

SimCoast: 大语言模型驱动的人类-海岸系统动态多代理模拟与可持续性评估

Yichen

July 9, 2025

Abstract

沿海地区由于快速城市化和资源开发, 海岸线生态面临着严重的压力。本文提出了一个基于大语言模型 (LLM) 的多智能体模拟系统——SimCoast。SimCoast 通过引入人类、生态系统和裁判 LLM 角色, 实现了对人类决策、生态响应及随机外部事件的动态交互模拟。实验结果表明, 人类社会发展与生态保护之间存在长期博弈关系, 短期激进发展往往以牺牲生态为代价, 而平衡发展需要更持久的调适过程。SimCoast 为沿海可持续治理初步提供了一种灵活高效的模拟平台, 并验证了 LLM 在复杂社会—生态系统建模中的潜力。模型的代码可见 <https://github.com/YichenWei0601/LLM-Adversarial-Modeling-of-Shoreline-Ecology>。

1 引言

沿海地区由于大城市众多, 经济较为发达, 很容易使得沿海区域的环境生态收到干扰。长期以来, 这里一直承受着人类开发活动与环境退化带来的双重压力 [9, 13]。随着城市化进程的加快、资源的过度开发以及气候变化的影响, 沿海生态系统面临的风险日益加剧, 生态韧性和可持续发展受到严重威胁。近年来, 尽管沿海可持续发展、社会—生态耦合、生态韧性等理念逐渐被提出, 并在不同尺度上得到应用 [2], 但目前仍然缺乏一种灵活、高效的工具, 能够模拟人类活动与生态系统之间的长期耦合动态, 难以为决策和治理创新提供有效支持。

大语言模型 (Large-Language Model, LLM) 因为其独特的语言理解能力, 在模拟复杂的系统以及涉及到长期博弈的领域能够表现出独特优势 [4]。本文通过构建一个 LLM 驱动的人类-海岸系统动态代理模拟系统, 引入多个 LLM 进行对抗发展。这样既能量化人类决策对生态韧性的影响, 又能通过反馈机制引导更合理的行为选择, 为沿海地区的可持续治理提供有力支持。

本文的贡献在于提出了一个基于 LLM 的沿海社会—生态耦合的动态模拟系统。这个框架能够：

- 量化人类决策对沿海生态韧性的长期影响；
- 根据系统状态的分数指标进行动态反馈，有助于实时的检测和调整；
- 为探索沿海城市的可持续治理路径提供实践参考和决策启示。

2 相关工作

海岸线管理与人类-自然交互模型的设计 在海岸线管理领域,代理建模 (Agent-Based Modeling, ABM) 被广泛用于刻画人类活动与自然过程之间的复杂互动关系。McDonald et al. [8] 借助 ABM 方法评估了生态保护方面的策略以及城市发展的政策对于对海洋生态系统的影响。Lawyer et al. [7] 提出了系统化建模,从而模拟沿海社区适应各种环境因素的不同策略。然而,这些模型并不能够根据实际的复杂情况进行跨领域的综合分析,从而选择更加全面的解决方法。

LLM 与世界模型模拟。 LLM 作为日益发展的大模型,在对于世界变化的潜在趋势的理解方面和对于多重信息的交叉与相互作用方面有着很强大的能力。Gao et al. [4] 对基于 LLM 的多智能体系统进行了调研,并且探讨了其在应对复杂问题与构建世界模拟方面的最新进展及面临的挑战。具体上,现在以及有研究在不同的领域成功实现了一定程度上的 LLM 模拟人类行为的成果。Joon et al. [11] 点名了 LLM 对于模拟人类行为有合理意义,通过扩展 LLM 的记忆来让它提供更有建议,并通过模拟世界加以验证。Wang et al. [17] 结合个人的活动模式和动机,成果模拟了在城郊的个人移动能力。由此可见,LLM 对于世界模型的模拟有很大的潜力。

LLM 间的对抗性交互。 LLM 之间的对抗性交互是实现最为贴近事实的模拟的关键。LLM 拥有着强大的语言理解和生成能力,因此可以为智能体之间的竞争与协作提供新的实现途径。Guo et al. [5] 梳理了 LLM 相关的多智能体,展示了其在博弈论场景中的潜力和实际应用。Pei et al. [12] 借助多智能体模拟教学,实现了对于教学质量的评估分析工作的完成。但是这些研究没有在本课题的领域有过实操。

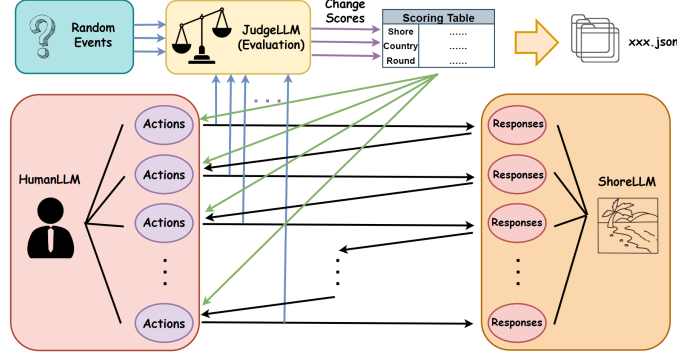


Figure 1: SimCoast 的模拟流程

3 SimCoast 模型介绍

因此，我们在此基础上提出了 SimCoast 模型。SimCoast 是一个创新的 LLM 驱动的多代理仿真系统，目的是模拟人类发展与海岸生态系统之间的复杂动态关系。这一模型通过利用多个 LLM 的交互和对抗来模拟人类决策与自然、社会事件的实时互动，同时利用 LLM 预训练积累的知识库，尽量真实地反映和推断现实世界中两者之间的平衡关系。SimCoast 设计的核心理念包括对抗性交互、智能决策、动态反馈、可持续性评估和人类决策输入，从而能够为研究人员和政策制定者提供强大的分析测试和模拟的平台。具体的流程参见 Figure 1。

3.1 状态管理分数系统

为了方便针对模型中的状态进行量化分析，我们引入了两个核心变量分数来表征状态：国家发展分数（country_score）和海岸线状态分数（shoreline_score）。两个分时分别在 0-100 的范围内衡量社会经济发展程度和生态健康状态。这两个分数会根据 1. 国家制定的策略 2. 随机事件（详细阐述在 Appendix C 中展示）的不同而产生不同程度的变化，而具体的分数变化量是经过 LLM 评估得到的。

3.2 多代理架构

SimCoast 的 LLM 多代理机制本质上经由三个代表着不同角色的 LLM 合作协同完成。

人类代理 (HumanLLM)。 HumanLLM 的工作是模拟沿海发展中国家的领导人。他需要基于国家发展分数、海岸线状态分数以及当前的生态环境的机遇与挑战制定策略, 依照自己的知识库和网络搜索, 选择合理的政策进行执行, 每年输出两个具体行动来代表“人类活动”。

海岸系统代理 (ShoreLLM)。 ShoreLLM 的工作是模拟海岸线生态系统。他会根据人类的政策行动, 生成海岸线生态系统可能会发生的变化, 更新生态环境带来的机遇和挑战, 体现环境的动态响应。例如人类政策选择“大力发展渔业”, ShoreLLM 返回的生态挑战可能是“海洋生物资源逐渐枯竭”、“海洋生态链产生破裂”等等。

裁判代理 (JudgeLLM) JudgeLLM 的工作是作为第三方的客观评估者。他会根据参考评分表量化人类行动对国家发展和海岸生态的影响 (其中, 构建参考评分表格的实现细节可见 Appendix A)。针对随机事件, JudgeLLM 则引入外部因素, 如自然灾害或积极事件, 在给定的可取区间内智能评估其对系统的综合影响。这些代理通过年度循环交互, 驱动系统的演化过程, 并不断自我迭代学习相关经验。

我们提供了一些现有的海岸线可持续性指标体系, 作为知识库的补充。这其中包括 NOAA 的官方指标体系报告 [10], 以及 CLSER 标准 [22], 从而使 LLM 能够做出更加准确的判断和分数赋值。

3.3 交互机制

SimCoast 的交互机制通过年度循环实现动态模拟。每年, HumanLLM 首先生成两个行动。JudgeLLM 随后评估其影响, 以及随机事件所带来的效应, 更新国家发展和海岸线分数。之后, ShoreLLM 响应新状态, 生成新的海岸生态机遇和挑战。最终的完整流程会以 JSON 格式存储, 方便进一步数据分析和处理。

同时, 人类专家为了测试某种可能的政策或行为对于海岸线长久的影响, 也可以加入来替代 HumanLLM 的生成, 从而使得回答进入交互系统进行进一步的决策交互中。

4 实验

在本文的实验阶段, 我们采用了 DeepSeek-R1, DeepSeek-V3 模型进行模拟 (各占一半模拟次数), 实验在 Python 3.10 上实现。具体的模型设定参数可见 Appendix B。

4.1 概况

我们一共运行了 20 次模拟测试，其中测试结果见 Table 1。

模拟结果	次数
国家分数达到 100 分	5
海岸线分数低于了 80 分	1
25 年内没有达到成功或失败条件	14

Table 1: Caption

可以发现只有 25% 的情况下可以在短期（25 年）内保证国家飞速发展的同时保持海岸线生态的稳定，而生态环境破碎的情况极少发生。相对而言，模拟结果展现出 70% 的情况无法在短期（25 年）内成功。这体现了人类的社会发展和海岸线生态的稳定是一个紧密相关的长期过程。

4.2 单局采样

我们分别从“国家分数达到 100 分”（Figure 2a）、“海岸线分数低于了 80 分”（Figure 2b）、“25 年内没有达到成功或失败条件”（Figure 2c, ??）三种情况中分别选取了一局、一局、二局进行采样分析。其中为了简化，超出 100 的都按照 100 计数。

如图，我们可以发现以下结论：

1. 对于成功在短期内快速发展的国家，其在海岸线生态保护方面一定是有所欠缺的。图中实例说明为了快速达到胜利条件，国家到后期并没有过多顾及生态的衰竭，只是保证着一定程度上的生态平衡。
2. 对于失败的模拟，由于 LLM 对于保障海岸线生态处于高水平的追求十分谨慎，仅有依次失败的模拟结果，而原因正是因为短期内采取了过多的激进政策，同时遇到了较多的负面随机事件，导致生态破碎。因此在任何时候都需要尽量避免义无反顾的激进政策。
3. 对于广泛存在的、不能在短期内获得成功发展的模拟，可以发现其社会发展都存在着阶段性的变化：在面对较差的生态环境情形时，其会选择放缓社会发展、甚至约束社会发展来保障生态回到良好的水平，之后再重新开始发展过程。这正是目前各个国家所采用的主流策略。由此可见，社会发展和海岸线生态的共同发展是一个长期的目标和过程，需要我们不断地付出思考、尝试和努力。

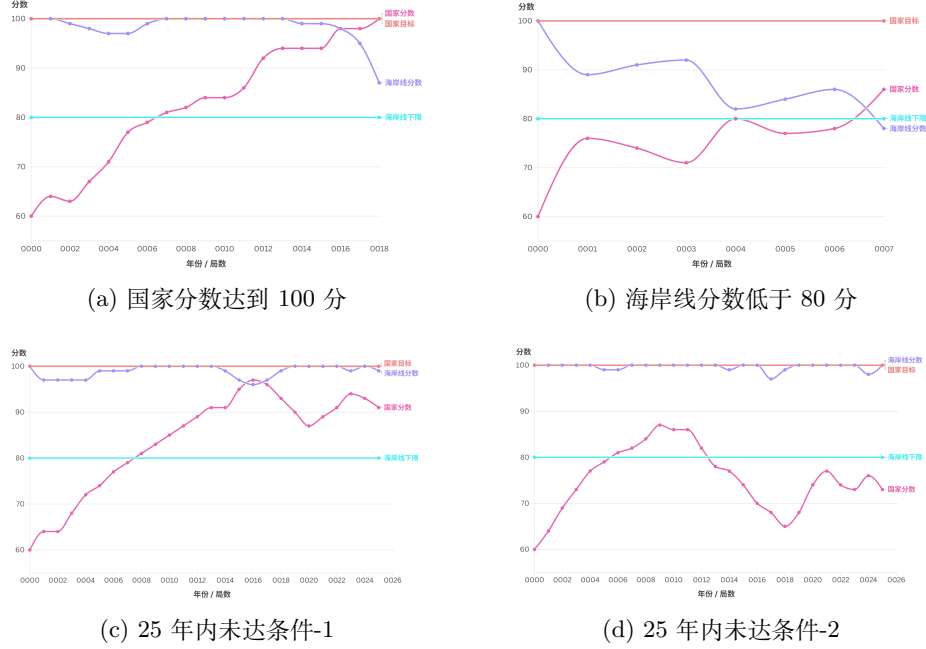


Figure 2: 四种模拟情景下的分数变化对比

5 总结与展望

本文构建并验证了 SimCoast——一个基于 LLM 的多智能体沿海社会—生态耦合动态模拟系统。通过 HumanLLM、ShoreLLM 与 JudgeLLM 之间的对抗性交互，SimCoast 不仅实现了对人类决策与生态响应的详细表现，还在实验中展示了不同策略下社会发展与生态韧性之间的状态。SimCoast 1. 首创性地将 LLM 引入沿海社会—生态系统模拟，利用自然语言作为策略表达与环境反馈的中介，突破了传统 ABM 静态规则的局限，使得模型能够自适应地生成多样化政策与生态响应 2. 引入随机事件，使得模拟更加具有鲁棒性 3. 支持生成完整的模拟结果，同时支持人类输入代替 HumanLLM 的回复，真正做到模型和现实的紧密结合，即——既可以将现实策略引入模型来模拟试验，又可以将模型的模拟创新思考应用到现实考量中。

然而，尽管 SimCoast 展现了显著潜力，仍在特色化地理与环境差异性、多“人类角色”主体动态交互、多样性数据和知识融合等方面仍有局限性，有待日后的进一步探索、完善。

References

- [1] Environmental Protection Agency. Industrial pollution and coastal water quality. <https://www.epa.gov/coastal-waters>, 2024. Accessed: 2025-07-08.
- [2] Carl Folke et al. Social-ecological resilience and biosphere-based sustainability science. *Ecology and Society*, 21(3), 2016. Accessed: 9 July 2025.
- [3] Food and Agriculture Organization. The state of world fisheries and aquaculture 2024. <https://www.fao.org/state-of-fisheries-aquaculture>, 2024. Accessed: 2025-07-08.
- [4] Chen Gao, Xiaochong Lan, Nian Li, Yuan Yuan, Jingtao Ding, Zhilun Zhou, Fengli Xu, and Yong Li. Large language models empowered agent-based modeling and simulation: a survey and perspectives. *Humanities and Social Sciences Communications*, 11(1):1259, 2024.
- [5] Taicheng Guo, Xiuying Chen, Yaqi Wang, Ruidi Chang, Shichao Pei, Nitesh V. Chawla, Olaf Wiest, and Xiangliang Zhang. Large language model based multi-agents: A survey of progress and challenges, 2024.
- [6] Intergovernmental Panel on Climate Change. Climate change 2023 synthesis report. <https://www.ipcc.ch/report/ar6/syr/>, 2023. Accessed: 2025-07-08.
- [7] Carly Lawyer, Li An, and Erfan Goharian. A review of climate adaptation impacts and strategies in coastal communities: From agent-based modeling towards a system of systems approach. *Water*, 15(14), 2023.
- [8] A. McDonald, L. Little, Randall Gray, Elizabeth Fulton, Keith Sainsbury, and Vincent Lyne. An agent-based modelling approach to evaluation of multiple-use management strategies for coastal marine ecosystems. *Mathematics and Computers in Simulation*, 78:401–411, 07 2008.
- [9] National Oceanic and Atmospheric Administration. State of the climate in 2021 report released. <https://www.aoml.noaa.gov/state-of-the-climate-in-2021-report/>, August 31 2022. Accessed: 2025-07-09.

- [10] National Oceanic and Atmospheric Administration. Coastal community resilience indicators and rating systems. <https://coast.noaa.gov/digitalcoast/training/resilience-indicators.html>, 2023. Accessed: 2025-07-09.
- [11] Joon Sung Park, Joseph C. O’Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior, 2023.
- [12] Jiahuan Pei, Fanghua Ye, Xin Sun, Wentao Deng, Koen Hindriks, and Junxiao Wang. Conversational education at scale: A multi-llm agent workflow for procedural learning and pedagogic quality assessment, 2025.
- [13] Andy Steven, Kwasi Appeaning Addo, Ghislaine Llewellyn, Vu Thanh Ca, Isaac Boateng, Rodrigo Bustamante, Christopher Doropoulos, Chris Gillies, Mark Hemer, Priscila Lopes, James Kairo, Munsur Rahman, Lalao Aigrette Ravaoarinorotsihoarana, Megan Saunders, U. Rashid Sumaila, Frida Sidik, Louise Teh, Mat Vanderklift, and Maria Vozzo. *Coastal Development: Resilience, Restoration and Infrastructure Requirements*, pages 213–277. Springer International Publishing, Cham, 2023.
- [14] UN-Habitat. Urbanization and environment. <https://unhabitat.org/urbanization-and-environment>, 2024. Accessed: 2025-07-08.
- [15] United Nations Environment Programme. Industrialization and environmental impacts. <https://www.unep.org/resources/report/industrialization-and-environment>, 2024. Accessed: 2025-07-08.
- [16] U.S. Green Building Council. Leed certification standards. <https://www.usgbc.org/leed>, 2024. Accessed: 2025-07-08.
- [17] Jiawei Wang, Renhe Jiang, Chuang Yang, Zengqing Wu, Makoto Onizuka, Ryosuke Shibasaki, Noboru Koshizuka, and Chuan Xiao. Large language models as urban residents: An llm agent framework for personal mobility generation, 2024.
- [18] World Bank. Climate-smart agriculture. <https://www.worldbank.org/en/topic/climate-smart-agriculture>, 2024. Accessed: 2025-07-08.

- [19] World Bank. Deforestation and climate change. <https://www.worldbank.org/en/topic/forests>, 2024. Accessed: 2025-07-08.
- [20] World Bank. Environmental impacts of mining. <https://www.worldbank.org/en/topic/extractiveindustries>, 2024. Accessed: 2025-07-08.
- [21] World Health Organization. Urban transport and health. <https://www.who.int/publications/i/item/urban-transport-and-health>, 2024. Accessed: 2025-07-08.
- [22] Wenting Wu, Yiwei Gao, and Chunpeng Chen. Clser: A new indicator for the social-ecological resilience of coastal systems and sustainable management. *Journal of Cleaner Production*, 435:140564, 2024.

附录 Appendix

A LLM 参考评分表格

我们从文献和各个政府的官方网站上收集了一些相关的政策措施、人类破坏行为以及其造成的正、负面影响，以及海岸线的变化规律。我们将这些政策以及其影响捆绑为数据集输送给 LLM，同时提供基础的游戏规则，让它为这些措施按照影响的好坏进行排序。我们依照它输出的排序进行了线性地赋值，正数是积极影响，负数是消极影响，取值范围是 $s \in [-5, 5]$ 。为了使得模拟系统能够有较高的成功可能性，符合现实，我们在经过并行的模拟实验、对比之后，选择将 Country 的分数变化量区间缩小，使得国家的发展分数更可能获得向上的发展趋势。

这个评分表格将会被传输给 JudgeLLM 和 HumanLLM 作为其在对抗模拟中的参考表格：针对给出的措施和行为，负责评价的 JudgeLLM 会参照表格寻找相似影响的行为作为评分参照，同时在考虑了所面对的情形和自身的行为的符合程度之后，从而给出恰当的评分；负责提供人类政策和行为的 HumanLLM 则会依照这个表格提供的分数加减变化量，来更好地理解模型的运行规律，从而更加精准地选择合适的措施来增加成功的可能性。具体的表格可见 Table 2。注意“城市扩展”被引入作为一个机遇选项，它会带来巨额的 country 加分和 shoreline 扣分，有助于让模型突破现有限制，去尝试一些风险决策。

行为	shoreline 加分	country 加分
关闭部分工厂 [1]	+5	-3
关闭渔场 [3]	+4	-3
使用有机肥 [18]	+3	-1
鼓励公共交通 [21]	+2	-1
建设绿色建筑 [16]	+1	-2
砍伐森林 [19]	-1	+1
支持化石燃料 [6]	-2	+3
矿场开采 [20]	-3	+3
发展渔业 [3]	-4	+4
发展工业 [15]	-5	+4
城市扩张 [14]	-10	+10

Table 2: LLM 参考评分表格

B 模型模拟的参数设定

在我们的实验中，SimCoast 的初始设定可见 Table 3。

项目	数据
初始国家发展分数	60
初始海岸线状态分数	100
国家胜利分数	100
海岸线生态破碎分数	80
总时间跨度（局数）	25
年际国家发展增长分数	+1
年际海岸线生态恢复分数	+1
随机事件发生概率范围	$[0, 0.05]$
随机事件影响分数范围	$[-3, 3]$

Table 3: 实验模型模拟的参数设定

C 随机事件

我们引入了随机事件系统，目的是模拟外部的不确定性变化。这包括了自然灾害（如海啸、台风）和积极事件（如珊瑚礁复苏、清洁技术突破）。具体的发生概率和分数分布范围参见 Appendix B。例如，海啸（概率 0.005）可能同时降低国家发展和海岸线分数 3 点，而珊瑚礁复苏（概率 0.02）则提升海岸线分数 3 点、国家发展分数 1 点。随机事件评估代理通过智能评估实现动态影响计算，增强模拟的真实性。