A Markov Decision Process Model for Socio-Economic Systems Impacted by Climate Change

Salman Sadiq Shuvo [1] **Yasin Yilmaz** [1] **Alan Bush** [2] **Mark Hafen** [3]

Abstract

Coastal communities are at high risk of natural hazards due to unremitting global warming and sea level rise. Both the catastrophic impacts, e.g., tidal flooding and storm surges, and the long-term impacts, e.g., beach erosion, inundation of low lying areas, and saltwater intrusion into aquifers, cause economic, social, and ecological losses. Creating policies through appropriate modeling of the responses of stakeholders, such as government, businesses, and residents, to climate change and sea level rise scenarios can help to reduce these losses. In this work, we propose a Markov decision process (MDP) formulation for an agent (government) which interacts with the environment (nature and residents) to deal with the impacts of climate change, in particular sea level rise. Through theoretical analysis we show that a reasonable government's policy on infrastructure development ought to be proactive and based on detected sea levels in order to minimize the expected total cost, as opposed to a straightforward government that reacts to observed costs from nature. We also provide a deep reinforcement learning-based scenario planning tool considering different government and resident types in terms of cooperation, and different sea level rise projections by the National Oceanic and Atmospheric Administration (NOAA).

1. Introduction

The consequences of global warming and sea level rise have been widely documented, examined, and forecasted (Sallenger Jr et al., 2012; Hapke et al., 2013; Pachauri et al., 2014; Plant et al., 2016; Hine et al.; Burke et al., 2019). The coastal communities will undergo multiple threats, including recurring hurricanes, storm surge, and heavy rainfall as a result of sea level rise (Wahl et al., 2015). This intensifies social vulnerability, especially in underprivileged neighborhoods (Kashem et al., 2016; Schrock et al., 2015); induces coastal hazards (Nicholls, 2011; Wahl et al., 2014); and affects regional economies by influencing property values, taxation, and the insurance cost (Fu et al., 2016). Due to sea level rise, coastal inhabitants are exposed to many of these consequences and need to develop the adaptive capability and resilience frameworks to counter these stressors through adequate planning and decision support (Beatley, 2012).

The governments, planners, coastal administrators, and personnel in a variety of agencies require substantial information for effective interaction, decision making, and adaptation planning (Beatley, 2012; Fu et al., 2017; Research Council, 2009). This demands the support of key actors to relate the science, the variability, and the hazard of various outlines to stakeholders (Tribbia & Moser, 2008). Explicitly modeling these agents' reaction to sea level rise scenarios can serve in the creation of strategies tailored to local impacts and resilience management and requires a mixture of social engagement and planning tools, including scenario planning (Berke & Stevens, 2016; Drogoul, 2015; Potts et al., 2017).

In this paper, for planning sea level rise scenarios, we study the interactions between a government, residents, and the nature around the sea level rise problem. Specifically, we use probabilistic models to simulate their behaviors and interactions between them. Focusing on the economic cost of the sea level rise-related natural events (e.g., flooding, hurricane) and government investments (e.g., infrastructure improvement) we propose a Markov decision process (MDP) framework to analyze the decision policies for government together with the reactions of the nature and residents.

We specialize the proposed MDP framework to the sea level rise problem and illustrate it using available economic data and sea level projections for the Tampa Bay region in Florida. However, the proposed MDP framework can be easily adapted to simulating other natural and socioeconomic

[*]Equal contribution [1]Department of Electrical Engineering, University of South Florida, Tampa, USA [2]Honors College, University of South Florida, Tampa, USA [3]School of Public Affairs, University of South Florida, Tampa, USA. Correspondence to: Salman Sadiq Shuvo <salmansadiq@usf.edu>, Yasin Yilmaz <yasiny@usf.edu>.

Proceedings of the 37^{th} International Conference on Machine Learning, Vienna, Austria, PMLR 108, 2020.

systems under the impacts of climate change.

MDP provides a suitable theoretical framework for creating agent-based scenarios (Howard, 1960; Bellman, 1957; Filar & Vrieze, 2012). The MDP agent (government in our case) interacts with the environment (nature and residents in our case) by taking action at each time and receiving a reward/cost from the environment in return. The objective of the agent is to optimize the return over time by choosing actions from a list of available actions. Each time, the system moves to a new state based on the agent's action according to a probability distribution. The optimal policy determines which action to take in which state by mapping system states to actions. It can be either found using dynamic programming or learned from experience using reinforcement learning (RL) (Sutton & Barto, 2018), which is also known as approximate dynamic programming. In deep reinforcement learning, a neural network structure is used to approximate the expected value for each action at each state. Deep RL is suitable for problems where the system states are many or continuous in nature (Mnih et al., 2015).

1.1. Related Works

As an impact of climate change, sea level change is neither globally nor regionally uniform (Hine et al.; Sallenger Jr et al., 2012), and can vary spatially and temporally (Ezer, 2013; Hine et al.; Wahl et al., 2015). This leads to a degree of uncertainty to the projection of sea level rise impacts at the local level, as well as planning for adaptation and resilience. The implications of this uncertainty for coastal community planning and adaptation are many, and have led to a variety of efforts to constrain, refine, and benchmark sea level rise projections for purposes of assessing risk (Buchanan et al., 2016; Hinkel et al., 2015; Le Cozannet et al., 2017; Sweet & Park, 2014; Tyler & Moench, 2012). Similarly, in this work, in addition to proposing a general scenario planning model, we also tailor our model to a specific region, Tampa Bay, FL, using NOAA projections and economic data for the region.

Agent-based modeling has been a popular choice for simulating complex systems including transportation systems (Faboya et al., 2018), disaster recovery (Eid & El-Adaway, 2018), and climate change (Patt & Siebenhüner, 2005). Different than existing works, in this paper we propose a novel scenario planning model based on MDP and deep RL for a realistic setup with arbitrary number of actions and continuous sea level rise values. Furthermore, as a major contribution, we theoretically analyze the optimal decision policy, and illustrate the proposed model for the Tampa Bay region.

1.2. Contributions

In our model, MDP agent is the government in an urban setup, which interacts with the nature and residents by making investment decisions for the sea level rise problem to minimize the incurred economic cost. Our contributions can be summarized as follows.

- We theoretically show that the government should base its investment decision on the observed sea level instead of the incurred cost from the nature. Making investment decisions based on the natural cost corresponds to the straightforward policy that any sensitive but shortsighted government would adopt. Through mathematical analysis of the proposed MDP model, we show that *proactive* actions triggered by the rising sea levels are more effective in reducing the cumulative cost than *reactive* actions triggered by the cost from the nature.
- Using the proposed MDP model we provide a deep RL-based scenario planning framework that takes into account arbitrary number of (discrete) government actions, continuous sea level rise values, different sea level projections by NOAA, and different government and resident prototypes in terms of being responsive to the sea level rise problem.
- Using the available economic data and the regionally adjusted NOAA sea level projections we present scenario simulations for Tampa Bay, FL. For each scenario, we consider different government and resident prototypes using cooperation indices that represent how responsive they are to the sea level rise problem. Finally, we show that the optimal (proactive) policy learned by the proposed deep RL algorithm achieves 40% less economic cost than the best reactive policy in 100 years in three different sea level projections by NOAA.

The rest of the paper is organized as follows. Section 2 explains the proposed MDP model. In Section 3.1, we characterize the optimal policy through function analysis. Then, we present a deep RL algorithm for finding the optimal policy in Section 3.2. The scenario simulations for Tampa Bay region are provided in Section 4. Discussions about the proposed model is given in Section 5, and the paper is concluded in Section 6. For interested readers, we provide the proofs in the accompanying supplementary file.

2. MDP Problem Formulation

We propose an MDP framework to model the behaviors of stakeholders (government, residents, and nature), and the interactions between them. As shown in Figure 1, government, the central decision maker in dealing with the sea level rise problem, at each time step n takes an action (i.e., investment decision) x_n and receives responses from the residents y_n and the nature z_n through the cost c_n. Then, this natural and

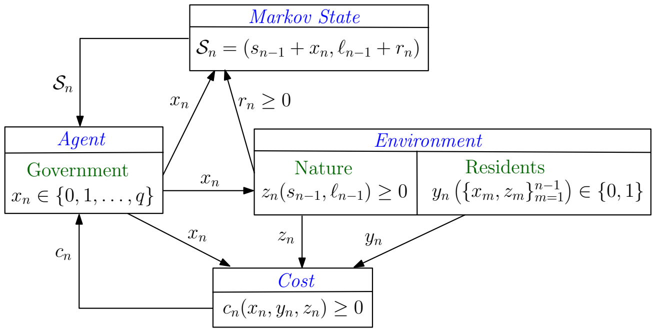

Figure 1. Proposed MDP framework.

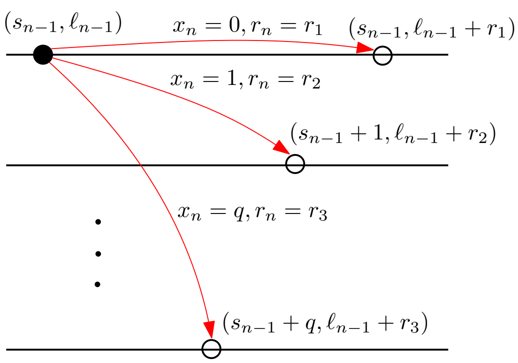

Figure 2. MDP state transition. The state space consists of a continuous horizontal dimension for sea level and discrete vertical dimension for infrastructure state. Three possible state transitions are shown in the figure. Note that r_n can take any nonnegative value.

socioeconomic system moves to a new state $\mathcal{S}_n$ based on the previous state $\mathcal{S}_{n-1}$, government's action x_n, and the sea level rise r_n. The system state consists of the pair of city's infrastructure state s_n and the sea level ℓ_n, i.e., $\mathcal{S}_n = (s_n, \ell_n)$. The government's investment decisions also determine the state of the city infrastructure s_n, which is modeled as $s_n = s_0 + \sum_{m=1}^{n} x_m = s_{n-1} + x_n$. Likewise, the sea level at step n is given by the cumulative sea level rise values: $\ell_n = \ell_0 + \sum_{m=1}^{n} r_m = \ell_{n-1} + r_n$. Here s_0 and l_0 are respectively the initial infrastructure state and the initial sea level of the city. In terms of simulations, these are two user-defined numbers representing the existing states at the beginning of the simulations. The system state clearly satisfies the Markov property: $\mathsf{P}(\mathcal{S}_n|\mathcal{S}_{n-1}, \dots, \mathcal{S}_0) = \mathsf{P}(\mathcal{S}_n|\mathcal{S}_{n-1})$. The state transition is illustrated in Figure 2, and the parameters of the proposed MDP framework are summarized in Table 1. We next explain our proposed models for the government, nature, and residents under the MDP framework.

Table 1. Model parameters.

Initial sea level	$\ell_0 \geq 0$
Sea level rise at time n	$r_n \geq 0$
Sea level at time n	$\ell_n = \ell_0 + \sum_{m=1}^{n} r_m$
Initial infrastructure state	$s_0 \in \{1, 2, ...\}$
Infrastructure decision at time n	$x_n \in \{0, 1, \dots, q\}$
Infrastructure state at time n	$s_n = s_0 + \sum_{m=1}^{n} x_m$
Residents' decision at time n	$y_n \in \{0, 1\}$
Cost from nature at time n	$z_n \geq 0$
Total cost at time n	$c_n = \alpha x_n - \beta y_n + z_n$

2.1. Government Model

At each time step, e.g., a year, the government decides the degree of its investment $x_n \in \{0, 1, 2, ..., q\}$ for infrastructure development, where q is a finite positive integer. $x_n = 0$ means no investment at step n. Hence, there are $q + 1$ possible actions for the government at each time step. The numerical value of $x_n = m$ can be interpreted as spending m unit money towards the infrastructure development or the mth action among q different actions with increasing cost and effectiveness. Possible government actions include but are not limited to building seawalls, raising roads, widening beach, building traditional or horizontal levees, placing stormwater pumps, improving sewage systems, relocating seaside properties, etc. (Sergent et al., 2014; mar, 2016; Xia et al., 2019). The range of x_n is designed to cover the real world costs from the cheapest investment like cleaning the pipes to the most expensive investment like buying lands and property to relocate the seaside inhabitants and businesses.

The total cost c_n to the agent at each time n consists of the investment cost, cost from nature, and the residents' contribution to the investment. Since the government's investment decision and the residents' contribution decision have integer values, we model the total cost as $c_n = \alpha x_n - \beta y_n + z_n$ using parameters α and β to map the decisions to monetary values. The three different entities in the cost definition are combined by adjusting the parameters α, β and the parameters of the probabilistic model introduced for the nature's cost z_n in Section 2.2. In Section 4, we discuss how to set these parameters to obtain realistic costs for the Tampa Bay region. The discounted cumulative cost for the government in N time steps is given by

$$C_N = \sum_{n=0}^{N} a_g^n [\alpha x_n - \beta y_n + z_n], \tag{1}$$

where the discount factor $a_g \in (0, 1)$ defines the weight of future costs in the agent's decisions. Apart from being a standard parameter in MDP cost function, a_g has an important contextual significance in this research. It indicates how much the government values the future costs due to sea level rise in its decision making process. Hence, in this

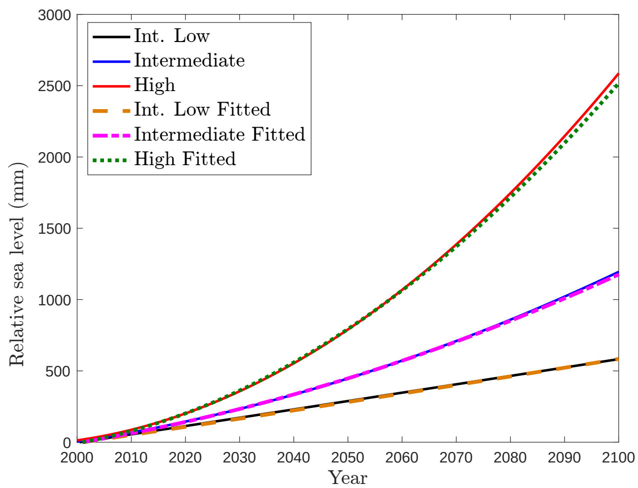

Figure 3. Adjusted NOAA projections (solid lines) and our fittings (dashed lines) for relative sea level change for Tampa Bay, FL.

work, a_g is termed as the *government's cooperation index*. The government's objective is to minimize the expected cumulative cost $\mathsf{E}[C_N]$ by taking investment actions $\{x_n\}$ over time.

2.2. Nature Model

Hurricanes, floods, and stagnant water are some of the many sea level rise-related natural events that cause cost in different ways such as loss of properties, jobs, taxes, and tourism incomes due to submerged areas. To model this cost z_n, we begin with modeling the sea level rise r_n. Recently, three different regionally adjusted NOAA projections for sea level rise have been proposed for the Tampa region (Burke et al., 2019), as shown in Fig. 3.

We take these projections as mean sea level rise values, and model the uncertainty around them using the Gamma distribution, i.e., $r_n \sim \text{Gamma}(\mu, \phi)$, since it is a flexible two-parameter probability distribution used for modeling positive variables including environmental applications, e.g., daily rainfall (Aksoy, 2000). We set the scale parameter $\phi = 0.5$ and vary the shape parameter μ in a range to match the mean relative sea level, given by $\sum_{m=1}^{n} \mathsf{E}[r_m]$, with the NOAA projection curves. The successful curve fitting, shown in Fig. 3, is achieved by increasing μ from 11.2 to 12.388 with 0.012 increments for the intermediate-low projection; from 13 to 34.78 with 0.22 increments for the intermediate projection; and from 14.6 to 87.86 with 0.74 increments for the high projection.

Then, we model nature's cost z_n using the generalized Pareto distribution, which is commonly used to model catastrophic losses, e.g., (Cebrián et al., 2003; Uchiyama & Watanabe, 2006; Daspit & Das, 2012). It is known that the storm- and flooding-related costs have been increasing with the sea level rise (nce, 2020). Thus, we model the scale parameter of generalized Pareto distributed z_n directly proportional to the most recent sea level ℓ_{n-1} and inversely proportional to the most recent infrastructure state s_{n-1}. The proposed model for z_n is given by

$$z_n \sim \text{GeneralizedPareto}(k, \sigma_n, \theta)$$
$$\theta \geq 0, \; \sigma_n = \frac{\eta(\ell_{n-1})^a}{(s_{n-1})^b}, \; k < 0, \qquad (2)$$

where θ, σ, k are the standard parameters of generalized Pareto: location, scale, and shape parameters, respectively; and $\eta > 0, a \in (0,1), b > 0$ are our additional model parameters. The parameters k, θ, η, a, b helps to regulate the impact of most recent sea level ℓ_{n-1} over nature's cost z_n relative to the most recent infrastructure state s_{n-1}. Choosing an appropriate set of parameters depends on the region considered for simulations. We explain how to set the parameters for the Tampa Bay region in Section 4. Our preference for modeling the scale parameter and not the location parameter is due to the fact that the scale parameter can control both the mean and the variance, whereas the location parameter appears only in the mean.

2.3. Residents Model

In our framework, government requests a contribution from residents to the sea level rise investments, e.g., through an additional tax referendum; and residents make a binary decision $y_n \in \{0, 1\}$ whether to support the government. The model is flexible in terms of the frequency of voting. For instance, residents may vote every 5 years and the voting decision remains constant until the next voting.

An appropriate model for y_n requires social modeling for the residents' voting behavior. While the residents of a region in general form a heterogeneous community in terms of social and psychological factors, it is sufficient to model the aggregate response in this work since we consider a referendum in which the majority vote wins. Similar to the government's cooperation index, we define the *residents' cooperation index* $a_r \in (0, 1)$ to quantify how responsive the residents are to the sea level rise problem. These cooperation indices are helpful in simulating different government and resident profiles, as illustrated in Section 4. As explained by the social exchange theory (Homans, 1974; Rasinski & Rosenbaum, 1987), some individuals' votes are motivated by what benefit they receive in exchange. A high percentage of such individuals in a community is represented by a low a_r value in our model. According to another perspective, voter behavior is determined by non-self-interested factors such as a sense of civic duty (Katosh & Traugott, 1982), moral obligation (Rasinski et al., 1985), and psychological sense of community (Davidson & Cotter, 1993). In our model, a high a_r value corresponds to a community which is dominantly composed of this type of individuals who value non-self-interested factors.

We use a_r to obtain a score function that gives the probability of support from the residents. In addition to internal factors, represented by a_r, voters' decision is also affected by external factors, government's actions $\{x_n\}$ and nature's cost $\{z_n\}$ in our case. We hypothesize that the probability of a positive voting outcome, $\mathsf{P}(y_n = 1)$, is high when both nature's cost and government's efforts are high (i.e., high x_n and z_n) in the recent history. In other words, when either nature's cost has been low or government is not responsive, residents lack the main reason for accepting an additional tax. Hence, we propose the following score function and the Bernoulli model for y_n:

$$h_n = \sum_{m=1}^{n-1} a_r^{n-m} x_m z_m, \tag{3}$$

$$y_n \sim \text{Bernoulli}(p_n),\ p_n = \sigma(h_n) = \frac{1}{1 + e^{-(h_n - h_0)}},$$

where $\sigma(\cdot)$ is the logistic sigmoid function used for mapping the score h_n to probability p_n. High h_n means a high likelihood of support from the residents. The sigmoid's midpoint h_0 is empirically obtained as the average value of h_n. In addition to being a cooperation index, another interpretation for a_r is the concept of discount (or forgetting) factor for past events, similar to the government's a_g coefficient. For high a_r values, residents have a longer memory, and they consider also the past disasters and government efforts with an exponentially decaying weight in their decision. Whereas, for a community with low a_r, residents have a short memory and only care about the most recent cost and government effort. In an extreme case where a_r is close to zero, they may not be even motivated to contribute by a very recent disaster and government's high efforts.

3. Optimal Policy

In this section, we first analyze the optimal policy for government investments, and then provide a practical algorithm to find the optimal policy.

3.1. Optimal Policy Analysis

In the proposed MDP framework, the objective of a rational government is to minimize the expected total cost $\mathsf{E}[C_N]$ in N time steps by following an optimal investment policy. Central to MDP is the optimal value function

$$V(s_n, \ell_n) = \min_{\{x_n\}} \mathsf{E}[C_N | \{x_n\}],$$

which gives the minimum expected total cost possible at each state (s_n, ℓ_n), denoted as the optimal value of that state, by choosing the best action policy $\{x_n\}$. To find the optimal policy, the Bellman equation

$$V(s_{n-1}, \ell_{n-1}) = \min_{x_n} \mathsf{E}[c_n + a_g V(s_n, \ell_n) | x_n]$$

provides a recursive approach by focusing on finding the optimal action x_n at each time step using the successor state value, instead of trying to find the entire policy $\{x_n\}$ at once. Using the cost definition given by (1) and considering possible q actions for x_n this iterative equation can be rewritten as

$$\begin{aligned} V(s_{n-1}, \ell_{n-1}) = \min \Big\{ & \\ & \underbrace{\mathsf{E}\big[-\beta y_n + z_n + a_g V(s_{n-1}, \ell_{n-1} + r_n)\big]}_{F_0(s_n, \ell_n)}, \\ & \underbrace{\mathsf{E}\big[\alpha - \beta y_n + z_n + a_g V(s_{n-1} + 1, \ell_{n-1} + r_n)\big]}_{F_1(s_n, \ell_n)}, \\ & \underbrace{\mathsf{E}\big[2\alpha - \beta y_n + z_n + a_g V(s_{n-1} + 2, \ell_{n-1} + r_n)\big]}_{F_2(s_n, \ell_n)}, \\ & \cdots, \\ & \underbrace{\mathsf{E}\big[q\alpha - \beta y_n + z_n + a_g V(s_{n-1} + q, \ell_{n-1} + r_n)\big]}_{F_q(s_n, \ell_n)} \Big\}, \end{aligned} \tag{4}$$

where $F_m(s_n, \ell_n)$ is defined as the expected total cost of taking action $x_n = m$ at state (s_n, ℓ_n).

At each time step n, the action x_n shapes the instant cost c_n and moves the system to the next state, which determines the discounted future cost $a_g V(s_n, l_n)$. The optimum policy chooses among the investment actions $x_n \in \{0, 1, 2, ..., q\}$ that has the minimum expected total cost, $\min_m \{F_m(s_n, \ell_n)\}$, as shown in (4). Since the functions $\{F_0(s_n, \ell_n), \ldots, F_q(s_n, \ell_n)\}$ determine the optimal policy, we next analyze them to characterize the optimal government policy.

Theorem 1. *For $m = 0, 1, \ldots, q$, $F_m(s_n, \ell_n)$ is nondecreasing and concave in ℓ_n for each s_n; and the derivative of $F_m(s_n, \ell_n)$ with respect to ℓ_n is lower than that of $F_{m-1}(s_n, \ell_n)$ for $m = 1, \ldots, q$.*

Proof is provided in the Supplementary file. For a specific infrastructure state s_n, expected costs $F_0(s_n, \ell_n), F_1(s_n, \ell_n), \ldots, F_q(s_n, \ell_n)$ for all the $q + 1$ actions are illustrated in Figure 4 according to Theorem 1. The optimum policy picks the minimum of the $q + 1$ curves at each time, which is shown with the solid curve in Figure 4. As a result of Theorem 1, we next give the outline of optimum policy in Corollary 1.

Corollary 1. *The optimum policy, at each infrastructure state s_n, compares the sea level ℓ_n with at most q thresholds where each threshold signifies a change of optimal action.*

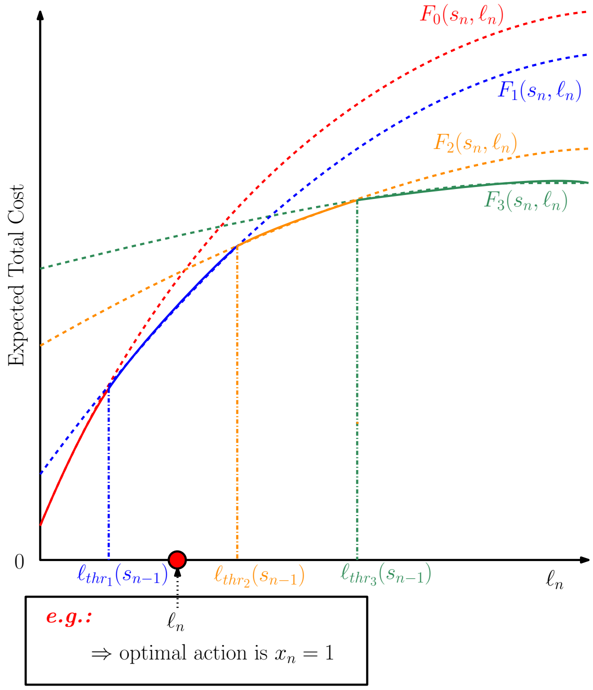

Figure 4. Expected total costs as a function of sea level for an example case with $q = 3$. The expected total costs intersect each other only once for any given infrastructure state according to Theorem 1.

To prove Corollary 1, note that

$$F_{m-1}(s_n, \ell_n = 0) < F_m(s_n, \ell_n = 0)$$

for $m \in \{1, 2, ..., q\}$ because $\ell_n = 0$ corresponds to the fictional case of zero sea level where there is no risk. That is, $F_{m-1}(s_n, \ell_n)$ starts at a lower point than $F_m(s_n, \ell_n)$, but increases faster than $F_m(s_n, \ell_n)$ since its derivative is higher (Theorem 1). Also from Theorem 1, it is known that both of them are concave and bounded, hence $F_{m-1}(s_n, \ell_n)$ and $F_m(s_n, \ell_n)$ intersect exactly at one point for $m \in \{1, 2, \dots, q\}$. While for ℓ_n less than the intersection point the action $x_n = m$ is less effective than the action $x_n = m - 1$ in terms of immediate cost and expected future cost, it becomes more effective when ℓ_n exceeds the intersection point.

Figure 4 gives an example case where there are $q = 3$ thresholds $\ell_{thr_1}(s_{n-1}), \ell_{thr_2}(s_{n-1}), \ell_{thr_3}(s_{n-1})$ which depend on s_{n-1} and indicate change points of optimal action. However, depending on the slopes of $\{F_m(s_n, \ell_n)\}$ curves at each infrastructure state s_{n-1}, there may be less than q change points.

To summarize, for a given state (s_{n-1}, ℓ_{n-1}), the optimum policy chooses x_n based on the relative value of the current sea level $\ell_n = \ell_{n-1} + r_n$, where r_n is the rise on top of ℓ_{n-1}, with respect to the existing infrastructure state $s_{n-1} = \sum_{m=1}^{n-1} x_m$.

The thresholds also depend on the cooperation indices a_g and a_r. Higher cooperation indices set the thresholds lower and vice versa. Notably, the government's cooperation index a_g is dominant in shaping the thresholds. Intuitively, as a_g grows, the government becomes more cautious about (i.e., sees more objectively without severely discounting) the expected future natural costs and sets a lower threshold for investment actions. On the contrary, small a_g implies underestimated future costs and thus overemphasized investment costs, which results in a high threshold for investment.

3.2. Finding the Optimal Policy

We showed that the optimal policy is given by a number of thresholds on the current sea level; however completely specifying the optimal policy with analytical expressions for thresholds does not seem tractable due to a large number of parameters involved in the problem. Hence, we next propose a reinforcement learning (RL) algorithm to find the optimal policy for simulated scenarios. Specifically, to learn the optimal policy here, a deep RL algorithm is needed due to the continuous sea level values, which cause infinite number of possible states. The deep Q network (DQN) algorithm (Mnih et al., 2015), which is a popular choice for deep RL, addresses well the infinite dimensional state space problem. It leverages a deep neural network to estimate the optimal action-value function

$$Q(s_n, \ell_n, x_n) = \mathsf{E}[c_n + a_g \min_x Q(s_n + x_n, \ell_{n+1}, x) \mid x_n].$$

where $V(s_n, \ell_n) = \min_{x_n} Q(s_n, \ell_n, x_n)$. In Algorithm 1, we present a DQN algorithm for learning the optimal policy on government's infrastructure investment actions.

Algorithm 1 operates many episodes to iteratively compute the weights of the DQN neural network that relates input states and actions with the output $Q(\mathcal{S}, x)$ values. Each episode consists of Monte-Carlo simulations in which several states are visited according to the current policy defined by the current value function. Selecting the actions always using equation (4) may keep some of the states not explored enough. To encourage the agent to explore the state space adequately, we use a technique that explores frequently at the beginning, but reduces exploration with gaining more experience over time (Singh et al., 2000). After the weight matrix values converge, the final weight matrix is used to generate scenario simulations, as described in the following section. The convergence of the weight matrix is depicted in Fig. 3.2.

4. Scenario Simulations

In this section, we present scenario simulations for the Tampa Bay region in Florida using our MDP model and deep RL algorithm. Tampa Bay, situated along the Gulf of

Algorithm 1 DQN algorithm for finding optimum policy

```
Input: a_g, a_r, α, β, θ, k, η, a, b
Initialize replay memory D to capacity N
Initialize action-value function Q with random weights
  w and target action-value function Q' with random
  weights w' = w
for episode = 1, 2, ... do
  Initialize state S_0 = (s_0, ℓ_0)
  for n = 1, 2, ..., N do
    With probability ε select a random action x_n, oth-
      erwise select x_n = arg min_x Q(S_n, x; w)
    Execute action x_n and observe cost c_n = αx_n −
      βy_n + z_n (see (2) and (3))
    Store transition (S_n, x_n, c_n, S_{n+1}) in D
    Sample random minibatch of transitions
      (S_j, x_j, c_j, S_{j+1}) from D
    Set target t_j = c_j + a_g min_x Q'(S_{j+1}, x; w')
    Perform a gradient descent step on [t_j −
      Q(S_j, x_j; w)] with respect to the weights w
    Every d steps reset w' = w
    if Q(S, x) converges for all S, x then
      break
    end if
  end for
end for
```

Mexico in Florida with more than 3 million residents, is listed number seven among the most at-risk areas in terms of risk of damaging floods on the globe by the World Bank (wor, 2013). We have three different sea level rise projections by NOAA until the year 2100 for the area (Burke et al., 2019), as shown in Fig. 3. The City of Tampa expects to collect \$14 M/year for 30 years from residents to improve its stormwater system (sto, 2019). Thus, we set $\beta = 14$ with the base unit of million dollars. According to the ongoing improvement projects and regular stormwater services of City of Tampa (sto, 2019), annual investment may range between \$ 25-75 M. Hence, we choose $\alpha = 25$ and the number of nonzero actions $q = 3$, i.e., $x_n \in \{0, 1, 2, 3\}$. A recent report by the Tampa Bay Regional Planning Council (tbr, 2017) gives "the cost of doing nothing" for years 2020-2060 due to the sea level rise impacts under the high sea level rise projections as \$162 billion. We obtain $\sum_{2020}^{2060} z_n = 162,000$ (in million dollars) by keeping the infrastructure state at the initial level $s_0 = 50$, increasing the sea level from the initial $\ell_0 = 100$ according to the gamma distribution following the high sea level projection (Section 2.2), and setting the parameters of generalized Pareto model as $k = -0.001, \theta = 1, \eta = 25, a = 0.9, b = 1.1$. All the costs in the following figures are normalized by taking the residents' contribution 14 M as one unit.

In Fig. 6, we analyze the effect of cooperation indices. We obtain different scenarios by varying the cooperation indices a_g and a_r for the government and residents, respectively, and by considering different NOAA projections for sea level rise. As expected, the average total cost decreases with growing cooperation index for both the government (a_g) and residents (a_r). The cost is significantly higher if both the government and residents are not cooperative ($a_g = a_r = 0.1$) compared to the cooperative case ($a_g = a_r = 0.9$).

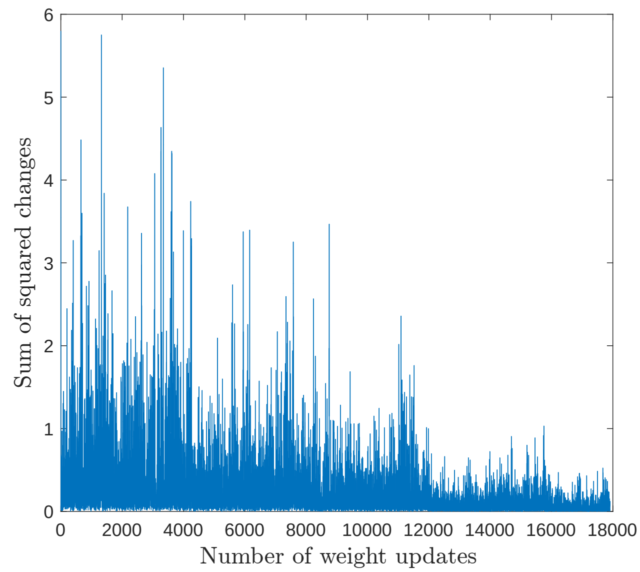

Figure 5. Convergence of weights in Algorithm 1.

We next compare the proposed MDP-based government policy, which proactively improves infrastructure based on observed sea levels, with the shortsighted policy which reacts to significant costs from the nature by improving the infrastructure. We model this shortsighted policy with a threshold parameter for the cost from nature z_n. When z_n exceeds the threshold, the shortsighted policy improves the infrastructure by three units ($x_n = 3$). Under the three sea level projections, we find the optimum thresholds that minimize the average total cost of shortsighted policy in 100 years as 34, 28, and 21 for the intermediate low, intermediate, and high projections, respectively (Figure 7). The optimal thresholds in Figure 7 show that the infrastructure improvement is more urgent under higher sea level projections, in accordance with the intuition.

Finally, in Fig. 8, we compare the best MDP-based policy ($a_g = a_r = 0.9$) with the best shortsighted policy, which uses the optimum threshold, in terms of average total cost in 100 years for all sea level projections. The same resident and nature models are used while simulating both policies. For all the three sea level projections, the MDP-based proactive policy, whose decisions are driven by the current sea level, achieves around 40% less cost than the shortsighted reactive policy, which decides based on the cost from nature.

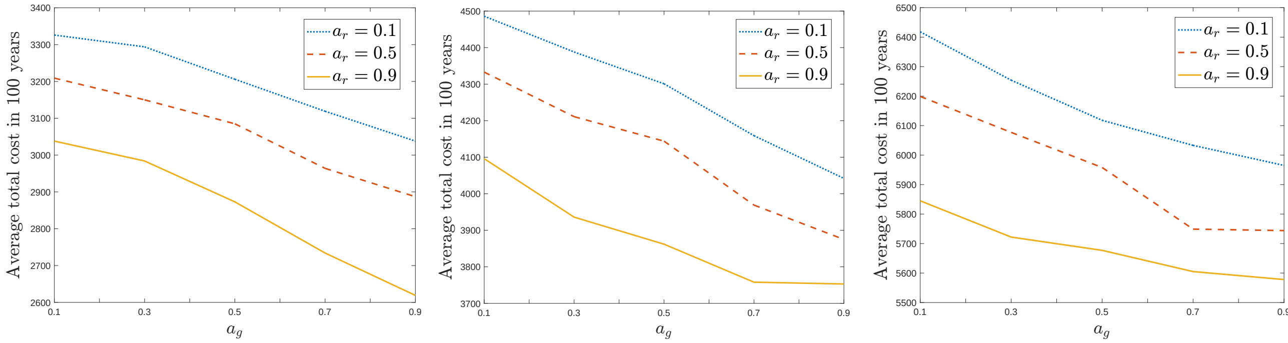

Figure 6. Average total cost as a function of government cooperation index a_g and resident cooperation index a_r for different NOAA projections: intermediate low (left), intermediate (middle), high (right).

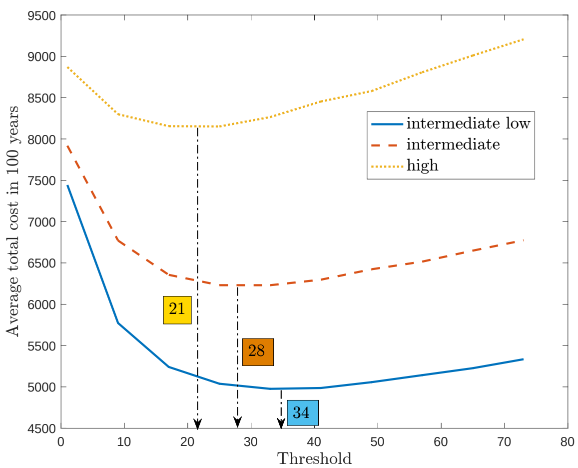

Figure 7. Average total cost for the shortsighted government policy as a function of investment threshold under the three sea level projections. The optimum thresholds for each projection are shown in colored boxes.

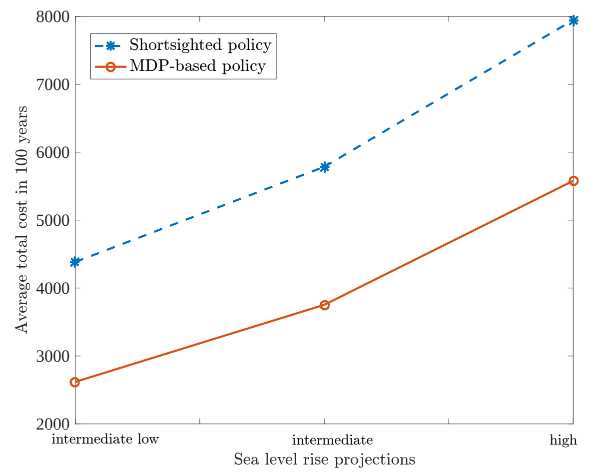

Figure 8. Comparison of the cumulative cost for threshold based and agent based optimal policy.

5. Discussions and Future Work

To the best of our knowledge, we presented the first comprehensive MDP framework with theoretical and numerical analysis for modeling socio-economic scenarios of the climate change-related sea level rise problem. Although we strived for a realistic framework, we know that the presented work can be improved in several aspects to obtain a more realistic framework. For instance, a continuous action space for the government can be integrated into the current model by considering the monetary value of investments instead of a discrete set of infrastructure improvement actions. Policy gradient methods such as the actor-critic method can be used instead of DQN to find the optimal policy with continuous actions. Moreover, multi-agent RL methods can be used to actively model the decisions of other stakeholders, such as residents and businesses, and interactions between them. Modeling of each stakeholder would ideally need a tailored approach that incorporates the characteristics of stakeholder into designing cost function, actions, etc. Last but not least, modeling the cooperation indices (a_g, a_r) is an interesting research direction that would require an interdisciplinary effort. By decomposing them into a number of fundamental traits, such as political, sociological, religious, and ethical traits of governments and residents, one can study how to obtain a_g, a_r, or another cooperation index for a new stakeholder, for a given community.

6. Conclusion

A novel Markov Decision Process (MDP) model has been proposed for simulating socio-economic scenarios of the sea level rise problem. The optimal government policy has been shown to rely on the observed sea levels through analyzing the expected cost functions. We also provided a deep reinforcement learning (RL) algorithm for finding the optimal policy, and presented scenario simulations for the Tampa Bay region using available economic data from the city government and regional sea level projections from NOAA. The simulations of several scenarios corroborated the importance of supportive residents and a proactive government that improves the infrastructure according to the observed sea level and does not undervalue the future cost of sea level rise. Finally, we discussed several ways of improving the proposed model for generating more realistic scenarios.

References

Which Coastal Cities Are at Highest Risk of Damaging Floods? New Study Crunches the Numbers, The World Bank. https://www.worldbank.org/en/news/feature/2013/08/19/coastal-cities-at-highest-risk-floods, 2013.

Marin County Community Development Agency, Game of Floods. https://www.marincounty.org/depts/cd/divisions/planning/csmart-sea-level-rise/game-of-floods, 2016.

The Cost Of Doing Nothing, Economic Impacts of Sea Level Rise in the Tampa Bay Area, Tampa Bay Regional Planning Council. http://www.tbrpc.org/wp-content/uploads/2018/11/2017-The_Cost_of_Doing_Nothing_Final.pdf, 2017.

City of Tampa Stormwater Utility Funding. https://www.tampagov.net/tss-stormwater/programs/assessment, 2019.

NOAA National Centers for Environmental Information (NCEI), U.S. Billion-Dollar Weather and Climate Disasters. https://www.ncdc.noaa.gov/billions/time-series, 2020.

Aksoy, H. Use of gamma distribution in hydrological analysis. *Turkish Journal of Engineering and Environmental Sciences*, 24(6):419–428, 2000.

Beatley, T. *Planning for coastal resilience: best practices for calamitous times*. Island Press, 2012.

Bellman, R. A Markovian decision process. *Journal of mathematics and mechanics*, pp. 679–684, 1957.

Berke, P. R. and Stevens, M. R. Land use planning for climate adaptation: Theory and practice. *Journal of Planning Education and Research*, 36(3):283–289, 2016.

Buchanan, M. K., Kopp, R. E., Oppenheimer, M., and Tebaldi, C. Allowances for evolving coastal flood risk under uncertain local sea-level rise. *Climatic Change*, 137(3-4):347–362, 2016.

Burke, M., Carnahan, L., Hammer-Levy, K., and Mitchum, G. Recommended projections of sea level rise in the tampa bay region (update). 2019.

Cebrián, A. C., Denuit, M., and Lambert, P. Generalized pareto fit to the society of actuaries? large claims database. *North American Actuarial Journal*, 7(3):18–36, 2003.

Daspit, A. and Das, K. The generalized pareto distribution and threshold analysis of normalized hurricane damage in the united states gulf coast. In *Joint Statistical Meetings (JSM) Proceedings, Statistical Computing Section, Alexandria, VA: American Statistical Association*, pp. 2395–2403, 2012.

Davidson, W. B. and Cotter, P. R. Psychological sense of community and support for public school taxes. *American Journal of Community Psychology*, 21(1):59–66, 1993.

Drogoul, A. Agent-based modeling for multidisciplinary and participatory approaches to climate change adaptation planning. In *Proceedings of RFCC-2015 Workshop, AIT*, 2015.

Eid, M. S. and El-Adaway, I. H. Decision-making framework for holistic sustainable disaster recovery: Agent-based approach for decreasing vulnerabilities of the associated communities. *Journal of infrastructure systems*, 24(3):04018009, 2018.

Ezer, T. Sea level rise, spatially uneven and temporally unsteady: Why the US East Coast, the global tide gauge record, and the global altimeter data show different trends. *Geophysical Research Letters*, 40(20):5439–5444, 2013.

Faboya, O. T., Figueredo, G. P., Ryan, B., and Siebers, P.-O. Position paper: The usefulness of data-driven, intelligent agent-based modelling for transport infrastructure management. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 144–149. IEEE, 2018.

Filar, J. and Vrieze, K. *Competitive Markov decision processes*. Springer Science & Business Media, 2012.

Fu, X., Song, J., Sun, B., and Peng, Z.-R. Living on the edge: Estimating the economic cost of sea level rise on coastal real estate in the Tampa Bay region, Florida. *Ocean & Coastal Management*, 133:11–17, 2016.

Fu, X., Gomaa, M., Deng, Y., and Peng, Z.-R. Adaptation planning for sea level rise: A study of US coastal cities. *Journal of Environmental Planning and Management*, 60 (2):249–265, 2017.

Hapke, C. J., Brenner, O., Hehre, R., and Reynolds, B. J. Assessing hazards along our Nation's coasts. Technical report, US Geological Survey, 2013.

Hine, A. C., Chambers, D. P., Clayton, T. D., Hafen, M. R., and Mitchum, G. T. *Sea Level Rise in Florida: Science,* *Impacts, and Options*. University Press of Florida. ISBN 9780813062891.

Hinkel, J., Jaeger, C., Nicholls, R. J., Lowe, J., Renn, O., and Peijun, S. Sea-level rise scenarios and coastal risk management. *Nature Climate Change*, 5(3):188–190, 2015.

Homans, G. C. Social behavior: Its elementary forms. 1974.

Howard, R. A. Dynamic programming and markov processes. 1960.

Kashem, S. B., Wilson, B., and Van Zandt, S. Planning for climate adaptation: Evaluating the changing patterns of social vulnerability and adaptation challenges in three coastal cities. *Journal of Planning Education and Research*, 36(3):304–318, 2016.

Katosh, J. P. and Traugott, M. W. Costs and values in the calculus of voting. *American Journal of Political Science*, pp. 361–376, 1982.

Le Cozannet, G., Manceau, J.-C., and Rohmer, J. Bounding probabilistic sea-level projections within the framework of the possibility theory. *Environmental Research Letters*, 12(1):14012, 2017.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540): 529, 2015.

Nicholls, R. J. Planning for the impacts of sea level rise. *Oceanography*, 24(2):144–157, 2011.

Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., Church, J. A., Clarke, L., Dahe, Q., and Dasgupta, P. *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. IPCC, 2014.

Patt, A. and Siebenhüner, B. Agent based modeling and adaption to climate change. *Vierteljahrshefte zur Wirtschaftsforschung*, 74(2):310–320, 2005.

Plant, N. G., Robert Thieler, E., and Passeri, D. L. Coupling centennial-scale shoreline change to sea-level rise and coastal morphology in the gulf of mexico using a bayesian network. *Earth's Future*, 4(5):143–158, 2016.

Potts, R., Jacka, L., and Yee, L. H. Can we Catch em All'? An exploration of the nexus between augmented reality games, urban planning and urban design. *Journal of Urban Design*, 22(6):866–880, 2017.

Rasinski, K., Tyler, T. R., and Fridkin, K. Exploring the function of legitimacy: Mediating effects of personal and institutional legitimacy on leadership endorsement and system support. *Journal of Personality and Social Psychology*, 49(2):386, 1985.

Rasinski, K. A. and Rosenbaum, S. M. Predicting citizen support of tax increases for education: A comparison of two social psychological perspectives. *Journal of Applied Social Psychology*, 17(11):990–1006, 1987.

Research Council, N. Informing decisions in a changing climate. Panel on Strategies and Methods for Climate-Related Decision Support of the Committee on the Human Dimensions of Global Change, National Research Council of the National Academies, 2009.

Sallenger Jr, A. H., Doran, K. S., and Howd, P. A. Hotspot of accelerated sea-level rise on the Atlantic coast of North America. *Nature Climate Change*, 2(12):884–888, 2012.

Schrock, G., Bassett, E. M., and Green, J. Pursuing equity and justice in a changing climate: Assessing equity in local climate and sustainability plans in US cities. *Journal of Planning Education and Research*, 35(3):282–295, 2015.

Sergent, P., Prevot, G., Mattarolo, G., Brossard, J., Morel, G., Mar, F., Benoit, M., Ropert, F., Kergadallan, X., Trichet, J.-J., et al. Adaptation of coastal structures to mean sea level rise. *La Houille Blanche*, (6):54–61, 2014.

Singh, S., Jaakkola, T., Littman, M. L., and Szepesvári, C. Convergence results for single-step on-policy reinforcement-learning algorithms. *Machine learning*, 38(3):287–308, 2000.

Sutton, R. S. and Barto, A. G. *Reinforcement learning: An introduction*. MIT press, 2018.

Sweet, W. V. and Park, J. From the extreme to the mean: Acceleration and tipping points of coastal inundation from sea level rise. *Earth's Future*, 2(12):579–600, 2014.

Tribbia, J. and Moser, S. C. More than information: what coastal managers need to plan for climate change. *Environmental science & policy*, 11(4):315–328, 2008.

Tyler, S. and Moench, M. A framework for urban climate resilience. *Climate and development*, 4(4):311–326, 2012.

Uchiyama, Y. and Watanabe, S. Modeling Natural Catastrophic Risk and its Application, 2006.

Wahl, T., Calafat, F. M., and Luther, M. E. Rapid changes in the seasonal sea level cycle along the US Gulf coast from the late 20th century. *Geophysical Research Letters*, 41(2):491–498, 2014.

Wahl, T., Jain, S., Bender, J., Meyers, S. D., and Luther, M. E. Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nature Climate Change*, 5(12):1093, 2015.

Xia, R., Kannan, S., and Castleman, T. The Ocean Game: The sea is rising. Can you save your town? https://www.latimes.com/projects/la-me-climate-change-ocean-game/, 2019.