Modeling and Simulating Adaptation Strategies against Sea-Level Rise using Multi-Agent Deep Reinforcement Learning

Salman Sadiq Shuvo, *Student Member, IEEE,* Yasin Yilmaz, *Senior Member, IEEE,* Alan Bush, Mark Hafen

***Abstract*—Sea-level rise (SLR) problem, which is a major outcome of climate change, has been well documented and studied. Although it is globally observed due to climate change, local projections are needed to plan SLR adaptation strategies accurately. Since SLR is a community-wide multi-stakeholder problem at the local level, adaptation strategies can be more successful if the main stakeholders, e.g., government, residents, businesses, collaborate in shaping them. Simulating the local socioeconomic system around SLR, including the interactions between essential stakeholders and nature, can be an effective way of evaluating different adaptation strategies and planning the best strategy for the local community. This work presents how such an SLR socioeconomic system can be modeled as a Markov decision process (MDP) and simulated using multi-agent reinforcement learning (RL). The proposed multi-agent RL framework serves two purposes. It provides a general scenario planning tool to investigate the cost-benefit analysis of natural events (e.g., flooding, hurricane) and agents' investments (e.g., infrastructure improvement). It also shows how much the total cost due to SLR can be reduced over time by optimizing the adaptation strategies. We demonstrate the proposed scenario planning tool using available economic data and sea-level projections for Pinellas County, Florida, in a case study.**

I. Introduction

Sea-level rise (SLR) is one of the most catastrophic outcomes of the global increase in greenhouse gas (GHG) emissions and climate change. While many policy makers have committed to reducing GHG emissions since the Paris agreement in 2015, coastal communities will require adaptation strategies to deal with SLR problems before harnessing the benefit of worldwide GHG emission reduction [1]. Due to climate change and SLR, storm surge, recurrent hurricanes, and permanent inundation pose significant challenges to most coastal cities, many of which are among the world's largest cities [2]. Underdeveloped areas will also face many social and financial crises apart from the property loss [3], [4].

Recently, in the literature, a quantification of present and future flood damages in 136 major coastal cities is presented in [5]. Population growth is also considered in [6] to assess the potential magnitude of future impacts in the continental US. The study in [7] proposes a coherent statistical model for coastal flood frequency analysis and validates a mixture model for 68 tidal stations along the contiguous United States coast with long-term observed data. [8] demonstrates a methodology to assess the economic impacts of climate change at city scale (Copenhagen, Denmark) and the benefits of SLR adaptation. [9] uses HAZUS-MH [10] coastal flood hazard modeling and loss estimation tools to determine flood extent and depth and the corresponding monetary losses to infrastructure in Miami-Dade County. A case study comparing the cost-effectiveness of nature-based and coastal adaptation for the Gulf Coast of the United States is presented in [11].

Several countries are already making significant investments toward reducing the catastrophic and long-term impacts of SLR. However, the progress in risk reduction is far behind the coastal development and population growth globally [12]. The need for appropriate planning and execution for hurricane and flood protection is becoming more prominent as SLR risks grow. The major challenge is the daunting cost of undertaking mega projects, and building megastructures [13]. The US government spends billions of dollars to fund agencies like the US Army Corps of Engineers and the US Department of Transportation for hazard mitigation. In 2020, the Federal Emergency Management Agency (FEMA) announced to grant up to \$660 million in grant funding, including a record-breaking \$500 million for the Building Resilient Infrastructure and Communities (BRIC) pre-disaster mitigation grant program and \$160 million for the Flood Mitigation Assistance program [14].

The success of such investments depends on understanding different risk drivers from a financial viewpoint, including SLR and the current state of infrastructure [15]. A disaster cost and investment benefit analysis for a region can provide a guideline to the government about budgeting its funds towards different mitigation programs. Furthermore, since SLR is not uniform across the globe [16], the adaptation planning for different regions may differ significantly. Risk assessment and investment planning for different regions might require separate analyses and consider different sea-level projections [17]. Governments, in particular administrators and officials in the corresponding agencies, need substantial information for better strategic vision and adaptation planning [1], [18], [19]. The challenging task of adaptation planning demands considering various stakeholder dynamics and SLR scenarios [20]. Explicitly modeling the stakeholders' reactions to SLR scenarios can help create strategies suitable for local impacts and resilience management, and requires planning mechanisms such as agent-based modeling and sequential decision making [21]–[23].

To this end, we here study the interactions between SLR

This work was supported in part by the U.S. National Science Foundation under the grant CNS-1737598.

The authors are with the University of South Florida, Tampa, FL 33620, USA. E-mail: {salmansadiq,yasiny,alanbush,mhafen}@usf.edu

stakeholders under different SLR scenarios using multi-agent deep reinforcement learning (RL). Specifically, we use a probabilistic model for nature's response to the collaborative policies of three local agents (government, residents, businesses). The proposed multi-agent RL framework serves two purposes. It provides a general scenario planning tool to investigate the cost-benefit analysis of natural events (e.g., flooding, hurricane) and agents' investments (e.g., infrastructure improvement), and also shows how much the total cost due to SLR can be reduced over time by optimized adaptation strategies. We demonstrate the proposed scenario planning tool using available economic data and sea-level projections for Pinellas County, Florida, in a case study. Although we here focus on the SLR problem, the proposed scenario planning framework can be adapted to other natural and socioeconomic systems. A preliminary version of this work was presented in [24]. This submission greatly enhances both the intellectual merit and the broader impacts of the work. The major improvements include a more realistic multi-agent setup, more effective state-of-the-art deep RL methods, and a case study with extensive experimental results based on real economic data from the Tampa Bay area.

The remainder of the paper is organized as follows. The proposed multi-agent RL framework is presented and analyzed in Sec. II. The case study is given in Sec. III, and the concluding discussions and remarks are provided in Sec. IV.

II. Multi-Agent RL Framework

A. Agent-based Modeling for Adaptation Strategies

The long-term effects of adaptation strategies can be effectively simulated using agent-based modeling, where an agent represents each stakeholder, and its actions are modeled through realistic policies. In the considered sequential decision-making setup, at the beginning of a year, residents and businesses decide on their additional tax contributions towards SLR adaptation; then the government decides on its own investment amount against SLR and implements an SLR adaptation strategy based on the total investment amount from all stakeholders. Finally, at the end of the year, the total cost from natural events is observed, which can serve as a feedback about the recent SLR adaptation strategies and inform the agents' future actions. A straightforward but realistic policy is the cost-based policy in which an agent decides whether to invest and the investment amount depending on the cost it experiences from the natural events. While different threshold levels on the natural cost can be used to model different agent prototypes (i.e., lower threshold for more reactive agents), it is hard to select such thresholds to link them to realistic prototypes. We here show that a multi-agent RL framework can be used to model realistic stakeholder policies in an easily controllable way. Our proposed RL framework provides an intuitive parameterization (cooperation indices between zero and one) to simulate different stakeholder prototypes conveniently. Moreover, the proposed RL framework illustrates how much cost can be saved by proactive and fully cooperative stakeholders with optimized decision policies.

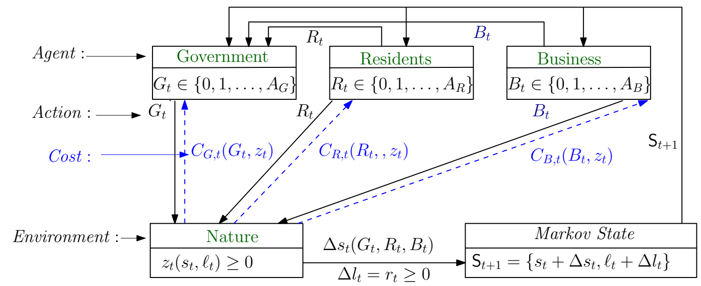

Fig. 1. Proposed multi-agent MDP framework.

B. MDP Formulation

Markov Decision Process (MDP) is an effective tool for modeling optimal sequential decision making problems. MDP provides a tractable mathematical formulation to model decision-making tasks in situations where outcomes are partly random and partly regulated by the agent actions [25]–[27]. RL is a field of machine learning which deals with iteratively learning optimal decision policies (i.e., which action to take at which state) from experience with the environment through the action-feedback mechanism. It provides a data-driven approach to solve MDP problems, which is of critical importance for the high-dimensional state and/or action spaces where exact dynamic programming solutions are not feasible [28]. Specifically, deep RL methods (e.g., deep Q-network (DQN)) utilize deep neural networks to handle significantly high-dimensional problems, which are too complex for traditional tabular RL methods (e.g., Q-learning).

We next explain the proposed cost models for the environment and the agents under the MDP framework. We propose a multi-agent MDP framework to model the behaviors of local SLR stakeholders (government, residents, and businesses), and their interactions among themselves and with nature. As shown in Fig. 1, government, which is the major decision-maker in dealing with the SLR problem, at each time step t takes an action (i.e., investment decision) G_t and receives feedback from nature through the cost $z_{G,t}$. The other two agents, residents and businesses, similarly take actions R_t, B_t and receives feedback from nature through the costs $z_{R,t}$ and $z_{B,t}$, respectively. Then, this natural and socioeconomic system moves to a new state S_{t+1} based on the current state S_t, agents' action G_t, R_t, B_t, and SLR r_t. The system state consists of the city's infrastructure state s_t and the sea level ℓ_t, i.e., $\mathsf{S}_t = \{s_t, \ell_t\}$. The agents' decisions determine the infrastructure state $s_t = s_0 + \sum_{m=1}^{t}(\Delta s_{G,m} + \Delta s_{R,m} + \Delta s_{B,m}) = s_{t-1} + \Delta s_t(G_t, R_t, B_t)$. Likewise, the sea level at time t is given by the cumulative SLR values: $\ell_t = \ell_0 + \sum_{m=1}^{t} r_m = \ell_{t-1} + r_t$, where r_m is the SLR value at time m. Here, s_0 and ℓ_0 are respectively the initial infrastructure state and the initial sea level of the region relative to a reference year. In terms of simulations, these are two user-defined numbers representing the existing states at the beginning of the simulations. The system state satisfies the Markov property: $\mathsf{P}(\mathsf{S}_{t+1}|\mathsf{S}_t, \dots, \mathsf{S}_0) = \mathsf{P}(\mathsf{S}_{t+1}|\mathsf{S}_t)$. We assume that the government observes the other agents' actions R_t, B_t, hence has the complete knowledge of multi-agent

MDP. However, the residents and businesses do not necessarily know the other agents' actions, hence the MDP is partially observable to them. The parameters of the proposed multi-agent MDP framework are summarized in Table I.

TABLE I
MODEL PARAMETERS.

Initial sea level	$\ell_0 \geq 0$
SLR at time t	$r_t \geq 0$
Sea level at time t	$\ell_t = \ell_0 + \sum_{m=1}^{t} r_m$
Initial infrastructure state	$s_0 \in \{1, 2, ...\}$
Infrastructure improvement at time t,	$\Delta s_t(G_t.R_t, B_t)$
Infrastructure state at time t,	$s_t = s_0 + \sum_{m=1}^{t} \Delta s_m$
Government's decision at time t	$G_t \in \{0, 1, \dots, A_G\}$
Residents' decision at time t	$R_t \in \{0, 1, \dots, A_R\}$
Businesses' decision at time t	$B_t \in \{0, 1, \dots, A_B\}$

C. Modeling Nature

We model nature's cost $z_t = z_{G,t} + z_{R,t} + z_{B,t}$ using the generalized Pareto distribution, which is commonly used to model catastrophic losses, e.g., [29]–[31]. It is known that the storm- and flooding-related costs for the stakeholders have been increasing with SLR [32]. Thus, we model the scale parameter of generalized Pareto distributed z_t directly proportional to the most recent sea level ℓ_t and inversely proportional to the most recent infrastructure state s_t. The cost from nature is distributed among the stakeholders through the multiplying factors ($m_G + m_R + m_B = 1$) for government ($z_{G,t} = m_G \times z_t$), residents ($z_{R,t} = m_R \times z_t$), and businesses ($z_{B,t} = m_B \times z_t$). These factors vary with regions; however, generally $m_G > m_R, m_B$ since typically government is faced with most of the cost from nature.

The probabilistic model for the cost from nature is given by

$$z_t \sim \text{GeneralizedPareto}(\xi, \sigma_t, \mu)$$
$$\mu \geq 0,\ \xi < 0,\ \sigma_t = \frac{\eta(\ell_t)^p}{(s_t)^q} \tag{1}$$

where μ, σ_t, ξ are the location, scale, and shape parameters of generalized Pareto, respectively; and $\eta > 0, p \in (0,1), q > 0$ are our additional model parameters. The parameters ξ, μ, η, p, q help to regulate the impact of the most recent sea level ℓ_t over the nature's cost z_t relative to the most recent infrastructure state s_t. Choosing an appropriate set of parameters depends on the region considered for simulations. 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. For certain values of shape parameter ξ, the expected value and range of the cost are as follows:

$$\mathsf{E}[z_t] = \mu + \frac{\eta \ell_t^p}{(1-\xi)s_t^q} \quad \text{for} \quad \xi < 1$$

$$\mu \leq z_t \leq \mu - \frac{\eta \ell_t^p}{\xi s_t^q} \quad \text{for} \quad \xi < 0.$$

The location parameter is set to be positive, $\mu > 0$, to generate positive disaster cost, $z_t > 0$. We choose $\mu = \$30$ million from historical data provided in Table A-1 in [33], which indicates that a year with no serious natural disaster might produce this cost, typically to cover maintenance. To get an upper bound on z_t, we need the shape parameter to be negative, $\xi < 0$. We select $\xi = -0.1$, which limits the upper bound to roughly 10 times the expected cost. The other parameters η, p, q are set according to the cost projections in Pinellas County presented in [34], and in Sec. III.

D. Modeling Stakeholders

In our model, the government is the biggest stakeholder and implementer of the investment decisions for other agents too. At each time step, e.g., a year, the government decides the degree of its investment $G_t \in \{0, 1, 2, ..., A_G\}$ for infrastructure development, where A_G is a finite positive integer. $G_t = 0$ means no investment at step t. Hence, there are $A_G + 1$ possible actions for the government at each time step. The numerical value of $G_t = m$ can be interpreted as spending m unit money towards infrastructure development or the $m + 1$ th action among $A_G + 1$ different actions with increasing cost and effectiveness. Possible government actions include but are not limited to building seawalls, raising roads, widening beaches, building traditional or horizontal levees, placing storm-water pumps, improving sewage systems, relocating seaside properties, etc. [35]–[37]. The range of G_t 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 $C_{G,t}$ to the agent at each time t consists of the investment cost and cost from nature. We assume most of the business and residential properties are insured by the government. So $f \in (0,1)$ fraction of their insurance payments, $f \times I_{R,t}$ and $f \times I_{B,t}$ respectively for residents and businesses go to the government, hence negatively contribute to $C_{G,t}$. We explain modeling $I_{R,t}$ and $I_{B,t}$ later in this section while presenting the models for residents and businesses. Since the government's investment decision has an integer value, we model the total cost as $C_{G,t} = \alpha_G G_t + z_{G,t} - f(I_{R,t} + I_{B,t})$ using parameter α_G to map the decision to monetary value. The discounted cumulative cost for the government in T time steps is given by $C_{G,T} = \sum_{t=0}^{T} \lambda_G^t [\alpha_G G_t + z_{G,t} - f(I_{R,t} + I_{B,t})]$, where the discount factor $\lambda_G \in (0,1)$ discounts the weight of future costs following the common practice in MDP. In our model, this discount factor also serves as a measure of the government's cooperation towards long-term welfare, hence, termed as the *government's cooperation index* in this paper. Higher λ_G corresponds to a more cooperative government which better recognizes the future SLR costs from nature, compared to a more short-sighted government represented by lower λ_G. The government's objective is to minimize the expected cumulative cost $\mathsf{E}[C_{G,T}]$ by taking investment actions $\{G_t\}$ over time.

Residents' community decides its own action based on its learning of the environment and hence modeled as an agent in our multi-agent setup. The community organization decides the degree of its investment $R_t \in \{0, 1, 2, ..., A_R\}$, i.e., how much additional tax they are going to pay to the authority to build infrastructure for them. The unit cost of residents' investment α_R is limited to some fraction of the government investment unit α_G since the government is expected to cover the majority of infrastructure investment costs. The change in infrastructure state by residents' investment decision R_t is set relative to the government: $\Delta s_{R,t} = R_t \times \alpha_R/\alpha_G$ where $\Delta s_{G,t} = G_t$. It is assumed that these coastal residents typically insure their vulnerable properties, so their cost from nature is mainly due to the raise of insurance premiums. Insurance premiums go up if the insurance had to pay more for recent catastrophic events. Hence, the insurance premium can be approximated based on the historical cost from nature as $I_{R,t} = I_{R,0} \times \rho_R^t + I_R \times \sum_{m=1}^{t-1} \rho_R^{t-m} z_{G,m}$, where $\rho_R \in (0,1)$ is the insurance company's memory factor for past events, and I_R is the coefficient that maps the total recent natural cost of the insurance company (i.e., the government) to insurance premium. Pre-existing insurance premium for the region, $I_{R,0}$, can serve as the initial value for the simulation. Apart from the insurance cost, the residents also endure a fraction of the cost from nature, represented by $z_{R,t} = m_R \times z_t$. The discounted cumulative cost for the residents in T time steps is given by $C_{R,T} = \sum_{t=0}^{T} \lambda_R^t(\alpha_R R_t + z_{R,t} + I_{R,t})$, where the discount factor $\lambda_R \in (0,1)$ can be interpreted as the *residents' cooperation index*, as in the government's model. The residents' objective is to minimize the expected cumulative cost $\mathsf{E}[C_{R,T}]$ by taking investment actions $\{R_t\}$ over time.

Businesses are another major stakeholder of SLR impacts. Businesses get monetary loss through inundation, loss of customers, property damages, and increasing insurance premiums. Similar to the residents' model, we consider a business association to implement their collective actions. The business association takes action $B_t \in \{0,1,2,...,A_B\}$, i.e., decides on their degree of monetary contribution towards infrastructure development. The unit cost of businesses investment α_B typically ranges between α_R and α_G. The change in infrastructure state by businesses' investment decision B_t is $\Delta s_{B,t} = B_t \times \alpha_B/\alpha_G$. Similar to the residents, businesses have insurance and non-insurance costs. Insurance premiums go up if the insurance had to pay more for recent catastrophic events. Hence, the insurance premium for businesses is modeled as $I_{B,t} = I_{B,0} \times \rho_B^t + I_B \times \sum_{m=1}^{t-1} \rho_B^{t-m} z_{G,m}$, where $\rho_B \in (0,1)$ is the insurance company's memory factor for past events, I_B and $I_{B,0}$ are the insurance coefficient and the initial insurance premium for businesses, respectively. The discounted cumulative cost for the business agent in T time steps is given by $C_{B,T} = \sum_{t=0}^{T} \lambda_R^t(\alpha_B B_t + z_{B,t} + I_{B,t})$, where discount factor λ_B can represent businesses' awareness and cooperation against SLR and is called *businesses' cooperation index*. The businesses' objective is to minimize the expected cumulative cost $\mathsf{E}[C_{B,T}]$ by taking investment actions $\{B_t\}$ over time.

E. Optimal Policy Analysis

In our proposed MDP structure, each agent tries to minimize its expected total cost $\mathsf{E}[C_T]$ in T time steps by following an optimal investment policy. At each time step, first, residents and businesses take their actions R_t and B_t respectively; then, the government collects their investments, makes its decision G_t, and implements the monetary investment towards developing infrastructure. So, the next state transition is fully observable to the government and at the same time partially observable to the other two agents. We begin with the optimal policy analysis for the government. Its optimal value function, which gives the minimum expected total cost possible at each state (s_t, ℓ_t), characterizes the best action policy $\{G_t\}$, and is written as

$$V_G(s_t, \ell_t, O_{G,t}) = \min_{\{G_t\}} \mathsf{E}[C_{G,t}^T | \{G_t\}],$$

where $C_{G,t}^T = \sum_{\tau=0}^{T} \lambda_G^\tau C_{G,t+\tau}$ is the cumulative cost starting from time t. We know from the main body of the paper

$$C_{G,t} = \alpha_G G_t + z_{G,t} - f(I_{R,t} + I_{B,t}) \tag{2}$$

and, $C_{G,T} = \sum_{t=0}^{T} \lambda_G^t[\alpha_G G_t + z_{G,t} - f(I_{R,t} + I_{B,t})]$.

Here, the government's observation $O_{G,t}$ represents its knowledge about other agents' actions at time t. To find the optimal policy, the Bellman equation

$$V_G(s_t, \ell_t, O_{G,t}) = \min_{G_t} \mathsf{E}[C_{G,t} + \lambda_G V_G(s_{t+1}, \ell_{t+1}) | G_t]$$

provides a recursive approach by focusing on finding the optimal action G_t at each time step using the successor state value instead of trying to find the entire policy $\{G_t\}$ at once. Using the cost expression given by (2) and considering possible $A_G + 1$ actions for G_t this iterative equation can be rewritten as

$$\begin{aligned} &V_G(s_t, \ell_t, O_{G,t}) = \\ &\min \Big\{ \underbrace{\mathsf{E}\big[z_{G,t} - f(I_{R,t} + I_{B,t}) + \lambda_G V(\hat{s}_t, \ell_t + r_t)\big]}_{F_0(\hat{s}_t, \ell_t)}, \\ &\underbrace{\mathsf{E}\big[\alpha_G + z_{G,t} - f(I_{R,t} + I_{B,t}) + \lambda_G V(\hat{s}_t + 1, \ell_t + r_t)\big]}_{F_1(\hat{s}_t, \ell_t)}, \\ &\underbrace{\mathsf{E}\big[2\alpha_G + z_{G,t} - f(I_{R,t} + I_{B,t}) + \lambda_G V(\hat{s}_t + 2, \ell_t + r_t)\big]}_{F_2(\hat{s}_t, \ell_t)}, \dots, \\ &\underbrace{\mathsf{E}\big[A_G\alpha_G + z_{G,t} - f(I_{R,t} + I_{B,t}) + \lambda_G V(\hat{s}_t + A_G, \ell_t + r_t)\big]}_{F_{A_G}(\hat{s}_t, \ell_t)} \Big\} \end{aligned} \tag{3}$$

where $\hat{s}_t = s_t + R_t \times \alpha_R/\alpha_G + B_t \times \alpha_B/\alpha_G$ is the deterministic next infrastructure state before government investment, termed as augmented infrastructure state. The knowledge of other agents' current actions provides the basis of our optimal policy analysis for the government and Theorem 1.

At each time step t, action G_t shapes the instant cost $C_{G,t}$ and moves the system to the next state, which determines the discounted future cost $\lambda_G V(s_{t+1}, l_{t+1})$. The optimum policy chooses among the investment actions $G_t \in \{0,1,2,....,A_G\}$ that has the minimum expected total cost, $\min_m\{F_m(\hat{s}_t, \ell_t)\}$, as shown in (3). Since the functions $\{F_0(\hat{s}_t, \ell_t), \dots, F_{A_G}(\hat{s}_t, \ell_t)\}$ determine the optimal policy, we next analyze them to characterize the optimal government policy.

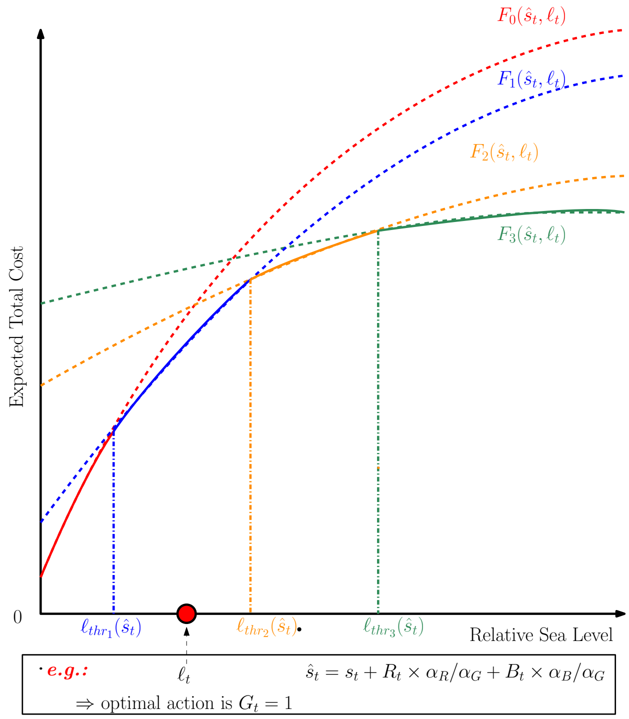

Fig. 2. Expected total costs as a function of sea level for an example case with $A_G = 3$. The expected total costs intersect each other only once for any given infrastructure state according to Theorem 1.

Theorem 1. *For $m = 0, 1, \ldots, A_G$, $F_m(\hat{s}_t, \ell_t)$ is nondecreasing and concave in ℓ_t for each $\hat{s}_t$; and the derivative of $F_m(\hat{s}_t, \ell_t)$ with respect to ℓ_t is lower than that of $F_{m-1}(\hat{s}_t, \ell_t)$.*

Proof is provided in the Appendix. For a specific infrastructure state $\hat{s}_t$, expected costs $F_0(\hat{s}_t, \ell_t), \ldots, F_3(\hat{s}_t, \ell_t)$ are illustrated in Figure 2 according to Theorem 1. The optimum policy picks the minimum of the $A_G + 1 = 4$ curves at each time, which is shown with the solid curve in Figure 2. As a result of Theorem 1, we next give the outline of optimum policy in Corollary 1.

Corollary 1. *The optimum policy, at each augmented infrastructure state $\hat{s}_t$, compares the sea level ℓ_t with at most A_G thresholds where each threshold signifies a change of optimal action.*

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

Figure 2 gives an example case where there are $A_G = 3$ thresholds $\ell_{thr_1}(\hat{s}_t), \ell_{thr_2}(\hat{s}_t), \ell_{thr_3}(\hat{s}_t)$, which depend on $\hat{s}_t$ and indicate change points of optimal action. However, depending on the slopes of $\{F_m(\hat{s}_t, \ell_t)\}$ curves at each augmented infrastructure state $\hat{s}_t$, there may be less than A_G change points. To summarize, for a given state $(\hat{s}_t, \ell_t)$, the optimum policy chooses G_t based on the relative value of the current sea level ℓ_t, with respect to the augmented infrastructure state $\hat{s}_t$.

The thresholds also depend on the cooperation index λ_G. Higher cooperation indices set the thresholds lower and vice versa. Intuitively, as λ_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 λ_G implies underestimated future costs and thus overemphasized immediate investment costs, which results in a high threshold for investment.

Optimal value functions for residents' and businesses' are similar to the government's. The Bellman equation for residents' is

$$V_R(s_t, \ell_t, O_{R,t}) = \min_{R_t} \mathsf{E}[C_{R,t} + \lambda_R V_R(s_{t+1}, \ell_{t+1}) | R_t]$$

Ideally, each agent would like to see other agents' actions. Nevertheless, in the considered problem, residents and businesses do not have the information of others' actions. Since the states do not change drastically, it is reasonable to approximate the agents' previous optimal action as their recent action. So, the residents approximate $G_t \approx G_{t-1}$ and $B_t \approx B_{t-1}$ in its observation $O_{R,t}$, where G_{t-1}, and B_{t-1} are the actions taken in previous time step by the corresponding agents. Similarly, the business approximate $G_t \approx G_{t-1}$ and $R_t \approx R_{t-1}$ in its observation $O_{B,t}$ for the value function

$$V_B(s_t, \ell_t, O_{B,t}) = \min_{B_t} \mathsf{E}[C_{B,t} + \lambda_B V_B(s_{t+1}, \ell_{t+1}) | B_t].$$

Due to such partial knowledge and approximations, functional analysis as in Theorem 1 is not tractable for residents and businesses.

F. Multi-Agent RL Algorithms

The continuous sea level values, which cause an infinite number of possible states, necessitate a deep RL algorithm instead of a traditional RL algorithm. We consider two deep RL approaches for comparison. The deep Q-network (DQN) algorithm [38], 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 for each of the three agents. The Advantage Actor-Critic (A2C) algorithm, a policy gradient-based algorithm [39], is a popular choice for multi-agent deep RL. A2C uses two neural networks:

(i) The actor network, also known as the policy network, outputs the probability for each action through a softmax function. It is updated using the gradient of expected return of the policy π_θ with respect to the weights θ of the neural network, e.g., for the government $\mathsf{E}[\nabla_{\theta_G} \log \pi_{\theta_G}(G_t|\mathsf{S}_t) D_G(\mathsf{S}_t, G_t)]$ where $\pi_{\theta_G}(G_t|\mathsf{S}_t)$ denotes the probability for action G_t at state S_t, and the advantage function is given by $D_G(\mathsf{S}_t, G_t) = C_{G,t} + \lambda_G V_{\psi_G}(\mathsf{S}_{t+1}) - V_{\psi_G}(\mathsf{S}_t)$, where V_{ψ_G} is the output of the critic network (see below) with ψ denoting the network weights.

(ii) The critic network, which is also known as the value network, is used to learn the value function for each state, e.g., $V_{\psi_G}(\mathsf{S}_t)$ for the government. It is updated using the gradient of the squared advantage function, $\mathsf{E}[\nabla_{\psi_G} D_G^2]$.

The cost from natural events, which is modeled with a generalized Pareto distribution, can have a high variance depending on the parameter settings, i.e., regular flooding costs in a typical year vs. major hurricane costs in another year. Our experiments in the case study, explained next, corroborate the previous findings that A2C in general deals with high variance more successfully than DQN [39].

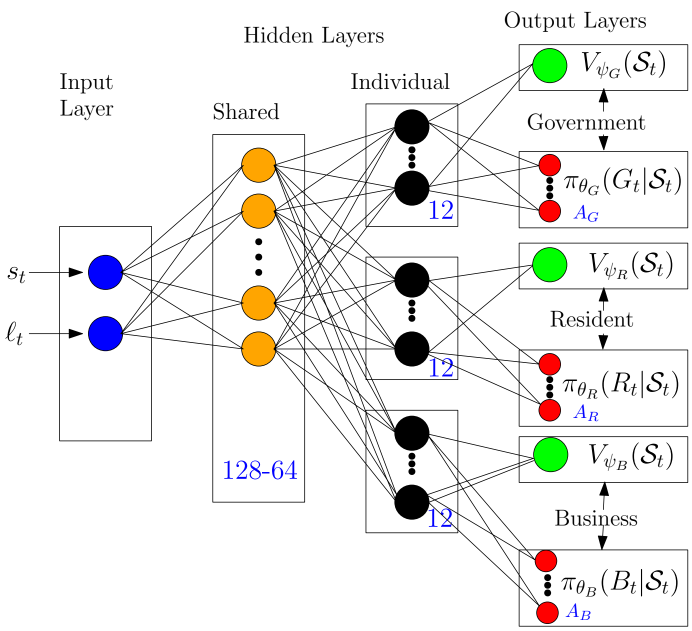

Fig. 3. Unified A2C structure for all three agents. Numbers inside the box give the neuron numbers in each layer.

For the multi-agent implementation of A2C, we consider three different structures. In the first one, a single deep neural network structure is used for all agents based on the similarities between their state and cost definitions. As shown in Fig. 3, the input layer, which represents the common system state $\mathsf{S}_t = \{s_t, \ell_t\}$, is the same for all agents. Since the cost functions for the agents have similarities, they also share some hidden layers. From there on, the agents have their individual hidden layers to output their state values $V_{\psi_G}, V_{\psi_R}, V_{\psi_B}$ (critic network) and action probabilities $\pi_{\theta_G}, \pi_{\theta_R}, \pi_{\theta_B}$ (actor network). In this unified A2C structure, the interaction between agents is not explicitly implemented through observations of other agents' actions $O_{G,t}, O_{R,t}, O_{B,t}$ at the input.

We next consider using a separate neural network for each agent with s_t, ℓ_t, and O_t at the input, as shown in Fig. 4. In this structure, the agents explicitly use the other agents' actions O_t in their input states. Government observes the residents' and businesses' actions before taking its own action, i.e., $O_{G,t} = \{R_t, B_t\}$. However, residents and businesses only know the previous actions of other agents, i.e., $O_{R,t} = \{G_{t-1}, B_{t-1}\}$ and $O_{B,t} = \{G_{t-1}, R_{t-1}\}$.

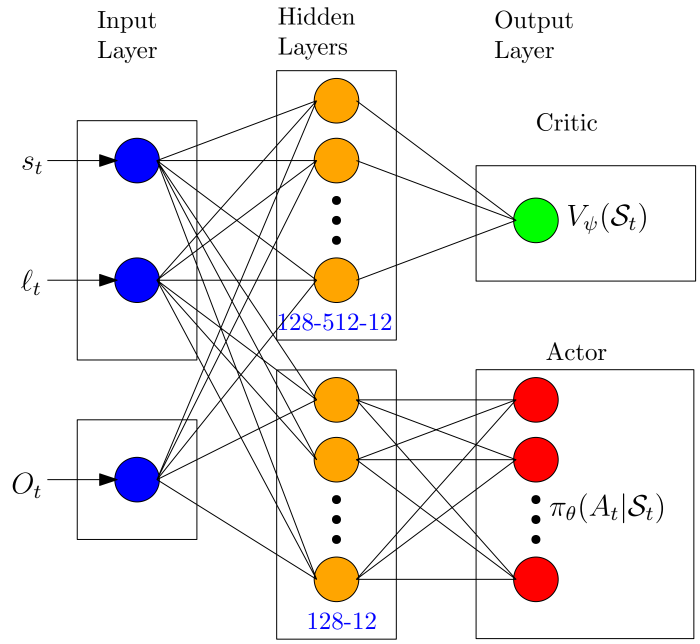

Fig. 4. Separate A2C structure for each agent. Numbers inside the box give the neuron numbers in each layer.

Algorithm 1 Multi-agent A2C algorithm (Fig. 4)

1: *Input:* $\mu, \epsilon, \eta, p, q, m_G, m_R, m_B, \alpha_G, \alpha_R, \alpha_B,$
2: *Input:* $\lambda_G, \lambda_R, \lambda_B, \rho_R, \rho_B, I_R, I_B$
3: *Initialize* policy network with random weights θ_G, θ_R, θ_B and critic network with random weights ψ_G, ψ_R, ψ_B.
4: **for** episode $= 1, 2, \ldots$ **do**
5:   *Initialize* state $\mathsf{S}_0 = (s_0, \ell_0)$
6:   **for** t $= 1, 2, \ldots, T$ **do**
7:     Sample action G_t, R_t, B_t from probability distribution generated by actor networks θ_G, θ_R, θ_B.
8:     Execute action G_t, R_t, B_t and observe costs $C_{G,t}$, $C_{R,t}$, $C_{B,t}$
9:   **end for**
10:   Update actor network θ_G (and similarly θ_R, θ_B) by back propagating $\mathsf{E}[\nabla_{\theta_G} \log \pi_{\theta_G}(G_t|\mathsf{S}_t) D_G(\mathsf{S}_t, G_t)]$.
11:   Update critic network ψ_G (and similarly ψ_R, ψ_B) by back propagating $\mathsf{E}[\nabla_{\psi_G} D_G^2]$.
12: **end for**

As a third (hybrid) structure, we also consider using a single critic network common to all agents. Specifically, each agent has its own actor network, as in Fig. 4, but shares a common critic network. According to our experimental results in the case study, among the three multi-agent A2C structures, the separate A2C structure (Fig. 4) performs the best.

Algorithm 1 summarizes the separate A2C algorithm (Fig. 4). Each episode consists of Monte-Carlo simulations in which several states are visited according to the current policy defined by the current actor network. Line 1 initializes the disaster cost and investment cost parameters. Line 2 sets up the discount factors and insurance parameters. An episode starts with the initial relative sea level and infrastructure state. Line 7 shows the action selection procedure for the A2C agents. Then, the

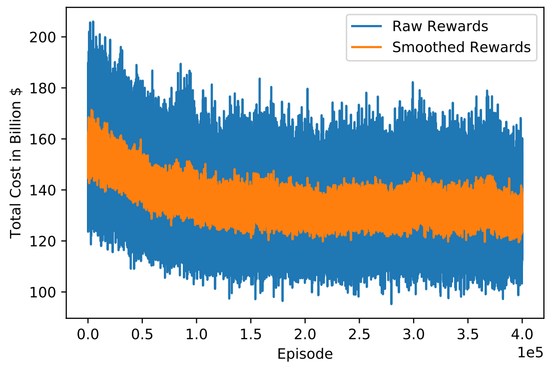

Fig. 5. Average episodic total cost of all agents in the separate A2C policy for the high SLR scenario.

TABLE II
RELATIVE SEA LEVEL (MM) FOR DIFFERENT SCENARIOS FOR ST. PETERSBURG.

Year	NOAA 2017 [41]			Simulation adjusted RSL, $\hat{\ell}_t$		
	Int-low	Intermediate	High	Int-low	Intermediate	High
2000	0	0	0	n/a	n/a	n/a
2010	50	70	110	n/a	n/a	n/a
2020	110	150	220	100	100	100
2030	170	240	380	160	190	260
2040	220	330	540	210	280	420
2050	290	440	780	280	390	660
2060	350	570	1060	340	520	940
2070	410	710	1390	400	660	1270
2080	470	860	1740	460	810	1620
2090	520	1030	2150	510	980	2030
2100	580	1190	2590	570	1140	2470
2120	670	1430	3460	660	1380	3340
2150	840	1980	5230	n/a	n/a	n/a
2200	1080	2970	8940	n/a	n/a	n/a

simulator calculates the costs. Actor and critic networks are updated at the end of an episode. The convergence of the separate A2C algorithm used in the experiments is shown in Fig. 5.

III. CASE STUDY

We here present a case study for Pinellas County, Florida, USA, using our multi-agent RL framework. Pinellas County, home to nearly one million residents, is the most visited destination on the US Gulf Coast. About 15 million tourists yearly spent over $20 Billion over the past five years [40]. The top two cities in the county, St. Petersburg and Clearwater, are ranked among the cities with a high risk of flooding [5].

A. Parameter Estimation

Hurricanes, floods, and stagnant water are some of the many SLR-related natural events that cause costs in different ways, such as loss of properties, jobs, taxes, and tourism incomes due to submerged areas. To model this cost z_t, we begin with modeling the SLR amount r_t. For the Tide Gauge #8726520 located in St. Petersburg, FL, we utilize the online sea level change calculator developed by the US Army Corps of Engineers (USACE) in [41], which uses the NOAA projections for Pinellas County [42]. Among the seven SLR projections, with relative sea level (RSL) zero for the year 2000, the Tampa Bay Climate Science Advisory Panel ruled out the very low, low, and extreme scenarios for planning purposes [43]. We also disregard the intermediate-high scenario and limit our simulations for intermediate-low, intermediate, and high SLR scenarios.

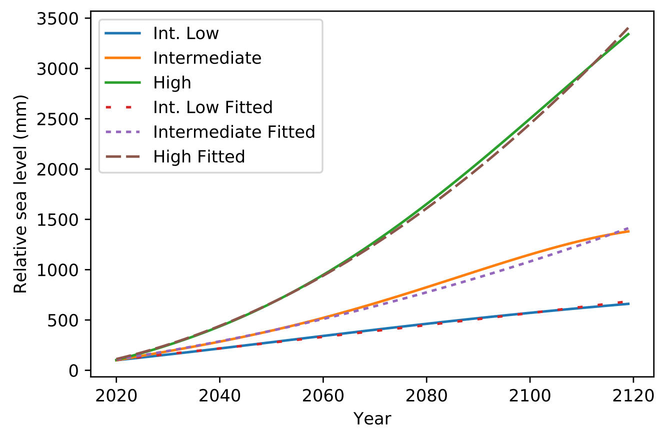

Fig. 6. SLR projections by NOAA [42] (solid lines) and our fittings (dashed lines) for relative sea level change for St. Petersburg, FL.

TABLE III
GENERALIZED PARETO PARAMETERS FOR PINELLAS COUNTY.

ξ	μ	η	p	q
-0.1	$ 30M	$ 100M	0.92	0.8

Our simulations target the hundred years 2020–2120, hence we adjust the initial sea level value for 2020 to $\hat{\ell}_0 = 100$mm for all scenarios. For the following years, we follow SLR projections in [42] till 2120. RSL for these scenarios from [41] and our adjusted values are shown in Table II. The randomness in SLR at each time step is modeled using the Gamma distribution, which is commonly used for modeling positive variables, including environmental applications, e.g., daily rainfall [44]. We use these adjusted projections $\{\hat{\ell}_t\}$ as the mean RSL values for the Gamma distribution, i.e., $r_t \sim \text{Gamma}(\alpha, \beta)$. Specifically, we set the scale parameter to $\beta = 0.5$ and vary the shape parameter α in a range to match the mean RSL, given by $\sum_{n=1}^{t} \mathsf{E}[r_n]$, with the adjusted NOAA projection curves. The successful curve fitting for mean RSL values shown in Fig. 6 is achieved by the following time series equations,

$$\begin{aligned} \text{Int-low:} \quad & \alpha_t = 11.102 + 0.012 \times t \\ \text{Intermediate:} \quad & \alpha_t = 15.8 + 0.211 \times t \\ \text{High:} \quad & \alpha_t = 24 + 0.85 \times t, \end{aligned} \tag{4}$$

where the subscript t represents calendar year $2020 + t$.

The generalized Pareto distribution used to model the cost from nature z_t, has three parameters, location, shape, and scale. The location parameter determines the range of z_t, in particular the lower limit. We set it as $30 million according to the data provided in Table A-1 in [33], which indicates that a year with no serious natural disaster might produce this cost, typically for maintenance. To get an upper limit on z_t, we need the shape parameter to be negative. We set it as -0.1 for the upper limit to be roughly ten times the expected cost. The scale parameter establishes the relation between sea level, infrastructure state, and disaster cost in our model. A recent report by the Tampa Bay Regional Planning Council [34] gives “the cost of doing nothing" due to SLR impacts under the NOAA’s high SLR projections for the Tampa Bay region, including Pinellas County. This report uses the widely accepted REMI PI+ economic modeling tool for their estimations. The following equations are obtained using the data provided in [34], where the cost unit is in million USD and subscripts represent the calendar year:

$$\mathsf{E}\left[z_{2060}\right] = \mu + \frac{\eta \ell_{2060}^{p}}{(1-\xi) s_{2020}^{q}} = 5057,$$

$$\mathsf{E}\left[\sum_{y=2020}^{2060} z_y\right] = 89000.$$

The report in [34] discusses the cost of doing nothing; hence we keep the infrastructure state $s_{2020} = 20$ constant between the years 2020 and 2060 in the above equations. We further set the relative sea level $\ell_{2020} = 100$, and following the high SLR scenario we obtain $\eta = 100, p = 0.92, q = 0.8$ as a set of values suitable for our simulations. Table III summarizes the values of the generalized Pareto parameters for the Pinellas County case study.

To determine the distribution of cost from nature, we use the economic data [40] and the cost projections [34] for tourism in Pinellas County. The tourism industry contributed \$9.25 billion annual spending impact to the Pinellas County local economy [40] in 2019. We consider only the tourism business in our model as they are the main business stakeholder of SLR impacts. The “cost of doing nothing" report gives the tourism loss in 2060 as \$898 million [34]. The principal cost for the business is the loss of net profit, which is estimated as the 5% of total tourism income. With the current sea level, we estimate the upper bound of business loss for the year 2019 as 10% of the net profit. This loss grows with 3% yearly growth for tourism business, in line with US GDP growth. With the high SLR scenario, business damage loss will increase to 100% of net profit in 2060, up from 10% in 2019, which is equivalent to saying that the tourism sector will lose all profit if no infrastructure is developed in the next 40 years. This cost model for tourism business in Pinellas County corresponds to the 22% of total cost from nature, i.e., $m_B = 0.22$ in (1). Together with the insurance cost explained below, it also gives similar costs in our simulations to the cost-of-doing-nothing estimation in [34].

Since the government is the major stakeholder with infrastructures, including buildings, roads, parks, etc. under its liability, we set the government’s portion within the cost from nature as 75%, i.e., $m_G = 0.75$ in (1). The majority of residents in coastal regions, in particular Pinellas County, have flood insurance, as explained next, hence most of the property inundation cost, which is estimated to be more than \$16 billion in the worst case scenario [34], is covered by the government. The direct cost to residents from nature is set as 3% of total cost from nature, i.e., $m_R = 0.03$, to account for the uninsured and uninsurable properties.

TABLE IV
ACTION AND COST PARAMETERS FOR PINELLAS COUNTY.

Agent	Action	Action multiplier	Portion of natural cost	Insurance factor	Insurance memory
Govt.	0,1,2,3,4	\$140M	$m_G = 0.75$	n/a	n/a
Resident	0,1,2,3,4	\$20M	$m_R = 0.03$	$I_R = 0.04$	$\rho_R = 0.9$
Business	0,1,2,3,4	\$50M	$m_B = 0.22$	$I_B = 0.006$	$\rho_B = 0.9$

For the insurance cost, we use a topology-based data set provided in [11] for exposed assets by ground heights for all the Gulf Coast. Pinellas County falls under a high-risk flood zone, with many of its 407,720 residential properties considered as exposed assets by ground heights [45]. The homeowners typically have the National Flood Insurance Program (NFIP) provided by the government. The residents of St. Petersburg paid an average insurance premium of \$950 and around \$33 million in total annually [46], which is the highest in Florida. Scaling this total insurance premium payment in St. Petersburg to the entire Pinellas County according to the almost 1/4 ratio of households [47], we set the base insurance premium by residents as $I_{R,0} = \$132$ million. The insurance cost and action parameters for each agent are given in Table IV. Although most of the cost from nature is covered by the insurance, increasing costs due to SLR is reflected to the residents through a higher premium rate in the future. We empirically determined the insurance factor as $I_R = 0.04$ and the insurance memory factor as $\rho_R = 0.9$ to match the insurance data stated above. Similarly, the structural properties of businesses are mostly covered by insurance, and the premiums increase with accumulating cost from nature, $\rho_B = 0.9$. Since the commercial land use and the number of commercial insurance policies are less than the residential ones [34], we set the initial insurance premium for business as $I_{B,0} = \$20$ million and the insurance coefficient as $I_B = 0.006$.

We assume the infrastructure improvement is proportionate to the total investment amount. Infrastructure development may include activities like beach and wetland restoration, home elevation, dykes, local levees, sandbags, etc. We estimate the cost of these actions to range between a couple of millions to billions of USD based on [11, supplementary]. The investment ranges for agents are determined such that the maximum continuous investment from all agents in 40 years prevents any cost from nature until 2060, e.g., 10-feet home elevation or 20-feet dyke all over the coastline. The investment cost parameters are given in Table IV along with the insurance and natural cost parameters.

B. Scenario Simulations

To benchmark the performance of the “proactive" deep RL framework, we also consider a straightforward "reactive" policy that makes an infrastructure improvement only after its need is proven by high natural cost. In a reactive community, the government and other stakeholders generally become active in infrastructure development after a major natural disaster. This trend can be portrayed through a threshold-based policy, where an agent invests in infrastructure if the cost from nature exceeds a predefined threshold. Although

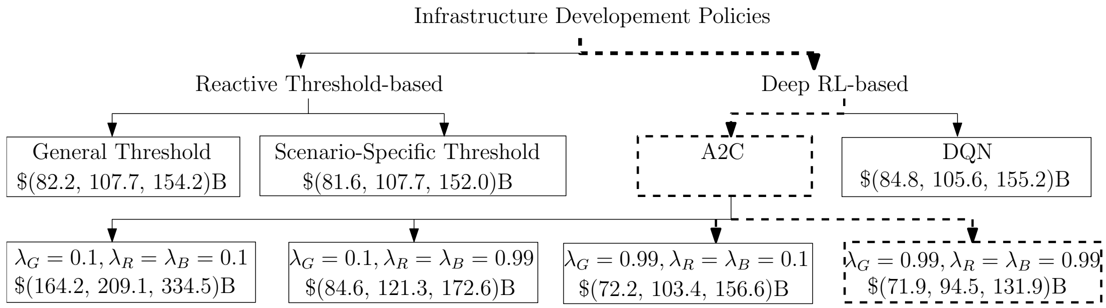

Fig. 7. Expected episodic cost under different policies for high SLR scenario.The values within the parenthesis indicate the cost for Intermediate Low, Intermediate, and High SLR, respectively.

simple, this policy is not tractable for generating simulations that represent realistic stakeholder behaviors because of the difficulty in selecting the thresholds for natural cost. Whereas, the cooperation indices in our simulation tool can be intuitively varied between zero and one to jointly simulate the adaptation strategies of different stakeholder prototypes.

Fig. 7 presents the total cost for all stakeholders over the 100-year period 2020-2120 when the threshold-based and deep RL policies are deployed. The three values in parentheses are the total cost over 100 years in billion US dollars considering the intermediate-low, intermediate, and high SLR projections of NOAA, respectively. The threshold-based policy is used in two forms: the general threshold policy represents the case where the agents are agnostic to SLR projection scenario in the simulations, and the scenario-specific threshold policy corresponds to the case where the best threshold is used for each SLR scenario. In both threshold-based policies, all three agents take the maximum investment action shown in Table IV once the cost from nature in a year exceeds the same threshold. This common threshold is optimized for each scenario in the scenario-specific policy, and for the average of three scenarios in the general threshold-based policy to demonstrate the best performance such threshold-based policies can attain (Fig. 8). As shown in Fig. 7, the proposed deep RL policy based on the A2C algorithm can intuitively simulate a variety of stakeholder prototypes by varying their cooperation indices between zero and one. While the fully non-cooperative case with all three cooperation indices equal to 0.1 results in huge costs, double the costs of the best threshold-based policy, the fully-cooperative case with cooperation indices equal to 0.99 reduces the total cost by 13% with respect to the best threshold-based policy. The costs presented for the DQN-based policy are for the fully cooperative case ($\lambda_G = \lambda_R = \lambda_B = 0.99$). They are significantly worse than their counterparts in the A2C policy due to the high variance in the cost from nature. Finally, Fig. 9 shows the cumulative yearly cost for the high SLR scenario for each policy.

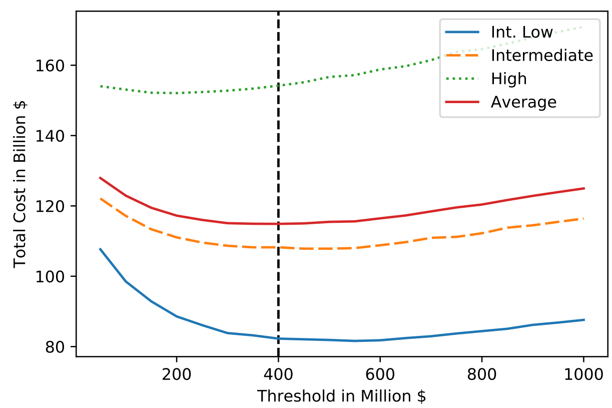

Fig. 8. 100-year total cost for the intermediate-low, intermediate, high scenarios of SLR, and their average. The vertical dashed line represents the best general threshold for the average SLR scenario.

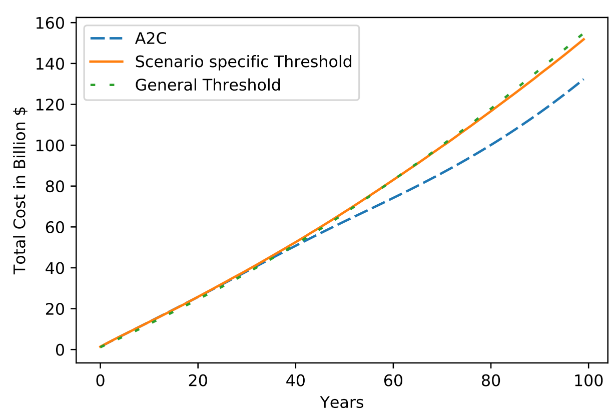

Fig. 9. Yearly total cost under different policies for the high SLR scenario.

IV. Discussions and Conclusion

In this work, we presented how a socioeconomic system around the sea-level rise (SLR) problem can be modeled as a Markov Decision Process (MDP) and simulated using Deep Reinforcement Learning (RL) algorithms. In addition to providing a general scenario planning tool to investigate the cost-benefit analysis of natural events and stakeholders' investments, the proposed framework also illustrates, through a case study for the Tampa Bay region based on real data, how optimizing the adaptation strategies can effectively minimize the total cost due to SLR. Being the first in the literature, the proposed MDP model relies on some simplifying assumptions.

For example, we assumed a uniform cost-benefit economic model for the adaptation actions to represent the natural disaster cost with a tractable model with respect to the taken adaptation actions so far (i.e., setting the scale parameter of the generalized Pareto distribution as a function of the sea level and infrastructure state). In the uniform model, the action level (Table IV) also determines the development level in the infrastructure. A detailed cost-benefit model for different actions can easily replace the considered uniform model. Specifically, for a set of actions such as beach restoration, raising roads, building seawalls, and relocating coastal properties, the cost levels and development levels can be non-uniformly set after a detailed study. Another feasible improvement over the proposed framework would be to represent residents and businesses with multiple independent agents and consider local infrastructure improvement actions for each subregion defined by a resident/business agent. Such an extension will increase the number of agents and the number of actions to decide for the government while the structure of the overall model remains the same. Note that the proposed multi-agent MDP model is not restricted to the RL policies; any action policy can be followed by the agents. The sequential and agent-based structure allows for a turn-taking game mode, where each agent decides on its action sequentially in a round, and at the end of the round nature imposes its cost on the agents. We developed a board game in which players can cooperate on adaptation strategies to mitigate SLR-related damages from nature [48]. A digital and improved version of the game is planned as a future work.

Appendix

In Theorem 1, we analyze the government's policy since it is the most dominant agent with the full observation of other agents' actions. In the first part of the proof, we will show that if $V(\hat{s_t},\ell_t)$=$V_G(s_t,\ell_t,O_t^G)$ is nondecreasing and concave in ℓ_t, then so is

$$F_m(\hat{s_t},\ell_t) = \mathsf{E}\left[m\alpha_G + z_{G,t} - f(I_{R,t}+I_{B,t}) + \lambda_g V(\hat{s_t}+m,\ell_t+r_t)\right],$$

for $m = 0,1,\ldots,A_G$. Government observes the residents' and businesses' decisions beforehand, thus R_t and B_t are considered constant for its decision making. We denote the next infrastructure state that includes the residents' and business' current action with $\hat{s}_t = s_t + R_t \times \alpha_R/\alpha_G + B_t \times \alpha_B/\alpha_G$. Assuming $V(\hat{s_t},\ell_t)$ is nondecreasing, i.e., $\frac{\partial}{\partial \ell_t}V(\hat{s_t},\ell_t) \geq 0$, and using the expected value of generalized Pareto-distributed $z_{G,t}$, we can write

$$\frac{\partial}{\partial \ell_t}F_m(\hat{s_t},\ell_t) = \frac{m_G\eta p \ell_t^{p-1}}{(1-\xi)s_t^q} + \lambda_G \mathsf{E}\left[\frac{\partial}{\partial \ell_t}V(\hat{s_t}+m,\ell_{t+1})\right] \geq 0.$$

Note that $I_{R,t}$ and $I_{B,t}$ are determined by past data, independent of ℓ_t. In the above equation, the derivative can be brought inside the integral due to the monotone convergence theorem. Assuming $V(\hat{s_t},\ell_t)$ is concave, i.e., $\frac{\partial^2}{\partial \ell_t^2}V(\hat{s_t},\ell_t) < 0$, for the second derivative we have

$$\frac{\partial^2}{\partial \ell_t^2}F_m(\hat{s_t},\ell_t) = \frac{m_G\eta p(p-1)\ell_t^{p-2}}{(1-\xi)s_t^q} + \lambda_G \mathsf{E}\left[\frac{\partial^2}{\partial \ell_t^2}V(\hat{s_t}+m,\ell_{t+1})\right] < 0$$

since $0 < p < 1$. Hence, it is sufficient to show that $V(\hat{s_t},\ell_t)$ is nondecreasing and concave.

Finding the value function iteratively (i.e., value iteration) is a common approach which is known to converge [28]:

$$\lim_{i\to\infty} V_i(\hat{s},\ell) = V(\hat{s},\ell),$$

where, for brevity, we drop the time index from now on. We will next prove that $V(\hat{s},\ell)$ is nondecreasing and concave iteratively. Initializing all the state values as zero, i.e., $V_0(\hat{s},\ell)=0, \forall s,\ell$, after the first iteration we get

$$\begin{aligned} V_1(\hat{s},\ell) &= \min_G \Big\{ \mathsf{E}[\alpha_G G + z_G(s,\ell) + \lambda_G V_0(\hat{s}+G,\ell+r) + \text{const.}] \Big\} \\ &= \text{const.} + \mathsf{E}[z_G(s,\ell)] = \text{const.} + \mu + \frac{m_G\eta\ell^p}{(1-\xi)s^q}. \end{aligned}$$

Differentiating with respect to ℓ, we get

$$\frac{\partial}{\partial \ell}V_1(\hat{s},\ell) = m_G\eta p\frac{\ell^{p-1}}{(1-\xi)s^q} \geq 0, \ \forall s, \tag{5}$$

$$\frac{\partial^2}{\partial \ell^2}V_1(\hat{s},\ell) = m_G\eta p(p-1)\frac{\ell^{p-2}}{(1-\xi)s^q} < 0, \ \forall s,$$

since $m_G,\eta > 0, p \in (0,1), q > 0, \xi < 0$. Thus, $V_1(\hat{s},\ell)$ is nondecreasing and concave in ℓ for all s. Similarly, the value function after the second iteration becomes

$$\begin{aligned} V_2(\hat{s},\ell) &= \min_G \Big\{ \mathsf{E}[\alpha_G G + z_G(s,\ell) + \lambda_G V_1(\hat{s}+G,\ell+r) + \text{const.}] \Big\} \\ &= \min_G \Big\{ \alpha_G G + \mu + \frac{m_G\eta\ell^p}{(1-\xi)s^q} + \mu\lambda_G + \lambda_G \mathsf{E}\left[\frac{m_G\eta(\ell+r)^p}{(1-\xi)(\hat{s}+G)^q}\right] + \text{const.}\Big\}. \end{aligned}$$

Denoting the optimum action with $\hat{G}$ we will show that $V_2(\hat{s},\ell)$ is nondecreasing and concave for any $\hat{G}$. Moreover, the pointwise minimum of nondecreasing and concave functions is also nondecreasing and concave. Taking the derivative with respect to ℓ we get

$$\begin{aligned} \frac{\partial}{\partial \ell}V_2(\hat{s},\ell) &= \frac{\partial}{\partial \ell}\left\{\frac{m_G\eta\ell^p}{(1-\xi)s^q} + \lambda_G\frac{m_G\eta\mathsf{E}[(\ell+r)^p]}{(1-\xi)(\hat{s}+\hat{G})^q}\right\} \\ &= m_G\eta p\frac{\ell^{p-1}}{(1-\xi)s^q} + \lambda_G m_G\eta p\frac{\mathsf{E}[(\ell+r)^{p-1}]}{(1-\xi)(\hat{s}+\hat{G})^q} \geq 0, \quad \forall s \\ \frac{\partial^2}{\partial \ell^2}V_2(\hat{s},\ell) &= m_G\eta p(p-1)\frac{\ell^{p-2}}{(1-\xi)s^q} \\ &\quad + \lambda_G m_G\eta p(p-1)\frac{\mathsf{E}[(\ell+r)^{p-2}]}{(1-\xi)(\hat{s}+\hat{G})^q} < 0, \quad \forall s. \end{aligned}$$

Hence, $V_2(\hat{s},\ell)$ is nondecreasing and concave. Now, for any i, given that $V_{i-1}(\hat{s},\ell)$ is nondecreasing and concave, we can write

$$\frac{\partial}{\partial \ell}V_i(\hat{s},\ell) = m_G\eta a\frac{\ell^{p-1}}{(1-\xi)s^q} + \lambda_G\mathsf{E}\left[\frac{\partial}{\partial \ell}V_{i-1}(\hat{s}+\hat{G},\ell)\right] \geq 0, \forall s$$

$$\frac{\partial^2}{\partial \ell^2}V_i(\hat{s},\ell) = m_G\eta p(p-1)\frac{\ell^{p-2}}{(1-\xi)s^q} + \lambda_G\mathsf{E}\left[\frac{\partial^2}{\partial \ell^2}V_{i-1}(\hat{s}+\hat{G},\ell)\right] < 0, \quad \forall s. \quad (6)$$

Consequently, by mathematical induction, $V(\hat{s},\ell)$ is nondecreasing and concave.

The second part of the proof is to show that $\frac{\partial}{\partial \ell}F_m(\hat{s},\ell) < \frac{\partial}{\partial \ell}F_{m-1}(\hat{s},\ell)$. Similar to the first part, if we show that

$$\frac{\partial}{\partial \ell}V(\hat{s}+m,\ell) < \frac{\partial}{\partial \ell}V(\hat{s}+m-1,\ell),$$

we can conclude that $\frac{\partial}{\partial \ell}F_m(\hat{s},\ell) < \frac{\partial}{\partial \ell}F_{m-1}(\hat{s},\ell)$ since

$$\frac{\partial}{\partial \ell_t}F_m(\hat{s_t},\ell_t) = \frac{m_G\eta p\ell_t^{p-1}}{(1-\xi)s_t^q} + \lambda_G\mathsf{E}\left[\frac{\partial}{\partial \ell_t}V(\hat{s_t}+m,\ell_t+r_t)\right]$$

$$\frac{\partial}{\partial \ell_t}F_{m-1}(\hat{s_t},\ell_t) = \frac{m_G\eta p\ell_t^{p-1}}{(1-\xi)s_t^q} + \lambda_G\mathsf{E}\left[\frac{\partial}{\partial \ell_t}V(\hat{s_t}+m-1,\ell_t+r_t)\right].$$

Starting again with $V_0(\hat{s},\ell)=0,\forall s,\ell$, from (5) we can write the following inequality for the first iteration

$$\frac{\partial}{\partial \ell}V_1(\hat{s}+m,\ell) = m_G\eta p\frac{\ell^{p-1}}{(1-\xi)(\hat{s}+m)^q} < \frac{\partial}{\partial \ell}V_1(\hat{s}+m-1,\ell) = m_G\eta p\frac{\ell^{p-1}}{(1-\xi)(\hat{s}+m-1)^q}.$$

For any i, given that $\frac{\partial}{\partial \ell}V_{i-1}(\hat{s}+m,\ell) < \frac{\partial}{\partial \ell}V_{i-1}(\hat{s}+m-1,\ell)$, from (6) we have

$$m_G\eta p\frac{\ell^{p-1}}{(1-\xi)(\hat{s}+m)^q} + \lambda_G\mathsf{E}\left[\frac{\partial}{\partial \ell}V_{i-1}(\hat{s}+m+,\ell)]\right] < m_G\eta p\frac{\ell^{p-1}}{(1-\xi)(\hat{s}+m-1)^q} + \lambda_G\mathsf{E}\left[\frac{\partial}{\partial \ell}V_{i-1}(\hat{s}+m-1+,\ell)\right], \text{i.e.}$$

$$\frac{\partial}{\partial \ell}V_i(\hat{s}+m,\ell) < \frac{\partial}{\partial \ell}V_i(\hat{s}+m-1,\ell)$$

As a result, by mathematical induction we conclude that $\frac{\partial}{\partial \ell}V(\hat{s}+m,\ell) < \frac{\partial}{\partial \ell}V\hat{s}+m-1,\ell)$.

Acknowledgement

The authors would like to acknowledge the valuable discussions and feedback from several policymakers in the Tampa Bay area, such as the Douglas Hutchens (Deputy City Manager, City of Dunedin), Melissa Zornitta (Executive Director, Hillsborough County Planning Commission), Vik Bhide (Director, Transportation and Stormwater Services at City of Tampa), and Alison Barlow (Executive Director, St. Petersburg Innovation District).

References

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

[2] T. Wahl, S. Jain, J. Bender, S. D. Meyers, and M. E. Luther, "Increasing risk of compound flooding from storm surge and rainfall for major US cities," *Nature Climate Change*, vol. 5, no. 12, p. 1093, 2015.

[3] S. B. Kashem, B. Wilson, and S. Van Zandt, "Planning for climate adaptation: Evaluating the changing patterns of social vulnerability and adaptation challenges in three coastal cities," *Journal of Planning Education and Research*, vol. 36, no. 3, pp. 304–318, 2016.

[4] G. Schrock, E. M. Bassett, and J. Green, "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*, vol. 35, no. 3, pp. 282–295, 2015.

[5] S. Hallegatte, C. Green, R. J. Nicholls, and J. Corfee-Morlot, "Future flood losses in major coastal cities," *Nature climate change*, vol. 3, no. 9, pp. 802–806, 2013.

[6] M. E. Hauer, J. M. Evans, and D. R. Mishra, "Millions projected to be at risk from sea-level rise in the continental united states," *Nature Climate Change*, vol. 6, no. 7, pp. 691–695, 2016.

[7] M. Ghanbari, M. Arabi, J. Obeysekera, and W. Sweet, "A coherent statistical model for coastal flood frequency analysis under nonstationary sea level conditions," *Earth's Future*, vol. 7, no. 2, pp. 162–177, 2019.

[8] S. Hallegatte, N. Ranger, O. Mestre, P. Dumas, J. Corfee-Morlot, C. Herweijer, and R. M. Wood, "Assessing climate change impacts, sea level rise and storm surge risk in port cities: a case study on copenhagen," *Climatic change*, vol. 104, no. 1, pp. 113–137, 2011.

[9] M. Ghanbari, M. Arabi, and J. Obeysekera, "Chronic and acute coastal flood risks to assets and communities in southeast florida," *Journal of Water Resources Planning and Management*, vol. 146, no. 7, p. 04020049, 2020.

[10] C. Scawthorn, N. Blais, H. Seligson, E. Tate, E. Mifflin, W. Thomas, J. Murphy, and C. Jones, "Hazus-mh flood loss estimation methodology. i: Overview and flood hazard characterization," *Natural Hazards Review*, vol. 7, no. 2, pp. 60–71, 2006.

[11] B. G. Reguero, M. W. Beck, D. N. Bresch, J. Calil, and I. Meliane, "Comparing the cost effectiveness of nature-based and coastal adaptation: A case study from the gulf coast of the united states," *PloS one*, vol. 13, no. 4, p. e0192132, 2018.

[12] P. Peduzzi, B. Chatenoux, H. Dao, A. De Bono, C. Herold, J. Kossin, F. Mouton, and O. Nordbeck, "Global trends in tropical cyclone risk," *Nature climate change*, vol. 2, no. 4, pp. 289–294, 2012.

[13] S. Hsiang, R. Kopp, A. Jina, J. Rising, M. Delgado, S. Mohan, D. Rasmussen, R. Muir-Wood, P. Wilson, M. Oppenheimer *et al.*, "Estimating economic damage from climate change in the united states," *Science*, vol. 356, no. 6345, pp. 1362–1369, 2017.

[14] "FEMA Opens $660 Million Grant Application Process," https://www.fema.gov/press-release/20200929/fema-opens-660-million-grant-application-process, September 29, 2020.

[15] E. R. Group *et al.*, *What Will Adaptation Cost?: An Economic Framework for Coastal Community Infrastructure*. National Oceanic and Atmospheric Administration Coastal Services Center, 2013.

[16] A. C. Hine, D. P. Chambers, T. D. Clayton, M. R. Hafen, and G. T. Mitchum, *Sea Level Rise in Florida: Science, Impacts, and Options*. University Press of Florida, 2016.

[17] M. K. Buchanan, R. E. Kopp, M. Oppenheimer, and C. Tebaldi, "Allowances for evolving coastal flood risk under uncertain local sea-level rise," *Climatic Change*, vol. 137, no. 3-4, pp. 347–362, 2016.

[18] X. Fu, M. Gomaa, Y. Deng, and Z.-R. Peng, "Adaptation planning for sea level rise: A study of US coastal cities," *Journal of Environmental Planning and Management*, vol. 60, no. 2, pp. 249–265, 2017.

[19] N. Research Council, "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.

[20] J. Tribbia and S. C. Moser, "More than information: what coastal managers need to plan for climate change," *Environmental science & policy*, vol. 11, no. 4, pp. 315–328, 2008.

[21] P. R. Berke and M. R. Stevens, "Land use planning for climate adaptation: Theory and practice," *Journal of Planning Education and Research*, vol. 36, no. 3, pp. 283–289, 2016.

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

[23] R. Potts, L. Jacka, and L. H. Yee, “Can we ‘Catch ‘em All’? An exploration of the nexus between augmented reality games, urban planning and urban design,” *Journal of Urban Design*, vol. 22, no. 6, pp. 866–880, 2017.

[24] S. S. Shuvo, Y. Yilmaz, A. Bush, and M. Hafen, “A markov decision process model for socio-economic systems impacted by climate change,” in *International Conference on Machine Learning*. PMLR, 2020.

[25] R. A. Howard, “Dynamic programming and markov processes.” 1960.

[26] R. Bellman, “A Markovian decision process,” *Journal of mathematics and mechanics*, pp. 679–684, 1957.

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

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

[29] A. C. Cebrián, M. Denuit, and P. Lambert, “Generalized pareto fit to the society of actuaries? large claims database,” *North American Actuarial Journal*, vol. 7, no. 3, pp. 18–36, 2003.

[30] Y. Uchiyama and S. Watanabe, “Modeling Natural Catastrophic Risk and its Application,” 2006.

[31] A. Daspit and K. Das, “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*, 2012, pp. 2395–2403.

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

[33] “PINELLAS COUNTY COASTAL MANAGEMENT PROGRAM SUMMARY PLANNING DOCUMENT,” http://www.pinellascounty.org/environment/coastalMngmt/pdfs/Pinellas_County_CMP_Sum_Document.pdf, 2018.

[34] “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.

[35] P. Sergent, G. Prevot, G. Mattarolo, J. Brossard, G. Morel, F. Mar, M. Benoit, F. Ropert, X. Kergadallan, J.-J. Trichet *et al.*, “Adaptation of coastal structures to mean sea level rise,” *La Houille Blanche*, no. 6, pp. 54–61, 2014.

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

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

[38] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, no. 7540, p. 529, 2015.

[39] P.-H. Su, P. Budzianowski, S. Ultes, M. Gasic, and S. Young, “Sample-efficient actor-critic reinforcement learning with supervised data for dialogue management,” *arXiv preprint arXiv:1707.00130*, 2017.

[40] “St. petersburg tourism,” https://www.stpete.org/economic_development/stpete_advantage/tourism.php#:~:text=Petersburg\%20offers\%2077\%20hotel\%2Fmotel,billion\%20to\%20the\%20local\%20economy., 2021.

[41] “Sea-Level Change Curve Calculator (Version 2021.12),” https://cwbi-app.sec.usace.army.mil/rccslc/slcc_calc.html, 2017.

[42] W. W. V. Sweet, R. Kopp, C. P. Weaver, J. Obeysekera, R. M. Horton, E. R. Thieler, C. E. Zervas *et al.*, “Global and regional sea level rise scenarios for the united states,” 2017.

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

[44] H. Aksoy, “Use of gamma distribution in hydrological analysis,” *Turkish Journal of Engineering and Environmental Sciences*, vol. 24, no. 6, pp. 419–428, 2000.

[45] “QuickFacts of Pinellas County, Florida,” https://www.census.gov/quickfacts/fact/table/pinellascountyflorida/RHI225219, 2020.

[46] “Cost of Flood Insurance in Florida and How Coverage Works,” https://www.valuepenguin.com/flood-insurance/florida, 2020.

[47] “QuickFacts of Pinellas County, Florida,” https://www.census.gov/quickfacts/fact/table/stpetersburgcityflorida/POP060210, 2020.

[48] Y. Yilmaz, A. Bush, and M. Hafen, “Atlantis: A Game of Sea Level Rise,” https://sites.google.com/site/nsfsccsealevelrise/atlantis-a-game-of-sea-level-rise.

Salman Sadiq Shuvo received his M.Sc. degree in Electrical Engineering from the University of South Florida and is currently a Ph.D. student at the same university. His research interests are in reinforcement learning based optimization methods that he has employed in several socioeconomic and energy systems.

Yasin Yilmaz (S’11-M’14-SM’20) received the Ph.D. degree in Electrical Engineering from Columbia University, New York, NY, in 2014. He is currently an Assistant Professor of Electrical Engineering at the University of South Florida, Tampa. His research interests include statistical signal processing, machine learning, and their applications to computer vision, cybersecurity, socioeconomic systems, energy systems, transportation systems, and communications systems.

Alan Bush is a faculty member in the Honors College who teaches courses in Geographies of Transformation, Cities of the Global South, Tampa in Transformation, Wicked Problems in Social Change, and Sustainability Practice. He holds a B.A. from Pomona College in Global Politics Economics, a M.S. in Sustainable Systems from the University of Michigan, and a Ph.D. from the University of Texas School of Architecture in Community and Regional Planning. His current research focuses on leadership as an emergent property of relationships, and the leadership practices that foster resilience for organizations and large-scale urban systems. Dr. Bush’s research interests include complex systems, urban agriculture, ensemble leadership, resilience, innovation in business model and organizational form, affordable housing, and bio-adaptive urban infrastructure.

Mark Hafen is a retired Master Instructor from the School of Public Affairs at the University of South Florida in Tampa, and former Director for the Master of Urban and Regional Planning (MURP) program, with an emphasis on climate change adaptation, and urban environmental policy and planning. He holds a B.S. in Business Logistics from Penn State University, as well as an M.A. in Geography and a Ph.D. in Marine Science (Geology), both from USF. He has professional experience in land use planning, and has lived in the Tampa Bay region since 1986. He has co-authored a book (with A.C. Hine, D.P. Chambers, T.D. Clayton, and G.T. Mitchum), Sea Level Rise in Florida: Science, Impacts and Options (2016, University Press of Florida), and actively serves as a member of the Tampa Bay Climate Science Advisory Panel, the Florida Water and Climate Alliance, and the USF Center for Brownfields Research and Redevelopment.