiple storms, weather, and tidal events that capture historical storm tide variability. It is also recent and short enough to lessen the likelihood of capturing any change in the frequency and severity of weather and tidal events that could occur over centuries(e.g., Tebaldi et al. 2012). The historical tide gauge record used is nearly complete with $< 3\%$ of the hourly values missing.

To isolate the effects of historical weather and tidal activity, the hourly time series is detrended using the long-term trend calculated with ordinary least squares:

$$\tilde{y}_i = y_i - \hat{\alpha}_1 t_i$$

where $\tilde{y}$ is the detrended value, y_i is the original value, $\hat{\alpha}_1$ is the estimated regression coefficient from ordinary least squares (i.e. the trend), and t_i is the time period. The detrending preserves the MHW level over the current National Tidal Datum Epoch. Daily maximum water levels are then calculated from the detrended hourly values.

2) EXTREME VALUE ANALYSIS

Extreme value analysis (EVA) is employed to assess flood height return levels. EVA is a statistical technique that is used to estimate the occurrence of events that are theoretically too rare to be found in observation records (Coles 2001). Specifically, a peak-over-threshold approach is applied to model the tail of the historical storm tide distribution with a generalized Pareto distribution (GPD). Following Tebaldi et al. (2012), the 99th percentile from the historical distribution of

---

[a]The Mean High Water (MHW) value is calculated as the average of all the high water heights occurring over the National Tidal Datum Epoch, currently 1983–2001 for the Atlantic City tide gauge station. For the current National Tidal Datum Epoch, the MHW for Atlantic City is 2.787 meters.

observations (0.6 m above MHW) is chosen as the threshold (Fig. 3). Since multiple consecutive measurements could be associated with the same event (i.e. a storm), values above the threshold are declustered using the *extRemes* module for R. In the case of successive exceedances, the declustering algorithm only uses the maximum value of the successive exceedances. Ideally, this helps assure statistical independence between observations.

The parameters of a GPD are fit to the declustered daily maximum observations using maximum likelihood estimation. The GPD models the distribution of the threshold exceedances:

$$P(z-\mu \leq y|z>\mu) = \begin{cases} 1-(1+\frac{\xi y}{\sigma})^{-\frac{1}{\xi}} \text{ for } \xi \neq 0^{\mathrm{b}} \\ \\ 1-\exp\left(-\frac{y}{\sigma}\right) \text{ for } \xi = 0^{\mathrm{c}} \end{cases}$$

where μ is the threshold above which exceedances are counted, and σ and ξ are the scale and shape parameters, respectively. The shape parameter determines the decay rate of the distribution's upper tail. A shape parameter of zero (ξ = 0) leads to an exponential distribution, a shape parameter greater than zero ($\xi > 0$) leads to a heavy tailed distribution, and a shape parameter greater than zero ($\xi < 0$) leads to a bounded distribution. The GPD parameters for the Atlantic City tide gauge location are given in Table 1. The shape parameter is approximately zero, leading to an exponential distribution.

The associated return levels for water height above MHW, x_T, for a given return period are calculated with:

$$x_T = \begin{cases} u + \frac{\hat{\sigma}}{\xi}\left\{\left[-\frac{\lambda\zeta_u}{\ln(1-\frac{1}{T})}\right]^{\xi} - 1\right\} \text{ for } \xi \neq 0 \\ \\ u - \hat{\sigma}\ln\left[-\frac{\ln(1-\frac{1}{T})}{\lambda\zeta_u}\right] \text{ for } \xi \neq 0 \end{cases}$$

where u is the GPD threshold of 0.6m, $\hat{\sigma}$ is the shape parameter, T is the return period (years), ζ_u is the probability of exceeding the threshold, ξ is the shape parameter, and λ is the number of events per year (in this case, 365 because daily data are used). Fig. 4 shows the historical return levels for Atlantic City.

---

[b] $y > 0$ and $(1+\frac{\xi y}{\sigma}) > 0$
[c] $y > 0$

b. Sea level projections

Kopp et al. (2014) makes available probabilistic sea level projections for the time slices of 2030, 2050, and 2100. Probability distribution values in between these periods are interpolated with a spline method. The probabilistic RSL projections for select quantiles for Atlantic City are plotted in Fig. 5. The climate forcing scenario RCP 8.5 is used and is analogous to a 'business as usual' greenhouse gas emissions scenario. Future return levels of flood heights at the Atlantic City tide gauge are generated using 10,000 Monte Carlo samples from the Kopp et al. (2014) distributions of RSL and the extreme value distribution of historical storm tide:

$$\mathrm{P}\{\mathrm{H}^f \leq \mathrm{h}\} = \mathrm{P}\{\mathrm{H} + \mathrm{S(t)} - \leq \mathrm{h}\} \Rightarrow \mathrm{P}\{\mathrm{H}^f > \mathrm{h}\} = 1 - \mathrm{P}\{\mathrm{H} + \mathrm{S(t)} - \leq \mathrm{h}\}$$

where H^f is the flood height, H is the storm tide (assumed to be time invariant), and S(t) is the RSL for year t. Although some studies suggest that future tropical cyclone (TC) intensity may increase in a warming climate (e.g. Knutson et al. 2010), in this study, TC characteristics is assumed constant. Both cumulative distribution functions and return periods of water levels for past and future years are presented in Fig. 6.

A flood occurs at the home when the flood height (H^f) is greater than the home elevation (determined from a USGS digital elevation map or DEM). In the case of flooding, the home elevation is a measure of the *vulnerability* component of risk. The elevations in the USGS DEM are all relative to a fixed point called the North American Vertical Datum of 1988 (NAVD88). The home in this study is roughly 3 meters above the NAVD88. The number of flood events per year at the home is determined using 100,000 Monte Carlo samplings from the joint distribution of the RSL projections and GPD distribution of historical storm tide. A large Monte Carlo sample size is used to capture the very low probability of flooding for the first two decades in the 21st century.

$$\mathrm{P}\{\mathrm{H}^f \leq \mathrm{E}_{elev}\} = \mathrm{P}\{\mathrm{H} + \mathrm{S(t)} \leq \mathrm{E}_{elev}\} \Rightarrow \mathrm{P}\{\mathrm{H}^f > \mathrm{E}_{elev}\} = 1 - \mathrm{P}\{\mathrm{H} + \mathrm{S(t)} \leq \mathrm{E}_{elev}\}$$

where E_{elev} is the home elevation. The probability of flooding at the home for each year from 2000—2100 is plotted in Fig. 7.

c. Depth-Damage Curves

Perhaps the most useful component of risk information for a decision maker is the *consequence* of the outcome, in this case, the cost of flood damage. To link expected flood events at the home with monetary loss, so-called depth damage functions are used. Depth-damage functions provide estimates of the fractional loss of a structure as a function of the flood depth at the location of the structure. This study uses an empirical depth-damage function that is derived from measurements taken by the U.S. Army Corps of Engineers after a flood event in near by Wilmington, Delaware, approximately 50 miles from the home in this study located in Linwood, New Jersey. A depth-damage function for a two story residential home and its contents is used (Fig. 8), and is retrieved from the *HAZUS* package for R. The depth-damage function provides expected values of the fractional loss for a given flood height, and when multiplied by the home value, gives the estimated expected monetary loss.

d. Expected Annual Damage

Following Lin and Shullman (submitted), this study uses the metric *expected annual damage/loss (EAD)* to calculate the expected monetary loss in a given year. To calculate the EAD for a specific year, one must consider multiple events arriving within that year when water levels exceed the home elevation. As such, the losses from these events are summed:

$$\mathbb{E}[A_i|N_i = n] = E\left[\sum_{i=1}^{N_i} L_i\right] = n\,\mathbb{E}[L_i]$$

where, N_i is the number of occurrences when water levels exceed the home elevation in year i (assumed to be Poisson distributed with mean λ_i) and $\mathbb{E}[L_i]$ is the expected loss estimated using the depth-damage function. In expectation, this becomes the product of the mean Poisson occurrence rate,λ_t, and the expectation of the loss during the year:

$$\mathbb{E}[A_t] = \mathbb{E}[N_t]\,\mathbb{E}[L_t] = \lambda_t\,\mathbb{E}[L_t], \text{ where } t \text{ is the year}$$

Assuming independence between loss events and the absence of risk mitigation measures (e.g. raising the home elevation, neighborhood levee construction), integration over a 30-year time period (i.e. t_1 to t_2) can give the total expected monetary loss:

$$\mathbb{E}[A] = \sum_{i=t_1}^{t_2} \lambda_i\,\mathbb{E}[L_i]$$

3. Results and Discussion

a. Flood Insurance Cost-Benefit

The monetized EAD for each year between 2000 and 2100 is shown in Fig. 9. In the near term, the EAD is less than \$1 yr^{-1}. However, at the end of the 21st century, the annual EAD grows to $>$ \$10,000 yr^{-1}. These amounts are in 2016 U.S. dollars.

An annual flood insurance quote for this property was obtained from *floodsmart.gov*. The home is currently in a FEMA zone with minimal flood risk (zone C), as such, the current premium is relatively low at \$779 yr^{-1} for a \$150,000 policy with a \$2,000 deductible. For comparison, neighboring homes in high-risk flood zones have quoted premiums of over \$10,000 yr^{-1} for the same policy. Assuming the homeowner buys at this present-day low rate and pays the same rate over the period from 2020–2050, the total cost of annual flood insurance premiums would be \$23,370. The integrated expected annual loss over this same period is about \$14. If the homeowner only considers the EAD, they clearly should *not* purchase flood insurance for this property. However, the EAD dramatically increases beyond the middle of the 21st century and could impact the valuation of the home at the end of the ownership period.

b. Cautions

Some import caveats of this assessment are in order. First, only the expected annual damage is considered. Insurance is generally not purchased to cover against expected annual damages, rather it is intended to cover losses from low probability, high consequence (i.e. large financial cost) events. Second, only the *expected* rate of SLR is considered. The actual rate of SLR could be much higher or lower than the expected value that is used in this assessment. Moreover, some recent studies suggest that glacier and ice shelf melt, which are significant contributors to SLR, could occur much faster than perviously thought (e.g. DeConto and Pollard 2016; Hansen et al. 2016), leading to faster rates of SLR over this century. The projections from Kopp et al. (2014) do not consider the findings of these recent studies and therefore could be considered as conservative estimates of SLR. Third, the depth-damage functions do not take account of building fragility, which could greatly affect structure vulnerability, and ultimately the EAD. Lastly, it is unlikely that the NFIP flood insurance premiums will remain constant over a 30-year period. The NFIP is a deeply politicized government program that must be periodically reviewed and re-authorized by the U.S. Congress (Knowles and Kunreuther 2014). The NFIP debt liabilities currently exceed \$20 billion and premiums continue to not reflect true flood risk due to the political unpopularity of placing higher costs on the homeowner.

4. Conclusion

A probabilistic framework for homeowner flood insurance cost-benefit is presented and applied to a \$725,000 home in Linwood, New Jersey that presently has very low flood risk, and, therefore, very low flood insurance premiums. But because of the low elevation and proximity to the coast, the home is vulnerable to future SLR over the remainder of this century. The framework in the presented assessment consists of the following: return levels for current storm tide are constructed from 1) a historical distribution of sea level measured at a tide gauge near the home and 2) extreme value analysis. Current flood height return levels are linearly shifted upward as a function of the expected amount of RSL for each year in the 21st century. Future flood heights and depth-damage functions are used to determine the annual expected damage the homeowner is expected to be burdened with each year, and the costs are integrated over a hypothetical 30-year ownership period (2020–2050).

The assessment concludes that the integrated expected annual damage is effectively negligible compared to the total cost of the insurance premiums over the 30-yr period, suggesting that the homeowner may be better off financially by not purchasing flood insurance. However, several critical caveats regarding this finding are discussed that imply this study's cost-benefit assessment could be improved to better reflect evolving flood risk at the home location.

Acknowledgments. I thank M. Buchanan, A. Hatzikyriakou, and N. Lin for helpful discussion.

References

Abel, N., R. Gorddard, B. Harman, A. Leitch, J. Langridge, A. Ryan, and S. Heyenga, 2011: Sea level rise, coastal development and planned retreat: analytical framework, governance principles and an australian case study. *Environmental Science Policy*, **14 (3)**, 279 – 288, doi:http://dx.doi.org/10.1016/j.envsci.2010.12.002, URL http://www.sciencedirect.com/science/article/pii/S146290111000167X.

Bin, O., B. Poulter, C. F. Dumas, and J. Whitehead, 2011: Measuring the impact of sealevel rise on coastal real estate: A hedonic property model approach. *Journal of Regional Science*, **51 (4)**, 751–767.

Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer.

DeConto, R. M., and D. Pollard, 2016: Contribution of antarctica to past and future sea-level rise. *Nature*, **531 (7596)**, 591–597, URL http://dx.doi.org/10.1038/nature17145.

Hansen, J., and Coauthors, 2016: Ice melt, sea level rise and superstorms: evidence from paleoclimate data, climate modeling, and modern observations that 2 °c global warming could be dangerous. *Atmospheric Chemistry and Physics*, **16 (6)**, 3761–3812, doi:10.5194/acp-16-3761-2016, URL http://www.atmos-chem-phys.net/16/3761/2016/.

Houser, T., and Coauthors, 2015: *Economic Risks of Climate Change: An American Prospectus*. Columbia University Press, New York, New York, 384 pp.

Knowles, S. G., and H. C. Kunreuther, 2014: Troubled Waters: The National Flood Insurance Program in Historical Perspective. *The Journal of Policy History*, **26 (3)**.

Knutson, T. R., and Coauthors, 2010: Tropical cyclones and climate change. *Nature Geosci*, **3 (3)**, 157–163, URL http://dx.doi.org/10.1038/ngeo779.

Kopp, R. E., R. M. Horton, C. M. Little, J. X. Mitrovica, M. Oppenheimer, D. J. Rasmussen, B. H. Strauss, and C. Tebaldi, 2014: Probabilistic 21st and 22nd century sea-level projections at a global network of tide-gauge sites. *Earth's Future*, 383–407, doi:10.1002/2014EF000239.

Lin, N., and E. Shullman, submitted: Dealing with Hurricane Surge Flooding in a Changing Climate: Part I. Risk Assessment.

Tebaldi, C., B. H. Strauss, and C. E. Zervas, 2012: Modelling sea level rise impacts on storm surges along us coasts. *Environmental Research Letters*, **7 (1)**, 014 032, URL http://stacks.iop.org/1748-9326/7/i=1/a=014032.

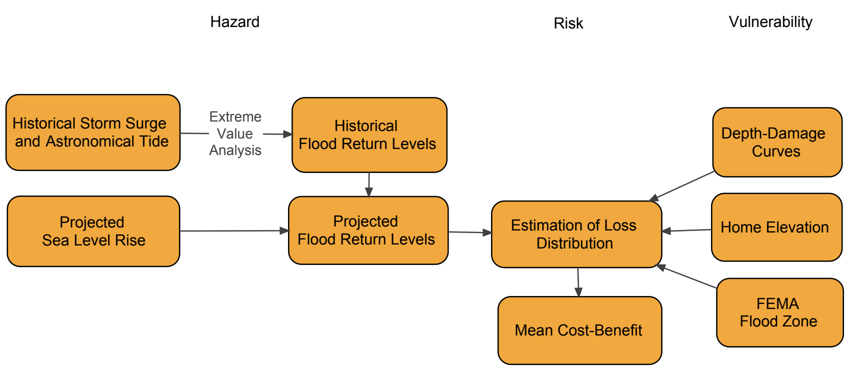


FIG. 1. Flow chart of this study's framework.

TABLE 1. Generalized Pareto Distribution parameters from maximum likelihood estimation for Atlantic City, New Jersey. Values in parenthesis are +/- 2 standard deviations.

Site	NOAA Station ID	Record Length (Years)	$\hat{\xi}$	$\hat{\sigma}$
Atlantic City, NJ	8534720	105	-0.0118 (-0.1471,0.1234)	0.1379 (0.1095, 0.1663)

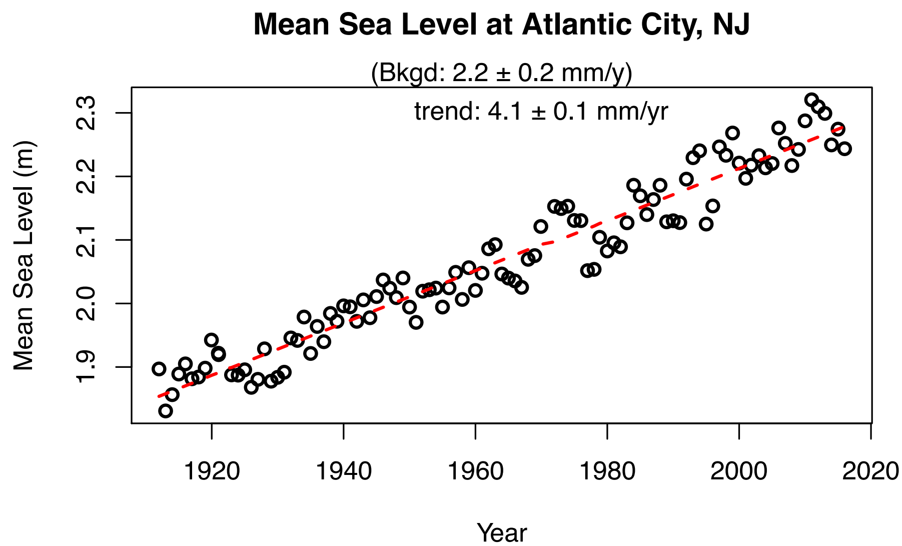


FIG. 2. Annual mean sea level trend at Atlantic City, New Jersey. The background estimate is from Kopp et al. (2014).

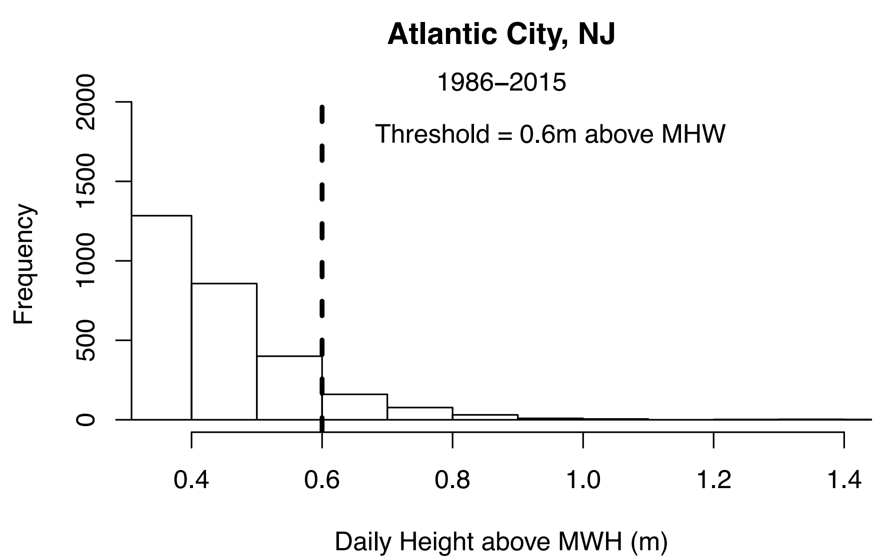


FIG. 3. Histogram of water level above MHW (1986-2015) for Atlantic City, New Jersey and the GPD threshold.

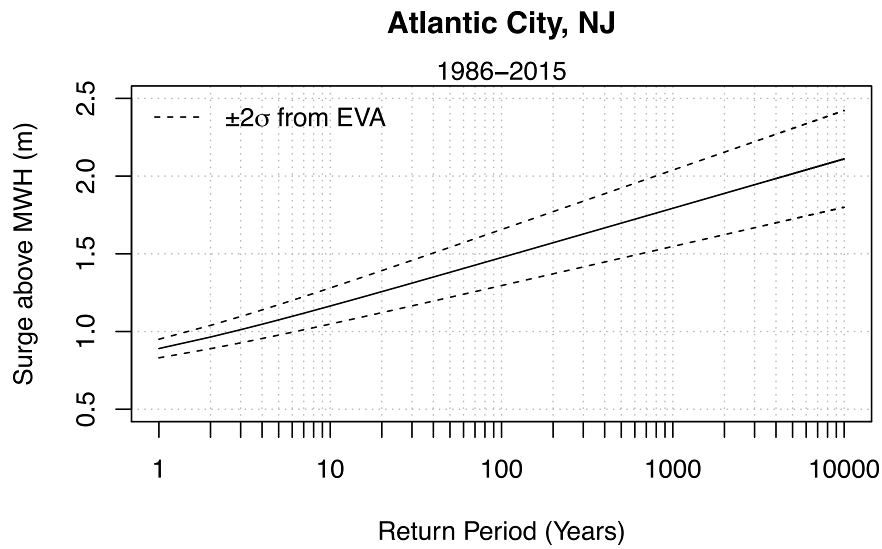


FIG. 4. Historical storm surge return level at Atlantic City, New Jersey. Confidence intervals are +/- 2 standard deviations from extreme value analysis (EVA)

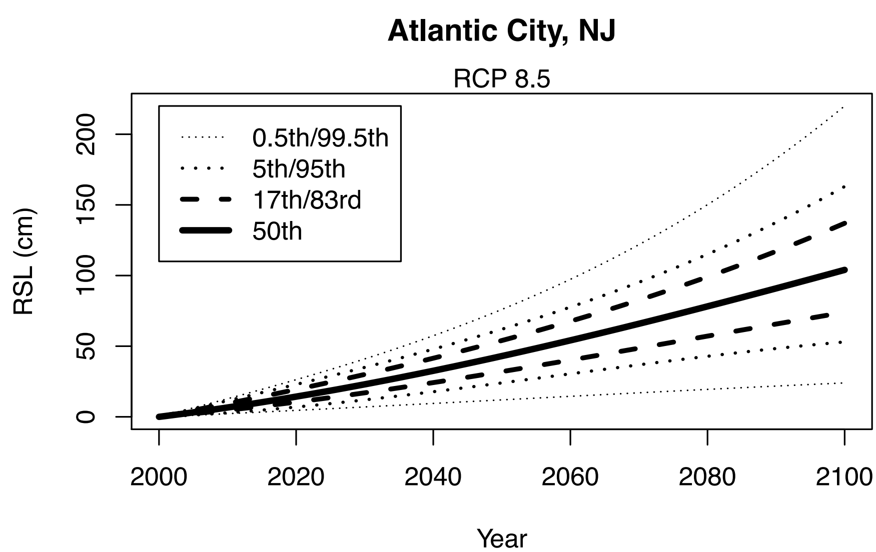
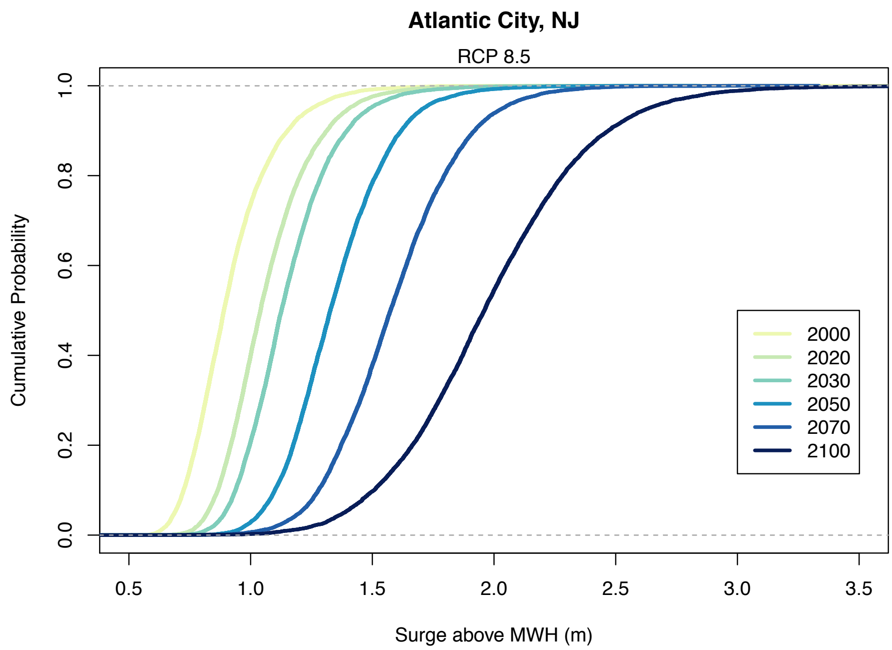


FIG. 5. Probabilistic relative sea level (RSL) trend at Atlantic City, New Jersey



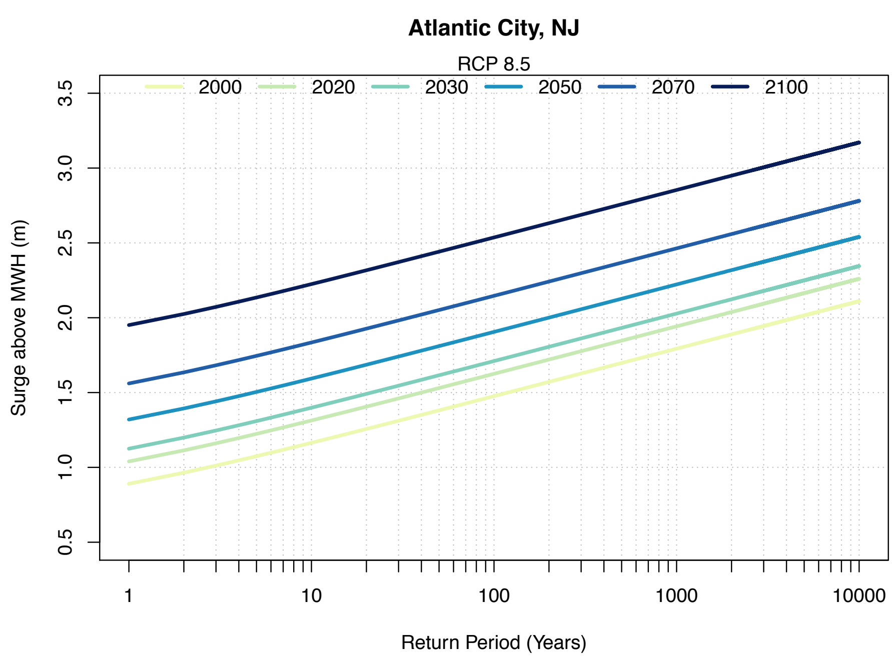


FIG. 6. (Top) CDF of water height above mean high water (MHW) at Atlantic City, New Jersey based on projected relative sea level rise from RCP 8.5 and historical storm surge events; (Bottom) As for Top, but flood height return level curves.

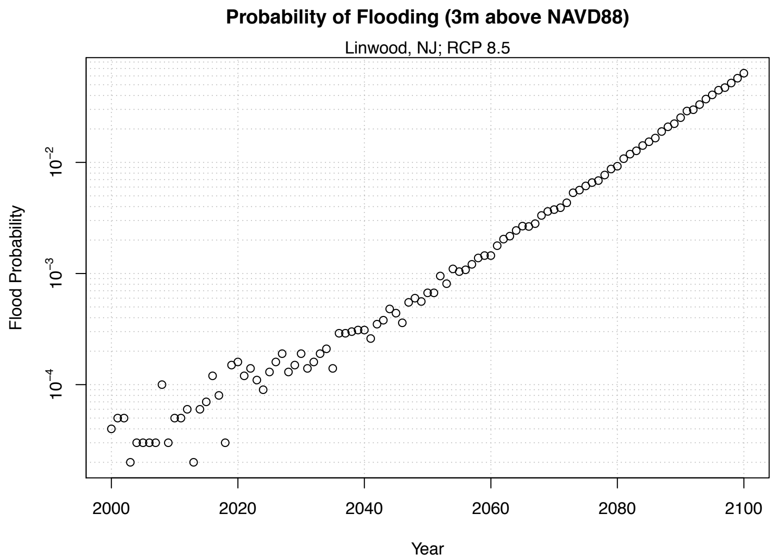


FIG. 7. Probability of home in Linwood, New Jersey flooding for years 2000–2100 under the RCP 8.5 climate forcing scenario. The home elevation is 3 meters above NAVD88.

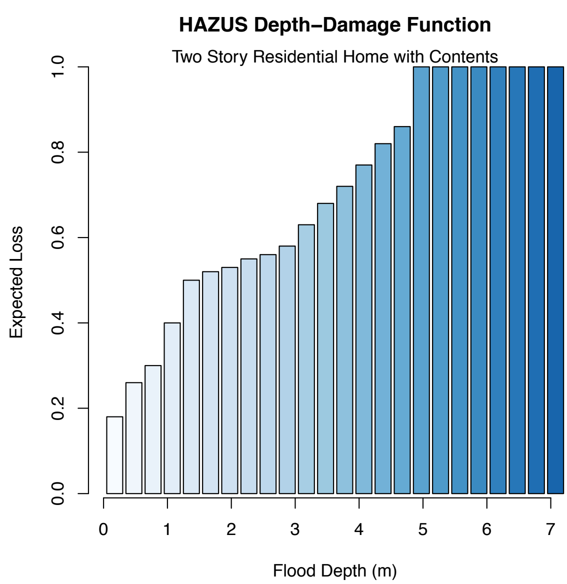


FIG. 8. Empirical depth-damage function for a two story residential home and its contents. Measurements were taken in Wilmington, Delaware by the U.S. Army Corps of Engineers.

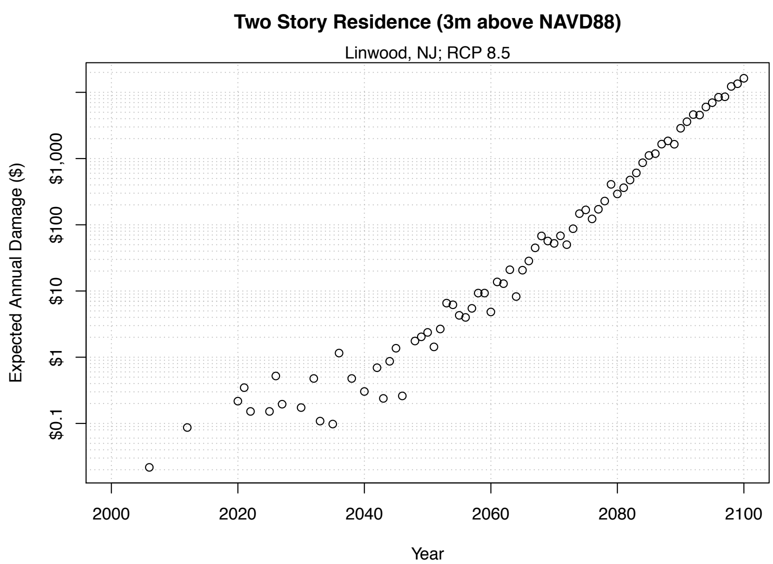


FIG. 9. Expected annual damage ($) for a $725,000 home in Linwood, New Jersey due to flooding for years 2000–2100 under the RCP 8.5 climate forcing scenario. The home elevation is 3 meters above NAVD88.

Computer codes for: “A probabilistic framework for homeowner flood insurance cost-benefit analysis under rising seas”

data_qa.R

```
rm(list=ls(all=TRUE))
setwd("/Users/dmr/Dropbox/Courses/CEE460/proj/preproc")

library(xts)

dat = read.table("atlantic_city_hr_sealevel_1911-2015.csv",sep=",",header=TRUE)

# Create POSIX time variable
date = as.character(paste(dat$year,dat$month,dat$day,sep="-"))
hr = as.character(paste(dat$hour,":00",sep=""))
time = as.character(paste(date,hr,sep=" "))
dates <- strftime(time, format = "%Y-%m-%d %H:%M")

# Create data structure of hourly sea level and time
d <- structure(list(dates = dates, level = dat$level),.Names = c("dates", "level"),
               row.names = c(NA, -857473L), class = "data.frame")
x <- xts(d[,-1], as.POSIXct(d[,1], format="%Y-%m-%d %H:%M"),
                                            colnames=c("dates", "level"))/1000

# Annual average
ann = apply.yearly(x, colMeans)

# plot time series with linear trend
pdf('annualtrend.pdf',width=6,height=4,paper='special')
plot(index(ann),ann,lwd=2,ylab="Mean Sea Level (m)",xlab="Year",
     main="Mean Sea Level at Atlantic City, NJ")
mtext("(Bkgd: 2.2  0.2 mm/y)")

# calculate annual trend
v = 1:length(ann)
ann.fit = lm(ann~v)
summary(ann.fit)

# add fitted values to plot
text(x=as.Date("1910-12-31 23:00:00 EST"),2.3,"trend: 4.1  0.1 mm/yr")
lines(index(ann),ann.fit$fitted.values,lty=2,col='red',lwd=2)
dev.off()

# Subtract MHW
x = x - 2.787

```

```r
# detrend houly sea level
v = 1:length(x)
dfit = lm(formula = x ~ v)
ll = x - dfit$coefficients[2]*v

# daily max from hourly
xgt86 = x[index(x)>="1986-01-01 00:00:00 EST"]
DailyMaxSL <- apply.daily(xgt86, max,na.rm=TRUE)

# Write to CSV file
df = data.frame(index(DailyMaxSL),DailyMaxSL)
write.table(df,"SL_daily_max_1986-2015.csv",sep = ",",row.names=FALSE)
```

slr_by_year.R

```r
setwd("/Users/dmr/Dropbox/Courses/CEE460/proj/preproc")
library('zoo')

dat = read.table("slr_atlantic_city_2000-2100_rcp85.csv",
                 sep=",",header=TRUE)

# interpolate values between time slices
z <- zoo(dat)
zaprox = na.spline(z)

# write interpolated values to a text file
write.zoo(zaprox,
          file="slr_atlantic_city_2000-2100_filled_rcp85.csv",sep=",")


# plot interpolated values

pdf('slr_atlantic_city_2000-2100.pdf',width=6,height=4,
    paper='special')
plot(dat$year,zaprox$q0.5,type="n",ylim=c(0,220),xlab="Year",
     ylab="RSL (cm)",
     main = "Atlantic City, NJ")
mtext("RCP 8.5")
lines(dat$year,zaprox$q0.5,lw=1,lty=3) # 0.5
lines(dat$year,zaprox$q5,lw=2,lty=3) # 0.5
lines(dat$year,zaprox$q17,lw=3,lty=2)
lines(dat$year,zaprox$q50,lw=4)
lines(dat$year,zaprox$q83,lw=3,lty=2)
lines(dat$year,zaprox$q95,lw=2,lty=3)
lines(dat$year,zaprox$q99.5,lw=1,lty=3) # 0.5
legend(2000,220,c("0.5th/99.5th","5th/95th","17th/83rd","50th"),
       lty=c(3,3,2,1),lwd=c(1,2,3,4))
dev.off()

pdf('slr_atlantic_city_2000-2050.pdf',width=6,height=4,
    paper='special')
plot(dat$year,zaprox$q0.5,type="n",ylim=c(0,80),xlab="Year",
     ylab="RSL (cm)",
     main = "Atlantic City, NJ",xlim=c(2000,2050))
mtext("RCP 8.5")
lines(dat$year,zaprox$q0.5,lw=1,lty=3) # 0.5
lines(dat$year,zaprox$q5,lw=2,lty=3) # 0.5
lines(dat$year,zaprox$q17,lw=3,lty=2)
lines(dat$year,zaprox$q50,lw=4)
```

```
lines(dat$year,zaprox$q83,lw=3,lty=2)
lines(dat$year,zaprox$q95,lw=2,lty=3)
lines(dat$year,zaprox$q99.5,lw=1,lty=3) # 0.5
legend(2000,80,c("0.5th/99.5th","5th/95th","17th/83rd","50th"),
       lty=c(3,3,2,1),lwd=c(1,2,3,4))
dev.off()
```

slr_gpd_routine.R

```
setwd("/Users/dmr/Dropbox/Courses/CEE460/proj/preproc")

library(evd)
library(extRemes)

u = 0.6 # threshold

# Open historical record of daily max water level (w.r.t MHW)
dat = read.table("SL_daily_max_1986-2015.csv",sep=",",header=TRUE)

pdf('surgehist_atlantic_city.pdf',width=6,height=4,paper='special')
hist(dat$DailyMaxSL,breaks=20,xlim = c(0.35,1.4),xlab="Daily Height above MWH (m)",
     main="Atlantic City, NJ")
mtext("1986-2015")
text(0.95,1800,"Threshold = 0.6m above MHW")
abline(v=0.6,lty=2,lwd=3)
dev.off()


# Decluster daily observations so that multiple
# observations from the same event are excluded
# This makes the data independent for MLE

# decluster above 0.3 m and decluster run length of 2 days
declusterSL = decluster(dat$DailyMaxSL,threshold = 0.3,clusterfun = "max",r=2)

# After declustering
hist(declusterSL,breaks=20,xlim = c(0.6,1.4),ylim=c(0,500),xlab="Daily Height above MWH (m)"
           main="Tide Gauge at Atlantic City, NJ (1979-2015)")

tcplot(declusterSL, tlim = c(0,.5), model = "gpd",npp = 365.25)
tcplot(dat$DailyMaxSL, tlim = c(0,.5), model = "gpd",npp = 365.25)

dmax.fit <- fpot(declusterSL, threshold = .6, model = "gpd",npp = 365.25)

sigma = dmax.fit$param[1] # scale
xi = dmax.fit$param[2] # shape
std_sigma = 2*dmax.fit$std.err[1]
std_xi = 2*dmax.fit$std.err[2]
prob_above = dmax.fit$pat

plot(dmax.fit,which=3)

# calculate the return level as a function of the return period (years)
```

```
i = 0

zp_lw = NULL
zp = NULL
zp_up = NULL

for (T in 1:10001){

  # shape .ne. 0
#   a = (-(prob_above/(log(1-1/T))))^(xi-std_xi)
#   zp_lw[i] = ((sigma-std_sigma)/(xi-std_xi))*(a - 1) + u

#   a = (-(prob_above/(log(1-1/T))))^xi
#   zp[i] = (sigma/xi)*(a - 1) + u

#   a = (-(prob_above/(log(1-1/T))))^(xi+std_xi)
#   zp_up[i] = ((sigma+std_sigma)/(xi+std_xi))*(a - 1) + u

  # shape ~ 0

  a = -log(1-1/T)/(365.25*prob_above)
  zp_lw[i] = -(sigma-std_sigma)*log(a) + u

  a = -log(1-1/T)/(365.25*prob_above)
  zp[i] = -(sigma)*log(a) + u

  a = -log(1-1/T)/(365.25*prob_above)
  zp_up[i] = -(sigma+std_sigma)*log(a) + u

  i = i + 1
}

# plot historical return levels
x <- seq(1,10000,1)

pdf('ac_hist_surge_returnlvl.pdf',width=6,height=4,paper='special')
plot(x,zp, type="n",xaxt="n",log="x",ylim=c(0.5,2.5), ylab="Surge above MWH (m)",
     xlab="Return Period (Years)", main="Atlantic City, NJ",
     panel.first=abline(v=c(2,3,4,5,6,7,8,9,10,20,30,40,50,60,70,80,90,100,
                            200,300,400,500,600,700,800,900,1000,2000,3000,4000,
                            5000,6000,7000,8000,9000,10000),h=c(.5,1,1.5,2,2.5),
                        lty=3,col="gray"))
mtext("1986-2015")
at.x <- outer(1:9, 10^(0:4))
lab.x <- ifelse(log10(at.x) %% 1 == 0, at.x, NA)
axis(1, at=at.x, labels=lab.x, las=1)
```

```
lines(zp_lw,lty=2)
lines(zp,lty=1)
lines(zp_up,lty=2)
legend("topleft", expression(paste("2",sigma," from EVA",sep="")),
       lty=c(2), bty='n', cex=1.0)
dev.off()
```

slr_by_decade.R

```
setwd("/Users/dmr/Dropbox/Courses/CEE460/proj/preproc")
library(RColorBrewer)

storm = read.table("SL_daily_max_1986-2015.csv",sep=",",header=TRUE)
lsl = read.table("LSLprojMCMC_180_rcp85.csv",sep=",",header=TRUE)/1000 # to m

# Expected RSL by decade
rsl = NULL

# sample mean approximates the expected value
rsl[1] = mean(lsl$X2000)
rsl[11] = mean(lsl$X2010)
rsl[21] = mean(lsl$X2020)
rsl[31] = mean(lsl$X2030)
rsl[41] = mean(lsl$X2040)
rsl[51] = mean(lsl$X2050)
rsl[61] = mean(lsl$X2060)
rsl[71] = mean(lsl$X2070)
rsl[81] = mean(lsl$X2080)
rsl[91] = mean(lsl$X2090)
rsl[101] = mean(lsl$X2100)

# interpolate expected RSL between decades
zfull <- zoo(rsl)
rsl = na.approx(zfull)

prob_above = 0.01557808 # probability above GPD Threshold
u = 0.6 # threshold in meters
sigma = 0.137974 # shape parameter

z2000 = NULL
z2010 = NULL
z2020 = NULL
z2030 = NULL
z2040 = NULL
z2050 = NULL
z2060 = NULL
z2070 = NULL
z2080 = NULL
z2090 = NULL
z2100 = NULL

for (i in 1:10000){
```

```r
# 2000
q = runif(1)
a = -log(1-q)/(365.25*prob_above)
h = -(sigma)*log(a) + u
q = runif(1)
s = as.numeric(quantile(lsl$X2000,probs=q))
z2000[i] = s + h

# 2010
q = runif(1)
a = -log(1-q)/(365.25*prob_above)
h = -(sigma)*log(a) + u
q = runif(1)
s = as.numeric(quantile(lsl$X2010,probs=q))
z2010[i] = s + h

# 2020
q = runif(1)
a = -log(1-q)/(365.25*prob_above)
h = -(sigma)*log(a) + u
q = runif(1)
s = as.numeric(quantile(lsl$X2020,probs=q))
z2020[i] = s + h

# 2030
q = runif(1)
a = -log(1-q)/(365.25*prob_above)
h = -(sigma)*log(a) + u
q = runif(1)
s = as.numeric(quantile(lsl$X2030,probs=q))
z2030[i] = s + h

# 2040
q = runif(1)
a = -log(1-q)/(365.25*prob_above)
h = -(sigma)*log(a) + u
q = runif(1)
s = as.numeric(quantile(lsl$X2040,probs=q))
z2040[i] = s + h

# 2050
q = runif(1)
a = -log(1-q)/(365.25*prob_above)
h = -(sigma)*log(a) + u
q = runif(1)
s = as.numeric(quantile(lsl$X2050,probs=q))
```

```
  z2050[i] = s + h

  # 2060
  q = runif(1)
  a = -log(1-q)/(365.25*prob_above)
  h = -(sigma)*log(a) + u
  q = runif(1)
  s = as.numeric(quantile(lsl$X2060,probs=q))
  z2060[i] = s + h

  # 2070
  q = runif(1)
  a = -log(1-q)/(365.25*prob_above)
  h = -(sigma)*log(a) + u
  q = runif(1)
  s = as.numeric(quantile(lsl$X2070,probs=q))
  z2070[i] = s + h

  # 2080
  q = runif(1)
  a = -log(1-q)/(365.25*prob_above)
  h = -(sigma)*log(a) + u
  q = runif(1)
  s = as.numeric(quantile(lsl$X2080,probs=q))
  z2080[i] = s + h

  # 2090
  q = runif(1)
  a = -log(1-q)/(365.25*prob_above)
  h = -(sigma)*log(a) + u
  q = runif(1)
  s = as.numeric(quantile(lsl$X2090,probs=q))
  z2090[i] = s + h

  # 2100
  q = runif(1)
  a = -log(1-q)/(365.25*prob_above)
  h = -(sigma)*log(a) + u
  q = runif(1)
  s = as.numeric(quantile(lsl$X2100,probs=q))
  z2100[i] = s + h

}

# save flood heights to CSV file
df = data.frame(z2000,z2010,z2020,z2030,z2040,z2050,z2060,z2070,z2080,z2090,z2100)
```

```r
write.table(df,"SL_surge_daily_max_2000-2100.csv",sep = ",",row.names=FALSE)

col = brewer.pal(9,"YlGnBu")

# plot historical and projected CDFs
pdf('ac_proj_surge_cdf.pdf',width=8,height=6,paper='special')
plot(ecdf(z2000),xlim=c(0.5,3.5),col=col[2],lwd=3,main="Atlantic City, NJ",
     xlab="Surge above MWH (m)",ylab="Cumulative Probability")
mtext("RCP 8.5")
lines(ecdf(z2020),col=col[3],lwd=3)
lines(ecdf(z2030),col=col[4],lwd=3)
lines(ecdf(z2050),col=col[6],lwd=3)
lines(ecdf(z2070),col=col[7],lwd=3)
lines(ecdf(z2100),col=col[9],lwd=3)
legend(3,.5,c("2000","2020","2030","2050","2070","2100"),
       lwd=c(3,3,3,3,3,3),col=c(col[2],col[3],col[4],col[6],col[7],col[9]))
dev.off()

z2000 = NULL
z2020 = NULL
z2030 = NULL
z2050 = NULL
z2070 = NULL
z2100 = NULL

exp2000 = sum(as.numeric(quantile(lsl$X2000,probs=seq(0.0001,1,.0001))))/10000
exp2020 = sum(as.numeric(quantile(lsl$X2020,probs=seq(0.0001,1,.0001))))/10000
exp2030 = sum(as.numeric(quantile(lsl$X2030,probs=seq(0.0001,1,.0001))))/10000
exp2050 = sum(as.numeric(quantile(lsl$X2050,probs=seq(0.0001,1,.0001))))/10000
exp2070 = sum(as.numeric(quantile(lsl$X2070,probs=seq(0.0001,1,.0001))))/10000
exp2100 = sum(as.numeric(quantile(lsl$X2100,probs=seq(0.0001,1,.0001))))/10000

i = 0
for (T in 1:10001){

  a = -log(1-1/T)/(365.25*prob_above)
  z2000[i] = -(sigma)*log(a) + u + exp2000
  z2020[i] = -(sigma)*log(a) + u + exp2020
  z2030[i] = -(sigma)*log(a) + u + exp2030
  z2050[i] = -(sigma)*log(a) + u + exp2050
  z2070[i] = -(sigma)*log(a) + u + exp2070
  z2100[i] = -(sigma)*log(a) + u + exp2100
    i = i + 1
}
```

```r
# plot historical and projected return levels
x <- seq(1,10000,1)

pdf('ac_proj_surge_returnlvl.pdf',width=8,height=6,paper='special')
plot(x,z2100, type="n",xaxt="n",log="x",ylim=c(0.5,3.5), ylab="Surge above MWH (m)",
     xlab="Return Period (Years)", main="Atlantic City, NJ",
     panel.first=abline(v=c(2,3,4,5,6,7,8,9,10,20,30,40,50,60,70,80,90,100,
                          200,300,400,500,600,700,800,900,1000,2000,3000,4000,
                          5000,6000,7000,8000,9000,10000),h=seq(1,4,0.5),
                        lty=3,col="gray"))
mtext("RCP 8.5")
at.x <- outer(1:9, 10^(0:4))
lab.x <- ifelse(log10(at.x) %% 1 == 0, at.x, NA)
axis(1, at=at.x, labels=lab.x, las=1)
lines(z2000,lwd=3,col=col[2])
lines(z2020,lwd=3,col=col[3])
lines(z2030,lwd=3,col=col[4])
lines(z2050,lwd=3,col=col[6])
lines(z2070,lwd=3,col=col[7])
lines(z2100,lwd=3,col=col[9])
legend(1,3.7,c("2000","2020","2030","2050","2070","2100"),
       lwd=c(3,3,3,3,3,3),col=c(col[2],col[3],col[4],col[6],col[7],col[9]),
       horiz=TRUE,bty="n")
dev.off()
```

damages.R

```r
setwd("/Users/dmr/Dropbox/Courses/CEE460/proj/preproc")

library('RColorBrewer')
library('hazus')
library('zoo')


# Retrieve depth-damage function
fl_dept = extract_hazus_functions(func_type = "depth", long_format = FALSE)
exp_dam = as.numeric(fl_dept[56,10:33])*.01 # Two story residential w/ contents
flood_depth = 1:length(exp_dam)*.3048 # feet to meters


# plot depth-damage function
col = colorRampPalette(brewer.pal(9,"Blues"))(30)

pdf('depth-damage.pdf',width=6,height=6,paper='special')
mp = barplot(exp_dam,axes=F,ylab="Expected Loss",xlab="Flood Depth (m)",col=col,
             main="HAZUS Depth-Damage Function")
mtext("Two Story Residential Home with Contents")
axis(1,at=c(0,4,8,12,16,20,24,28),labels=c("0","1","2","3","4","5","6","7"))
axis(2,seq(0,1,.2))
dev.off()


mhw = 2.787 # Mean High Water (meters)
nav88 = 2.308 # North American Vertical Datum of 1988 (meters)

lsl = read.table("LSLprojMCMC_180_rcp85.csv",sep=",",header=TRUE)/1000 # to m

prob_above = 0.01557808
u = 0.6
sigma = 0.137974

# Expected RSL by decade
rsl = NULL

rsl[1] = mean(lsl$X2000)
rsl[11] = mean(lsl$X2010)
rsl[21] = mean(lsl$X2020)
rsl[31] = mean(lsl$X2030)
rsl[41] = mean(lsl$X2040)
rsl[51] = mean(lsl$X2050)
rsl[61] = mean(lsl$X2060)
```

```
rsl[71] = mean(lsl$X2070)
rsl[81] = mean(lsl$X2080)
rsl[91] = mean(lsl$X2090)
rsl[101] = mean(lsl$X2100)

# interpolate expected RSL between decades
zfull <- zoo(rsl)
rsl = na.approx(zfull)


# 100,000 Monte Carlo samples of joint RSL and Surge distribution
zp = matrix(0, 101, 100000)

for (j in 1:101){
  i = 0
  print(1999+j)
for (T in 1:100001){
  p = runif(1)
  a = -log(1-p)/(365.25*prob_above)
  zp[j,i] = -(sigma)*log(a) + u + rsl[j] + mhw - nav88
  i = i + 1
}
}


nsim = 100000
home_elev = 3 # meters above NAV88
home_val = 725000 # USD (est. from Zillow.com)

# determine probability home will be flooded
p = NULL

for (j in 1:101){
    p[j] = 1 - sum(zp[j,]<home_elev)/nsim
}

x = seq(2000,2100,1)
pdf('ac_prob_flood.pdf',width=8,height=6,paper='special')
plot(x,p,log="y",yaxt="n",ylab="Flood Probability",xlab="Year",
     panel.first=abline(h=c(1e-05,2e-05,3e-05,4e-05,5e-05,6e-05,7e-05,8e-05,9e-05,
                              1e-04,2e-04,3e-04,4e-04,5e-04,6e-04,7e-04,8e-04,9e-04,
                              1e-03,2e-03,3e-03,4e-03,5e-03,6e-03,7e-03,8e-03,9e-03,
                              1e-02,2e-02,3e-02,4e-02,5e-02,6e-02,7e-02,8e-02,9e-02,
                              1e-01,2e-01,3e-01,4e-01,5e-01,6e-01,7e-01,8e-01,9e-01),
                              v=seq(2000,2100,20),
                          lty=3,col="gray"),main = "Probability of Flooding (3m above MSL)")
```

```
mtext("Linwood, NJ; RCP 8.5")
aty = c(1e-05,1e-04,1e-03,1e-02,1e-01)
atylab = c(-5,-4,-3,-2,-1)
labels <- sapply(atylab,function(i)
  as.expression(bquote(10^ .(i)))
)
axis(2, at=aty,labels = labels)
dev.off()

# Calcuate expected annual damages
ead = NULL
for (j in 1:101){
xx = zp[j,] - home_elev
counts = as.numeric(table(cut(xx, flood_depth)))
exp_loss = (counts[1]*(1-exp_dam[1])*home_val + counts[2]*(1-exp_dam[2])*home_val +
         counts[3]*(1-exp_dam[3])*home_val + counts[4]*(1-exp_dam[4])*home_val)/100000

# arrival rates of flood events
lambda = 365*p[j] # frequency of events where home elevation is flooded
ead[j] = lambda*exp_loss
}

# plot expected annual damage
x = seq(2000,2100,1)
pdf('ac_ead.pdf',width=8,height=6,paper='special')
plot(x,ead,log="y",yaxt="n",ylab="Expected Annual Damage ($)",xlab="Year",
     panel.first=abline(h=c(10,20,30,40,50,60,70,80,90,100,
                        200,300,400,500,600,700,800,900,1000,
                        2000,3000,4000,5000,6000,7000,8000,9000,10000,
                        20000,30000,40000,50000,60000,70000,80000,90000,100000),
                        v=seq(2000,2100,20),
                        lty=3,col="gray"),main = "Two Story Residence (3m above MSL)")
mtext("Linwood, NJ; RCP 8.5")
aty = c(10,100,1000,10000,100000)
atylab = c(1,2,3,4,5)
axis(2, at=aty,labels = c("$10","$100","$1,000","$10,000","$100,000"))
dev.off()
```