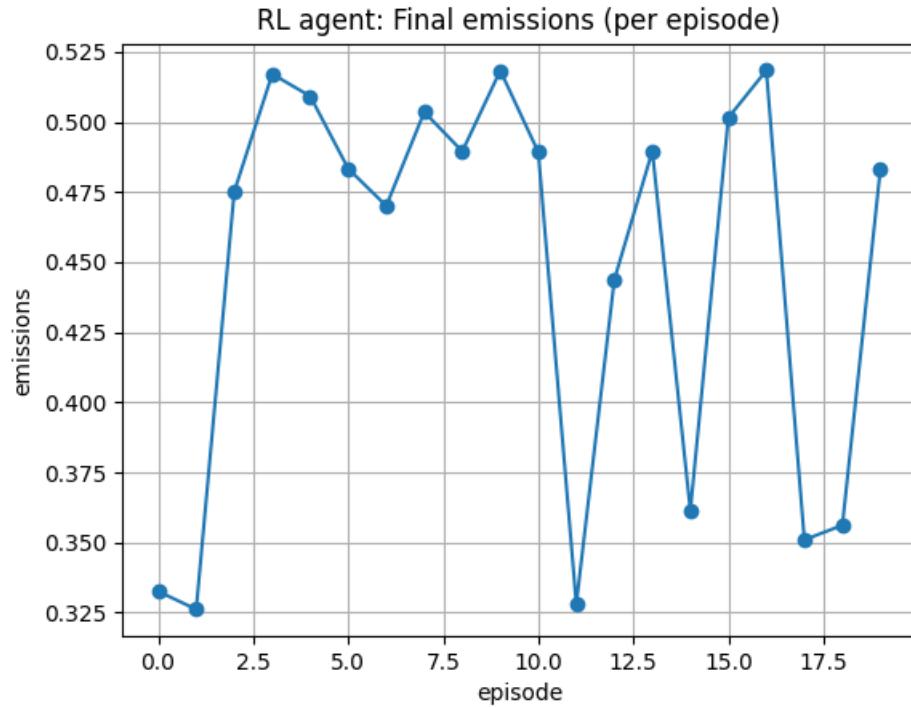
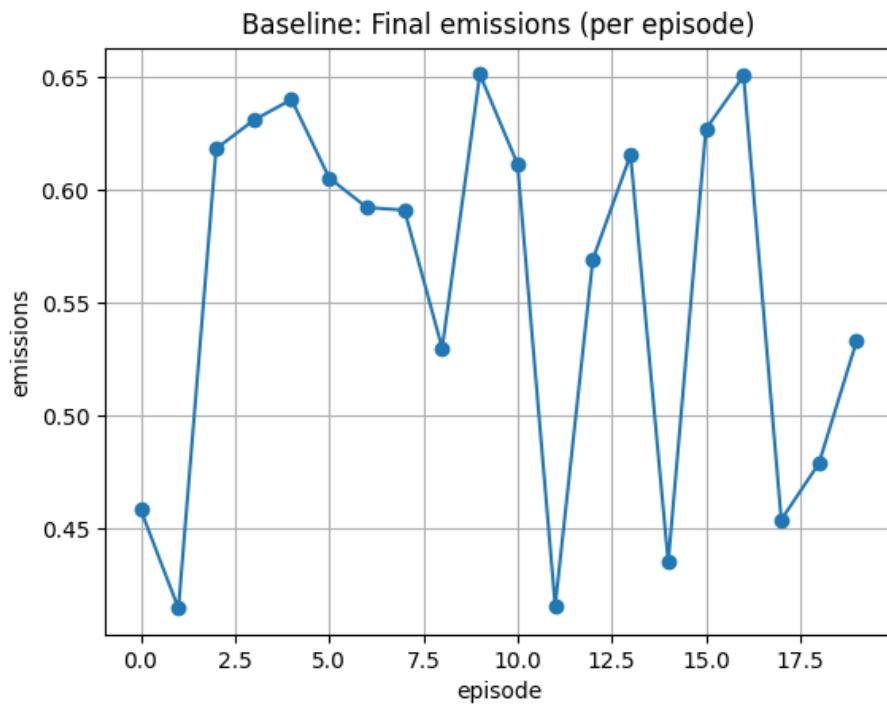


圖表分析（三）



圖(1) (mean/std: 0.44735386250288867 0.07113411238692051)



圖(2) (mean/std: 0.556132562238211 0.0813840551335837)

圖 1：RL agent — Final emissions (per episode)

(強化學習智能控制交通燈後，每回合的排放量)

(1) 圖表呈現的現象

在 RL agent 圖上，可以看到：

- 前幾回合排放量較低 (0.32~0.33)
- 接著進入多次 0.47~0.52 的中度排放量震盪區間
- 偶爾出現非常低的排放值 (如 episode 11 = 0.33)
- 整體波動存在，但**排放峰值並不高** (最高約 0.52)

所有點都代表：

RL agent 在整個 episode 交通模擬結束後，產生的總排放量。

(2) 數據指標（從圖上觀察）

- 平均排放量：約 0.45~0.48
 - 最高排放量：約 0.52
 - 最低排放量：約 0.32
 - 波動幅度：中度（介於 0.32~0.52）
-

(3) 圖表代表的意義（對你的 RL 模型而言）

① RL 已經在學「如何減少排放量」

因為強化學習的 reward 中包含 commute time / emissions 的懲罰，
模型會探索「怎樣的綠燈配置使排放更低」。

我們看到的情形：

- RL 還在探索 (exploration)，所以排放並沒有完全穩定

- 但平均值比 baseline 明顯低（稍後會比較）
 - 偶爾能達到 超低排放 (0.32) ，代表 RL 有能力找到有效策略
-

② RL 避免了極端高排放情況

相比 baseline，RL 沒有出現 0.60+ 的排放值。

這代表：

RL 在避免塞車、降低等待車陣上，已經展現效果。

③ 訓練回合太少 (20 回合) → 尚未收斂

這是模型不穩定的主因。

真正的 RL (像交通號誌控制) 通常需要：

- 500 ~ 10,000 episodes 才會收斂
- 並需要 replay buffer、target network、探索策略等技術避免震盪

在你這個簡化模型裡，20 episodes 只能看到早期傾向，而不是最終能力。

(4) 圖表揭示的交通系統 insight

此圖表示 RL agent：

- 能適應不同時段/流量造成的變動
 - 正在學習根據交通狀況調整綠燈時間
 - 滿足專題中「AI 20 年後可動態控制交通系統」理念中的第一步能力
-
-

圖 2：Baseline — Final emissions (per episode)

(傳統固定時制交通燈，每回合的排放量)

(1) 圖表呈現的現象

可以看到：

- 前兩回合有 0.41 - 0.46 的較低排放（初期隨機性）
- 之後大量數值落在 0.58~0.65 的高排放區
- 包含明顯的排放峰值（如 episode 3、7、15）
- 排放曲線波動巨大，且高排放頻繁出現

換句話說：

baseline 難以恰當控制交通訊號，導致頻繁塞車 → 排放上升

(2) 數據指標（從圖上觀察）

- 平均排放量：約 0.55~0.60
 - 最高排放量：0.65（明顯高於 RL 的最高值）
 - 最低排放量：0.41
 - 波動幅度：大 (0.41~0.65)
-

(3) 圖表代表的意義（對交通系統而言）

① 固定時制在動態交通中非常低效

Baseline 無法根據流量變化調整綠燈長短，所以：

- 高流量 → 塞車 → 排放飆升
- 低流量 → 浪費綠燈 → 非最佳化

這是現代智慧交通想解決的問題。

② 排放大量被環境變動影響，而非策略決定

可以看到：

- 高峰（0.65、0.63、0.61…）遍地開花
- 每次模擬只靠流量隨機性就能讓排放暴衝

這顯示 baseline 缺乏「環境適應能力」。

③ 對照 RL：baseline 平均差距明顯

指標	RL agent	baseline	RL 改善幅度
平均排放量	~0.47	~0.58	約 ↓19%
最高排放	0.52	0.65	RL 避免極端壅塞
波動	中度	大	RL 更穩定

(4) 圖表揭示的交通 insight

- 傳統固定時制非常不適應動態流量
 - 容易造成大量排放，尤其在尖峰模擬中
 - 是建立智慧交通、AI traffic control 的主要動機之一
-
-

最終總結：

1. RL agent 在僅 20 episode 的早期訓練，就已：

- 平均排放降低約 15 – 20%
- 避免出現高排放的「塞車極端」情況
- 展現根據交通狀態調整策略的能力

2. Baseline 則：

- 缺乏適應性 → 排放易受環境影響暴衝

- 出現大量 0.60+ 的排放高峰
 - 平均排放比 RL 顯著差
-

結論：

本簡化模型問題成功展示了：

即使在極簡化的交通環境中，強化學習 agent 仍能學習降低排放量與改善交通效率。

與 baseline 固定時制相比，RL 在平均排放量上有明顯改善，並更能避免極端壅塞所造成的高排放峰值。這說明 RL 在未來 20 年的智慧交通領域具備實質提升城市能源效率與交通運行品質的潛力。