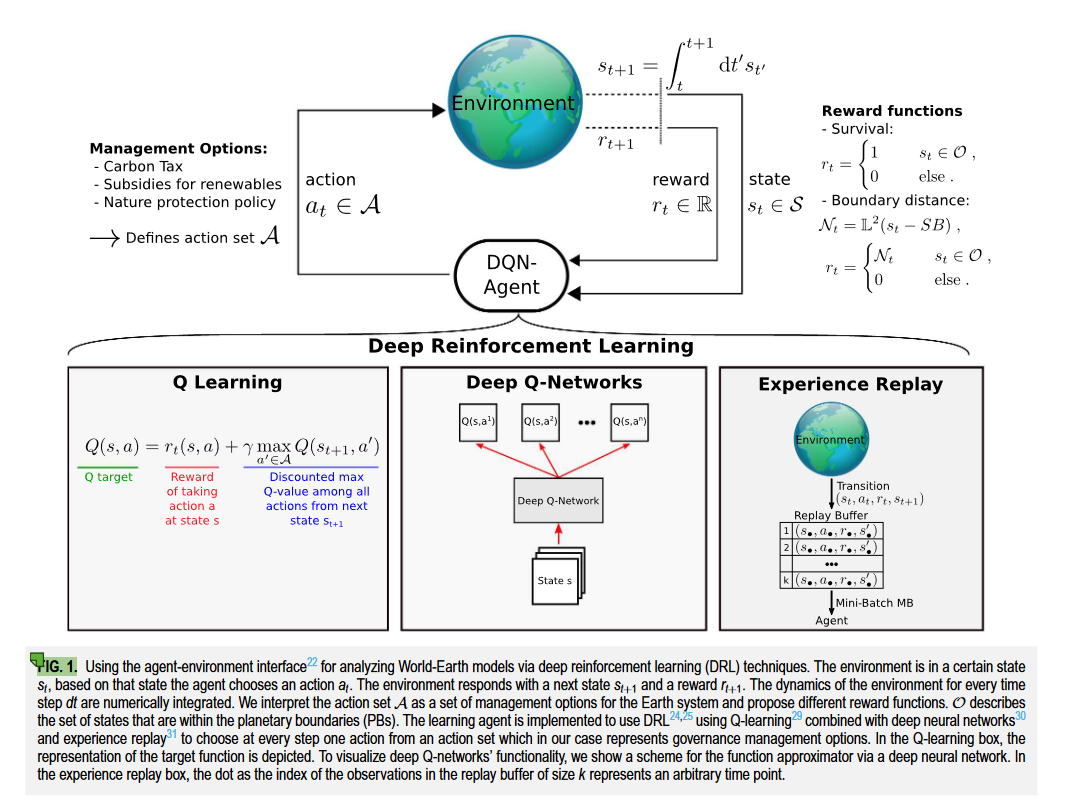
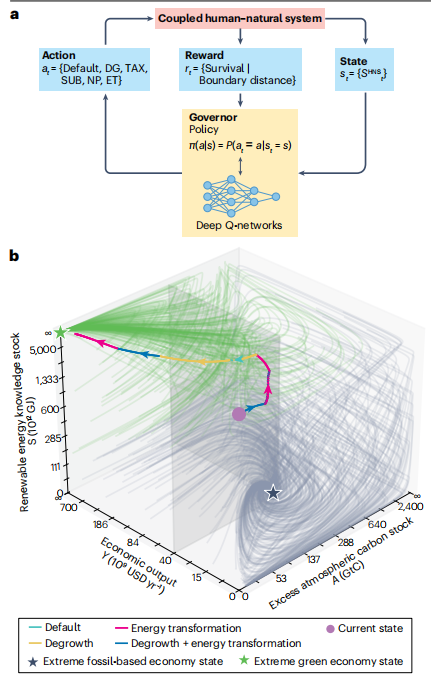
# 模型结构framework图

### 【AYS 结构示意图】

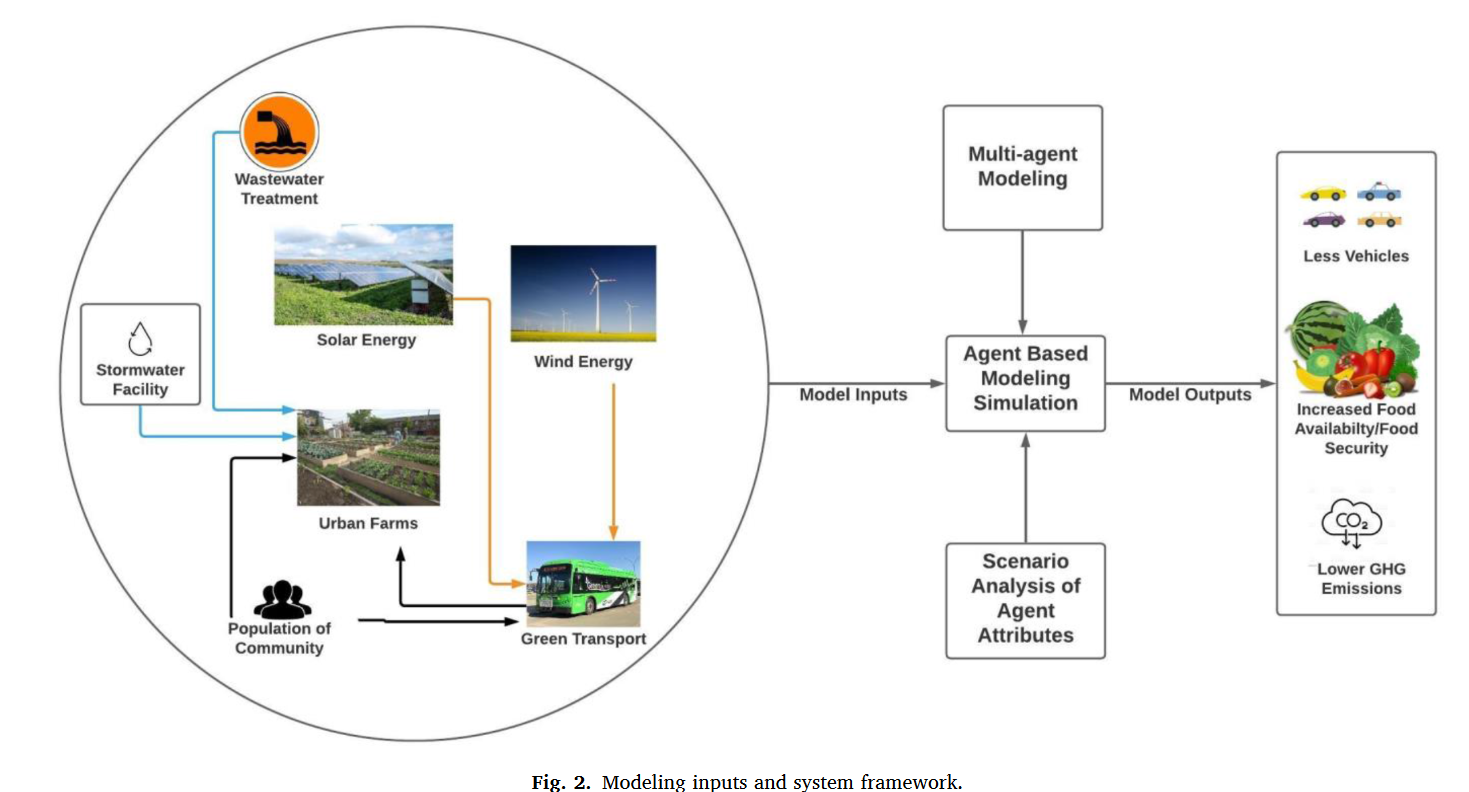
2F. M. Strnad, W. Barfuss, J. F. Donges, and J. Heitzig, “Deep reinforcement learning in World-Earth system models to discover sustainable management strategies,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 12, p. 123122, Dec. 2019, doi: [10.1063/1.5124673](https://doi.org/10.1063/1.5124673).





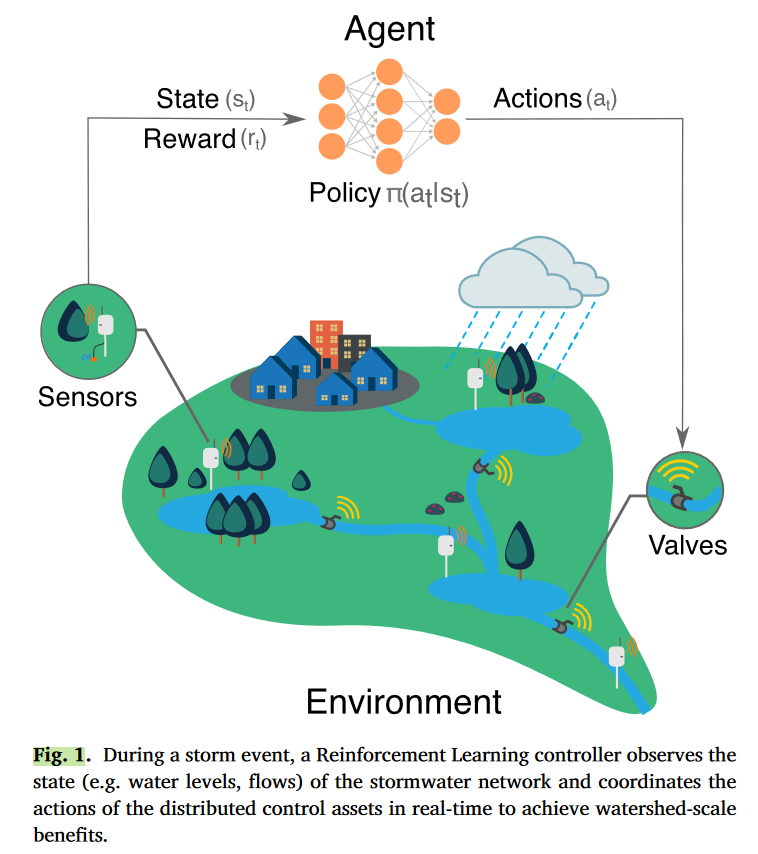
### 【建模输入和系统模块】

M. Elkamel, A. Valencia, W. Zhang, Q. P. Zheng, and N.-B. Chang, “Multi-agent modeling for linking a green transportation system with an urban agriculture network in a food-energy-water nexus,” Sustainable Cities and Society, vol. 89, p. 104354, Feb. 2023, doi: .



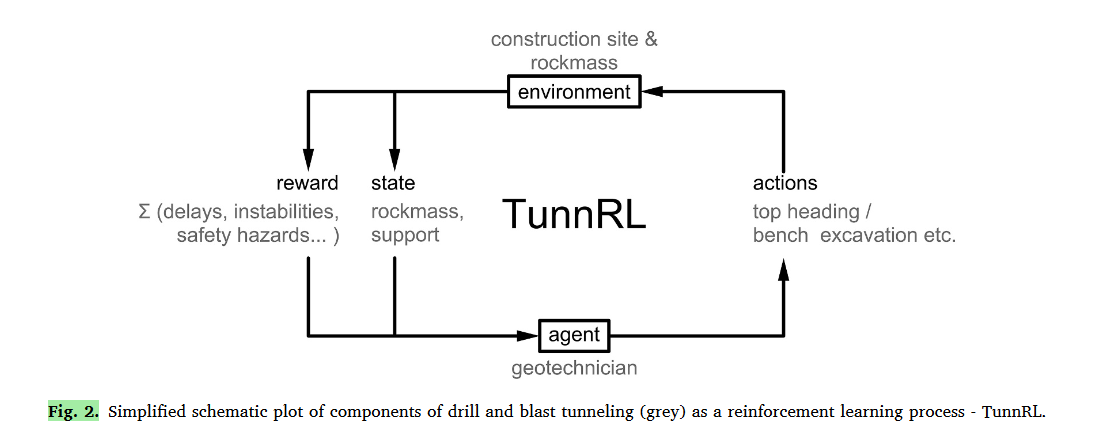
### 【整体示意图：RL-ENV 环境中的 controller 控制环境部分】

A. Mullapudi, M. J. Lewis, C. L. Gruden, and B. Kerkez, “Deep reinforcement learning for the real time control of stormwater systems,” Adv. Water Res., vol. 140, p. 103600, Jun. 2020, doi: .



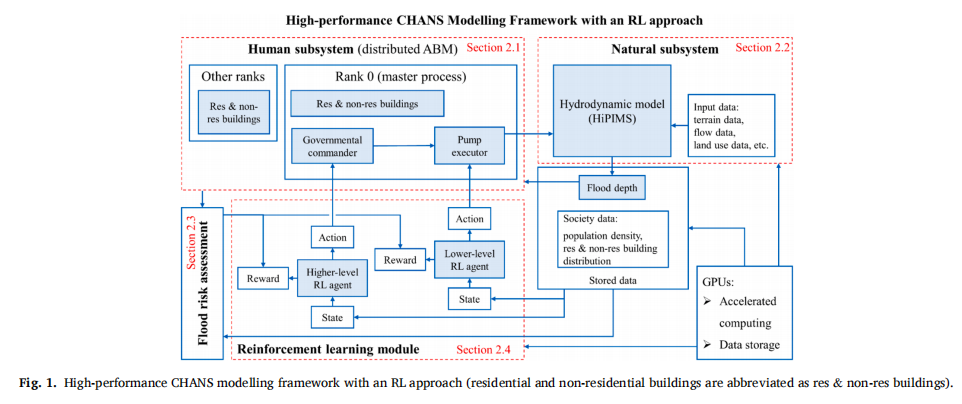
### 【简化的结构示意图】

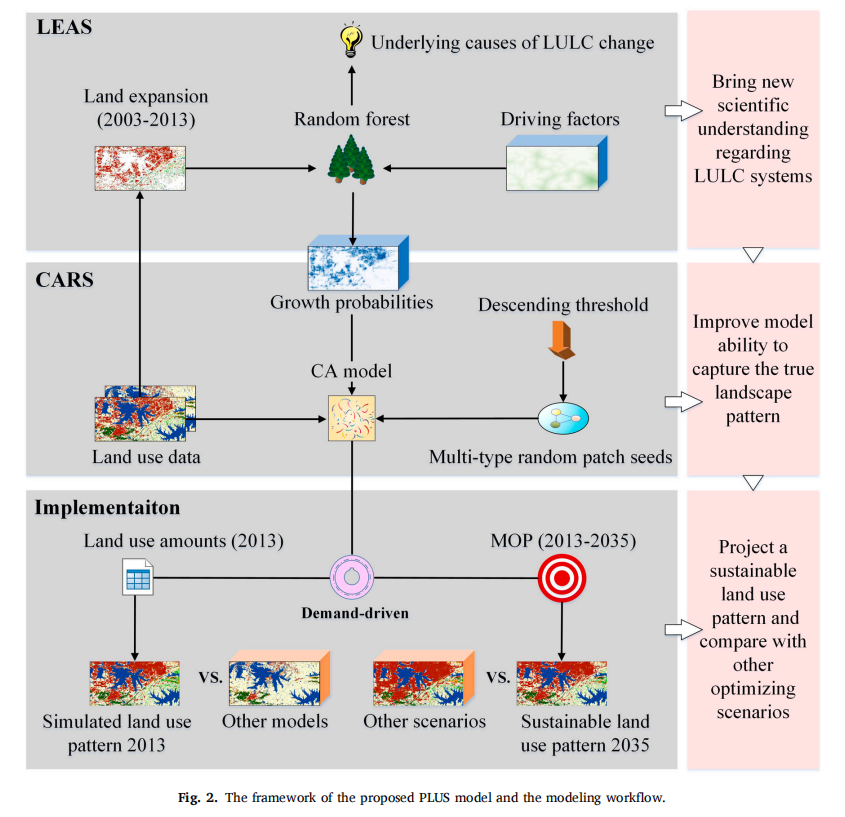
G. H. Erharter, T. F. Hansen, Z. Liu, and T. Marcher, “Reinforcement learning based process optimization and strategy development in conventional tunneling,” Automation in Construction, vol. 127, p. 103701, Jul. 2021, doi: 10.1016/j.autcon.2021.103701.



### 【结构框架和建模流程】【结构框架和建模流程】Understanding the drivers of sustainable land expansion using a patch-generating land use simulation (PLUS) model: A case study in Wuhan, China

### 【利用RL方法的高性能人地耦合建模框架】【利用RL方法的高性能人地耦合建模框架】A Coupled Human and Natural Systems (CHANS) framework integrated with reinforcement learning for urban flood mitigation

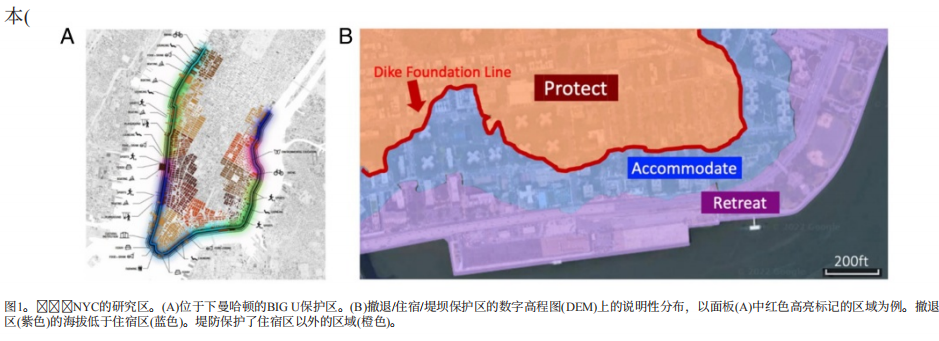




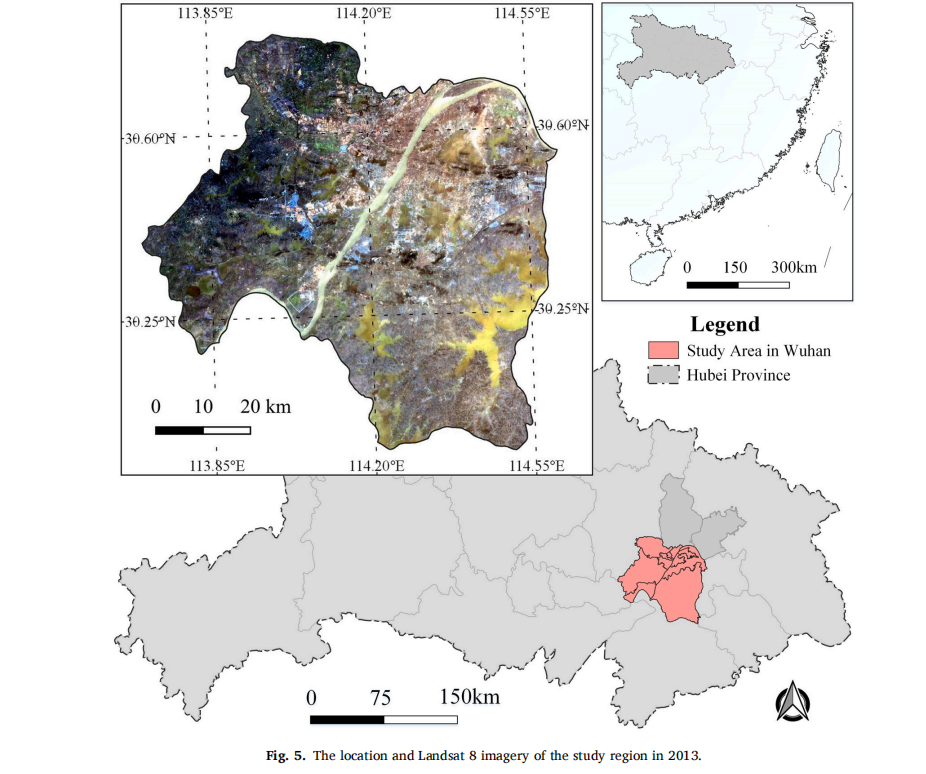
# 研究区

### 【图2 不同 SSP-2 和 SSP 8.5 下，SLR和风暴潮对应的告诉分析，RL设计结果对应】

【图2 不同 SSP-2 和 SSP 8.5 下，SLR和风暴潮对应的告诉分析，RL设计结果对应】PNAS——基 于 强 化 学 习 的 气 候 变 化 适 应 策 略 :在 沿 海 洪 水 风 险 管 理 中 的 应 用



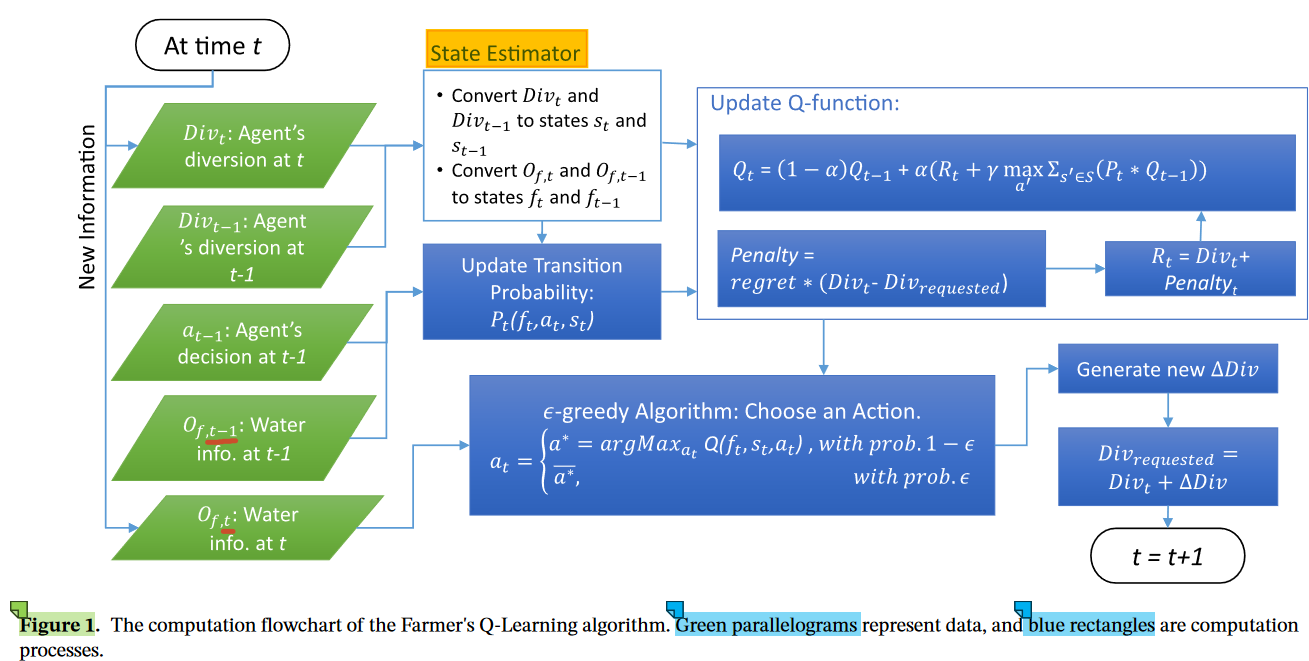
### 【图5-2013 年 landsat 8影响】【图5-2013 年 landsat 8影响】Understanding the drivers of sustainable land expansion using a patch-generating land use simulation (PLUS) model: A case study in Wuhan, China



# 算法示意

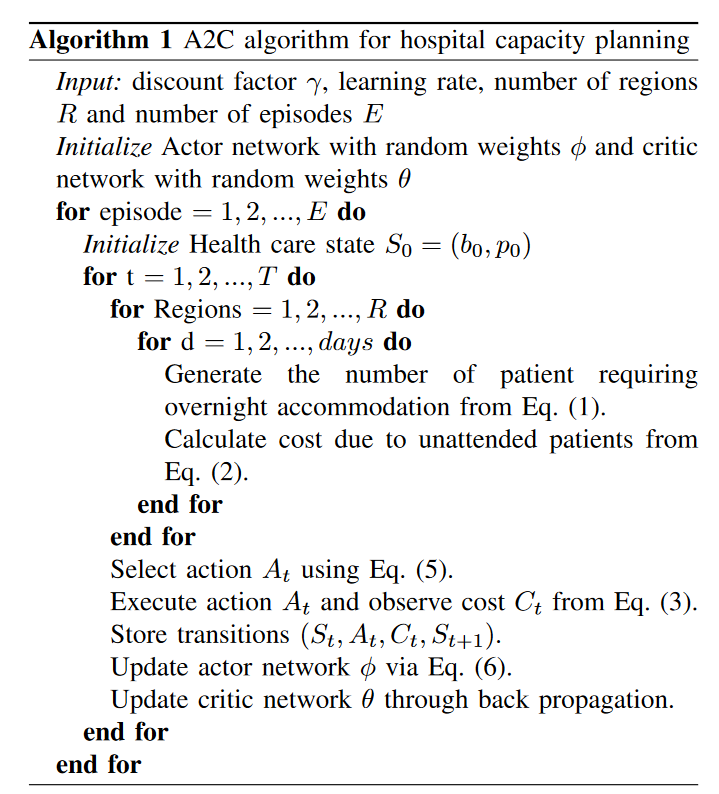
### 【流程图展示-agent rl -t 时刻对应的部分】

F. Hung and Y. C. E. Yang, “Assessing Adaptive Irrigation Impacts on Water Scarcity in Nonstationary Environments—A Multi‐Agent Reinforcement Learning Approach,” Water Resources Res., vol. 57, no. 9, p. e2020WR029262, Sep. 2021, doi: .



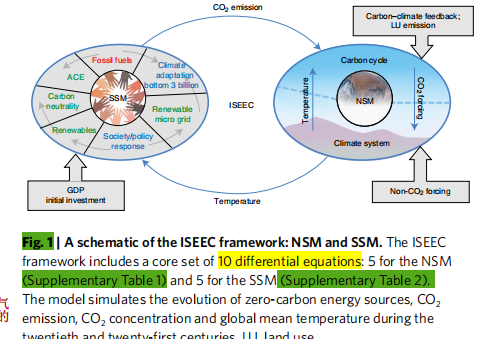
### 【算法部分】

S. S. Shuvo, M. R. Ahmed, H. Symum, and Y. Yilmaz, “Deep Reinforcement Learning Based Cost-Benefit Analysis for Hospital Capacity Planning,” in 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China: IEEE, Jul. 2021, pp. 1–7. doi: .



# 模型结构——Env

### 【图 1 ISEEC 框架】【图 1 ISEEC 框架】ISEEC 部分——Modelling human–natural systems interactions with implications for twenty-first-century warming

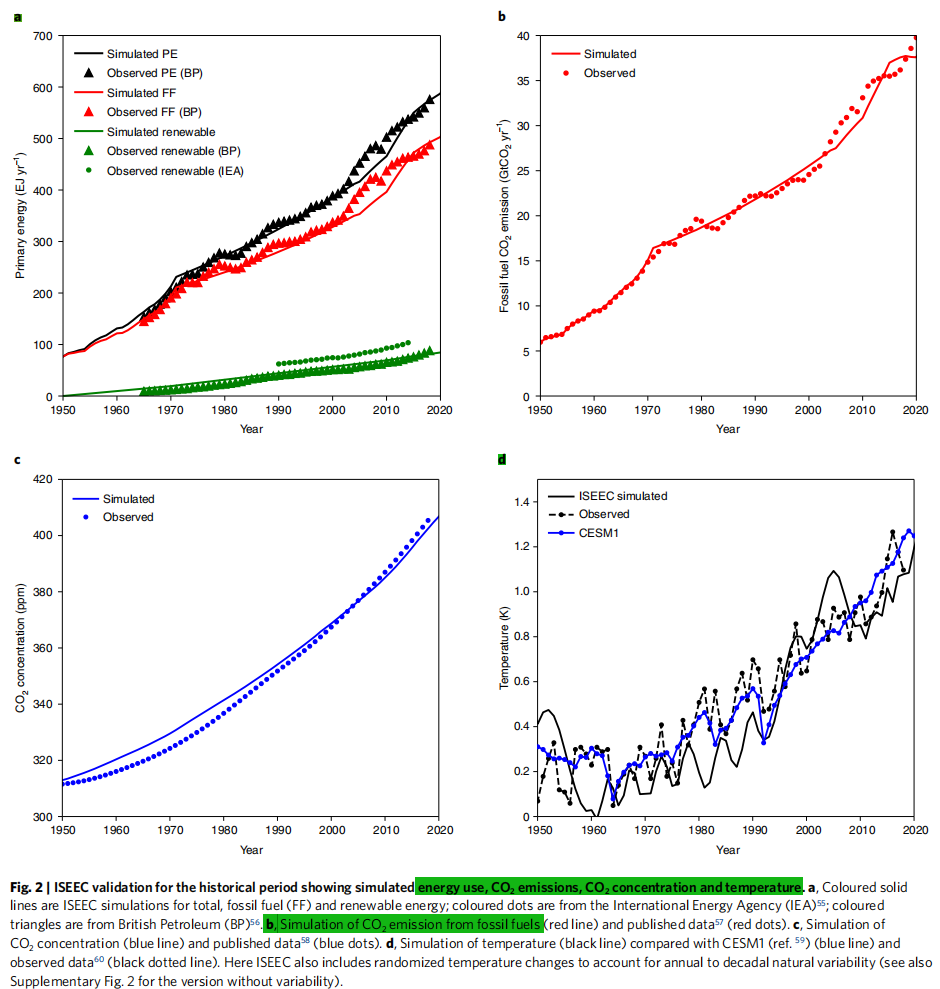


# 模型使用数据——Env

### 【表3 研究中用的数据】【表3 研究中用的数据】Understanding the drivers of sustainable land expansion using a patch-generating land use simulation (PLUS) model: A case study in Wuhan, China

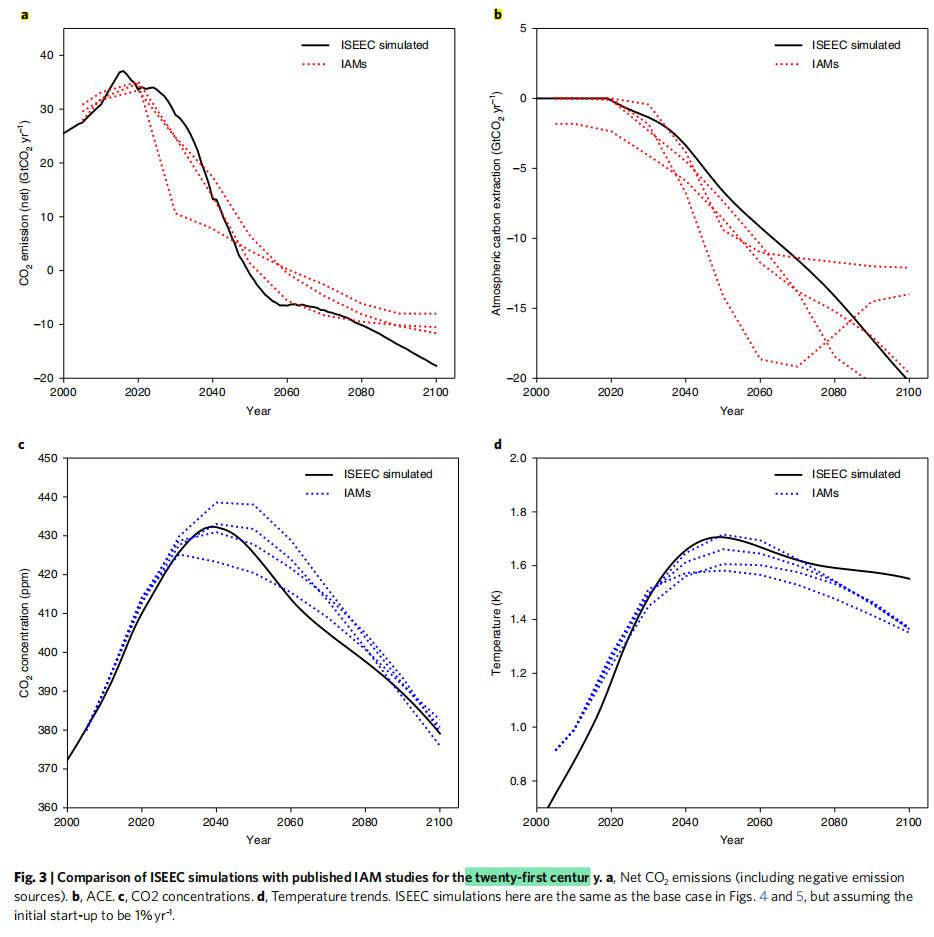
# 模型模拟结果——Env

### 【历史模拟的相关对比和验证】【历史模拟的相关对比和验证】ISEEC 部分——Modelling human–natural systems interactions with implications for twenty-first-century warming



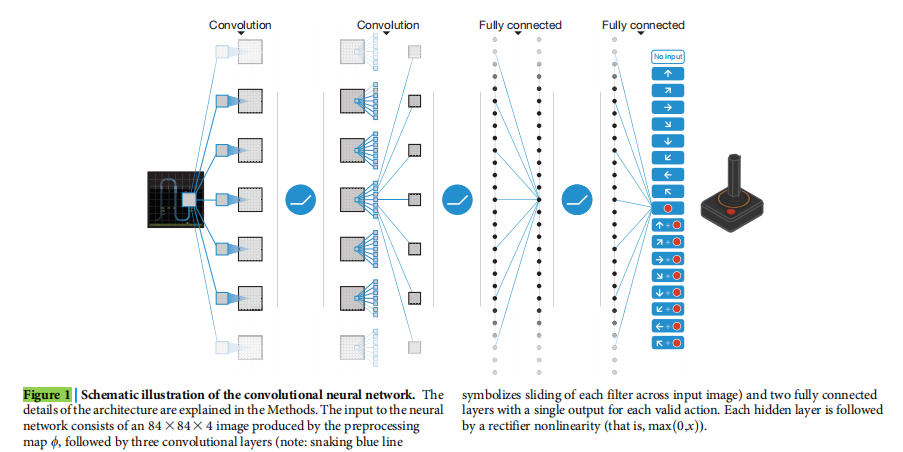
# 模型效果对比主流模型——Env

### 【在21世纪的与出版的IAM研究对比 ISEEC】【在21世纪的与出版的IAM研究对比 ISEEC】ISEEC 部分——Modelling human–natural systems interactions with implications for twenty-first-century warming



# 模型结构—— DRL

### 【卷积神经网络的结构示意图】【卷积神经网络的结构示意图】Human-level control through deep reinforcement learning



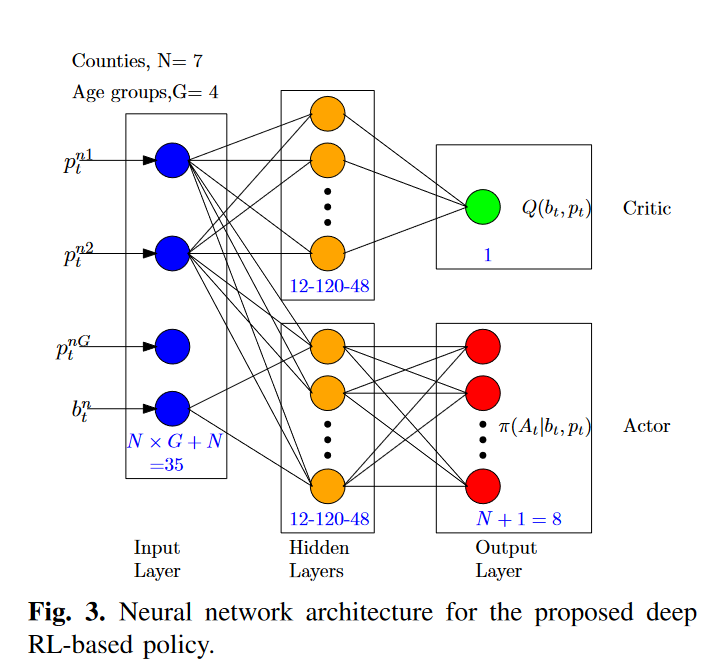
### 【DQN 和 ANN 架构示意图】

G. H. Erharter, T. F. Hansen, Z. Liu, and T. Marcher, “Reinforcement learning based process optimization and strategy development in conventional tunneling,” Automation in Construction, vol. 127, p. 103701, Jul. 2021, doi: 10.1016/j.autcon.2021.103701.



### 【RL结构部分】

S. S. Shuvo, M. R. Ahmed, H. Symum, and Y. Yilmaz, “Deep Reinforcement Learning Based Cost-Benefit Analysis for Hospital Capacity Planning,” in 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China: IEEE, Jul. 2021, pp. 1–7. doi: .



# 训练过程图

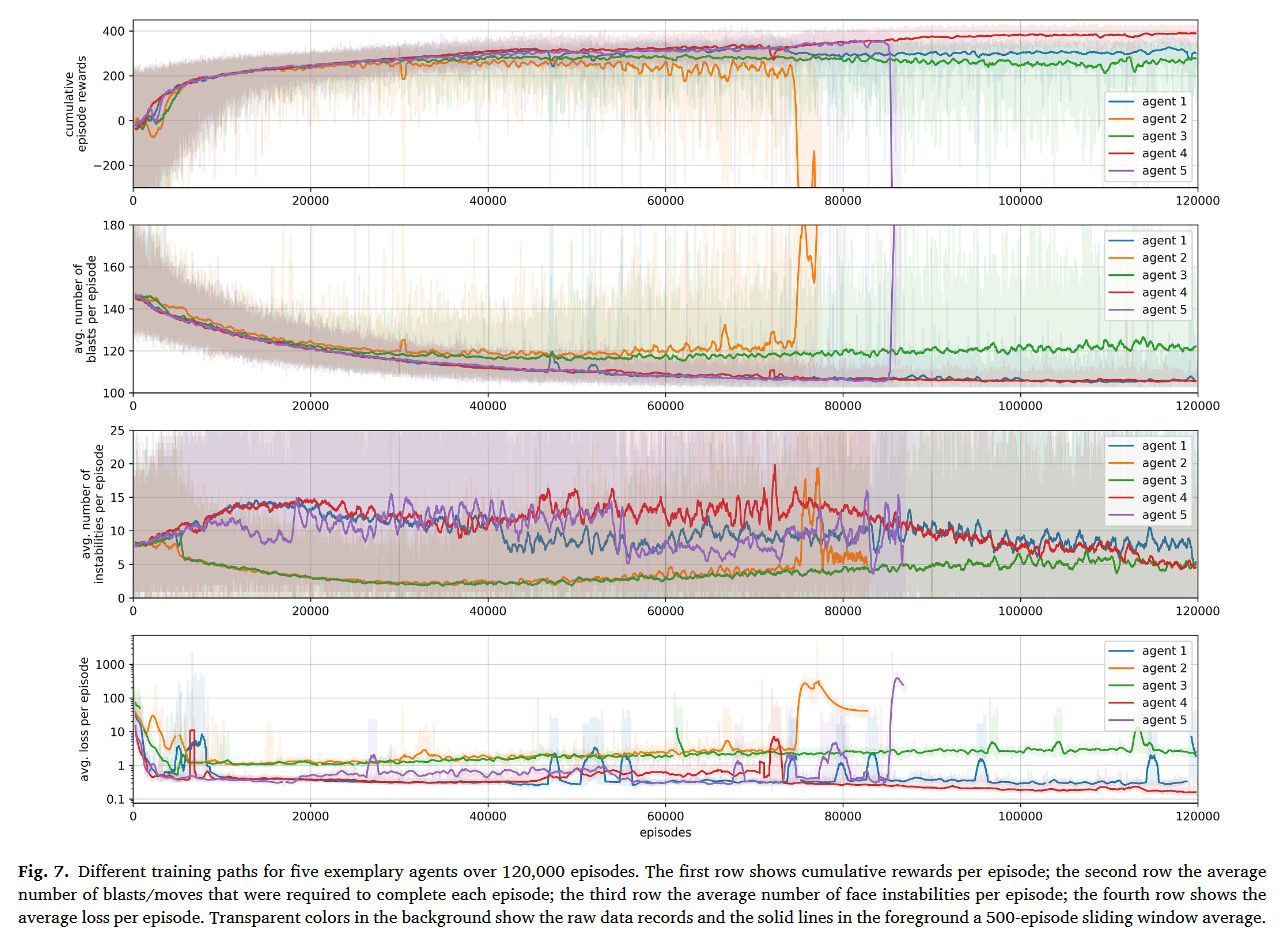
### 【single-agent-不同 reward 的结果图-随着时间——包括控制与没有控制】

A. Mullapudi, M. J. Lewis, C. L. Gruden, and B. Kerkez, “Deep reinforcement learning for the real time control of stormwater systems,” Adv. Water Res., vol. 140, p. 103600, Jun. 2020, doi: .



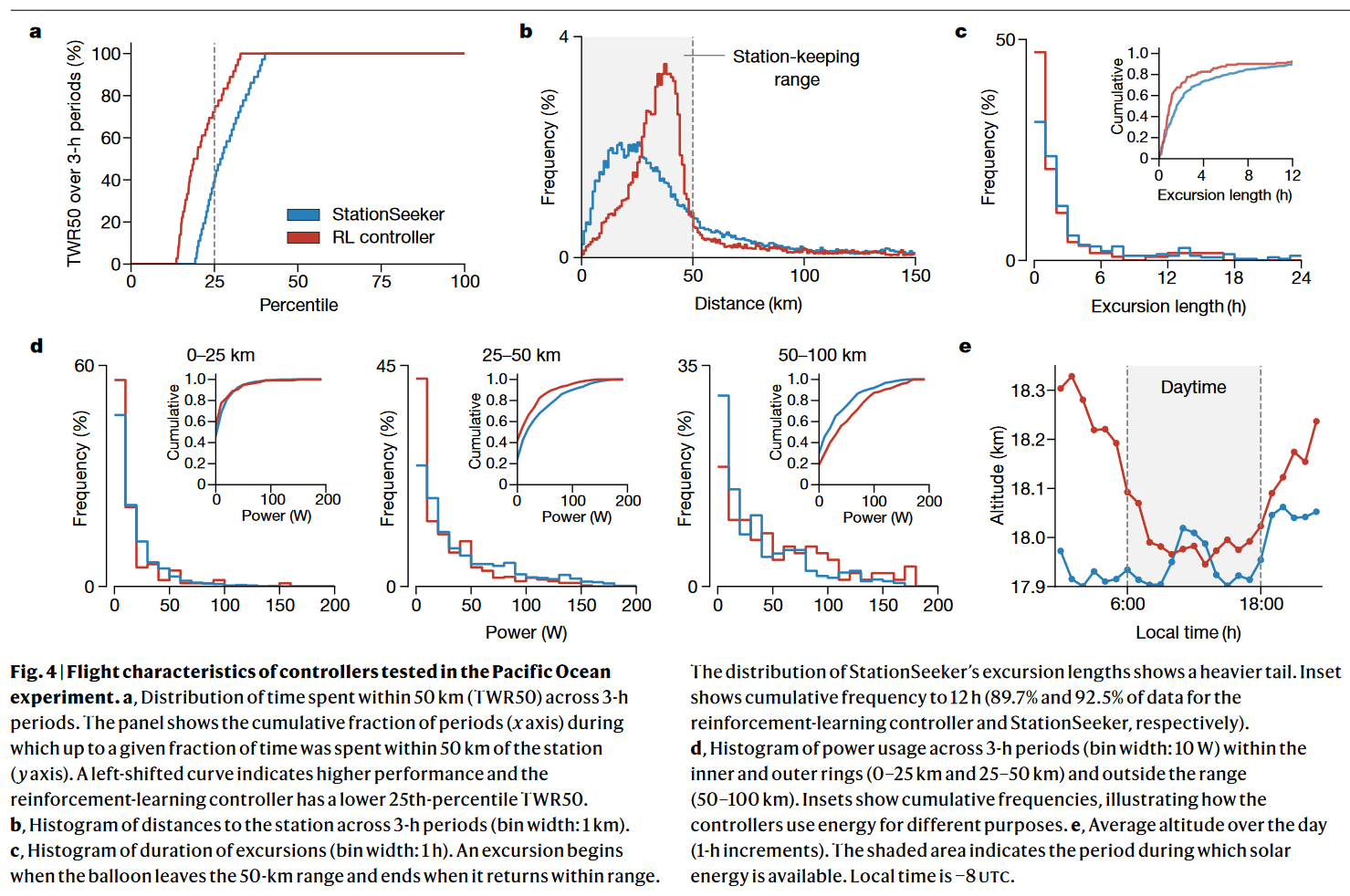
### 【不同 agent 中的训练奖励展示】

G. H. Erharter, T. F. Hansen, Z. Liu, and T. Marcher, “Reinforcement learning based process optimization and strategy development in conventional tunneling,” Automation in Construction, vol. 127, p. 103701, Jul. 2021, doi: 10.1016/j.autcon.2021.103701.



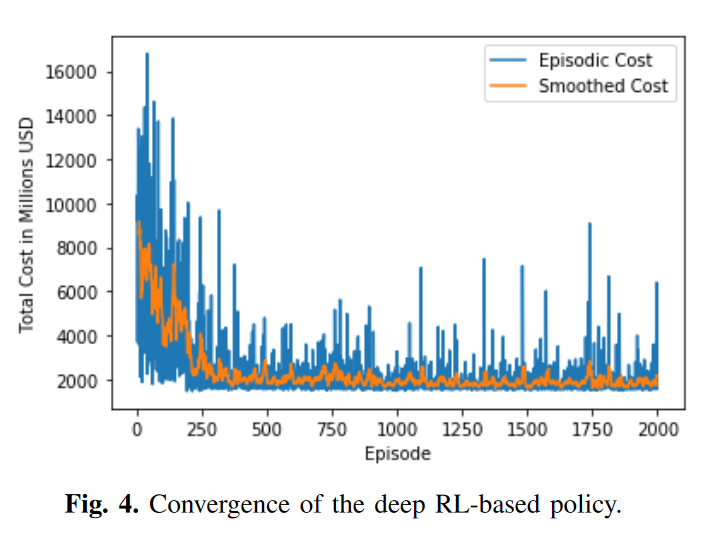
### 【在具体地区的实验飞行特征——rl 和 station seeker】

M. G. Bellemare et al., “Autonomous navigation of stratospheric balloons using reinforcement learning,” Nature, vol. 588, no. 7836, pp. 77–82, Dec. 2020, doi: .



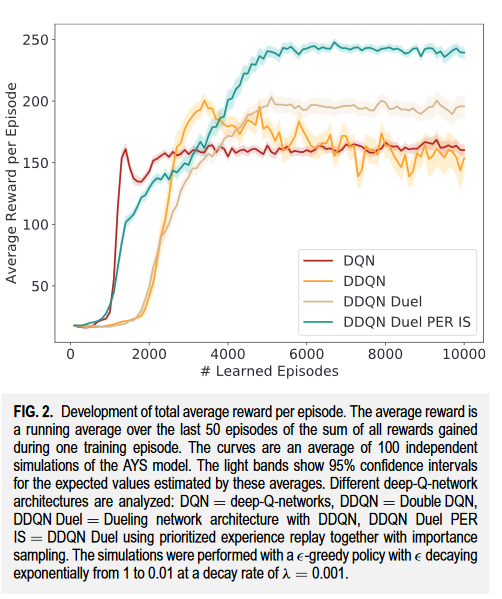
### 【基于RL结果的 reward 训练对应图】

S. S. Shuvo, M. R. Ahmed, H. Symum, and Y. Yilmaz, “Deep Reinforcement Learning Based Cost-Benefit Analysis for Hospital Capacity Planning,” in 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China: IEEE, Jul. 2021, pp. 1–7. doi: .



### 【训练次数结果】

2F. M. Strnad, W. Barfuss, J. F. Donges, and J. Heitzig, “Deep reinforcement learning in World-Earth system models to discover sustainable management strategies,” Chaos: An Interdisciplinary Journal of Nonlinear Science, vol. 29, no. 12, p. 123122, Dec. 2019, doi: .



# Reward 设计

### 【reward-1】

A. Mullapudi, M. J. Lewis, C. L. Gruden, and B. Kerkez, “Deep reinforcement learning for the real time control of stormwater systems,” Adv. Water Res., vol. 140, p. 103600, Jun. 2020, doi: .



### 【reward-2】

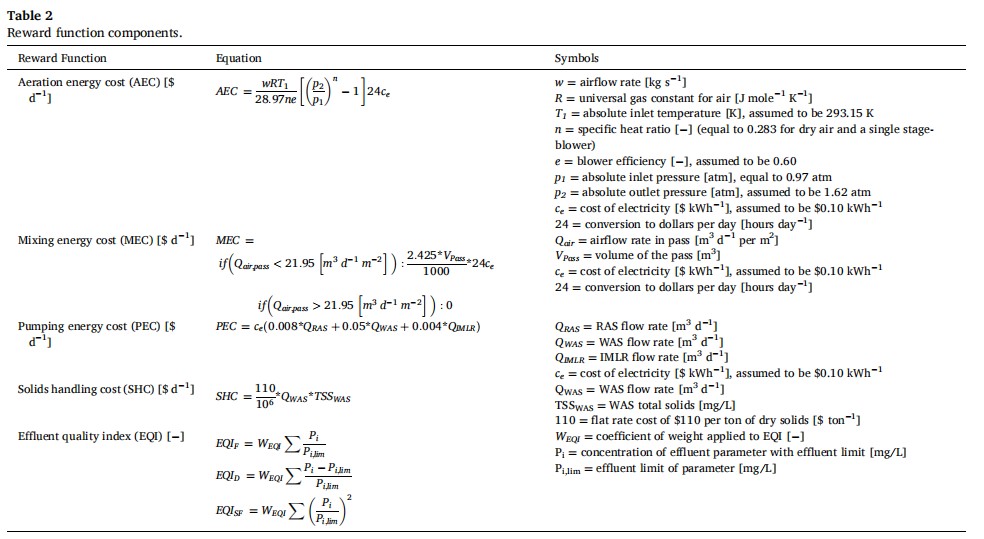
A. Mullapudi, M. J. Lewis, C. L. Gruden, and B. Kerkez, “Deep reinforcement learning for the real time control of stormwater systems,” Adv. Water Res., vol. 140, p. 103600, Jun. 2020, doi: .



# Action 设计

### 【奖励函数组件】【奖励函数组件】Reinforcement learning optimization of a water resource recovery facility: Evaluating the impact of reward function design on agent training, control optimization, and treatment risk

Reinforcement learning optimization of a water resource recovery facility: Evaluating the impact of reward function design on agent training, control optimization, and treatment risk

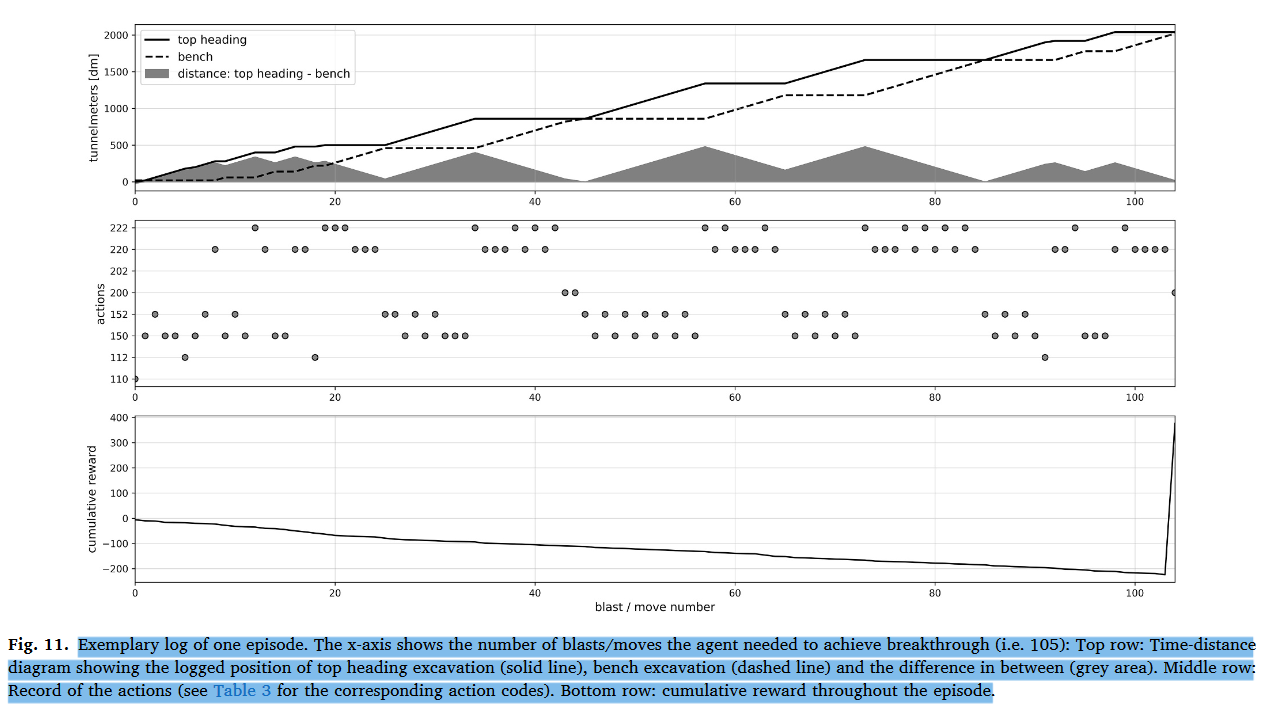


奖励函数设计是强化学习中常被提及的挑战，试错法是确定奖励函数的常用方法[1,20]。设施运营成本（OC）包括了许多与BSM1相同的组成部分，如曝气能耗（AEC）、混合能耗（MEC）、泵送能耗（PEC）和固体处理成本（SHC）[16]。与BSM1不同的是，OC直接以每天美元（$d−1）为单位表示。此外，还使用了出水质量指数（EQI）来平衡处理成本与处理水平，以激励出水合规。然而，在本研究中，EQI未纳入最终设施成本的考量，因此需要定义EQI，使得操作过程在满足监管要求的同时最小化OC。代理奖励函数利用这些组件之一或多个构建密集的奖励函数，提供接近实时的处理反馈，因为处理过程的时间尺度与代理时间步长（15 min）相似。·

# 状态展示图

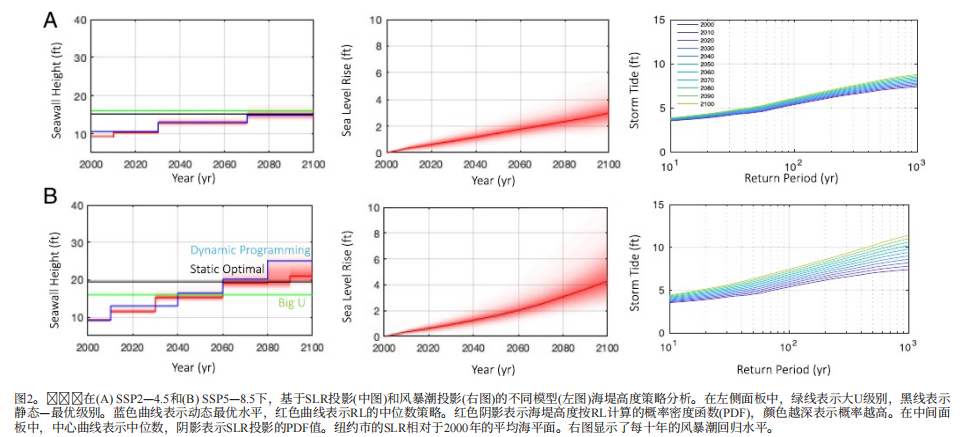
### 【一次 episode的 105次 Step 的action-reward-展示】

G. H. Erharter, T. F. Hansen, Z. Liu, and T. Marcher, “Reinforcement learning based process optimization and strategy development in conventional tunneling,” Automation in Construction, vol. 127, p. 103701, Jul. 2021, doi: 10.1016/j.autcon.2021.103701.



# 管控结果比较

### 【SSP5-8.5 情景下不同 SLR的视线，给出了三种case对应的路径变化部分】【SSP5-8.5 情景下不同 SLR的视线，给出了三种case对应的路径变化部分】PNAS——基 于 强 化 学 习 的 气 候 变 化 适 应 策 略 :在 沿 海 洪 水 风 险 管 理 中 的 应 用

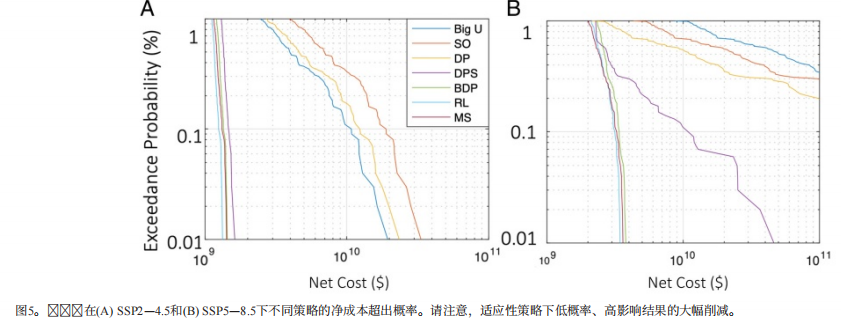


注意其中的颜色对应部分

### 【RL寻找的具体对 SSP2 和 SSP5对应的SLR变化】【RL寻找的具体对 SSP2 和 SSP5对应的SLR变化】PNAS——基 于 强 化 学 习 的 气 候 变 化 适 应 策 略 :在 沿 海 洪 水 风 险 管 理 中 的 应 用



### 【SSP2和SSP5情景下对应的**成本概率对应**——非时间维度】【SSP2和SSP5情景下对应的成本概率对应——非时间维度】PNAS——基 于 强 化 学 习 的 气 候 变 化 适 应 策 略 :在 沿 海 洪 水 风 险 管 理 中 的 应 用

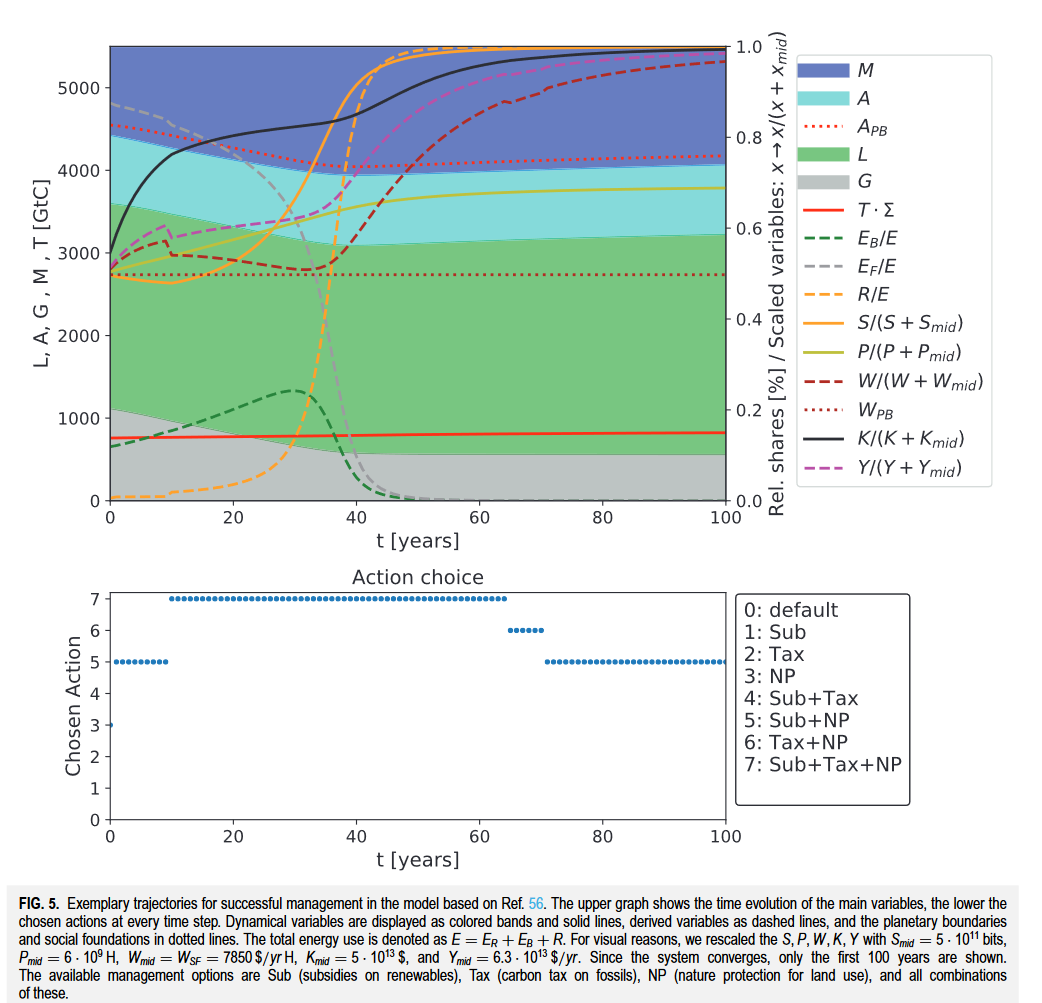


### 【表 定量地对不同场景下不同方法计算的成本部分】【表 定量地对不同场景下不同方法计算的成本部分】PNAS——基 于 强 化 学 习 的 气 候 变 化 适 应 策 略 :在 沿 海 洪 水 风 险 管 理 中 的 应 用

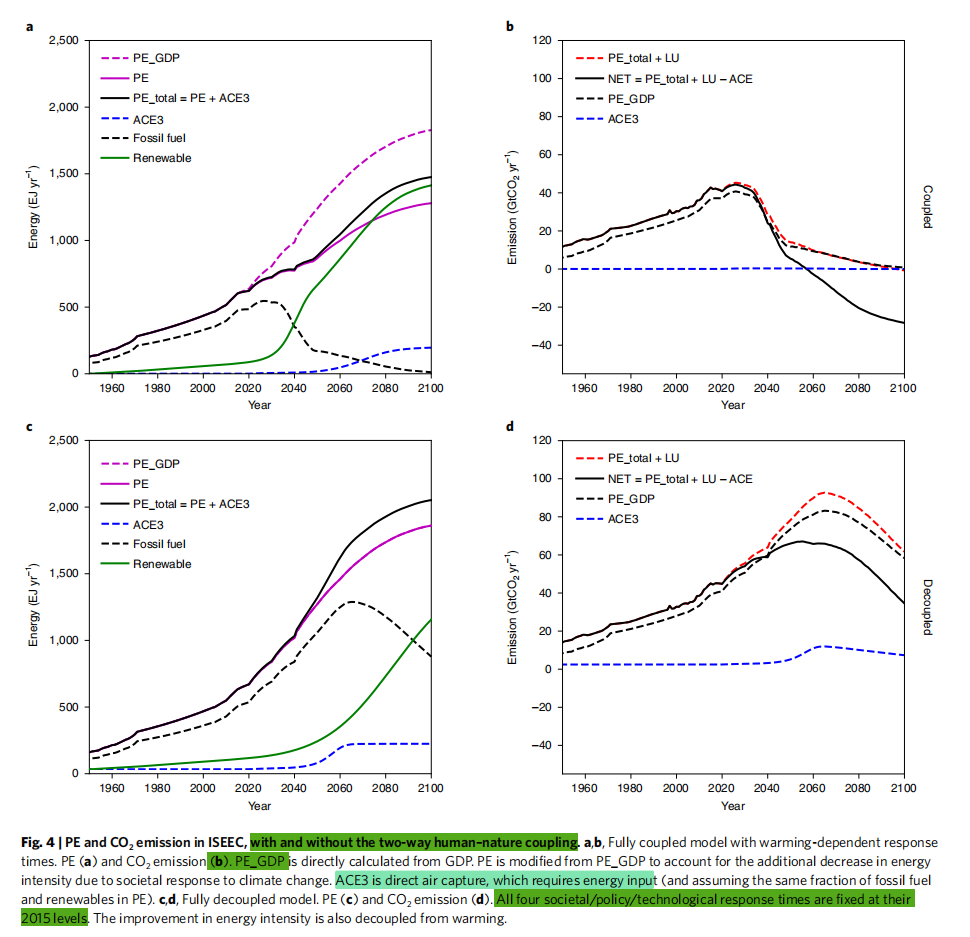


### 【训练结果展示-copan】

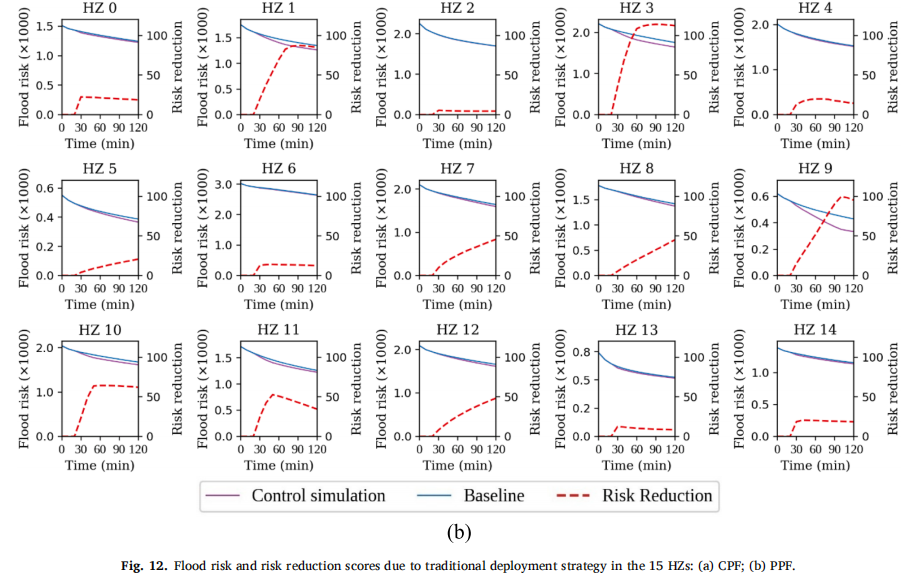
2F. M. Strnad, W. Barfuss, J. F. Donges, and J. Heitzig, “Deep reinforcement learning in World-Earth system models to discover sustainable management strategies,” Chaos: An Interdisciplinary Journal of Nonlinear Science, vol. 29, no. 12, p. 123122, Dec. 2019, doi: .



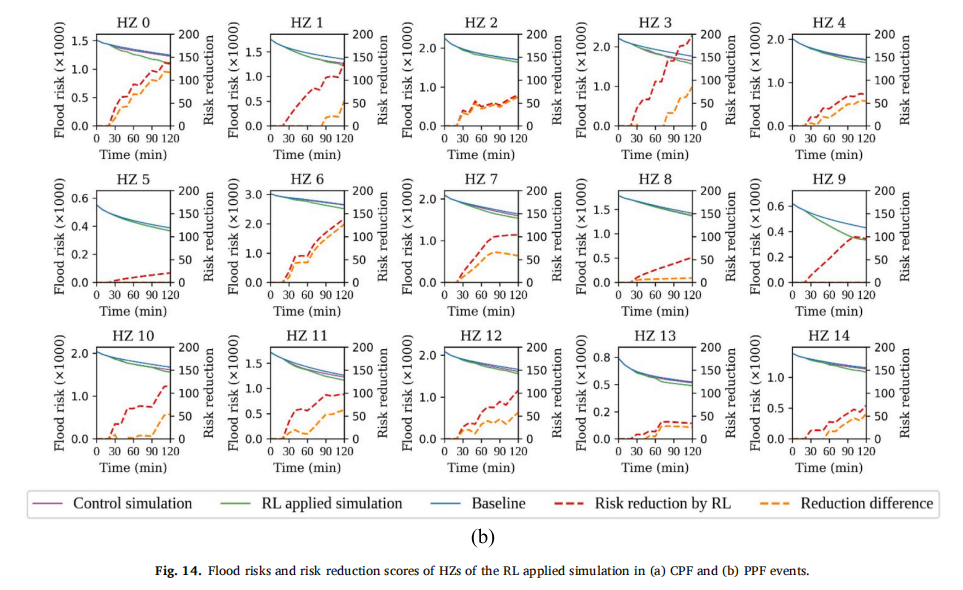
### 【有和没有 two-way 人类反馈耦合的对比 PE 和Co2排放】【有和没有 two-way 人类反馈耦合的对比 PE 和Co2排放】ISEEC 部分——Modelling human–natural systems interactions with implications for twenty-first-century warming



### 【Baseline 与多种管控结果对比】【Baseline 与多种管控结果对比】A Coupled Human and Natural Systems (CHANS) framework integrated with reinforcement learning for urban flood mitigation



### 【多种方法对比，同时对比了RL方法减少的部分】【多种方法对比，同时对比了RL方法减少的部分】A Coupled Human and Natural Systems (CHANS) framework integrated with reinforcement learning for urban flood mitigation



# 管理解释——讨论部分

### 【agent 训练结果模拟】

2F. M. Strnad, W. Barfuss, J. F. Donges, and J. Heitzig, “Deep reinforcement learning in World-Earth system models to discover sustainable management strategies,” Chaos: An Interdisciplinary Journal of Nonlinear Science, vol. 29, no. 12, p. 123122, Dec. 2019, doi: .

