要有效地适应气候变化，就必须在复杂、相互关联、具有多重不确定性的系统中制定强 有力的政策并进行公共投资。因此，规划者需要能够根据不断变化的现实情况调整计划 ( 1 )。 动态响应、灵活的适应战略通常优于僵化、固定的战略，因为它纳入了不断发展的风险和不 确定性知识。首先，可以规划分阶段的投资，促进成本相对较低的初期行动。其次，动态投 资可根据未来的意外状况调整行动或计划，避免灾难性的失败。第三，在当前决策中可以考 虑未来可能采取的行动，以避免过高估计终生风险。 灵活适应框架被称为 "适应路径" ( 2 , 3 ) 、"动态适应" ( 4 , 5 ) 或 "实际选择分析" ( 6 , 7 ) 。为模拟这些政策框架，已开发了几种分析方法 ( 8 - 10 ) 并用于评估适应措施的效益和 成本，但这些方法并未考虑灵活适应的全部潜力。表 1 将应用于环境政策设计的定量方法 按其能力分类：a) 制定动态政策；b) 纳入观测数据；c) 在当前决策中系统地考虑未来观测 和战略调整。 (a)类方法可以设计出随着时间推移的动态决策路径。例如，参考文献例如，参考文献 13 使用动态规划（DP）方法（一种经典的连续决策框架），根据目前对未来气候的预测，估算 出海岸保护的最佳路径。参考文献 14 参考文献 14 采用启发式算法，随机生成成千上万条 潜在的海岸保护路径，并选择更好的路径。这些启发式算法提高了 DP 处理多步决策维度 诅咒的能力 ( 14)。然而，这些方法假定信息基础是静态的，不能直接解决灵活政策设计的 一个关键优 点：灵活政策设计的一个关键优势：学习能力，以及在收集到外源信息后更新 和改进决策的能力。 (b)类方法可以设计出适应新观测结果的动态决策路径。在极端情况下，它们会考虑到 新信息可能会使科学信念偏离正确的气候结果，这种现象被称为负学习（21）。具体来说， 贝叶斯动态规划（BDP）方法（15，16）将新的观测数据和预测纳入最优路径的估计中，从 而扩展了传统的 DP 方法。虽然贝叶斯动态编程 可以将观察和学习纳入当前决策，但却不 考虑未来的学习和更新，因此可能无法准确估计需要应对的终生风险。换句话说，未来决策 调整的可能性会影响当前决策的最优性。 决策树或实际选择方法通过搜索情景树来生成灵活的计划，可以克服这一局限性 ( 7 , 17 , 18 )。然而，实际方案方法涉及到一棵事件树，其中的方案随着政策路径中时间步骤总 数的增加而呈指数增长。只有在潜在解决方案和情景数量有限的情况下，真实选项分析才是 可行的。直接政策搜索（DPS）方法（19，20）通过将每个时间步骤的决策建模为观察结果 的特定函数，并通过模拟优化函数参数，从而降低了计算成本。因此，错综复杂的随机顺序 决策被视为一个参数优化问题。尽管 DPS 方法计算效率高，但在适应性气候决策中仍可能 无法实现真正的最优化。 RL 是机器学习的一个领域，涉及代理在不断变化的环境状态下应如何行动，以最大化 其累积回报 ( 22 )。RL 方法系统地纳入观察结果，考虑未来的结果和反应，并在连续的未来 环境心理状态范围内提供政策设计。此外，为了提高数值效率，RL 还可以进行各种近似（例 如，在描述状态和奖励时）。RL 已在国际象棋 ( 23 , 24 ) 、自动驾驶 ( 25 ) 和机器人控制 ( 26 ) 等领域取得成功，并已被用于处理具有较大决策空间的连续环境决策，包括电力存储5 ( 27 ) 和水资源管理 ( 28 ) 。然而，它还未被用于解决气候变化风险管理中的不确定性问题。 我们研究了 RL 应用于适应性气候决策的潜力和性能。从广义上讲，我们研究了系统学习和 更新在气候适应中的价值。 为了说明这一点，我们采用 RL 方法来模拟应对沿海洪水风险的适应战略。潜在的沿海 适应策略包括："保护"，如在沿海岸线修建海堤或在高地修建堤坝；"适应"，如改造建筑物 （通过激励措施和保险法规加以鼓励）；以及 "撤退"，或从危害中撤出（可通过补贴 "买断 "加以鼓励） ( 29)。我们考虑了热带气旋（TC）和海平面上升（SLR）的预测变化，热带气 旋可能会在气候变 化的情况下引起更高的风暴潮（29-33），而海平面上升一直是并将继续 是沿海洪水增加 的主要原因。然而，未来海平面上升的预测具有巨大和深刻的不确定性， 特别是与人类排放和冰盖物理相关的不确定性，目前阻碍了最佳风险缓解战略的建模 ( 29 , 34 - 38 ) 。我们开发了 RL 方法，以确定美国纽约市曼哈顿的最优沿海风险减缓策略（包括 适应性海堤建设， 以及涉及撤出、改造和筑堤的组合策略），并将 21 世纪持续的 SLR 观 测纳入其中（材料与方法）。RL 通过状态和奖励近似方法有效地处理了计算成本（随着 SLR 情景数量和决策更新时间分辨率的增加，传统算法的计算成本呈指数级增长）[材料与方法； ( 39 )]。 我们以沿海洪水风险管理为重点，评估了 RL 在更广泛的 气候适应战略优化框架中的 有效性。通过与 DP、BDP 和 DPS 的比较，我们发现 RL 在制定灵活的战略以最大限度地 降低成本和尾端风险方面具有优势。当实际气候条件与最初设想相差甚远时，RL 框架还能 最大限度地减少 "遗憾"。这些结果凸显了持续学习和系统适应在应对气候项目中的巨大不 确定性方面的重要性，以及 RL 在模拟最佳气候适应策略方面的潜力。