



強化學習智慧電網電價策略 Reinforcement Learning Electricity Price Controller Design Smart Grid

組別: B80 學生: 電機系21 林奕馨 指導教授: 邱偉育 教授

研究目的

將電網達到供需平衡是一個迫切的問題。根據研究[1][2]，全球的裝置容量有20%被用在供應占總時段5%的尖峰時段用電，如此製造了許多發電建設上的浪費。因此，如何在尖峰用電時段，除了增加供給，需量反應是其中一種方法。很多研究提出若干設計方法，來實現需量反應系統。

例如，有些研究使用負載係數、尖峰對平均比例(peak to average ratio)，對系統做最佳化。然而，在現存方法中，使用最佳化方法，須已知道系統相關參數，較不易處理系統不確定性。隨著再生能源使用提高，如何降低電網/微電網系統不確定性，是一重大課題。

而強化學習 (Reinforcement Learning, RL) 透過直接與環境互動而學習的方式，不需要人為事先決定模型變數，能更有效因應市場變化，降低市場模型不準確的影響。

研究方法

此研究使用的系統架構，包含一非再生(P_g)與再生能源供給(RES)、用電需求(P_d)、儲能(E)。在電力市場模型部分，本研究使用Alvarado model [3][4] 並採用美國PJM地區data模擬

$$\begin{aligned} \tau_g P_g \dot{} &= -b_g - c_g P_g + \lambda - kE \\ \tau_d P_d \dot{} &= b_d + c_d P_d - \lambda \\ \dot{E} &= P_g + RES - P_d \end{aligned}$$

在演算法部分，強化學習的概念可簡單區分為四塊: 狀態 (State)、環境 (Environment)、動作(Action) 與獎勵(Reward)。狀態是我們想要觀察的變數 $S = [p_g \ p_d \ e]$ ，環境則會依據不同的狀態和動作，計算下一個時間的狀態，並產出獎勵。透過獎勵的設計，我們能讓agent學習達到特定目標的狀態。因為狀態空間和電價的動作空間連續且過大，這裡採用瓦片編碼(Tile Coding)將連續空間轉換成離散，以及價值函數估測來進行演員評論家法(Actor-critic method)。下圖是流程的Block Diagram，詳細Matlab程式碼可掃描左上方Qrcode。

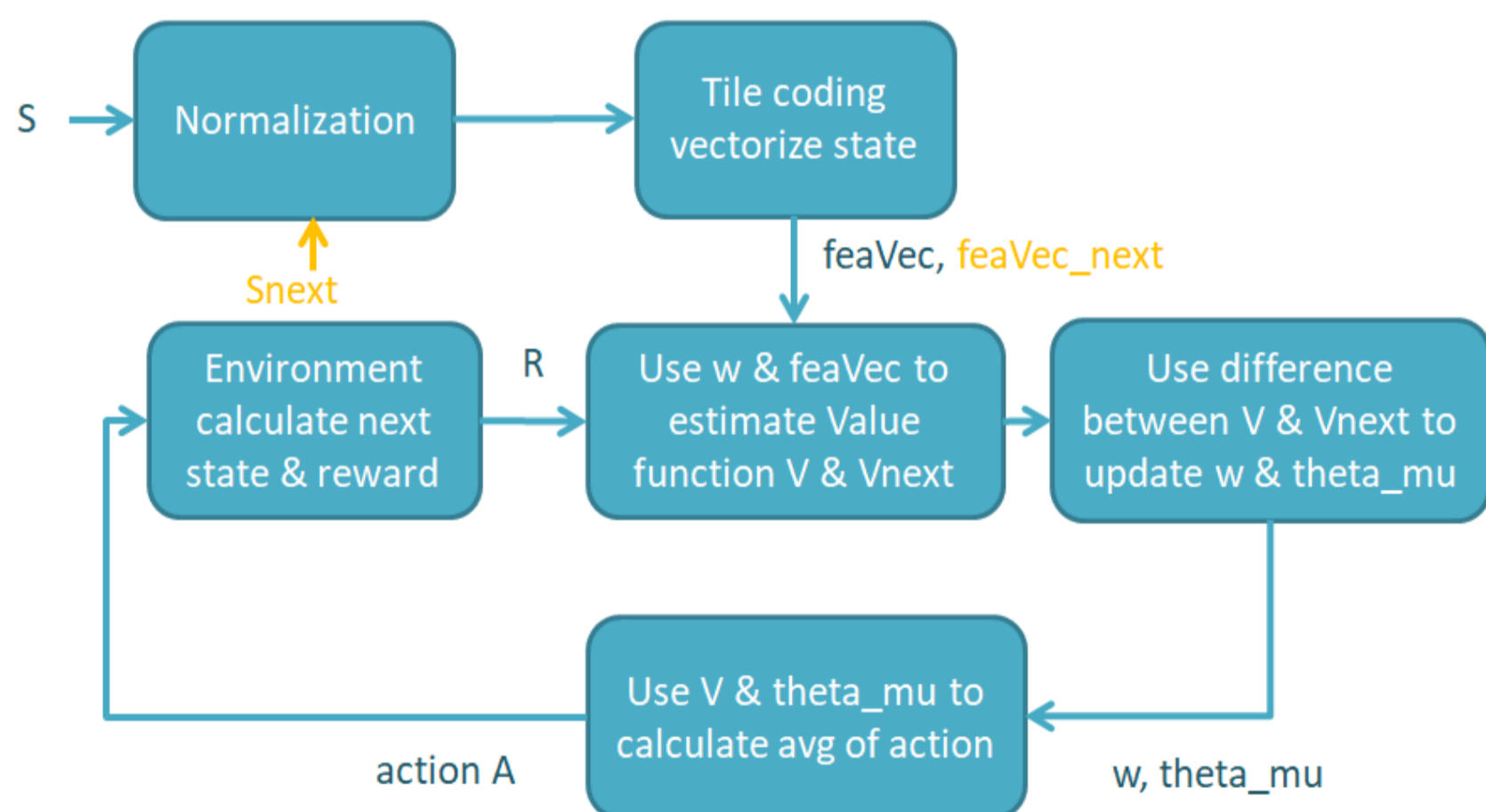


Fig1. Block Diagram 演算法流程圖

實驗結果

此研究呈現三種實驗，分別探討在不同市場參數下(如: 再生能源比例，儲能成本，以及電價變化率)，此強化學習演算法是否能因應環境變遷，達到model free的預期目標。結果顯示多數情況下能在t=2000th sample time達到收斂。證實在一定範圍下，市場參數對演算法穩定下的影響不大。但超出太大範圍就需要再行調整演算法設定。另外，結果證實電價能夠變化的速度越高，供需越快能達到平衡，然而目前最快電價變化為5分鐘一次，因此純粹透過電價達到供需平衡是長程手段，短程內仍需要藉助一定的儲能系統，或是在硬體上提高發電與需求系統對於電價的反應速度。

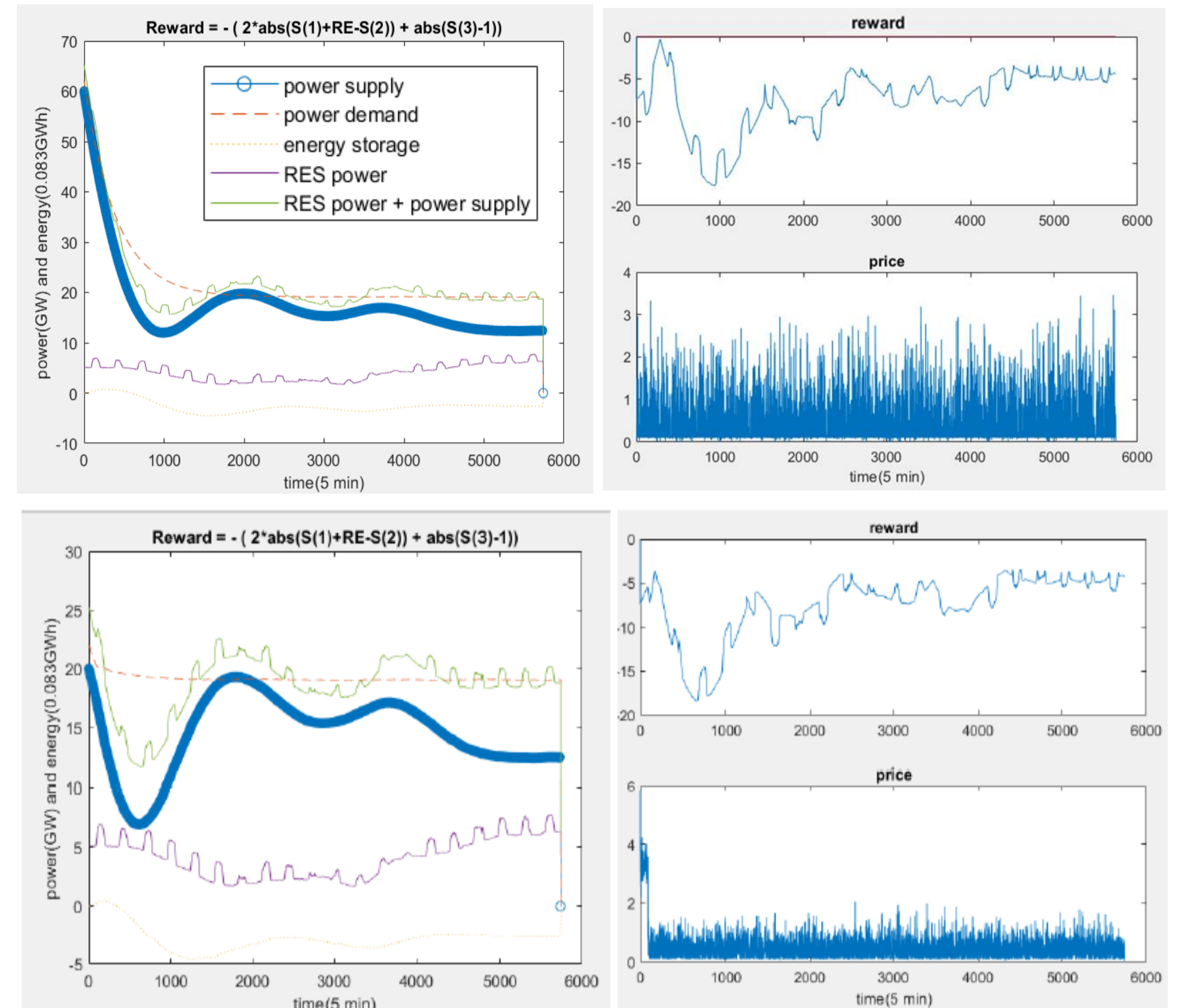


Fig2. 其中一組實驗: 再生能源佔比= 1/10(上兩圖) v.s 1/4(下兩張)

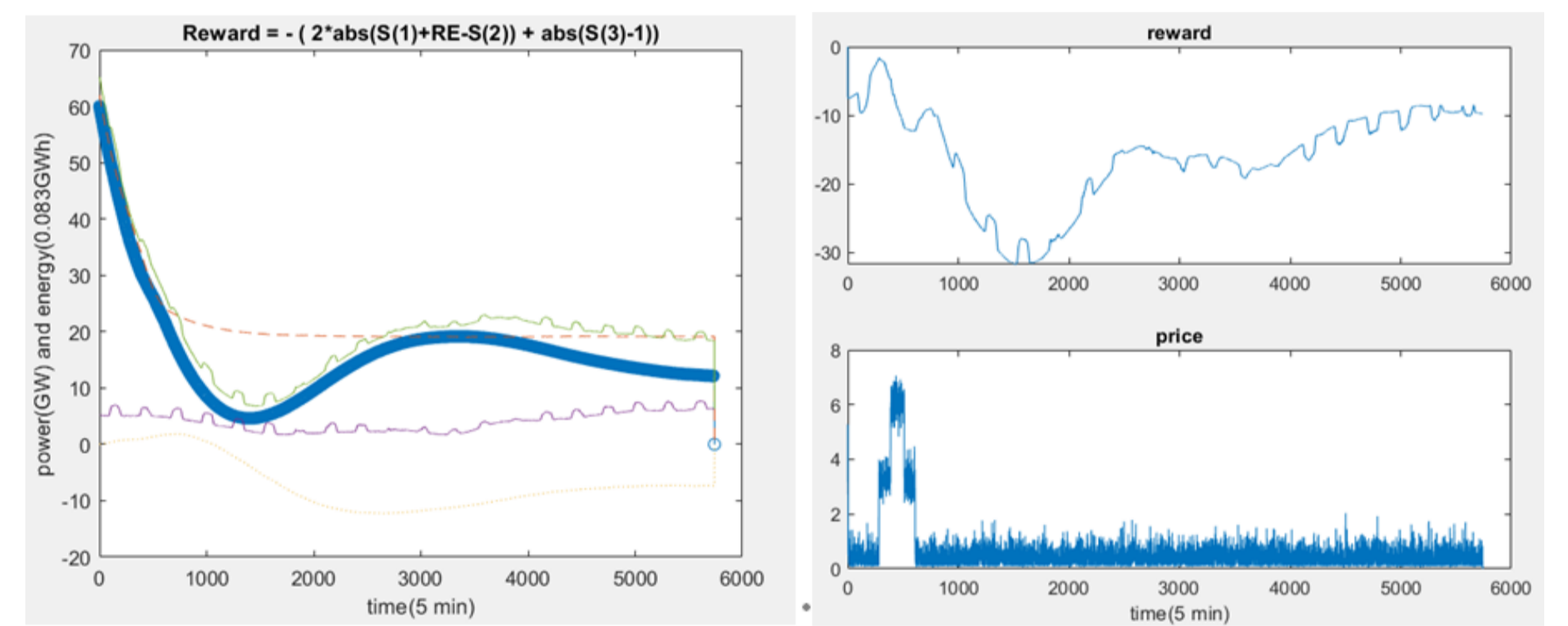


Fig3. 其中一組實驗: 儲能成本參數k=3降為k=1

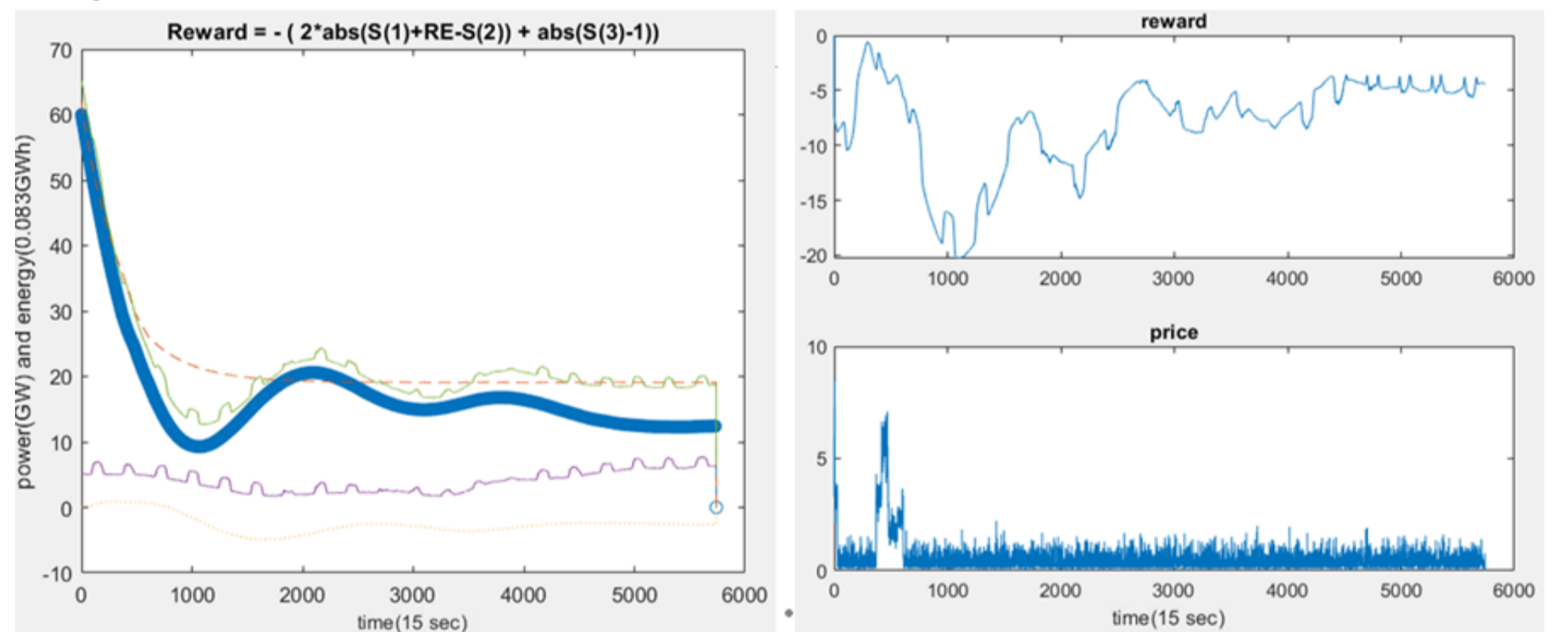


Fig4. 其中一組實驗: 取樣頻率從5min/次提升為15sec/次

預期應用&效益

1. 台電對於需量反應系統的設計: 依義電在全 球推廣需量反應的經驗，發展成熟的國際市場上，需量反應具有電網規模10%-12%的關鍵影響力。台灣夏日用電高峰約3600萬瓩，以此初估，需量反應總潛力可上看360萬瓩，相當於6.7部中火機組的裝置容量

2. 微電網: 至今越來越多民間企業或組織會設置微電網，在不同電力價格時，選擇將能源 儲存或賣電到台電電力網路，若能將控制項 結合最大化利益，將有效提升民間設置再生 能源和儲能裝置的意願。另外，儲能裝置除了一般的電池，辦公大樓的冷卻槽加熱槽也可以結合。

3. 電動車電力網絡設計: 目前系統架構只有考量再生 能源的變數，未來在推動交通減碳成 為主流時，電動車也是一龐大的儲能商機。透過強化學習 控制價格，讓使用者提升充放 電意願，便能減少 額外電網端儲能裝置的設置。