



國立臺灣大學工學院土木工程學系

碩士論文

Department of Civil Engineering

College of Engineering

National Taiwan University

Master's Thesis

使用同步擾動隨機近似演算法最佳化路網號誌

Optimizing Traffic Signals in Road Networks Using
Simultaneous Perturbation Stochastic Approximation

陳建豪

Chien-Hao Chen

指導教授：陳彥向 博士

Advisor: Yen-Hsiang Chen, Ph.D.

中華民國 114 年 10 月

October 2025



誌謝



回首剛進入研究所的自己，也讓我感慨在這兩年多以來，無論是在學業上、經驗上、生活上都有了不小的成長。而這篇論文，更是我在這段期間累積成果的具體展現。從面對艱澀難懂的數學、程式碼的修正、軟體操作的挫折與數次反覆實驗的精神消耗，每一步都讓我深刻體會到整個過程並不容易。正因如此，我更加感激在這段過程中，所有給予指導、支持與伸出援手的人們。

首先必須感謝指導教授陳彥向老師。本研究能夠順利完成，最大的功勞是來自於老師的悉心指導，特別是暑假結束後，老師依然不辭辛勞地撥冗幫我審閱論文。在整個研究過程中，老師總能點出我思慮不周的地方，也教會我如何更主動、更積極地面對研究上的種種挑戰，這些都讓我受益良多。另外，我也非常感謝口試委員李家岩教授與陳彥佑教授。兩位老師在學期中忙碌之際，仍願意抽空協助口試，並給予許多切中要點的提問與建議，真的十分感謝兩位老師的幫助。

接著，我要感謝 MoSOR 實驗室的夥伴們。在整個研究期間，我獲得了許多協助與支持，也特別感謝以下幾位同學。感謝詩涵，從上學期到口試結束，給予我非常多的實質幫助，包含提供研究意見、協助驗證結果與口試當天紀錄等。感謝欣濤不僅從口試前給予意見，更在畢業後願意特地回來幫我練習簡報，並且在口試當天也給予了很大的支持。感謝涵儀，由於安排在同一天口試，透過互相提醒口試事宜與練習簡報，有效緩和了我的緊張情緒。感謝浩哲在本研究前期給予了很多幫助。當然，也想感謝所有不論在口試前或口試後，給予我支持與鼓勵的其他 MoSOR 實驗室的夥伴們，以及曾在 314 與 114 陪我聊天、一起努力的各位。

最後感謝我的家人，感謝您們給予各方面的協助，讓我能無後顧之憂地完成這兩年多的學業，並完成這本論文。



摘要

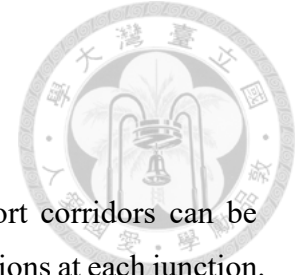


市區道路或運輸走廊之號誌連鎖主要藉由控制每個號誌路口的綠燈開始時間（時差）與綠燈長以達成。找出最適號誌連鎖路網可視為一個最佳化問題，目標為最小化負效用。本研究選擇作為目標負效用之延滯的解析式較難取得，因此使用模擬最佳化方法來直接取得目標函數值。本研究使用基於梯度的同步擾動隨機近似演算法（SPSA）作為最佳化器（optimizer），其相比於傳統基於梯度的最陡下降法在 N 維變數情況下每次迭代至少需 $N+1$ 次，SPSA 在每次迭代只需 2 次的模擬。本研究使用 TRANSYT 17 作為模擬器（simulator），希望能得到一個基於梯度的速解法。本研究基於 SPSA 設計了一個「區別變數型態 SPSA」，將決策變數的號誌時差與綠燈長分階段做迭代，考慮了迭代中止條件、階段交替條件，也另外提出綠燈長初始解的搜尋方式與限制式可行域的處理方法。演算法分別在三個不同規模的路網上進行實證，並各選取 2、3、5 個初始解。小型路網取自既有文獻，故可用來比較先前結果與本研究結果差異；中型路網顯示本研究之演算法在不同隨機擾動下具有收斂性；大型路網中則展示 SPSA 應於於多幹道路網的實務可行性。同時，小型與中型路網也做了最陡下降法作為比較基準。結果顯示，小型路網可節省 58.5% 的求解時間；中型路網雖未能改善求解時間，但最佳解可改善 7.08% 以上；大型路網可在實證下得到與 TRANSYT 17 最佳化相似之結果，展現未來應用之潛力。

關鍵字：號誌最佳化、路網連鎖、模擬最佳化、同步擾動隨機近似



Abstract



Traffic signal coordination in urban road networks or transport corridors can be achieved by controlling the green start times (offsets) and green durations at each junction. Designing an effective coordinated network can be viewed as an optimization problem, with the objective of minimizing disutility. Since a closed-form expression for delay, the disutility adopted in this study, is difficult to obtain, a simulation-based optimization framework is employed to evaluate the objective function. This framework consists of an optimizer, which adjusts the decision variables to improve the objective, and a simulator, which provides the function values as feedback to the optimizer. The optimizer is the gradient-based Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm, which requires only two simulations per iteration regardless of the number of variables, in contrast to the steepest descent method that needs at least $N+1$ simulations for N variables. TRANSYT 17 is adopted as the traffic flow simulator to achieve a gradient-based yet computationally efficient approach. A “variable-type distinguished SPSA” is proposed, in which offsets and green times are updated in separate stages, with termination rules and stage-switching criteria considered. Additional procedures are provided for searching initial green times and handling feasible domains under constraints. The algorithm is tested on three networks of different scales, using two, three, and five initial points respectively. The small network is drawn from prior literature, enabling comparison between prior results and those of this study; the medium network demonstrates convergence of the algorithm under different random perturbations; and the large network illustrates the practical applicability of SPSA to multi-arterial networks. For comparison, the steepest descent method is also applied to the small and medium networks. Results show that in the small network, solution time is reduced by 58.5%; in the medium network, solution time is not reduced but the best solution improves by at least 7.08%; and in the large network, the results are comparable to those produced by the TRANSYT 17 optimizer, demonstrating the potential for future practical applications.

Keywords: signal optimization, signal coordination, simulation optimization, SPSA



目次



誌謝	i
摘要	iii
Abstract.....	v
目次	vii
圖次	xi
表次	xv
第一章 緒論	1
1.1 研究背景與動機	1
1.2 研究目的	4
1.3 章節安排	5
第二章 文獻回顧	7
2.1 最陡下降法 (Steepest Descent)	7
2.2 隨機近似 (Stochastic Approximation)	9
2.3 號誌最佳化	10
2.4 文獻小結	12
第三章 模型建立	13
3.1 模型架構	14
3.2 變數與參數	16
3.3 目標式與限制式	22
3.4 最佳化問題	23
3.5 最佳解搜尋	24
3.5.1 SPSA 方法	24



3.5.2 SPSA 方法中的參數.....	26
3.5.3 迭代終止條件	27
3.5.4 區別變數型態之 SPSA.....	28
3.5.5 限制式處理方法	31
3.5.6 變數初始值選擇	33
 第四章 案例分析	 35
4.1 案例路網簡介	35
4.2 初始參數設定	36
4.2.1 系統參數設定	36
4.2.2 演算法參數設定	63
4.2.3 初始解搜尋	63
4.3 分析結果—旺角路網	66
4.3.1 結合不同變數型態之 SPSA—不同迭代步長參數結果.....	66
4.3.2 區別變數型態之 SPSA.....	68
4.3.3 不同演算法間的比較	72
4.4 分析結果—現代圓環	73
4.4.1 迭代步長參數搜尋	73
4.4.2 各初始解的迭代過程	75
4.4.3 不同演算法間的比較	78
4.4.4 隨機擾動與演算法的收斂性	79
4.5 分析結果—雙幹道路網	84
4.5.1 迭代步長參數搜尋	85
4.5.2 各初始解的迭代過程	85
4.5.3 不同演算法間的比較	92
4.6 小結	92

第五章 結論與建議	93
5.1 結論	93
5.2 建議	94
參考文獻	95
附錄 A 最陡下降法	101
A.1 最陡下降法之步驟	101
A.2 最陡下降法結果	103
附錄 B SIGMIX 迭代步長測試	107
附錄 C 口試委員意見回覆表	111





圖次



圖 1-1 簡單的 3×3 路網示意圖.....	1
圖 1-2 以啟發式演算法為最佳化器之模擬最佳化模型架構示例.....	3
圖 1-3 本研究模擬最佳化之模型架構.....	4
圖 1-4 研究流程圖.....	6
圖 2-1 最陡下降法示意圖（本研究繪製，函數參考 Himmelblau（1972））.....	8
圖 3-1 模型架構圖.....	13
圖 3-2 總流程圖.....	14
圖 3-3 本研究迭代過程簡圖（以 5 個變數為例）.....	15
圖 3-4 號誌週期圖.....	17
圖 3-5 一個三時相號誌路口示意圖（標線等未按實際情況繪製）.....	18
圖 3-6 時相圖（圖 3-5 路口）.....	18
圖 3-7 時差與綠燈長（各路口之第 1 個時相）.....	19
圖 3-8 時差與綠燈長（各路口之第 1 個時相， $\theta_1 = 0$ ）.....	23
圖 3-9 SPSA 範例圖（初始點相同，不同隨機擾動）（圖源：本研究繪製）.....	26
圖 3-10 區別變數型態之 SPSA 演算法邏輯流程圖（續上頁）.....	30
圖 3-11 最小綠燈平衡流程圖.....	32
圖 4-1 旺角路網基本配置圖（未按比例繪製，畫面截自 TRANSYT 17）.....	37
圖 4-2 現代圓環基本配置圖（未按比例繪製，畫面截自 TRANSYT 17）.....	38
圖 4-3 雙幹道路網全路口基本配置圖（未按比例繪製，畫面截自 TRANSYT 17）	40
圖 4-4 雙幹道路網路口 1 基本配置圖（未按比例繪製，畫面截自 TRANSYT）	41
圖 4-5 雙幹道路網路口 2 基本配置圖（未按比例繪製，畫面截自 TRANSYT）	42

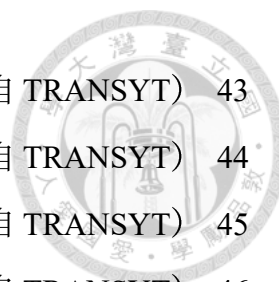


圖 4-6 雙幹道路網路口 3 基本配置圖（未按比例繪製，畫面截自 TRANSYT）	43
圖 4-7 雙幹道路網路口 4 基本配置圖（未按比例繪製，畫面截自 TRANSYT）	44
圖 4-8 雙幹道路網路口 5 基本配置圖（未按比例繪製，畫面截自 TRANSYT）	45
圖 4-9 雙幹道路網路口 6 基本配置圖（未按比例繪製，畫面截自 TRANSYT）	46
圖 4-10 雙幹道路網路口 7 基本配置圖（未按比例繪製，畫面截自 TRANSYT）	47
圖 4-11 雙幹道路網路口 8 基本配置圖（未按比例繪製，畫面截自 TRANSYT）	48
圖 4-12 雙幹道路網路口 9 基本配置圖（未按比例繪製，畫面截自 TRANSYT）	49
圖 4-13 旺角路網時相圖（可參考圖 4-1 配置）	60
圖 4-14 現代圓環路網時相圖（可參考圖 4-2 配置）	61
圖 4-15 雙幹道路網時相圖（可參考配置圖 4-3 及圖 4-4 至圖 4-12 配置）	62
圖 4-16 旺角路網：結合不同變數型態之 SPSA（點 1）— $a = 1.44$	67
圖 4-17 旺角路網：結合不同變數型態之 SPSA（點 1）— $a = 1.8$	67
圖 4-18 旺角路網：結合不同變數型態之 SPSA（點 1）— $a = 2.16$	67
圖 4-19 旺角路網：區別變數型態之 SPSA 迭代過程（結果 1）	69
圖 4-20 旺角路網：區別變數型態之 SPSA 迭代過程（結果 2）	69
圖 4-21 旺角路網：結果 1 各變數值隨迭代變化（正擾動）	70
圖 4-22 旺角路網：結果 1 各變數值隨迭代變化（負擾動）	70
圖 4-23 旺角路網：區別變數型態之 SPSA 迭代過程（結果 3）	71
圖 4-24 旺角路網：結果 3 各變數值隨迭代變化（正擾動）	71
圖 4-25 旺角路網：結果 3 各變數值隨迭代變化（負擾動）	72
圖 4-26 現代圓環：初始解 1 迭代過程	75
圖 4-27 現代圓環：初始解 2 迭代過程	76
圖 4-28 現代圓環：初始解 3 迭代過程	76
圖 4-29 現代圓環：初始解 3 各變數值隨迭代變化（正擾動）	77
圖 4-30 現代圓環：初始解 3 各變數值隨迭代變化（負擾動）	77

圖 4-31 現代圓環：不同演算法之求解時間與迭代最小值（總延滯）比較.....	79
圖 4-32 現代圓環：初始解 3 不同隨機擾動第 20 次迭代後最佳結果直方圖.....	82
圖 4-33 現代圓環：初始解 3 不同隨機擾動最終迭代結果直方圖.....	82
圖 4-34 現代圓環：初始解 3 不同隨機擾動之 KDE 結果比較.....	83
圖 4-35 雙幹道路網：初始解 1 迭代過程.....	86
圖 4-36 雙幹道路網：初始解 2 迭代過程.....	87
圖 4-37 雙幹道路網：初始解 3 迭代過程.....	87
圖 4-38 雙幹道路網：初始解 4 迭代過程.....	88
圖 4-39 雙幹道路網：初始解 5 迭代過程.....	88
圖 4-40 雙幹道路網：初始解 4 時差變數值隨迭代變化（正擾動）	90
圖 4-41 雙幹道路網：初始解 4 時差變數值隨迭代變化（負擾動）	90
圖 4-42 雙幹道路網：初始解 4 綠燈長變數值隨迭代變化（正擾動）	91
圖 4-43 雙幹道路網：初始解 4 綠燈長變數值隨迭代變化（正擾動）	91
圖 A-1 黃金分割法說明.....	102
圖 B-1 旺角路網：表 B-1 迭代過程（ $a = 3, 4, 5$ ）	108
圖 B-2 旺角路網：表 B-1 迭代過程（ $a = 6, 7$ ）	109



表次



表 3-1 集合對照表.....	20
表 3-2 變數對照表.....	20
表 3-3 參數對照表.....	21
表 3-4 五時相路口最小綠燈平衡示例.....	33
表 4-1 各路網簡介.....	36
表 4-2 旺角路網路段資料表.....	39
表 4-3 現代圓環路段資料表.....	39
表 4-4 雙幹道路網路段資料表（路口 1）	41
表 4-5 雙幹道路網路段資料表（路口 2）	42
表 4-6 雙幹道路網路段資料表（路口 3）	43
表 4-7 雙幹道路網路段資料表（路口 4）	44
表 4-8 雙幹道路網路段資料表（路口 5）	45
表 4-9 雙幹道路網路段資料表（路口 6）	46
表 4-10 雙幹道路網路段資料表（路口 7）	47
表 4-11 雙幹道路網路段資料表（路口 8）	48
表 4-12 雙幹道路網路段資料表（路口 9）	49
表 4-13 旺角路網起訖矩陣.....	50
表 4-14 旺角路網轉向量表.....	51
表 4-15 現代圓環起訖矩陣.....	52
表 4-16 現代圓環轉向量表.....	52
表 4-17 雙幹道路口起訖矩陣.....	54
表 4-18 雙幹道路網轉向量表.....	55
表 4-19 旺角路網初始點列表（2 個點）	64

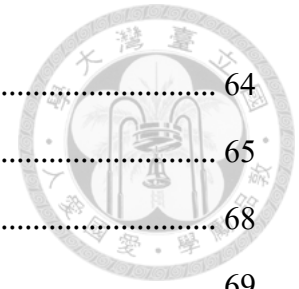


表 4-20 現代圓環初始點列表 (3 個點)	64
表 4-21 雙幹道路網初始點列表 (5 個點)	65
表 4-22 旺角路網：3 種不同的 α 值得到的最小值與求解時間表	68
表 4-23 旺角路網區別變數型態之 SPSA 結果整理	69
表 4-24 旺角路網：不同演算法之比較	73
表 4-25 現代圓環：不同組合 α 值的比較	74
表 4-26 現代圓環：三個不同初始解之迭代結果	75
表 4-27 現代圓環：不同演算法之比較	78
表 4-28 現代圓環：初始解 3 各不同隨機擾動與其迭代結果	81
表 4-29 現代圓環：初始解 3 不同隨機擾動之平均數、標準差與變異係數	83
表 4-30 現代圓環：初始解 3 不同隨機擾動之四分位數與四分位距	83
表 4-31 雙幹道路網：不同組合 α 值的比較	85
表 4-32 雙幹道路網：五個不同初始解之迭代結果	85
表 4-33 雙幹道路網：初始解 4 之最佳時制	89
表 4-34 雙幹道路網：不同演算法之比較	92
表 A-1 旺角路網初始解 1 第一次迭代估計梯度需模擬之變數組合	102
表 A-2 旺角路網：初始解 1 最陡下降法迭代過程	103
表 A-3 旺角路網：初始解 1 之第一次迭代梯度估計與單位向量	103
表 A-4 旺角路網：初始解 1 之第一次迭代步長搜尋結果	104
表 A-5 旺角路網：初始解 2 最陡下降法迭代過程	105
表 A-6 現代圓環：初始解 1 最陡下降法迭代過程	106
表 A-7 現代圓環：初始解 2 最陡下降法迭代過程	106
表 A-8 現代圓環：初始解 3 最陡下降法迭代過程	106
表 B-1 旺角路網：SIGMIX α 值搜尋 (初始解 1，結合不同變數型態 SPSA) .	107

第一章 緒論



1.1 研究背景與動機

道路路網之號誌連鎖 (coordinated) 為市區道路或運輸走廊之重要議題。連鎖的概念為當號誌間距離夠近時¹，車輛以車隊 (platoon) 形式抵達下游路口，此時必須透過號誌連鎖來避免上游路口綠燈後車隊到達下游卻又轉為紅燈的情況 (Roess 等人，2004)。換言之，良好的連鎖可以減少車輛號誌停等的時間。然而，如圖 1-1 所示，每個號誌需要考慮東西向與南北向號誌的連鎖，讓整個系統變得複雜，也將使求解時間變得較長。除此之外，在考慮號誌連鎖的同時，也應該考慮路網中各個路口的綠燈長，如何調整路網號誌連鎖與各路口的綠燈長，使整個路網達到一個最佳狀態，可視為一個最佳化問題。

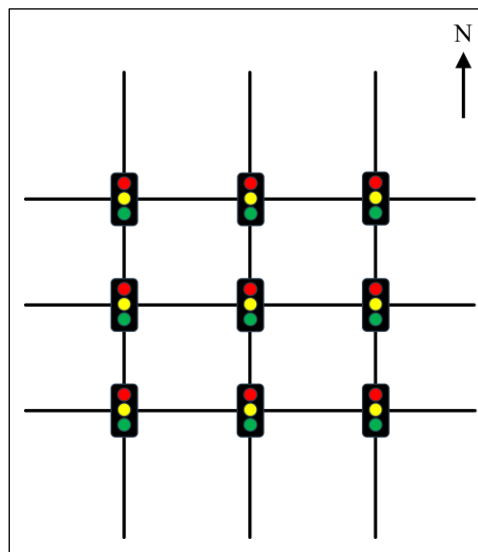
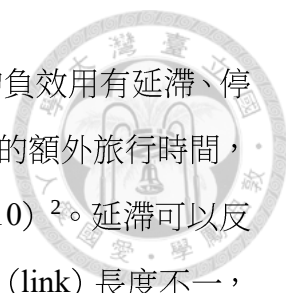


圖 1-1 簡單的 3×3 路網示意圖

¹ 根據 Roess 等人 (2004)，當號誌間距離 1 英里 (1.6 公里) 以內時須考慮號誌連鎖。



號誌最佳化問題的目標通常由最小化負效用來達成。常見的負效用有延滯、停等次數、旅行時間、燃油、廢氣等。其中延滯是指用路人所經歷的額外旅行時間，即真實旅行時間與以期望速度完成旅程所需時間的差 (TRB, 2010)²。延滯可以反映駕駛人實際感受到的不便，若僅考慮旅行時間，因為每個路段 (link) 長度不一，旅行時間不同可能造成失真情況；停等次數則無法考慮停等時間的差異以及加減速損失的時間。另外當延滯越大時，期望旅行時間不變，所需旅行時間會變長，其燃油消耗量與廢氣排放也會自然變大。基於以上原因，本研究將目標設為最小化總延滯，其決策變數為各路口號誌間綠燈開始時間差 (即時差 (offset)) 與綠燈長。

將目標式訂為最小化延滯後，必須想辦法得到延滯對時差與綠燈長的函數。而由於延滯的解析式較難以得知，其中一種方法為利用車流模擬，輸入變數值並直接得到目標函數值。這種使用模擬的最佳化技巧也稱為模擬最佳化 (Simulation Optimization, Fu, 2002)。在模擬最佳化的架構中，需要透過最佳化器 (Optimizer) 與模擬器 (Simulator) 不斷互動。最佳化器透過調整決策變數改善目標函數；模擬器則輸出目標函數值並回饋予模擬器。由於相同模擬器的速度差異不大，找到一個理想的最佳化器，以降低模擬次數來加快整體求解速度就非常重要。

目前在實務上常使用啟發式演算法 (如基因演算法 (Genetic Algorithm)、爬山法 (Hill Climbing)、模擬退火 (Simulated Annealing) 等) 作為最佳化器 (如 GATRANPFE (Ceylan & Bell, 2004)、VISGAOST (Stevanovic 等人, 2007) 與 TRANSYT-7F (McTrans, 2008) 等)，架構如圖 1-2 所示。然而在交通工程實證上，這些方法大多缺乏收斂性保證，且在高維問題中往往伴隨龐大的計算需求。本研究希望尋求一種能有一定的收斂性保證，且在高維問題中也能快速求解的演算法。該

² 延滯 (delay) 又可分為停等延滯 (stopped delay)、加減速延滯 (acceleration/deceleration delay)、排隊延滯 (queue delay) 等，以上可合稱為控制延滯 (control delay)，本研究所稱之延滯即為控制延滯 (TRB, 2010)。

種演算法可從具收斂性保證的以數學理論為基礎的方法中加以探討。

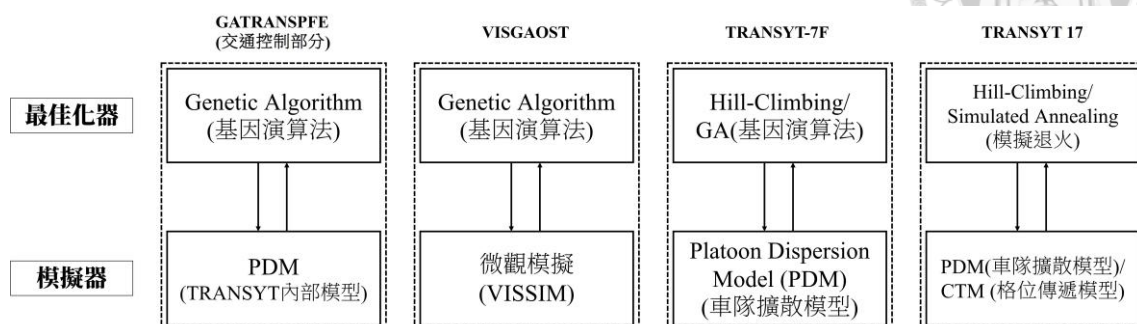


圖 1-2 以啟發式演算法為最佳化器之模擬最佳化模型架構示例

當考慮使用基於數學理論之演算法時，傳統的最陡下降法（Steepest Descent, Cauchy, 1847 ; Nocedal & Wright, 1999）使用負梯度³方向為函數下降最快方向的概念，藉由一次次的梯度更新迭代來找到最佳解，然而卻會遇到一次迭代的模擬次數隨變數數目增加而增加的問題。若能使模擬執行次數降低，且能得到相同效果之梯度，則最為理想。因此，本研究將使用 SPSA（Simultaneous Perturbation Stochastic Approximation，同步擾動隨機近似）方法（Spall, 1987）作為最佳化器，模擬器部分則使用 TRANSYT 17 內部的 PDM 與 CTM 模型，其架構如圖 1-3 所示。在模擬最佳化中，此方法可節省每次迭代的模擬次數至 2 次（無論變數數目多寡），期望能得到加快求解速度的效果。

³ 梯度 (gradient) 為一個可微多變數函數 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 在點 $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]^T$ 的所有偏導數組成的向量，即 $\nabla f(\mathbf{x}) = \left[\frac{\partial f}{\partial x_1}(\mathbf{x}) \ \frac{\partial f}{\partial x_2}(\mathbf{x}) \ \cdots \ \frac{\partial f}{\partial x_n}(\mathbf{x}) \right]^T$ 。

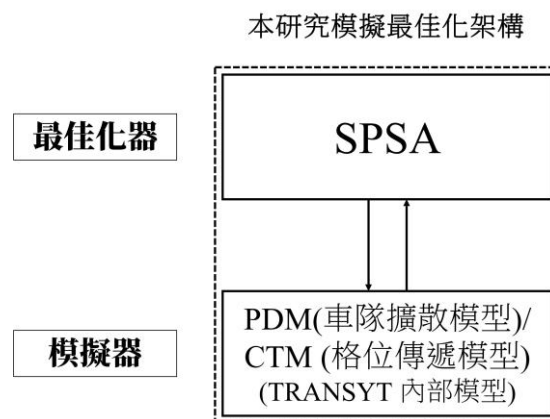


圖 1-3 本研究模擬最佳化之模型架構

1.2 研究目的

考量到最陡下降法等傳統基於梯度之方法在每一次迭代中所需的模擬次數會隨變數增多而變多，本研究希望使用一個能較快速且績效更佳的基於梯度估計演算法來最佳化路網號誌，且可以在變數增多時維持同樣的迭代速度。因 **SPSA** 方法可以每次迭代保持同樣的模擬次數，故本研究將使用此方法期望達到該目標。本研究的目的如下：

- 一．針對號誌時制最佳化問題，建立一套適用且具效率的 **SPSA** 演算法架構。
- 二．實際驗證 **SPSA** 在路網號誌最佳化的最適性，以確保演算法所得結果在實務上的可接受性。
- 三．與最陡下降法比較，探討 **SPSA** 在節省模擬次數與求解時間上的優勢。



1.3 章節安排

本研究在後續章節安排如下：

- 第二章「文獻回顧」：
從最陡下降法、隨機近似開始回顧，探討 **SPSA** 相關理論，同時了解號誌最佳化中使用的各種模擬器與最佳化器，說明 **SPSA** 於本研究之適用性與優勢，作為後續模型建立之基礎。
- 第三章「模型建立」：
說明 **SPSA** 的迭代過程，並建立一個用於本研究的 **SPSA** 演算法流程，考慮區別時差與綠燈長變數做迭代的情況下，如何交替迭代與迭代的停止條件；另外，若當變數值超過可行解時，將透過適當調整將其投影回可行域內；最後也將說明如何找到決策變數的初始值。
- 第四章「案例分析」：
藉由分析不同尺度的路網，分別由少至多的初始解做 **SPSA**，並設定其迭代步長參數，最後比較 **SPSA** 與其他標竿演算法的結果，並在現代圓環案例中分析 **SPSA** 的收斂性，以及在雙幹道路網中實證其實務可行性。
- 第五章「結論與建議」：
綜合研究成果，探討本研究方法之優勢與貢獻，並提出本研究之限制與未來可延伸的研究方向。

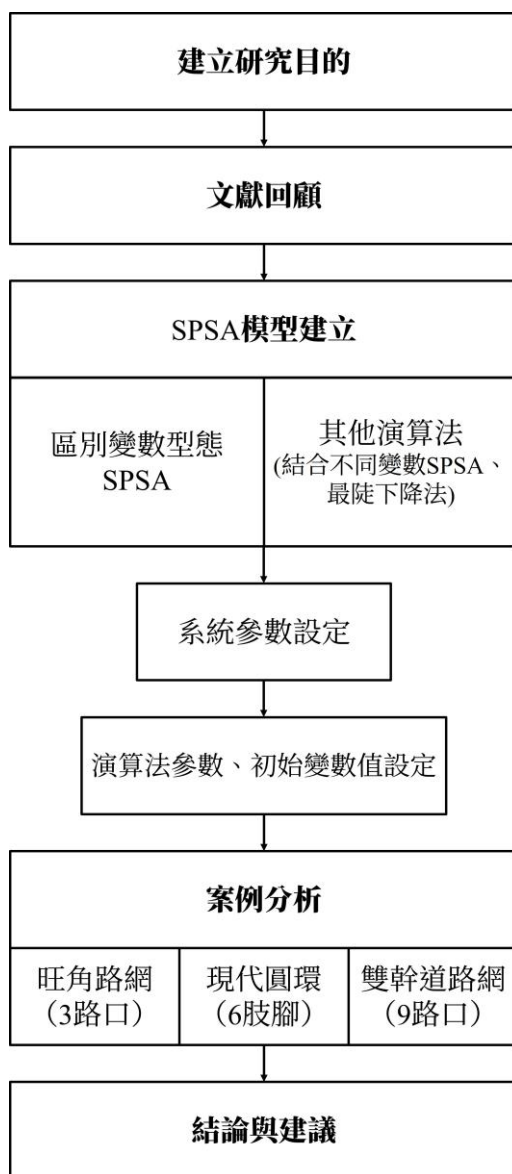


圖 1-4 研究流程圖

第二章 文獻回顧



在對 SPSA 方法進行回顧之前，須先回顧基於梯度的傳統最陡下降法 (steepest descent) 並接續回顧隨機近似 (stochastic approximation)。有了前兩者的最佳化問題求解概念後，再進一步回顧 SPSA，並探討其應用於號誌最佳化之可行性。

2.1 最陡下降法 (Steepest Descent)

最陡下降法的架構主要是利用梯度的資訊更新迭代以找到最佳解。欲求一可微函數 $f(\mathbf{x})$ 之最小值，透過尋找一個函數下降方向，並乘上一個步長決定其下降長度，最終可以得到一個區域 (local) 或全域 (global) 的最佳解 (即最小值)，如式(2.1)。

$$\mathbf{x}^{k+1} = \mathbf{x}^k - a_k \cdot \mathbf{g}_k \quad (2.1)$$

其中 $\mathbf{x}^k$ 為第 k 次迭代產生的變數向量 ($\mathbf{x}^0$ 為初始變數)， a_k 為第 k 次迭代之步長 (step size)， $\mathbf{g}_k$ 為函數 $f_k(\mathbf{x}^k)$ 的梯度，其定義如式(2.2)所示。

$$\mathbf{g}_k \equiv \nabla f_k(\mathbf{x}^k) = \left[\frac{\partial f_k(x_1^k)}{\partial x_1^k}, \frac{\partial f_k(x_2^k)}{\partial x_2^k}, \dots \right]^T \quad (2.2)$$

透過不斷迭代的過程可以收斂到一組最佳的 $\mathbf{x}_k$ 。讓函數朝下降最大的方向 (負梯度方向) 更新，也會幫助找到一個最佳解。

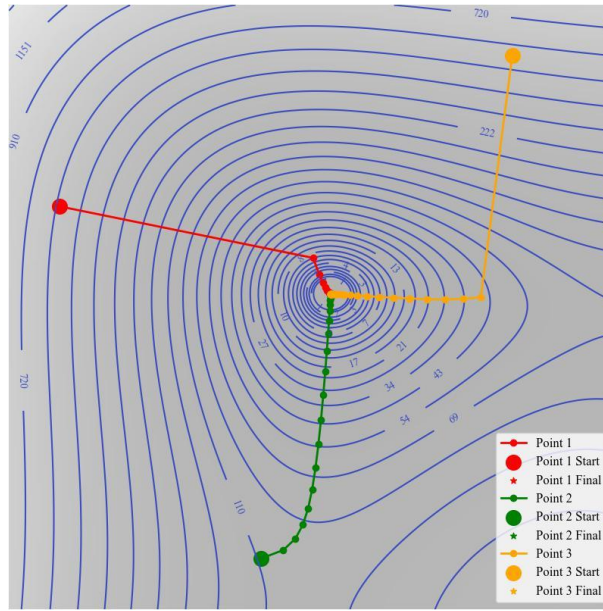


圖 2-1 最陡下降法示意圖（本研究繪製⁴，函數參考 Himmelblau（1972））

圖 2-1 隨機取三個初始解進行最陡下降，其中每一個圓點是一次迭代的結果，步長 a_k 設為定值，計算每次迭代中函數的負梯度作為方向，可以看出每一次的迭代都是朝等高線垂直方向下降，也是變化最大的方向。追溯其發展，此方法最早或由 Cauchy（1847）所提出，其隱含使用負梯度方向來讓函數值下降的概念，也是最早有梯度下降架構的文獻。最陡下降在現代的收斂理論可以追溯至 Zoutendijk（1970），在滿足 Armijo（1966）或 Wolfe（1969）步長、梯度 Lipschitz 連續且目標函數有下界下，可保證收斂到駐點（stationary point）。然而在實務上，真實的梯度 $\nabla f(\mathbf{x})$ 可能無從得知或計算成本高昂，使得最陡下降的可用性受限。為解決此問題，可以尋找可近似梯度之函數 $\hat{\mathbf{g}}(\mathbf{x}) \approx \nabla f(\mathbf{x})$ （如前向差分（forward difference gradient approximation, FDGRAD, Dennis & Schnabel, 1996）方法，如式(2.3)所示）。

$$\hat{g}_n^k(\mathbf{x}^k) = \frac{D(\mathbf{x}^k + h_n^k \mathbf{e}_n) - D(\mathbf{x}^k)}{h_n^k} \quad (2.3)$$

其中 $\hat{g}_n^k$ 為梯度向量 $\hat{\mathbf{g}}_k(\mathbf{x}^k)$ 的第 n 個分量值，其中上方標有帽型（hat）符號「^」

⁴ 使用 Python 的 Matplotlib 函式庫繪製。

者為估計值，即 $\hat{\mathbf{g}}_k(\mathbf{x}^k) = [\hat{g}_1^k(\mathbf{x}^k) \ \hat{g}_2^k(\mathbf{x}^k) \ \dots]^T$ ； h_n^k 為一個小變化量； $\mathbf{e}_n$ 為第 n 個單位向量，即除了第 n 個分量為 1 之外其餘分量皆為 0。



2.2 隨機近似 (Stochastic Approximation)

前述之最陡下降法假設函數可準確獲知，然而系統觀測時會有不準確的狀況發生，即真實訊號 $f(\mathbf{x})$ 與雜訊 (noise) ε 混為觀測值 $h(\mathbf{x}) = f(\mathbf{x}) + \varepsilon$ 。隨機近似將在函數不準確的限制下發展解方。Robbins 與 Monro (1951) 提出在樣本含有隨機雜訊的情況下，使用迭代方法逼近解，也提出數列 $\mathbf{a}_k$ (即迭代步長) 的收斂性質，也為使用迭代方法處理隨機問題之濫觴。Kiefer 與 Wolfowitz (1952) 結合此方法與有限差分方法估計導數，不須知道函數真正的梯度，其中每次的迭代需要兩倍維度的函數取值進而得到估計的梯度 (即在 N 維問題下每次迭代須 $2N$ 次取值)。然而，此方法在效率上隨著變數的維度變大也會需要更多的估計才能夠得到最佳解。Spall (1987) 提出的 SPSA 方法有效地將這個問題排除。SPSA 方法給定一個隨機擾動，同時決定變數各個維度的方向，如此一來在估計梯度上，因為方向已經被隨機擾動所決定，每一次迭代不管變數維度，只需要兩次的目標函數取值即可得到估計的梯度。

Spall (1992) 提到在 SPSA 中，數列 $\mathbf{a}_k$ 與 $\mathbf{c}_k$ 有個別的收斂性質，且使用隨機擾動來估計梯度。若該隨機擾動的值為 +1 或 -1，每次取到 +1 或 -1 的機率各為 50% 的情形下，且隨機擾動向量內各元素彼此獨立，當擾動幅度趨近於 0 時，可以證明得到估計梯度值是無偏的 (unbiased)。Spall (1998) 則較完整地提供了數列 $\mathbf{a}_k$ 與 $\mathbf{c}_k$ 的公式，另外也提供了 $\mathbf{a}_k$ 與 $\mathbf{c}_k$ 公式中部分常數計算或取值的經驗技巧。Zhu 與 Spall (2002) 又將 SPSA 擴展到二階，透過近似 Hessian 矩陣以改善收斂效率。

在 SPSA 方法與一系列微分方程理論中，迭代數列最後會收斂到微分方程 $\frac{d\mathbf{x}(t)}{dt} = -\mathbf{g}(\mathbf{x}(t))$ 的一個漸進穩定解 (asymptotically stable solution)，其中 $\mathbf{x}$ 為變數， $\mathbf{g}$ 為目標函數之梯度 (Kushner & Yin, 1997; Spall, 1992)。當目標函數為凸 (convex)



時，該解就會代表全域(global)最佳解，而函數非凸時，就只能代表一個局部(local)最佳解。由此可知，SPSA 方法不須保證函數的凸性，且能確定收斂。在固定演算法參數 (a_k 、 c_k) 的情況下，Gerencsér 等人 (1999) 驗證其迭代穩定與有界性 (boundedness)，另外對整數變數也能維持可行性並穩定、有界。

2.3 號誌最佳化

一個道路路網中往往會有許多號誌，且同一路口內互相衝突之轉向不能同時為綠燈，且路口間之號誌需互相協調方能保持順暢。這樣的情境可視為一個最佳化問題，目標式可為最小化負效用，例如延滯、停等次數、旅行時間、廢氣、燃油等；決策變數則可為各轉向之綠燈長與協調各路口之時差等號誌變數。因此，號誌最佳化之方法論不出於作業研究領域之求解方法，為須針對運輸系統的問題本質深入研究並於模型內考慮周到。當求解此類最佳化問題時，可能因為梯度不明或者無法得知目標函數的明確解析式，需要使用模擬方法來得到目標函數值，這種方法即為模擬最佳化 (simulation optimization) (Fu, 2002)。在此方法中，需要模擬器 (simulator) 與最佳化器 (optimizer) 不斷互動以找到最佳解。由於模擬器可用來呈現交通流動並提供目標函數值，因此在最佳化器之前，有必要先回顧模擬器的發展。

其中一個經典案例是 Robertson (1969) 提出的 TRANSYT 系統，其核心為車隊擴散模型 (Platoon Dispersion Model, PDM)，用以模擬上游路口號誌紅燈期間形成的車隊，在綠燈時沿下游路段擴散，並以到達流率的形式陸續抵達下游路口。PDM 以此計算路網的績效指標，包含路網總延滯 (Delay) 與停等次數 (Number of Stops)。然而，PDM 模擬僅著重向下游流動之流率樣態，無法有效考慮壅塞時車輛隊列回堵或隊列在路段上的傳遞等問題。Daganzo (1994;1995) 提出的格位傳遞模型 (Cell Transmission Model, CTM) 補足了這些限制，其以更高效的方式模擬車輛隊列在路段間的演變過程。此外，Ben-Akiva 等人 (1998) 提出了 DynaMIT (Dynamic Network Assignment for the Management of Information to Travelers)，結合需求模擬與供給模擬，除了交通流模擬以外同時考慮起訖與路徑選擇等因素。另

外，若使用微觀模擬（如 VISSIM 等）可更精細地逐車模擬，考慮每輛車的行為以及隨機性，但所需時間成本也會增加。



使用模擬器得到績效指標後，必須使用最佳化器來搜尋更好的解。TRANSYT 使用爬山法（Hill Climbing）來解最小化績效指標的問題（Robertson, 1969）。因其僅需比較附近點的績效是否進步，計算簡單，也可以快速找到一個可接受的解。但也因此，爬山法容易陷入局部最小值（local minimum），在後續的 TRANSYT-7F（McTrans, 2008）中也將基因演算法（Genetic Algorithm, GA）加入最佳器中。相關研究如 Lo 等人（2001）使用了 GA 結合 CTM 的數學模式來做最佳化，並發展出一個動態路口號誌控制最佳化（DISCO）架構。Stevanovic 等人（2007）則提出 VISGAOST，結合 VISSIM 與 GA 以最佳化綠燈長、週期時間、時差與時相順序等號誌變數。然而，Hill Climbing 與 GA 等啟發式演算法在數學上難以解釋其收斂性，若使用一些基於數學概念的最佳化方法求解號誌最佳化問題時（例如最陡下降法、SPSA 等），可以在數學理論基礎上給予該結果支持，當收斂條件完善時，也能找到區域甚至全域最佳解。此外，Stevanovic 等人（2007）亦指出，使用 VISSIM 與 GA 方法需要多台電腦同時計算，並且需要花上約 10 天完成一個情境（該文中考慮三種情境），非常耗時。因此，除了嘗試啟發式演算法之外，具有數學理論基礎的演算法也逐漸被採用，以期在求解效率與解的品質上能有所進步。

此類方法使用梯度資訊作為函數下降方向，例如 Allsop (1971) 證明了 Webster (1958) 提出的延滯公式在特定假設下會是一個凸函數，即可以找到一個全域最佳解，並進一步利用該函數的導數配合限制式決定每次迭代的下降方向。Chiou (1999) 亦使用梯度方法進行最佳化，其梯度自 Wong (1995) 推導而來，再對步長做線性搜尋（line search）。其中 Wong (1995) 對 TRANSYT 績效指標的導數近似，延續該成果，Wong (1996) 建立了一個整數化的搜尋過程，利用這些導數判斷時差與綠燈長較佳的方向，並做線性搜尋找到步長。



除了以上一階方法外，若能進一步獲得函數的曲率資訊，則可使下降方向更為精確。如 Lan 與 Chang (2016) 提出的方法基於一個汽機車混合流數學模型，調整綠燈時間與週期使用邏輯推導，在綠燈時間部分則是使用 BFGS 方法。BFGS 方法的二階方法相對梯度下降的一階方法，另外還可獲得二階（曲率）資訊。二階方法需要使用 Hessian Matrix，即二階偏導數資訊。BFGS 方法的特點在 Hessian Matrix 使用估計方式取得。

2.4 文獻小結

綜合以上所述，基於數學理論的號誌最佳化方法在理解問題結構與收斂特性方面具有明確優勢，模擬器的發展也讓模擬最佳化方法漸趨成熟。另一方面，隨機近似克服了梯度難以精確取得的問題，而 SPSA 更將梯度估計的成本有效降低。值得注意的是，SPSA 最初的設計目的在於緩解隨機性，使含雜訊的函數仍能穩定收斂，然而一些相關研究也顯示 SPSA 在號誌最佳化上的可行性。

Spall 與 Chin (1997) 將 SPSA 方法結合神經網路，用於訓練網路以學習不同交通情境下的號誌時制設計，然其對參數的描述相對模糊，後續研究難以完整重現；Chen 與 Hartono (2023) 則採用 SPSA 搭配 VISSIM 對時差 (offset) 做最佳化，但仍未涵蓋綠燈長等其他決策變數。這些成果雖肯定了 SPSA 的可行性，但在方法透明度與研究範疇上仍有不足。

相較於最陡下降法在 N 維問題下須 $N + 1$ 次模擬，KW 方法 (Kiefer-Wolfowitz 方法) 在 N 維下須 $2N$ 次模擬等，SPSA 無論維度高低僅需 2 次模擬即可完成梯度估計。延續以上脈絡，本研究將進一步延伸，期望將 SPSA 作為一種高效率的梯度近似方法，探討其在號誌最佳化中的決策變數一時差與綠燈長的應用。

第三章 模型建立



本研究希望將 **SPSA** 方法應用於路網號誌的最佳化中。透過車流模擬器與最佳化器的架構運作可以進行求解，整體模式架構如圖 3-1 所示。本章將介紹研究方法，於 3.1 小節介紹本研究模型架構；3.2 小節介紹變數與參數；3.3 小節將介紹目標式與限制式；3.4 小節為最佳化過程；3.5 小節則是介紹最佳解的搜尋。本研究將使用 **SPSA** 方法進行迭代，先將路網建立在 **TRANSYT** 車流模擬器中，輸入所需參數後，先將決策變數初始值代入車流模擬器進行模擬，之後到達迭代標準後找到最佳的號誌時制與總延滯，研究流程如圖 3-2 所示。

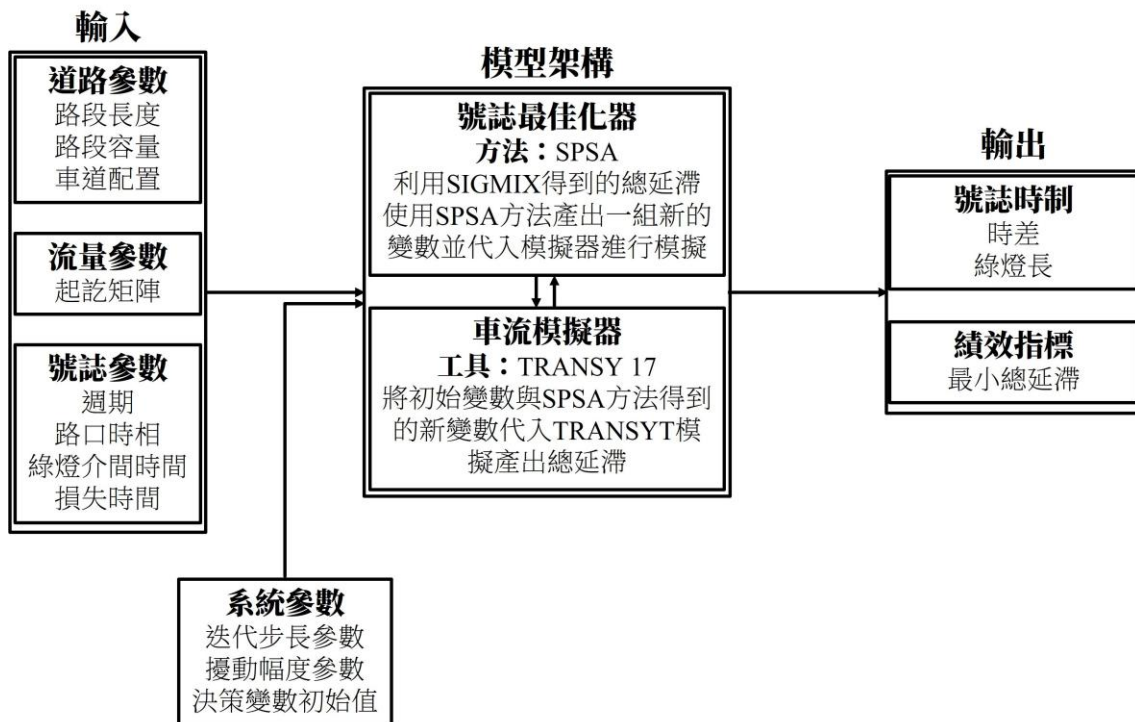


圖 3-1 模型架構圖

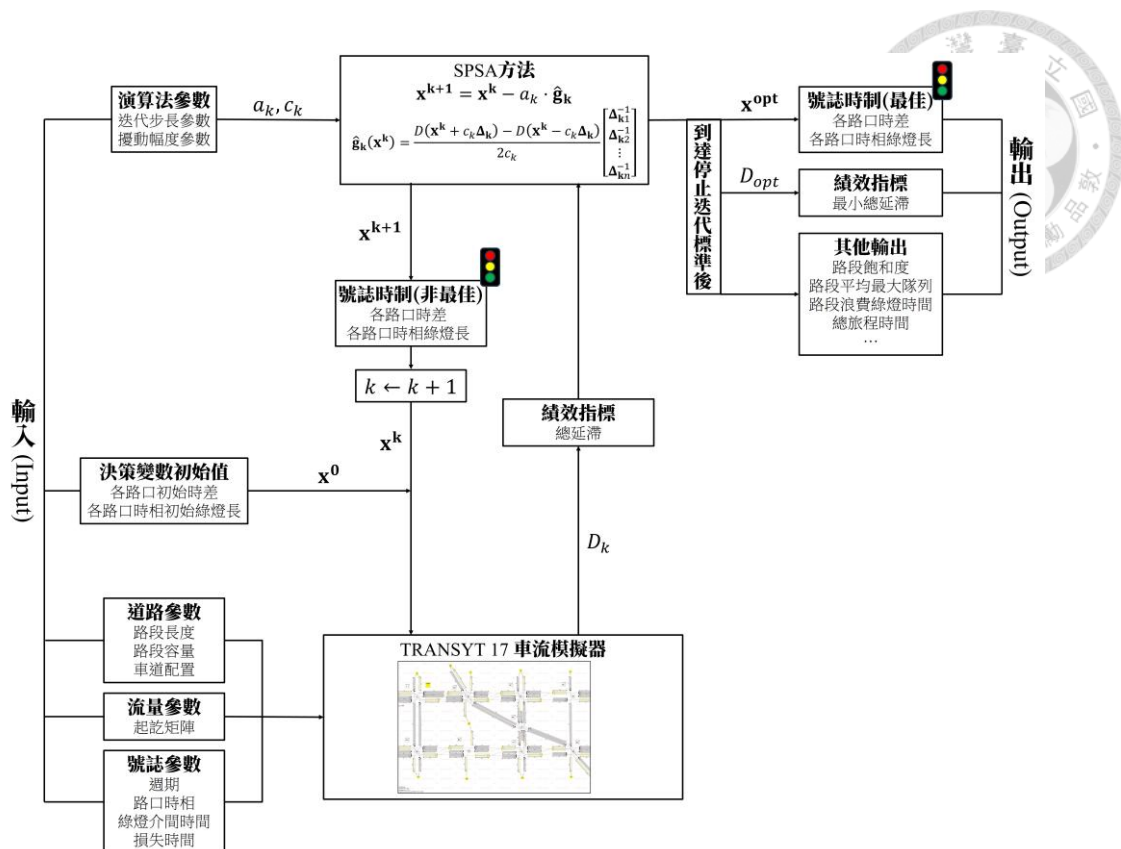


圖 3-2 總流程圖

3.1 模型架構

本研究欲使用基於模擬的最佳化方法來最佳化路網號誌。選用 SPSA 方法求解配合 TRANSYT 車流模擬。如圖 3-2，道路參數、流量參數與號誌參數等屬於系統參數。道路參數主要是在建立路網時決定，有路段⁵長度、路段容量（capacity）、車道配置（lane configuration）⁶等。流量參數則為路網端點間的起訖矩陣（O-D Matrix），藉由該矩陣可得知兩兩端點間的流量流動，進一步可將流量分配到路網上。號誌參數則是根據每個路網中的號誌系統做設定，如號誌週期、路口時相、綠燈介間時間與損失時間等固定參數。另外非系統參數的則有變數初始值與一些演

⁵ 路段（Link）指一段單向、連續的道路區間，位於兩個相鄰交流道或路口之間（FHWA, 2015）。

⁶ 如車道數、各車道之轉向通行權等，例如 3 個車道含一個左轉專用車道與最外側直行左轉共用等。



算法的參數（詳見 3.5 小節）。藉由最佳化器與模擬器的不斷迭代，最終我們可以得到一個最好的績效 (performance)，在本研究中定義為最小總延滯 (Total Delay)，迭代過程簡圖如圖 3-3 所示。

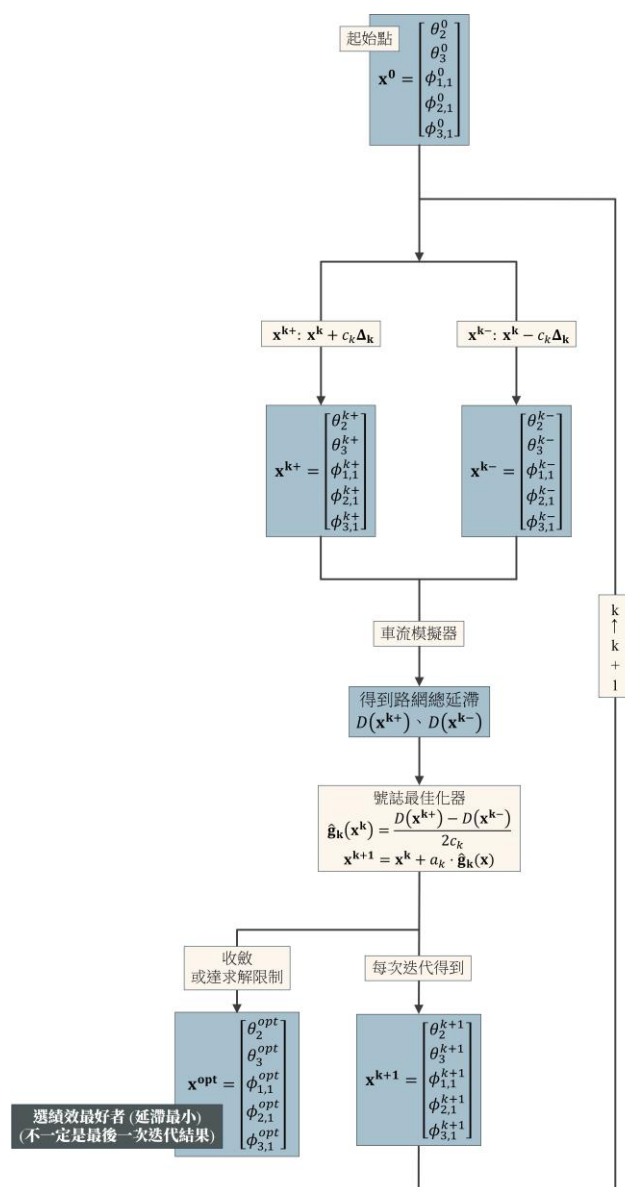


圖 3-3 本研究迭代過程簡圖（以 5 個變數為例）

從圖 3-3 可得知，在完成一次迭代後，並不會直接代回車流模擬器進行新的模擬，而是先計算 $\mathbf{x}^{k+} = \mathbf{x}^k + c_k \Delta_k$ 與 $\mathbf{x}^{k-} = \mathbf{x}^k - c_k \Delta_k$ 兩組變數再代回模擬器模擬。這是因為演算法需要，因若直接將迭代完成的 $\mathbf{x}^k$ 代回模擬，後續還是要計算 $\mathbf{x}^{k+}$

與 $\mathbf{x}^{k-}$ 的模擬結果，有關相關變數與方法在 3.5 小節會詳細介紹。這也代表 $\mathbf{x}^k$ 的模擬結果並非必要過程，若省略此步驟可以節省總模擬時間。



3.2 變數與參數

在號誌系統中，流動 (movement) 指的是前往路口各個車隊，其中以行車方向 (direction)、車道配置 (lane allocation) 與是否給予路權 (right-of-way provision) 做區分。而當一個或多個流動 (movement) 獲得路權時，這一段時間稱為時相 (phase) (Akçelik, 1981)。

在時相被給予路權時，號誌會顯示綠燈。該綠燈之持續時間即為 (顯示) 綠燈長 ((displayed) green time)。在兩個綠燈時相之間會有一段黃燈與全紅 (即路口各時相皆為紅燈) 時間，這段時間稱為後時相的綠燈介間時間 (intergreen time)。

參考圖 3-4，假設一個路口共有 3 個時相，若最上方的點 F_A 是時相 C 轉換為時相 A 的點 (F_A 代表 A 的時相轉換時間 (phase change time))，該點設為基準點 0 秒，則順時針跑完一圈回到該基準點的時間即為週期，如式 (3.1) 所示。因此，週期即為所有的綠燈介間時間與綠燈長相加得來。

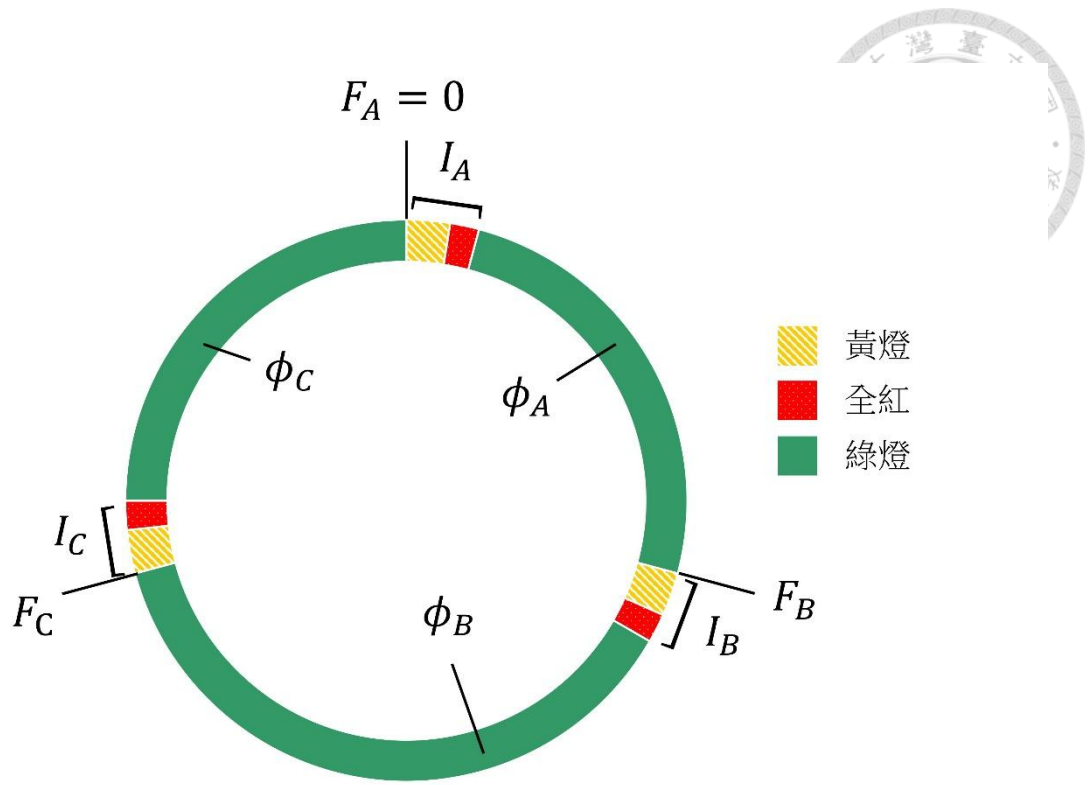


圖 3-4 號誌週期圖

$$\text{重回基準點時間} = I_A + \phi_A + I_B + \phi_B + I_C + \phi_C$$

$$= \sum_{i=A}^C (I_i + \phi_i) = \text{週期} \quad (3.1)$$

本研究所用到的變數主要為時差 θ_i 與綠燈長 $\phi_{i,j}$ 。綠燈長 $\phi_{i,j}$ 代表的是在第 i 個路口的第 j 個時相的綠燈長。假設有三個皆為三時相的連續路口，路口配置皆如圖 3-5 所示。其三個時相的時相順序如圖 3-6 所示，其中時相圖中的虛線則因與其他流動有所衝突，必須等待其衝突流動的車輛通過才能進行轉向。

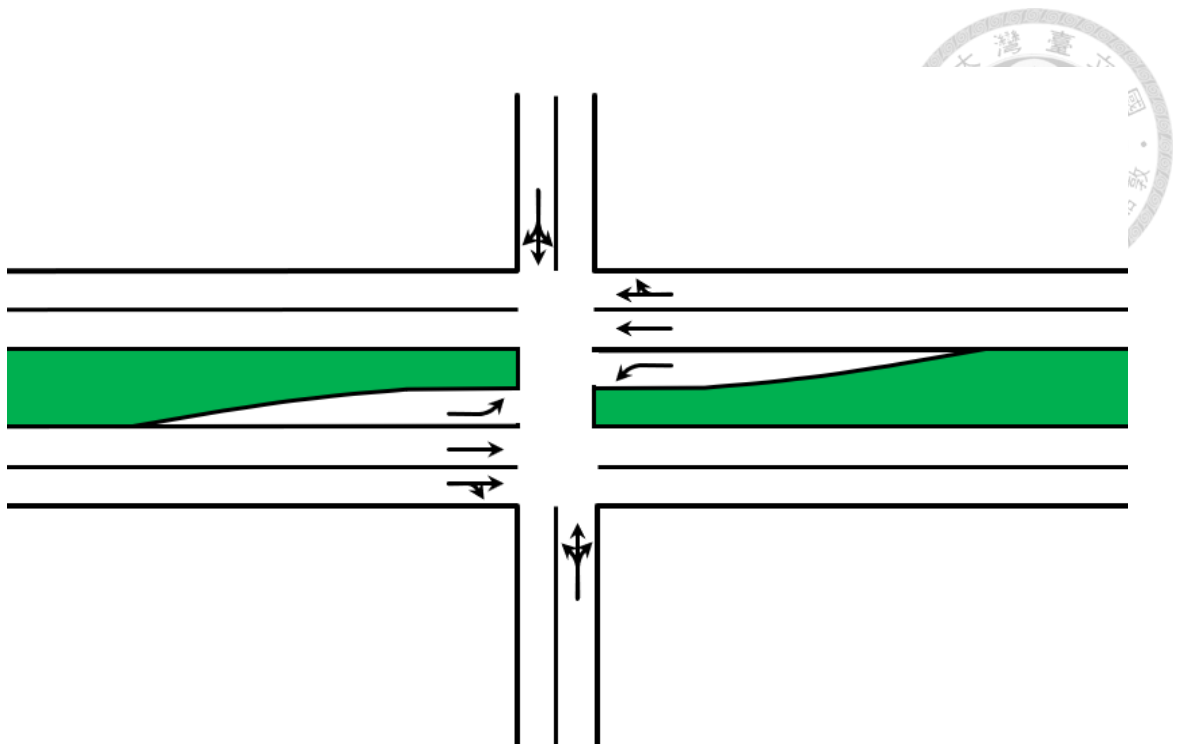


圖 3-5 一個三時相號誌路口示意圖（標線等未按實際情況繪製）

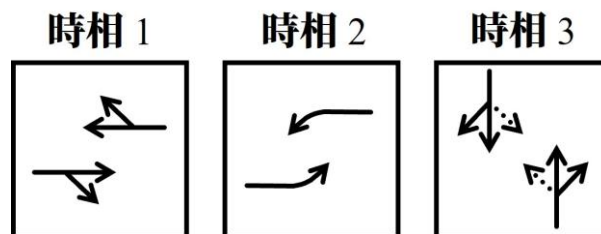


圖 3-6 時相圖（圖 3-5 路口）

在圖 3-7 中，時差 θ_i ，即第 i 個路口第一個時相的時差，可以被得到。在考慮號誌連鎖（coordinated）的情形下，為了能讓整個路網的總延滯降到最低，會有不同路網有不同時差的情況。為表示此種情況，我們可以一個主時鐘（master clock）為基準來得到每個路口第一個時相的時差，導致幹道或路網上的車流更為順暢。另外，將每個路口的號誌週期長度 C_i 設為相同或相差一個倍數（例如 60 秒與 120 秒）以讓號誌連鎖能夠持續，圖 3-7 將週期長度設為相同的 C 。

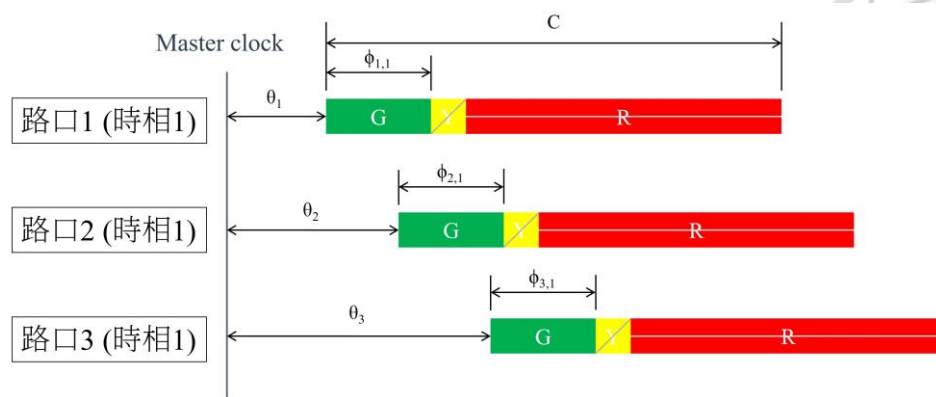


圖 3-7 時差與綠燈長（各路口之第 1 個時相）

為便於後續討論，以下將本研究會使用到的集合、變數與參數綜整於表 3-1、表 3-2 與表 3-3。

表 3-1 集合對照表

集合	說明
$\mathbb{I}$	所有路口的集合
$\mathbb{J}_i$	路口 i 的所有時相集合
$\mathbb{K}$	所有「迭代次數 - 1」的集合 ⁷
$\mathbb{L}_{i,j}$	路口 i 的第 j 個時相的所有車道集合

表 3-2 變數對照表

變數	說明
$\phi_{i,j}$	路口 i 的第 j 個時相的顯示綠燈長
θ_i	路口 i 的時差
θ_i^0	路口 i 的時差初始值（未經餘數處理）
$\phi_{i,j}^0$	路口 i 的第 j 個時相的顯示綠燈長初始值
θ_i^k	路口 i 在第 k 次迭代得到的時差（未經餘數處理）
$\tilde{\theta}_i^k$	路口 i 在第 k 次迭代得到的時差（經餘數處理）
$\phi_{i,j}^k$	路口 i 的第 j 個時相在第 k 次迭代得到的顯示綠燈長
$\mathbf{x}^0$	迭代前的初始變數向量（包含 $\phi_{i,j}$ 與 θ_i ）
$\mathbf{x}^k$	第 k 次迭代得到的變數向量
$\mathbf{x}^{k+}$	第 $k + 1$ 次迭代正擾動之變數向量估計值
$\mathbf{x}^{k-}$	第 $k + 1$ 次迭代負擾動之變數向量估計值
$\hat{\mathbf{g}}_k$	第 $k + 1$ 次迭代的梯度估計值
D_k^+	第 $k + 1$ 次迭代後變數正擾動產生的總延滯
D_k^-	第 $k + 1$ 次迭代後變數負擾動產生的總延滯
$D(\cdot)$	總延滯函數

⁷ 真正的迭代次數從 1 開始，但這裡的 $\mathbb{K}$ 從 0 開始，故標示為「迭代次數 - 1」。

表 3-3 參數對照表

參數	說明
C_i	路口 i 的週期長度
$(\phi_{min})_i$	路口 i 的最小綠燈長
Δ_k	第 $k + 1$ 次迭代的隨機擾動向量
Δ_{k,θ_i}	路口 i 的時差在第 $k + 1$ 次迭代的隨機擾動
$\Delta_{k,\phi_{i,j}}$	路口 i 的第 j 個時相在第 $k + 1$ 次迭代的隨機擾動
a_k	第 $k + 1$ 次迭代的迭代步長
c_k	第 $k + 1$ 次迭代的擾動幅度
a 、 A 、 α	迭代步長常數
c 、 γ	擾動幅度常數
$I_{i,j}$	路口 i 的第 j 個時相的綠燈介間時間
$EG_{i,j}$	路口 i 的第 j 個時相的初始有效綠燈長
$l_{i,j}$	路口 i 的第 j 個時相的損失時間
$y_{i,j}$	路口 i 的第 j 個時相的流量比
$q_{i,j,l}$	路口 i 的第 j 個時相的第 l 個車道的抵達流率
$s_{i,j,l}$	路口 i 的第 j 個時相的第 l 個車道的飽和流率
Y_i	路口 i 的所有流量比總和

3.3 目標式與限制式

本研究欲使用 **SPSA** 方法來最佳化路網的綠燈長與時差，因此我們的目標是將總延滯降到最低。在本研究中，我們將使用 **SPSA** 方法來降低總延滯，在模擬出的總延滯最小的情況下，會得到一組變數，即本研究欲得到的最佳綠燈長與時差，目標式如式 (3.2) 所示。

$$\min_{\theta, \phi} D = \min_{\mathbf{x}} D \quad (3.2)$$

其中 D 為路網總延滯， θ 為時差， ϕ 為綠燈長， $\mathbf{x} = [\theta^T, \phi^T]^T$ 。在本研究中，路網總延滯 D 為 $\mathbf{x}$ 的函數。對本研究來說，總延滯的值代入車流模擬器進行估計，而非一解析式，為基於模擬之最佳化方法的常態。當給定一組 $\mathbf{x}$ ，可以使用 **TRANSYT** 模擬器得到一個總延滯值 $D = f(\mathbf{x})$ 。

限制式部分本研究僅列出與號誌時制較相關的限制。第一個限制式為時差的限制。時差 θ_i 必須控制在 0 與週期長度（定值）之間。由於本研究假設各路口週期皆相同，各路口的估計出的時差如果超過週期，代表的就會是超過週期的那部分，例如當週期為 120 秒，估計出的時差為 145 秒時，我們可以說時差是 25 秒，因為整個號誌系統有其週期性。亦即，時差會等於時差除以週期的餘數。如式 (3.3)、式 (3.4) 所示。

$$\theta_i \in [0, C_i), \quad \forall i \in \mathbb{I} \setminus \{1\} \quad (3.3)$$

$$\tilde{\theta}_i^k = \hat{\theta}_i^k \bmod C_i, \quad \forall i \in \mathbb{I}, \forall k \in \mathbb{K} \quad (3.4)$$

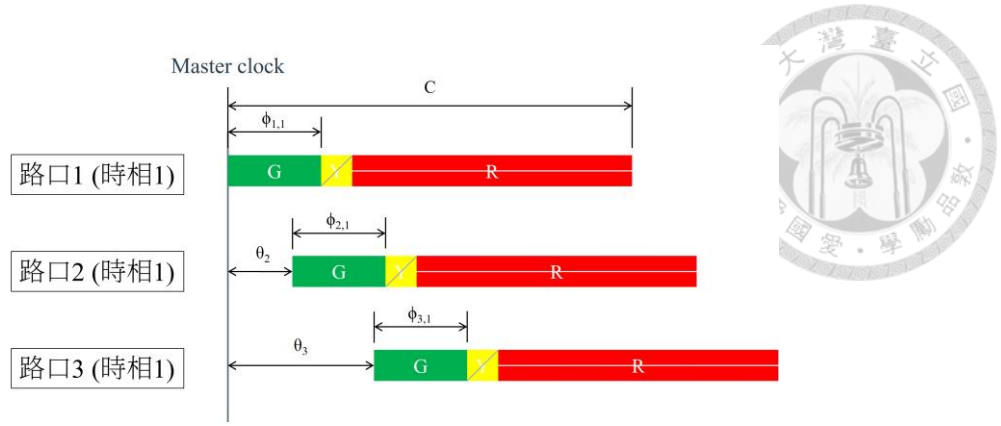


圖 3-8 時差與綠燈長（各路口之第 1 個時相， $\theta_1 = 0$ ）

考慮圖 3-7 情況，若將主時鐘的基準點移到路口 1 時相 1 的時差，則路口 1 時相 1 的時差可固定為 0，如圖 3-8 所示，將可以減少時差變數的數量。也就是說，若路網中有 I 個路口，我們僅需要定義路口 2 至路口 I 共 $I-1$ 個路口的時差即可，如式 (3.5) 所示。

$$\theta_1^k \triangleq 0, \quad \forall k \in \mathbb{K} \quad (3.5)$$

有關綠燈長的限制式，首先是綠燈長必須不小於最小綠燈長 ϕ_{min} 且不超過週期 C_i 。限制式如式 (3.6) 與式 (3.7) 所示。

$$\phi_{i,j} \geq \phi_{i,j}^{min}, \quad \forall i \in \mathbb{I}, \forall j \in \mathbb{J}_i \quad (3.6)$$

$$\phi_{i,j} \leq C_i, \quad \forall i \in \mathbb{I}, \forall j \in \mathbb{J}_i \quad (3.7)$$

由於在本研究中週期長度 C_i 與綠燈介間時間 $I_{i,j}$ 皆屬於參數，若一個路口有 J 個時相，則僅需 $J-1$ 個時相的綠燈長即可得知第 J 個時相的綠燈長，如式 (3.8) 所示。

$$\sum_{j \in \mathbb{J}_i} \phi_{i,j} + \sum_{j \in \mathbb{J}_i} I_{i,j} = C_i, \quad \forall i \in \mathbb{I} \quad (3.8)$$

3.4 最佳化問題

基於以上敘述，可以定義出本研究的最佳化問題如下：



$$\begin{aligned} & \min D \\ & D = f(\mathbf{x}) \\ & s. t. \\ & \text{式(3.3)~ (3.8)} \end{aligned}$$

其中 $D = f(\mathbf{x})$ 單純表達 D 為 $\mathbf{x}$ 的函數，其他部分 $D = f(\mathbf{x})$ 本研究直接以 $D(\mathbf{x})$ 來表示。

3.5 最佳解搜尋

3.5.1 SPSA 方法

同其他梯度方法，由負梯度方向與步長來得到下一次迭代的變數更新值 (Spall 1987 ; 1992 ; 1998)，如式 (3.9) 所示。

$$\mathbf{x}^{k+1} = \mathbf{x}^k - a_k \cdot \hat{\mathbf{g}}_k(\mathbf{x}^k) \quad (3.9)$$

其中由於梯度 $\hat{\mathbf{g}}_k$ 為估計值，以置於變數上方的帽型 (hat) 符號 $\hat{\cdot}$ 表示估計。

由於本研究之總延滯 $D = f(\mathbf{x})$ 使用模擬取得，在梯度解析式未知的情況下，我們需要使用一些方法來估計梯度。在 SPSA 方法中，提供了一個估計梯度的方式，首先引入了一個隨機擾動向量 Δ_k 。其中該向量元素與變數數量相同，此擾動的值常取+1 或-1，而每次取擾動值時取到+1 或-1 的機率各為 50%。梯度估計公式如式 (3.10) 所示。

$$\hat{\mathbf{g}}_k(\mathbf{x}^k) = \frac{D(\mathbf{x}^k + c_k \Delta_k) - D(\mathbf{x}^k - c_k \Delta_k)}{2c_k} \begin{bmatrix} \Delta_{k,\theta_1}^{-1} \\ \vdots \\ \Delta_{k,\theta_i}^{-1} \\ \Delta_{k,\phi_{1,1}} \\ \vdots \\ \Delta_{k,\phi_{i,j}} \end{bmatrix} \quad (3.10)$$

其中 c_k 代表擾動幅度 (perturbation size)， a_k 與 c_k 等參數會在之後的 3.1.3 節進行說明，基本上這兩個參數僅與迭代次數有關，隨著迭代次數越多，這兩個參數會越

來越小。



在式 (3.10) 中，本研究稱公式中 $D(\cdot)$ 前項稱為正擾動（標示為 D_k^+ ），即在原本的變數加上一個擾動後運行模擬，後者稱為負擾動（標示為 D_k^- ），即將變化減去一個擾動後運行該變數模擬結果。⁸

從式 (3.9) 可以看出，我們使用總延滯加上擾動與總延滯減去擾動進行相減，兩者的差再除以 2 倍 c_k 來當作梯度的估計值。由於 $\frac{D(\mathbf{x}_k + c_k \Delta_k) - D(\mathbf{x}_k - c_k \Delta_k)}{2c_k}$ 為一個純量 (scalar)，且 $[\Delta_{k,\theta_1}^{-1} \cdots \Delta_{k,\theta_i}^{-1} \Delta_{k,\phi_{1,1}} \cdots \Delta_{k,\phi_{i,j}}]^T$ 向量中僅可能有 +1 或 -1 兩值，此方法估計出的梯度向量的所有元素絕對值應皆相同。這也是此方法相較其他梯度法操作較容易的部分，僅需得到兩個 D (D_k^+ 與 D_k^-) 即可估計整個梯度向量，而其他梯度方法可能會需要針對每個變數分量得到個別的 $\frac{\partial D}{\partial x_i}$ 來進行梯度估計。

SPSA 方法在變數較多時會顯現出其優勢。若是使用 KW 方法等有限差分 (FDSA) 方法，可能需要在每次迭代中針對每個不同變數分量進行目標函數的取值 (Kiefer & Wolfowitz, 1952 ; Spall, 2003)，但在 SPSA 方法中，不管變數多寡，每次迭代僅需兩次取值（在本研究中即為 D_k^+ 與 D_k^- 的模擬）即可得到一個梯度估計。

若用更直觀的幾何觀念說明 SPSA 方法，假設有一個可微 (differentiable) 的凸函數 (convex function)，此函數會在其梯度的負方向下降最快。另外從式 (3.9) 中也可看出，當梯度越小時， $\|\mathbf{x}^{k+1} - \mathbf{x}^k\|$ （或者說變數變化量）也會變小。當梯度越大時，變數的更新量也會變得比較大。因為我們需要找到的 a_k 與 c_k 參數僅與

⁸ 因此，本研究並不需要 $D(\mathbf{x}^k)$ （該次迭代產生的 $\mathbf{x}^k$ 直接運行模擬）的結果，在取最佳解時也是取最小延滯對應到的該次 $D(\cdot)$ ，也就是某次正擾動或負擾動的結果。

迭代次數有關，且估計出來的梯度絕對值都相同，在 SPSA 方法中，我們能用方向與大小來做一個幾何解釋。當時差與綠燈長共有 n 個變數，進行第 $k+1$ 迭代時，選出的 Δ_k 中會有 n 個元素，這些元素代表著每一個變數移動的方向。這些方向會隨著不同的變數改變，但變化大小每個變數都會相同。用簡單的方式解釋，當找到的方向接近梯度方向時， D_k^+ 與 D_k^- 的差會變得比較大（因梯度概念上來講就是兩點間的斜率），進而讓變數變化量也變大，相反地，當找到的函數與梯度方向有很大差異時，正擾動與負擾動的差會比較小，也因此變數的變化量就會比較小。

舉一個簡單的範例，在一個三維空間中，一個函數 $f(x, y)$ 共有兩個變數，其等高線如圖 3-9 所示，在此範例中， Δ_k 的選取有不同，可以看到右圖的第一步朝向梯度方向（與等高線接近垂直），而左圖朝著與梯度不同方向前進，故右圖的第一步變化量也會比較大。另外紅色線則是使用梯度下降法進行迭代產生的結果。

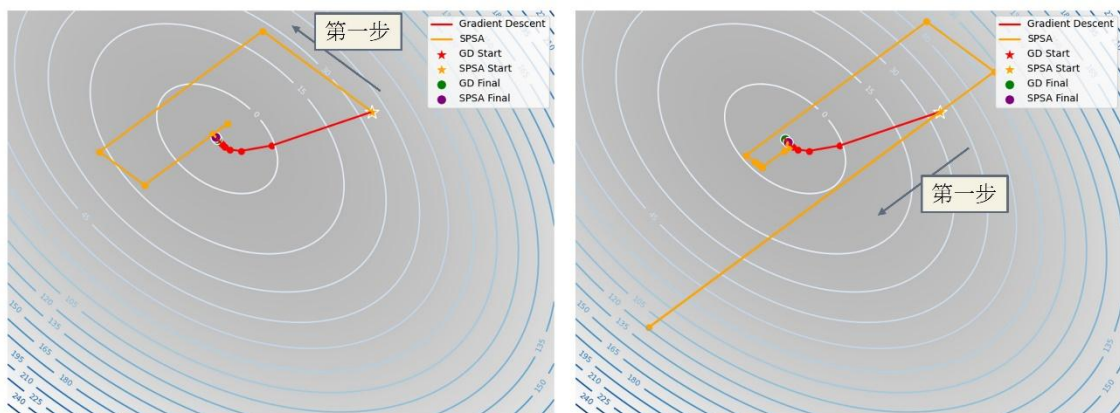


圖 3-9 SPSA 範例圖（初始點相同，不同隨機擾動）（圖源：本研究繪製）

3.5.2 SPSA 方法中的參數

SPSA 演算法中，在選取迭代步長 a_k 與擾動幅度 c_k 上也非常重要，並可能影響結果。Spall（1998）也提供了 a_k 與 c_k 的計算公式。其中迭代步長 a_k 的計算公式如式（3.11）。

$$a_k = \frac{a}{(A + k + 1)^\alpha}, \quad \forall k \in \mathbb{K} \quad (3.11)$$

另外在擾動幅度 c_k 的部分，同樣也有一個計算公式如式 (3.11)。

$$c_k = \frac{c}{(k + 1)^\gamma}, \quad \forall k \in \mathbb{K} \quad (3.12)$$

其中 a_k 與 c_k 須滿足

1. $a_k > 0, c_k > 0$
2. $a_k \rightarrow 0, c_k \rightarrow 0$
3. $\sum_{k=0}^{\infty} a_k = \infty$
4. $\sum_{k=0}^{\infty} \left(\frac{a_k}{c_k}\right)^2 < \infty$ 。

第一個條件能避免變數朝相反方向更新；第二個條件避免讓變數在後期的更新太大，幫助穩定收斂；第三個條件讓整個步長級數發散，避免步長加總在一個定值，而導致後續無法再更新；第四個條件考慮整體的雜訊 (noise) 影響。因雜訊變異數與 $\left(\frac{a_k}{c_k}\right)^2$ 成正比，需控制其收斂至有限值以避免過程中越來越偏離真實解。

在本研究中，決策變數時差與綠燈長兩者皆為整數。在這個情況下，變數並非連續，且最小單位為 1 (秒)。當考慮 $c_k \rightarrow 0$ 時，到後期 c_k 會變得越來越小，並會出現正擾動與負擾動之變數沒有差異的狀況，導致演算法無法進行為避免此狀況，因此，本研究固定 c_k 值為 1。

3.5.3 迭代終止條件

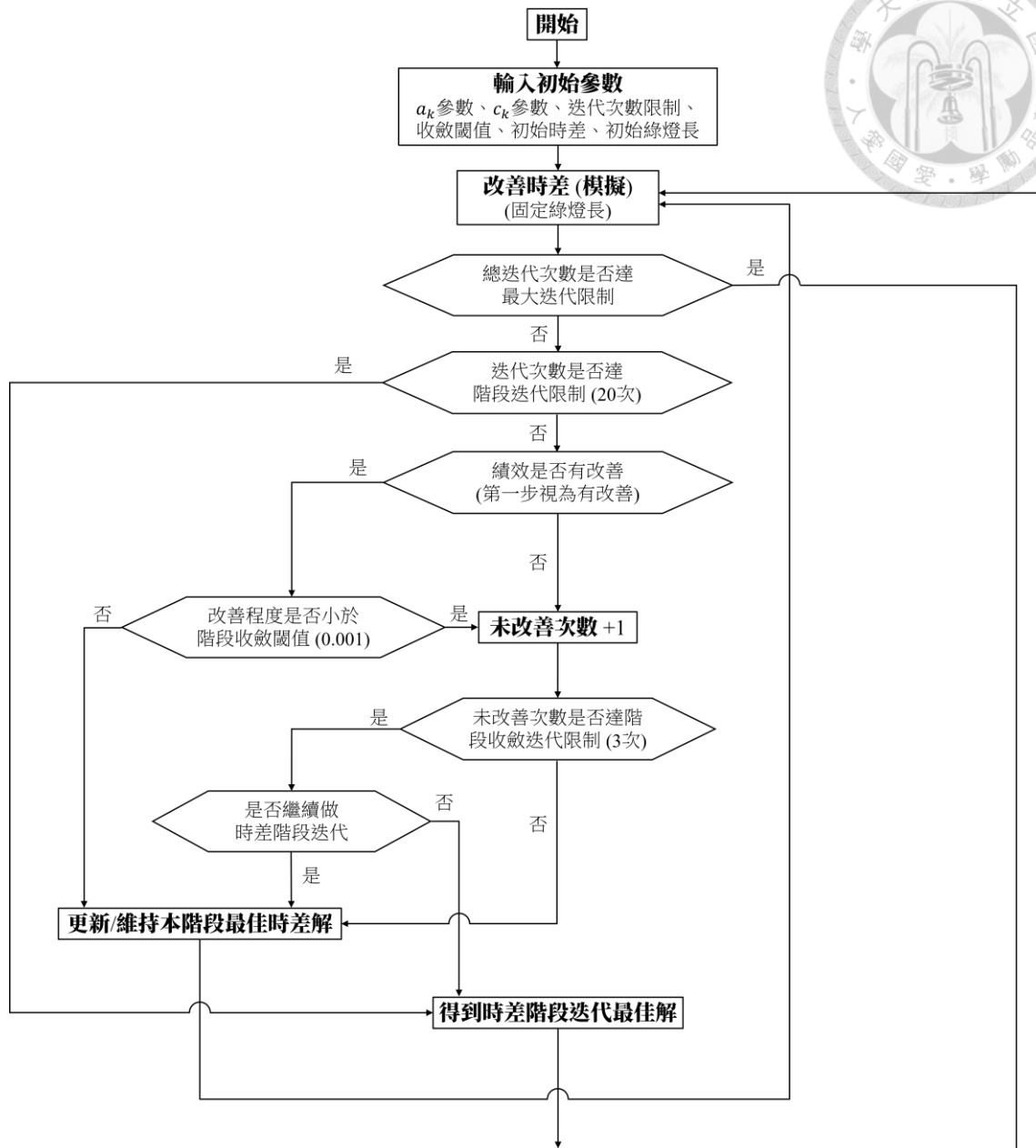
何時停止迭代在隨機估計法中一直是一個難題 (Kiefer & Wolfowitz, 1952; Spall, 2003)。本研究也不例外。在沒有一個完美的迭代終止條件的情況下，只能使用一個收斂閾值或迭代次數限制等方法來決定。本研究將會使用迭代次數限制的方式，限制其最大迭代次數，並在 3.5.4 小節之方法中輔以收斂閾值與連續無改善次數等條件，以完成演算法。

3.5.4 區別變數型態之 SPSA

本研究亦將另外提出以區別變數型態進行階段式最佳化的 SPSA，使用交互迭代的方式，先固定初始綠燈長進行數步時差最佳化，再固定時差進行數步綠燈長最佳化，其演算法流程圖如圖 3-10 所示。⁹



⁹ 參考 TRANSYT Hill Climbing 的作法，Robertson（1969）使用交互迭代的方式最佳化時差與綠燈長。



(續下頁)



從圖 3-10 可得知區別變數型態 SPSA 加入了迭代次數限制、收斂閾值限制，以及結合兩者的收斂閾值改善次數限制。

本演算法在進行階段切換時，會使用上一階段的最佳解進行下一階段的迭代。這也代表迭代過程因隨機擾動或參數關係導致逐漸偏離最佳點時，在到達停止標準後會使用原本還沒偏離的點。此法可避免多餘的迭代，以使績效足以被改善。然而，這也可能造成搜尋到的解逐漸往一個區域限縮。因此，演算法中採的是「上一階段」最佳解而非「所有階段」最佳解，儘量不讓一個次佳的解被埋沒，讓每個階段還有可能找到一個區域外的最佳解。

另外，演算法中也有一個彈性處理來避免停滯在區域最佳解的情況，即在迭代達到這些停止標準時，可選擇繼續重複原本的迭代階段或停止迭代前往下一步。迭代過程因隨機擾動或參數導致偏離最佳點時，因為可能落在不同區域導致績效變差，但也有可能在多迭代幾步之後又變得比原本更好，亦即找到另一個區域內的更好的最佳解。然而因過程中可能因為績效未改善到達標準而自動停止，為使演算法能夠更容易找到其他的解，在演算法中亦加入此彈性機制。此機制不一定要使用，且因目前演算法為半手動（手動輸入績效、模擬時間與是否繼續迭代），若未來可將此演算法自動化，還是要重新探討此機制的必要性。

3.5.5 限制式處理方法

當迭代值超出可行域時，必須想辦法將該值投影回可行域上。3.3 節的式(3.3)、(3.4)表示時差需要控制在週期時間內，但在超過該範圍的情況下，只要取餘數讓該值落在式(3.3)的範圍即可。另外因為在一個 P 時相路口中，僅有 $P-1$ 個決策變數，第 P 個時相由式(3.8)可直接算出，故限制式(3.8)必滿足。在這個小節中，主要探討的是式(3.6)與式(3.7)未被滿足的情況下調整變數且讓式(3.8)重新被滿足的情況。換句話說，當一或多個綠燈長太長或太短時，如何調整其他變數，使所有時相綠燈長



的和保持不變。例如在一個二時相的路口中，當變數自演算法得到的變數值大於或小於最小綠燈長，因為綠燈介間時間為固定，僅能靠調整另一個時相來滿足限制式。

舉一個簡單的例子說明，假設二時相一路口其最小綠燈長為 5 秒，週期時間為 120 秒，兩者的綠燈介間時間皆為 5 秒時，兩時相的綠燈長和須為 110 秒，此時演算法得到的值須在 5 至 105 秒之間。若得到的綠燈長小於 5 秒時，將綠燈長鎖定為 5 秒，而當綠燈長大於 105 秒時，則鎖定為 105 秒。

以上例子代表二時相的情況，但如果在三或更多時相的路口中，有一個時相未能滿足最小綠燈長時，必須衡量要如何扣減剩餘的時相綠燈長才能有較好的績效。本研究中提出平衡法，其流程圖如圖 3-11 所示。

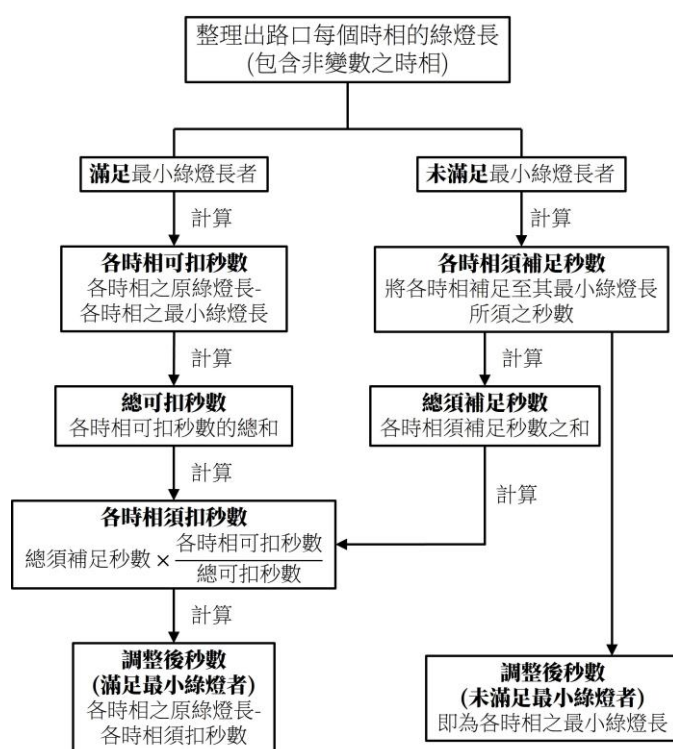


圖 3-11 最小綠燈平衡流程圖

從圖 3-11 可以看到，本研究使用的比例法並非直接將總補足秒數比例分配給其他有滿足最小綠燈長的時相，這是為了預防有滿足最小綠燈長的時相直接按照



比例被扣除時，反而小於最小綠燈長，便還要再做一次比例分配。在這個方法中，一次性計算每個時相可以被扣除的秒數之後，再從這些有餘裕的秒數中扣除，可以避免再重做一次的情形。

表 3-4 提供了一個 5 時相路口之示例，算出各路口在有時相未滿足最小綠燈長時的流程。其中其週期長為 150 秒，各時相之綠燈介間時間皆為 5 秒，由式(3.8)，可得知可分配的綠燈長為 125 秒。

表 3-4 五時相路口最小綠燈平衡示例

時相	最小綠燈 (秒)	某次覆算結果 (秒)	是否滿足 最小綠燈	須補足秒數	可扣秒數	調整後秒數
1	5	4	否	$(5-4) = 1$		5
2	5	13	是		$(13-5) = 8$	$13-5*(8/75) = 12.5 \approx 12$
3	5	12	是		$(12-5) = 7$	$12-5*(7/75) = 11.5 \approx 12$
4	20	80	是		$(80-20) = 60$	$80-5*(60/75) = 76$
5 (非變數)	20	16	否	$(20-16) = 4$		20
加總		$(4+13+12+80+16)$ $= 125$		$(1+4) = 5$	$(8+7+60) = 75$	$(5+12.5+11.5+76+20)$ $= 125$

在每次演算法覆算的結果中，都有可能出現不足最小綠燈的問題。每當發現有該問題時，就直接使用以上方法將綠燈長拉回可行域，也方便後續代回模擬器模擬。

3.5.6 變數初始值選擇

為了更有效率地進行路網號誌最佳化，如何取初始變數值也相當重要。在綠燈長的部分，由於獨立路口的號誌最佳化（亦即假設各路口無互動）目前早已有許多文獻與方法可以參考，本研究採用 Akçelik（1981）方法大略計算獨立路口的綠燈長。藉由將每個路口視為獨立路口取得的綠燈長結果，可以當作每個路口的綠燈長初始值，使用該值開始迭代。

因本研究所使用的綠燈長 ϕ 為顯示綠燈長 (displayed green time)，Akçelik 方法需要先將顯示綠燈長改為有效綠燈長 (effective green time)，如式 (3.13) 所示。

$$EG_{i,j} = \phi_{i,j} + I_{i,j} - l_{i,j}, \quad \forall i \in \mathbb{I}, \forall j \in \mathbb{J}_i \quad (3.13)$$

另外我們需要計算各流動 (movement) 的流量比 (flow ratio)，由於本研究沒有時相與流動之分，故可計算各時相的流量比公式如 (3.14) 所示。

$$y_{i,j} = \max_l \frac{q_{i,j,l}}{s_{i,j,l}}, \quad \forall i \in \mathbb{I}, \forall j \in \mathbb{J}_i, \forall l \in \mathbb{L}_{i,j} \quad (3.14)$$

流量比可以說在流量最大的車道上，到達流率 (arrival flow) 除以飽和流率 (saturation flow) 的值。從各路口各時相的 $y_{i,j}$ ，相加即可得到各路口的流量比和 Y_i ，如式 (3.15) 所示。

$$\sum_j y_{i,j} = Y_i, \quad \forall i \in \mathbb{I}, \forall j \in \mathbb{J}_i \quad (3.15)$$

將每個時相依 $\frac{y_{i,j}}{Y_i}$ 的比例分配，即可得到一個有效時間，計算公式如式 (3.16)。

$$EG_{i,j} = \left(\frac{y_{i,j}}{Y_i} \right) \cdot C_i \quad (3.16)$$

計算出有效綠燈長後再使用式 (3.13) 轉為顯示綠燈長 $\phi_{i,j}$ 即可得到一個初始的綠燈長解。

第四章 案例分析



本研究將使用 3 個路網來分析 SPSA 方法搭配 TRANSYT 車流模擬產生出的結果。三路網由小到大分別為一個 3 路口的幹道（或稱旺角路網，因路網取自香港旺角）、一個 6 肢腳的號誌化現代圓環以及一個 9 路口的雙幹道路網。本章將分析三路網使用 SPSA 產生的結果。其中本章安排 4.1 節為案例路網介紹；4.2 節為初始參數設定；4.3 節為旺角路網的分析結果；4.4 節為現代圓環之分析結果；4.5 節為雙幹道之分析結果；4.6 節為小結。

4.1 案例路網簡介

為更完整地探討使用 SPSA 得到的解之最適性（optimality）與收斂速度等議題，本研究取了三個不同大小的路網，各路網簡介如表 4-1 所示。各路網設計由簡至繁，可以用來驗證 SPSA 在不同複雜程度的路網中的表現。其中旺角路網參考自 Lo (2001)，該篇文章使用了文內提出的 DISCO 模型最佳化這個 3 路口幹道的號誌，此路網將該研究以基因演算法做出的結果做為標竿（benchmark），分析其結果與本研究結果的差異。號誌化的現代圓環¹⁰路口則考慮一個環狀路口的聯鎖，共有 6 個肢腳（leg）（FHWA, 2000），此路網取了特定一個初始解並使用多個不同的隨機擾動來分析，確保演算法的可行性與正確的求解方向。雙幹道路網則考慮了一個 9 路口的大範圍路網，由於演算法目標是希望能應用於實務上，藉由一個較複雜的路網，也為實務上的使用展示。本研究使用 TRANSYT 17.0.1（2025）進行模擬與 Hill Climbing 最佳化（作為與本研究方法之比較）。

¹⁰ 現代圓環（modern roundabout）為一種環內車輛享有優先權的單向環狀路口，其標誌與標線設計可減少車輛在環內的交織（weaving）行為，並降低衝突風險（DMRB，2023）。

表 4-1 各路網簡介

	旺角路網	現代圓環	雙幹道路網
路網說明	3 路口幹道	6 肢腳現代圓環	9 路口雙幹道路網
路網規模	小	中	大
路段數 ¹¹	16	18	83
路口/肢腳數目	3	6	9
變數數目	5	11	26
時差數目	2	5	8
綠燈長數目	3	6	18

4.2 初始參數設定

4.2.1 系統參數設定

系統參數主要分為道路參數、流量參數與號誌參數。道路參數主要為路段長度、路段飽和流率、車道配置等；流量參數則因 TRANSYT 為起訖矩陣自動分配，故起訖矩陣即為流量參數；號誌參數則主要為週期長度、路口時相順序、綠燈介間時間與損失時間等。

有關道路參數之圖表部分，旺角路網之基本配置圖如圖 4-1 所示，路段資料表如表 4-2 所示；現代圓環之基本配置圖如圖 4-2 所示，路段資料表如表 4-3 所示；雙幹道路網之全路口基本配置圖如圖 4-3 所示，路口別基本配置圖如圖 4-4 至圖 4-12 所示，路段資料表如表 4-4 至表 4-12 所示。

¹¹ 配合 TRANSYT 標示（稱為 arm）與製表便利，當一路段上有專用道時另切為 3 路段，分別為岔出專用道前、岔出專用道後之專用道、岔出專用道後之非專用道等，如圖 4-1、圖 4-3 所示。

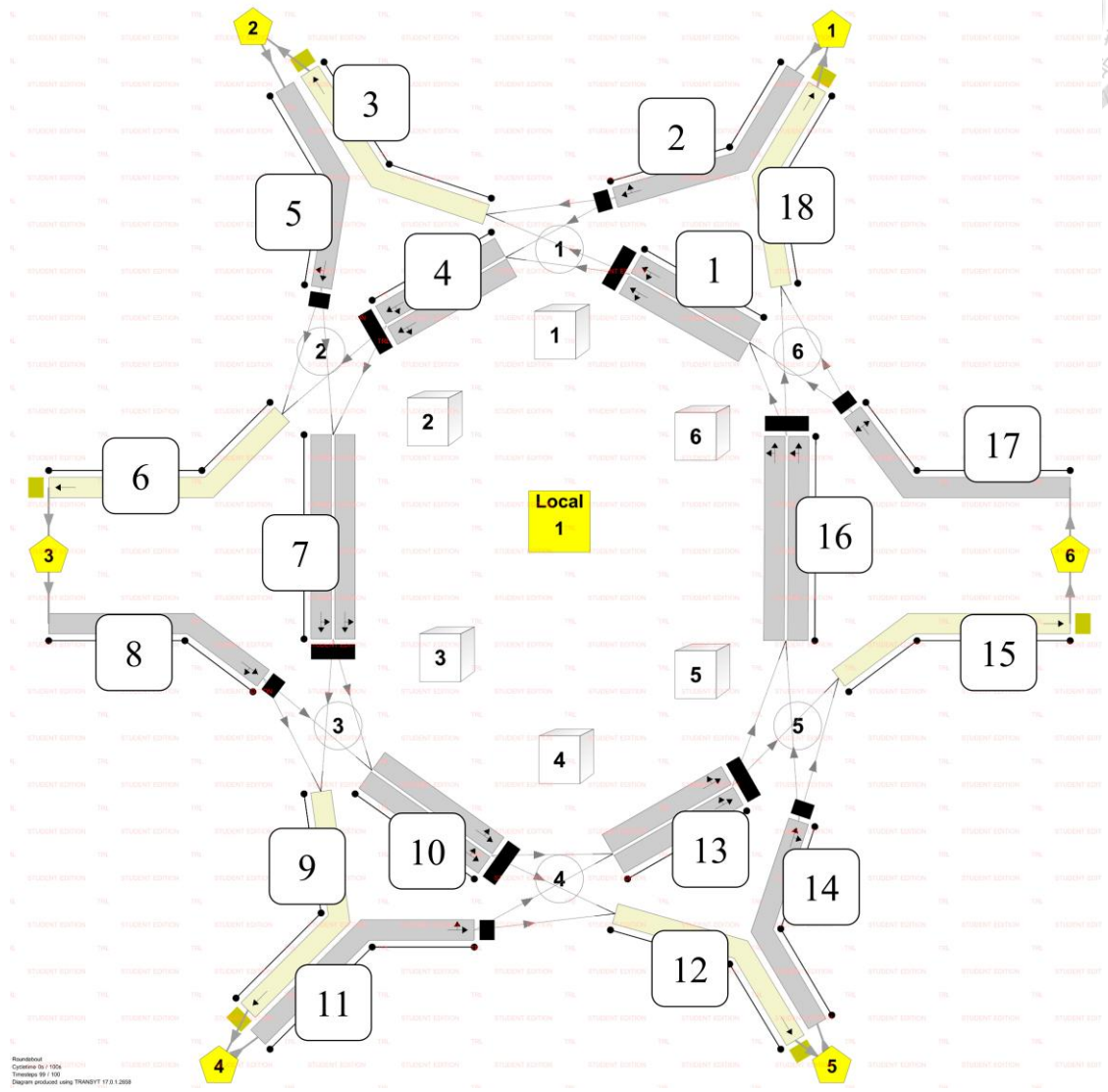


圖 4-2 現代圓環基本配置圖（未按比例繪製，畫面截自 TRANSYT 17）



表 4-2 旺角路網路段資料表

路段編號	路段長度 (m)	飽和流率 (PCU/hr)
1	74	6000
2	74.4	2000
3	26	--
4	26	6000
5	26	2000
6	74.4	2000
7	74.4	2000
8	26	--
9	70.8	--
10	26	6000
11	26	2000
12	74.4	2000
13	74.4	2000
14	70.8	--
15	70.8	--
16	36.67	--

表 4-3 現代圓環路段資料表

路段編號	路段長度 (m)	飽和流率 (PCU/hr)
1	82.63	3600
2	82.35	1800
3	127.4	--
4	86.77	3600
5	102.35	1800
6	132.01	--
7	160.55	3600
8	76.91	1800
9	131.25	--
10	94.36	3600
11	86.18	1800
12	124.33	--
13	91.73	3600
14	90.12	1800
15	132.32	--
16	123.24	3600
17	84.42	1800
18	120.08	--

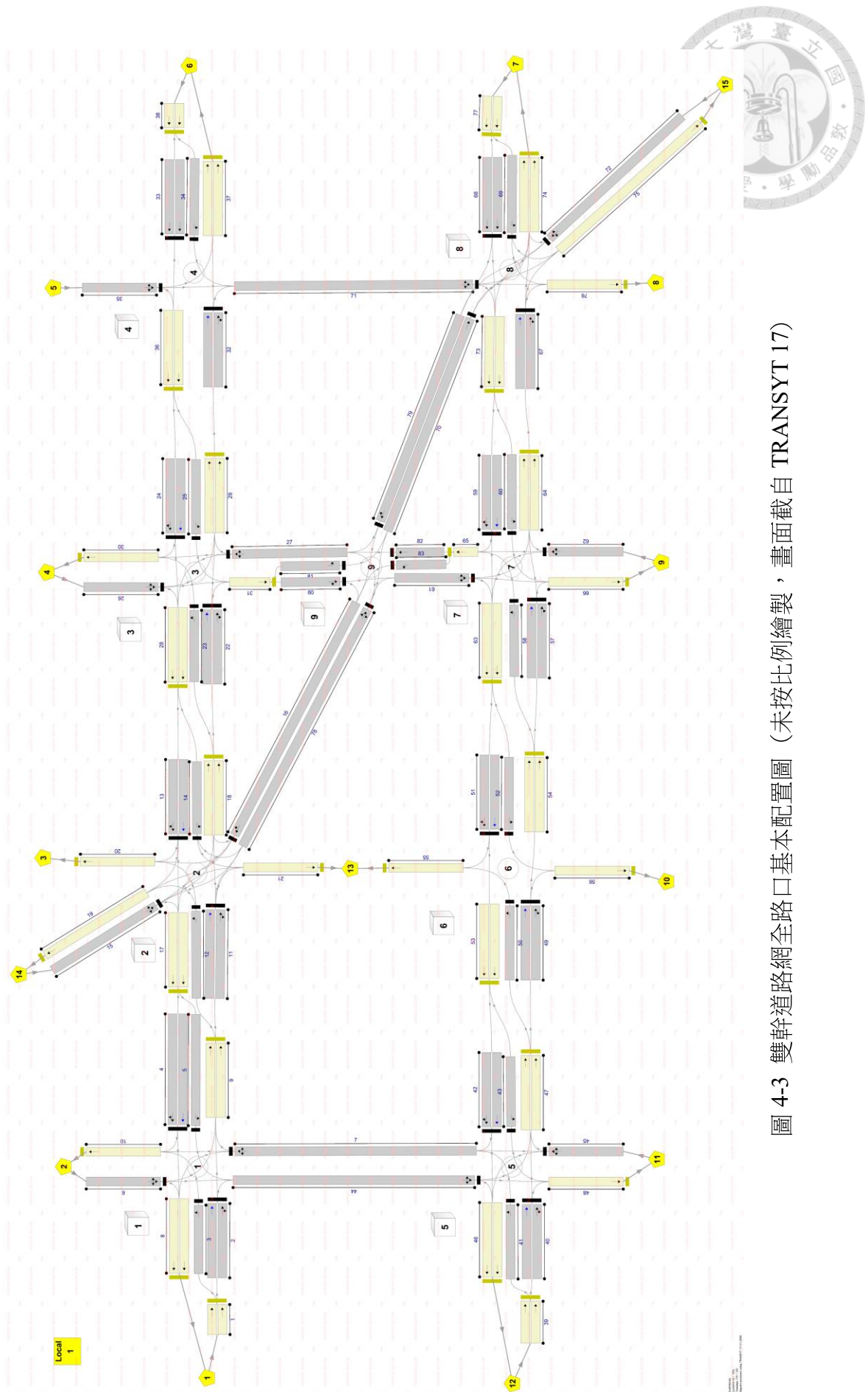


圖 4-3 雙幹道路網全路口基本配置圖（未按比例繪製，畫面截自 TRANSYT 17）

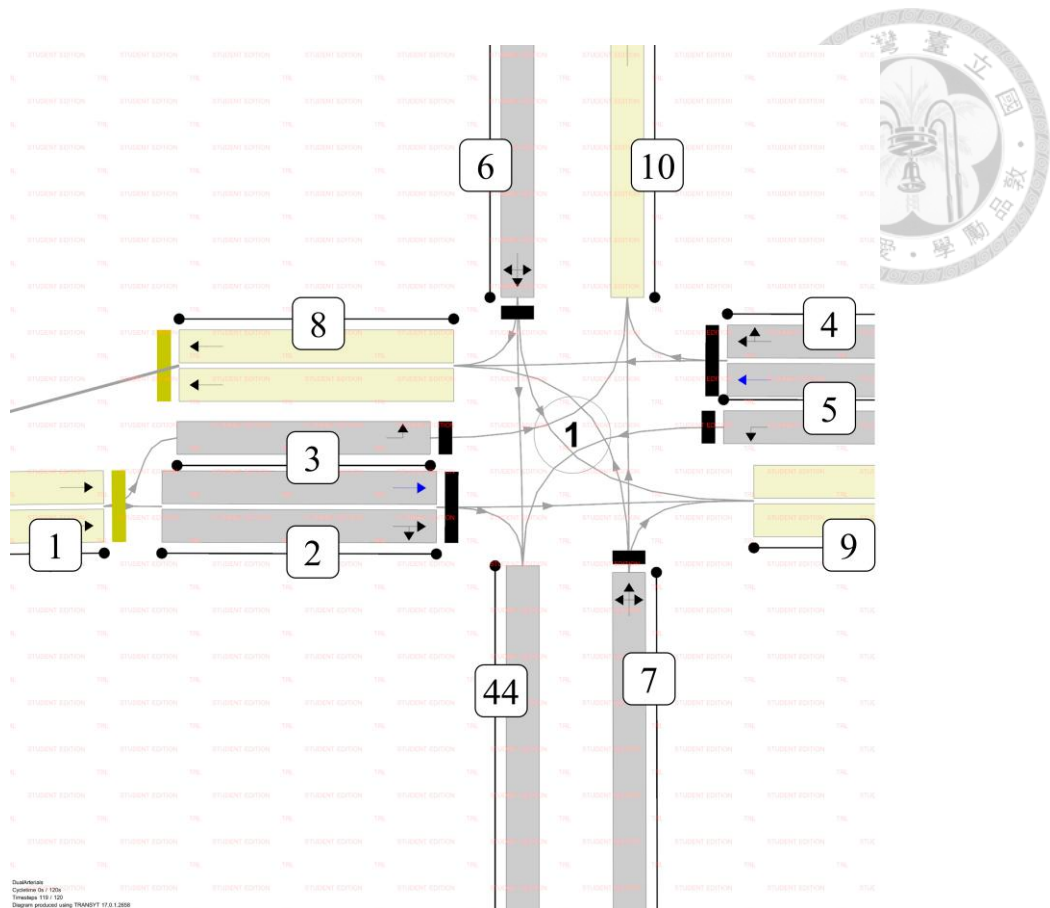


圖 4-4 雙幹道路網路口 1 基本配置圖（未按比例繪製，畫面截自 TRANSYT）

表 4-4 雙幹道路網路段資料表（路口 1）

路段編號	路段長度 (m)	飽和流率 (PCU/hr)
1	36.67	--
2	74.4	4000
3	74.4	2000
4	74.4	4000
5	74.4	2000
6	50	1904
7	260	2000
8	70.8	--
9	70.8	--
10	92.23	--
44	260	--

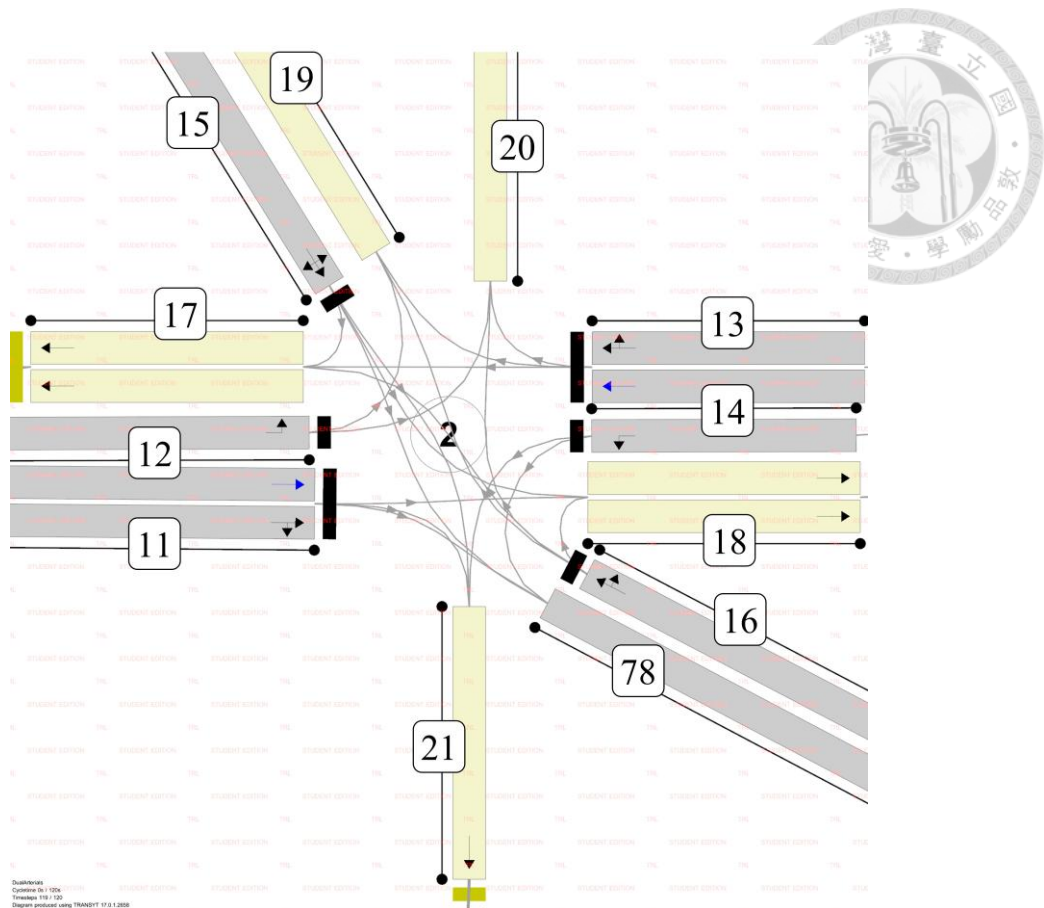


圖 4-5 雙幹道路網路路口 2 基本配置圖（未按比例繪製，畫面截自 TRANSYT）

表 4-5 雙幹道路網路路段資料表（路口 2）

路段編號	路段長度 (m)	飽和流率 (PCU/hr)
11	74.4	4000
12	74.4	2000
13	75.3	4000
14	75.3	2000
15	82.03	1800
16	360	1800
17	70.8	--
18	250	--
19	135.42	--
20	96.73	--
21	150	--
78	360	--

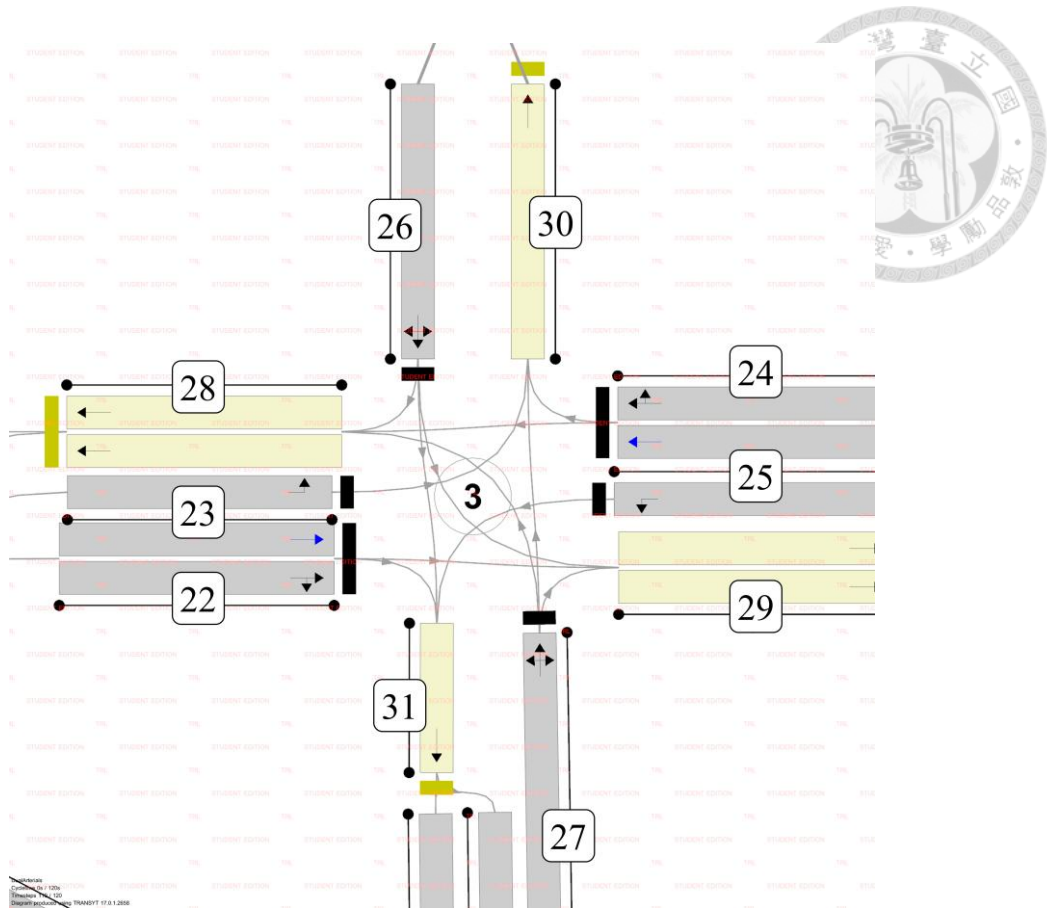


圖 4-6 雙幹道路網路路口 3 基本配置圖（未按比例繪製，畫面截自 TRANSYT）

表 4-6 雙幹道路網路路段資料表（路口 3）

路段編號	路段長度 (m)	飽和流率 (PCU/hr)
22	75.3	4000
23	75.3	2000
24	70	4000
25	70	2000
26	50	2000
27	250	2000
28	250	--
29	120	--
30	91.05	--
31	170	--

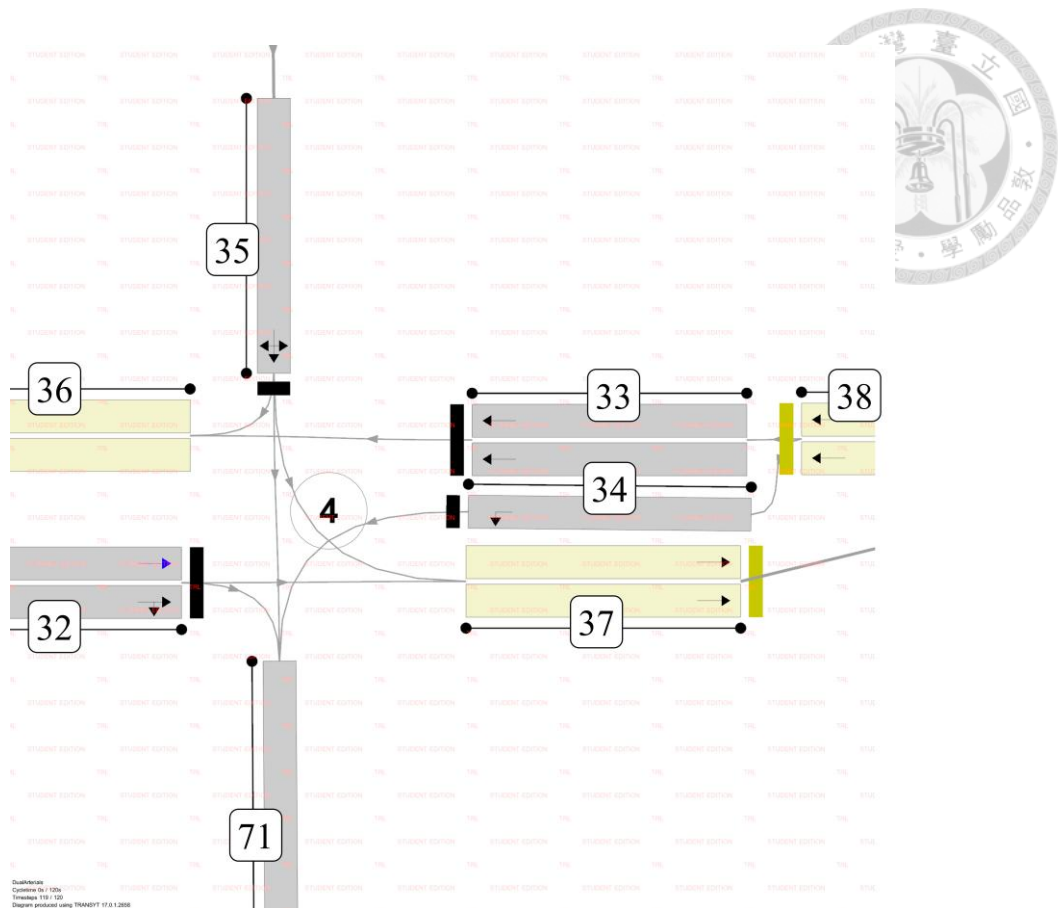


圖 4-7 雙幹道路網路口 4 基本配置圖（未按比例繪製，畫面截自 TRANSYT）

表 4-7 雙幹道路網路段資料表（路口 4）

路段編號	路段長度 (m)	飽和流率 (PCU/hr)
32	70	4000
33	74.4	4000
34	74.4	2000
35	50	2000
36	120	--
37	105.77	--
38	36.67	--
71	400	--

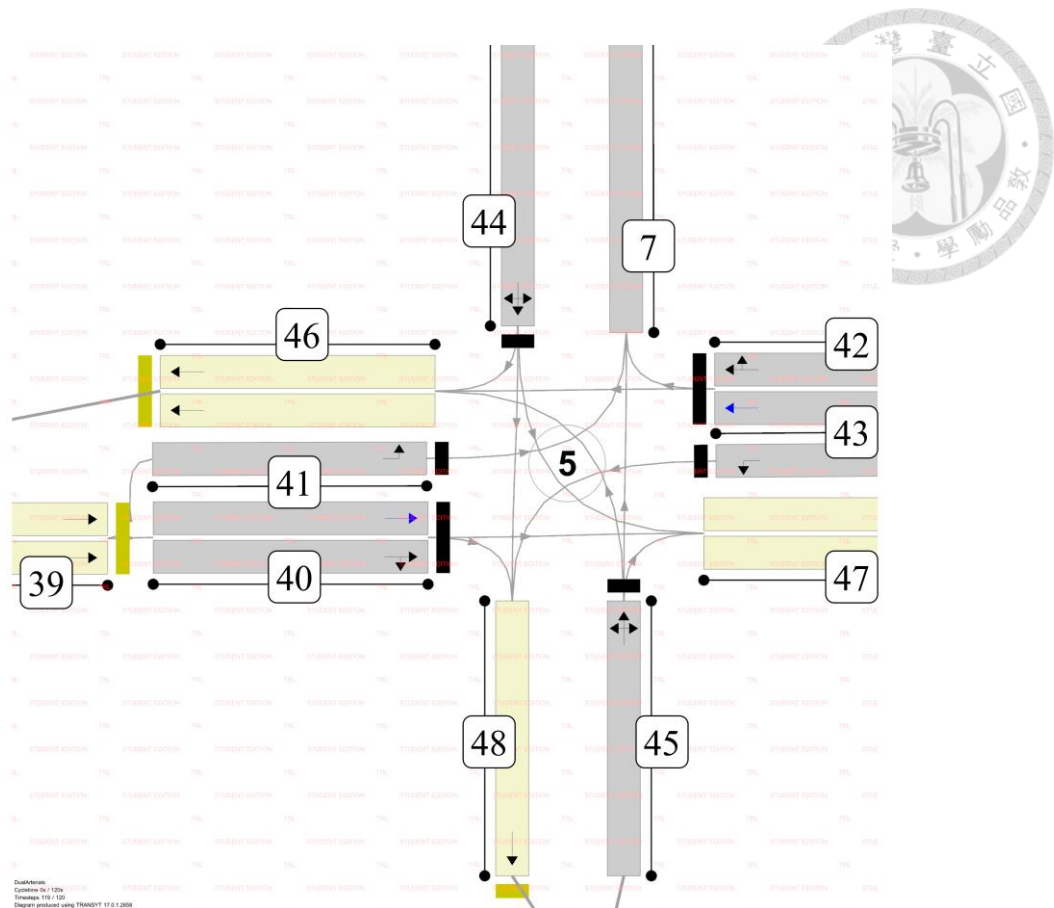


圖 4-8 雙幹道路網路路口 5 基本配置圖（未按比例繪製，畫面截自 TRANSYT）

表 4-8 雙幹道路網路路段資料表（路口 5）

路段編號	路段長度 (m)	飽和流率 (PCU/hr)
7	260	--
39	36.67	--
40	69.43	4000
41	71.6	2000
42	80	4000
43	80	2000
44	260	2000
45	50	2000
46	94.23	--
47	175	--
48	91.47	--

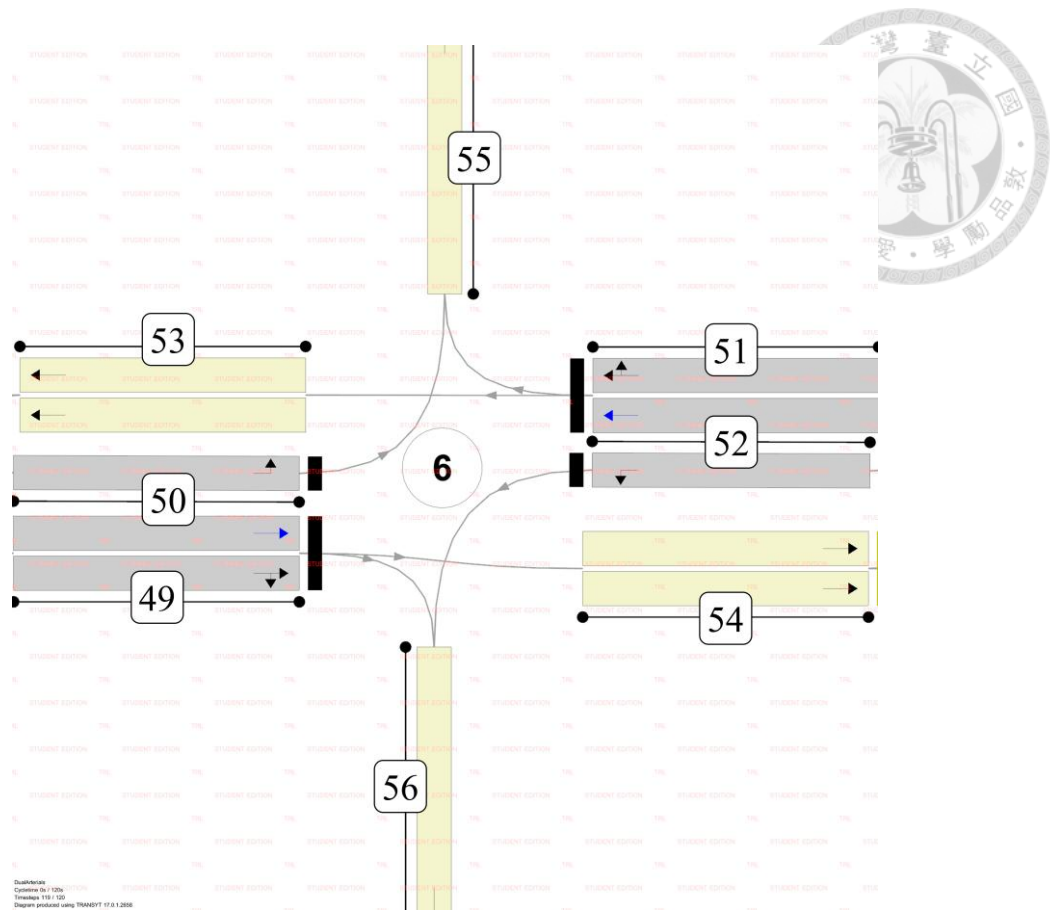


圖 4-9 雙幹道路網路路口 6 基本配置圖（未按比例繪製，畫面截自 TRANSYT）

表 4-9 雙幹道路網路路段資料表（路口 6）

路段編號	路段長度 (m)	飽和流率 (PCU/hr)
49	80	4000
50	80	2000
51	80	4000
52	80	2000
53	175	--
54	135	--
55	150	--
56	90.96	--

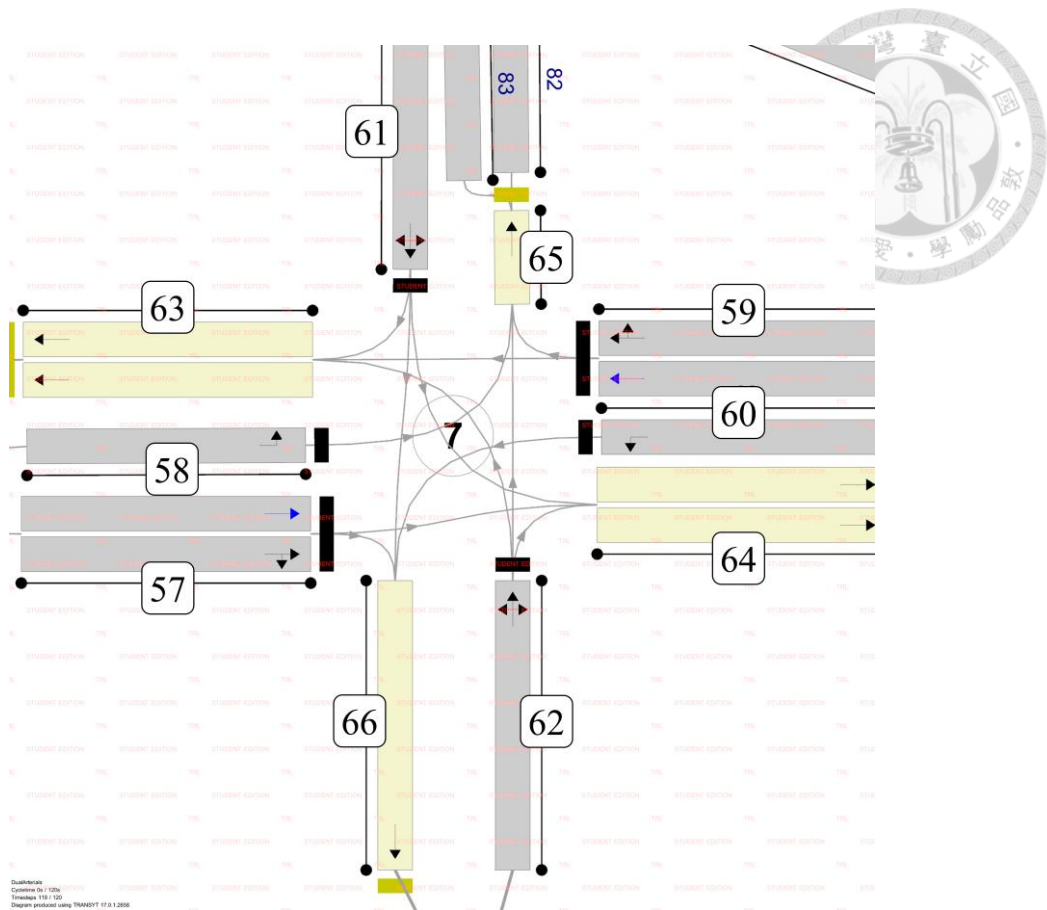


圖 4-10 雙幹道路網路路口 7 基本配置圖（未按比例繪製，畫面截自 TRANSYT）

表 4-10 雙幹道路網路路段資料表（路口 7）

路段編號	路段長度 (m)	飽和流率 (PCU/hr)
57	80	4000
58	80	2000
59	80	4000
60	80	2000
61	100	1800
62	50	2000
63	135	--
64	260	--
65	20	--
66	91.21	--

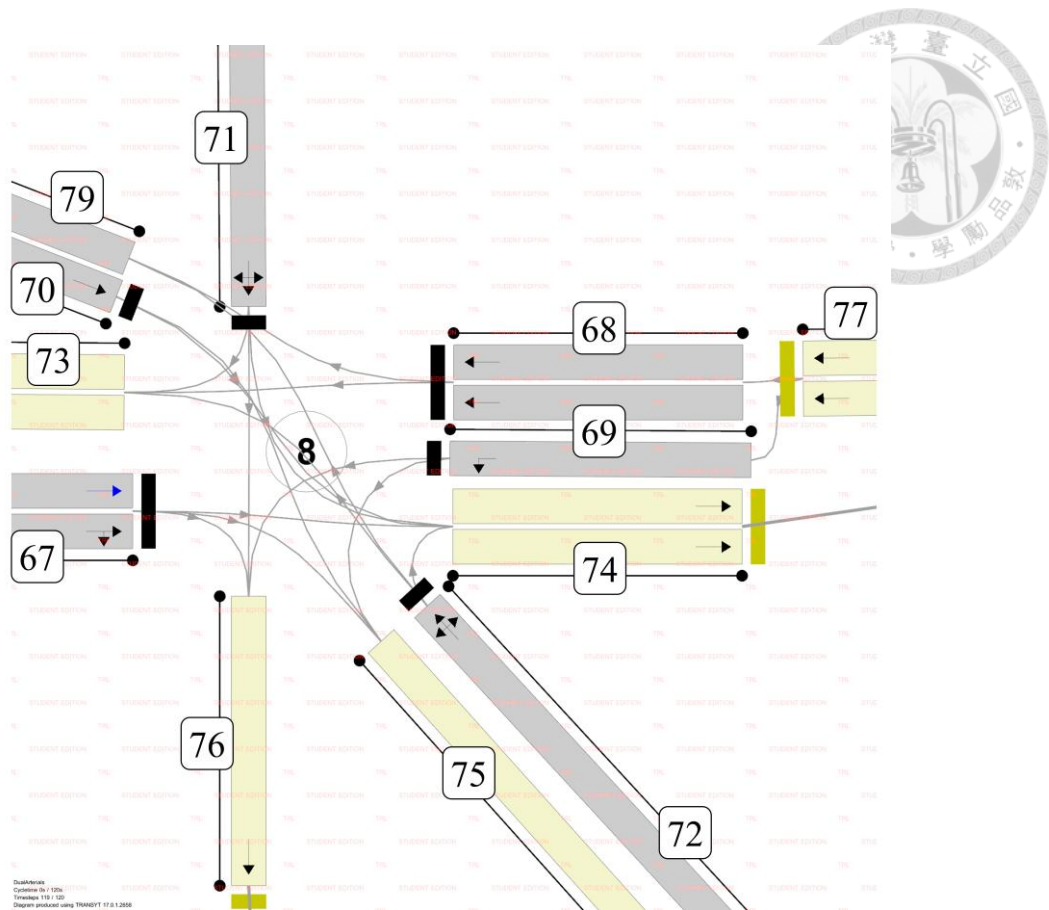


圖 4-11 雙幹道路網路路口 8 基本配置圖（未按比例繪製，畫面截自 TRANSYT）

表 4-11 雙幹道路網路路段資料表（路口 8）

路段編號	路段長度 (m)	飽和流率 (PCU/hr)
67	80	4000
68	80	4000
69	80	2000
70	420	1800
71	400	2000
72	150	1800
73	260	--
74	100	--
75	150	--
76	92.63	--
77	36.67	--
79	420	--

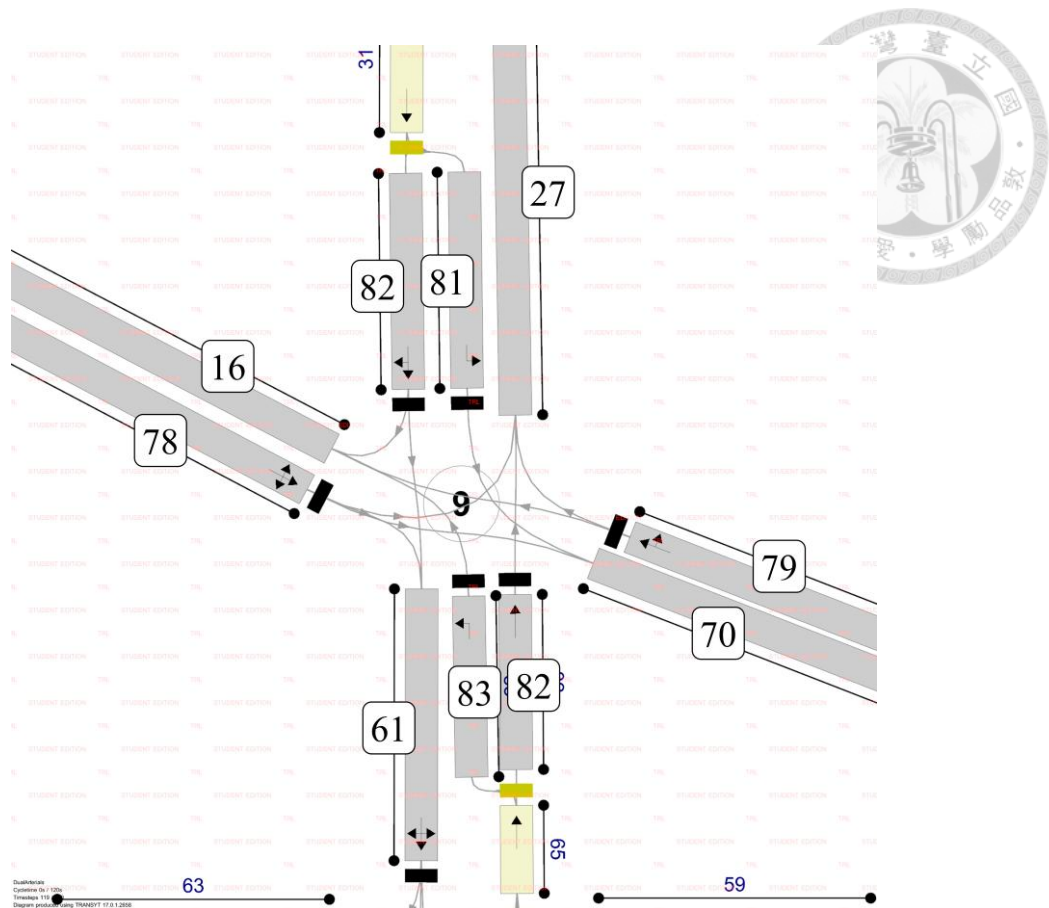


圖 4-12 雙幹道路網路路口 9 基本配置圖（未按比例繪製，畫面截自 TRANSYT）

表 4-12 雙幹道路網路路段資料表（路口 9）

路段編號	路段長度 (m)	飽和流率 (PCU/hr)
16	360	--
27	250	--
61	100	--
70	420	--
78	360	1800
79	420	1800
80	47.21	1800
81	80	1800
82	100	1800
83	100	1800

由以上的表可以得到各路網的路段配置與長度、飽和流率¹²等。由於雙幹道路網範圍較大，較難看出路段代表數字，故將每個路口獨立出來會比較清楚。另外，一般在飽和流率部分每車道皆設為 1800 PCU/hr 或 2000 PCU/hr，在雙幹道路網的路段 6（路口 1）因建置路網的操作失誤設為 1904 PCU/hr，如表 4-4 所示。

在流量參數部分，考慮某個小時內的車輛數，建置起訖矩陣，如表 4-13、表 4-15、所示。TRANSYT 可以藉由起訖矩陣分配各個路口的流量，在此亦將列出，各路口的轉向量如所示。其中表格中的數值皆為某一小時內的車輛數。

表 4-13 旺角路網起訖矩陣

旺角路網		終點			總出發
		5	6	7	
起點	1	194	300	1126	1620
	2	48	40	308	396
	3	984	10	158	1152
	4	--	114	114	228
總抵達		1226	464	1706	3396

¹² 對於每個進入路口的車道，號誌週期分為有效紅燈時間與有效綠燈時間。在有效綠燈期間內，只要有排隊車輛，車輛便能以一個穩定速率離開路口，此速率即為飽和流率（saturation flow）（Allsop, 1971）。意即一個路口只有進入路口的路段有飽和流率的概念。



表 4-14 旺角路網轉向量表

(a) 旺角路網路口 1 轉向量

路口1轉向量		離開路口路段	進入交通量合計
		3	
進入路口路段	1	1620	1620
	2	396	396

(b) 旺角路網路口 2 轉向量

路口2轉向量		離開路口路段		進入交通量合計
		8	9	
進入路口路段	4	1774	--	1774
	5	--	242	242
	6	--	984	984
	7	168	--	168

(c) 旺角路網路口 3 轉向量

路口3轉向量		離開路口路段		進入交通量合計
		14	15	
進入路口路段	10	1592	--	1592
	11	--	350	350
	12	--	114	114
	13	114	--	114

表 4-15 現代圓環起訖矩陣

現代圓環		終點						總出發
		1	2	3	4	5	6	
起點	1	0	50	50	50	100	50	300
	2	200	0	100	50	50	50	450
	3	200	150	0	100	100	100	650
	4	300	200	100	0	100	100	800
	5	300	200	200	100	0	100	900
	6	100	50	200	100	100	0	550
總抵達		1100	650	650	400	450	400	3650

表 4-16 現代圓環轉向量表

(a) 現代圓環路口 1（肢腳 1）轉向量

路口1轉向量		離開路口路段		進入交通量合計
		3	4	
進入路口路段	2	50	250	300
	1	600	800	1400

(b) 現代圓環路口 2（肢腳 2）轉向量

路口2轉向量		離開路口路段		進入交通量合計
		6	7	
進入路口路段	5	100	350	450
	4	550	500	1050

(c) 現代圓環路口 3（肢腳 3）轉向量

路口3轉向量		離開路口路段		進入交通量合計
		9	10	
進入路口路段	8	100	550	650
	7	300	550	850

(d) 現代圓環路口 4（肢腳 4）轉向量

路口4轉向量		離開路口路段		進入交通量合計
		12	13	
進入路口路段	11	100	700	800
	10	350	750	1100

(e) 現代圓環路口 5（肢腳 5）轉向量

路口5轉向量		離開路口路段		進入交通量合計
		15	16	
進入路口路段	14	100	800	900
	13	300	1150	1450

(f) 現代圓環路口 6（肢腳 6）轉向量

路口6轉向量		離開路口路段		進入交通量合計
		18	1	
進入路口路段	17	100	450	550
	16	1000	950	1950

表 4-17 雙幹道路路口起訖矩陣

雙幹道路網		終點															總出發
		1	2	3	4	6	7	8	9	10	11	12	13	14	15		
起點	1	0	120	60	65	530	90	65	45	50	60	2	39	28	36	1190	
	2	80	0	1	1	120	48	39	37	59	70	36	51	0	44	586	
	4	65	1	1	0	58	37	65	71	34	42	39	42	0	28	483	
	5	103	2	1	1	68	57	39	30	28	42	44	36	3	26	480	
	6	623	97	64	75	0	1	92	41	36	49	24	43	32	7	1184	
	7	23	19	13	18	0	0	50	51	17	30	324	23	12	50	630	
	9	94	82	26	88	34	62	1	0	1	0	71	49	22	3	533	
	11	43	71	29	31	65	90	5	5	1	0	90	51	21	45	547	
	12	0	23	18	58	50	296	29	29	22	48	0	27	5	36	641	
	14	10	1	0	1	52	32	40	39	13	19	11	55	0	39	312	
	15	67	54	39	29	5	30	0	0	3	5	51	30	18	0	331	
	總抵達		1108	470	252	367	982	743	425	348	264	365	692	446	141	314	6917



表 4-18 雙幹道路網轉向量表

(a) 雙幹道路網路口 1 轉向量

路口1轉向量	離開路口路段					進入交通量合計
	8	9	10	44		
2	--	905	--	165		1070
3	--	--	120	--		120
4	920	--	200	--		1120
5	--	--	--	181		181
6	80	302	--	204		586
7	108	175	150	--		433

(b) 雙幹道路網路口 2 轉向量

路口2轉向量	離開路口路段							進入交通量合計
	17	18	19	20	21	78		
11	--	998	--	--	90	131		1219
12	--	--	108	54	--	--		162
13	1050	--	50	90	--	--		1190
14	--	--	--	--	121	0		121
15	45	123	--	--	55	89		312
16	206	0	37	54	--	--		297

(c) 雙幹道路網路路口 3 轉向量

路口3轉向量	離開路口路段				進入交通量合計
	28	29	30	31	
22	--	828	--	181	1009
23	--	--	111	--	111
24	1057	--	76	--	1133
25	--	--	--	274	274
26	150	103	--	231	484
27	104	96	180	--	380

(d) 雙幹道路網路路口 4 轉向量

路口4轉向量	離開路口路段				進入交通量合計
	36	37	71		
32	--	914	114		1028
33	1091	--	--		1091
34	--	--	93		93
35	316	68	96		480



(e) 雙幹道路網路路口 5 轉向量

路口5轉向量	進入路口路段	離開路口路段				進入交通量合計
		46	47	7	48	
40	進入路口路段	--	493	--	48	541
41		--	--	100	--	100
42		488	--	121	--	609
43		--	--	--	86	86
44		114	206	--	231	551
45		90	245	212	--	547

(f) 雙幹道路網路路口 6 轉向量

路口6轉向量	進入路口路段	離開路口路段				進入交通量合計
		53	54	55	56	
49	進入路口路段	--	729	--	136	865
50		--	--	78	--	78
51		695	--	102	--	797
52		--	--	--	128	128

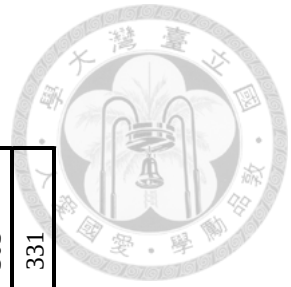


(g) 雙幹道路網路路口 7 轉向量

路口7轉向量	離開路口路段					進入交通量合計
	63	64	65	66	66	
進入路口路段	57	--	566	--	61	627
	58	--	--	102	--	102
	59	516	--	130	--	646
	60	--	--	--	51	51
	61	200	264	--	236	700
	62	209	66	258	--	533

(h) 雙幹道路網路路口 8 轉向量

路口8轉向量	離開路口路段							進入交通量合計
	73	74	79	75	76	76	76	
進入路口路段	67	--	542	--	182	--	172	896
	68	490	--	40	--	--	--	530
	69	--	--	--	50	50	50	100
	70	--	71	--	82	--	--	153
	71	0	100	--	--	203	303	303
	72	207	30	94	--	--	--	331



(i) 雙幹道路網路路口 9 轉向量

路口9轉向量		離開路口路段				進入交通量合計
		16	70	27	61	
進入路口路段	78	--	55	0	165	220
	79	54	--	80	--	134
	80	53	--	--	535	588
	81	--	98	--	--	98
	82	--	--	300	--	300
	83	190	--	--	--	190



在號誌參數部分，旺角路網與雙幹道路網週期時間定為 120 秒，現代圓環則為 100 秒，損失時間中的起始時移 (start displacement) 與結束時移 (end displacement)¹³在各路網中的各時相皆設定為 2 秒，綠燈介間時間在各路網的各時相皆為 5 秒。另有時相順序圖如所示。

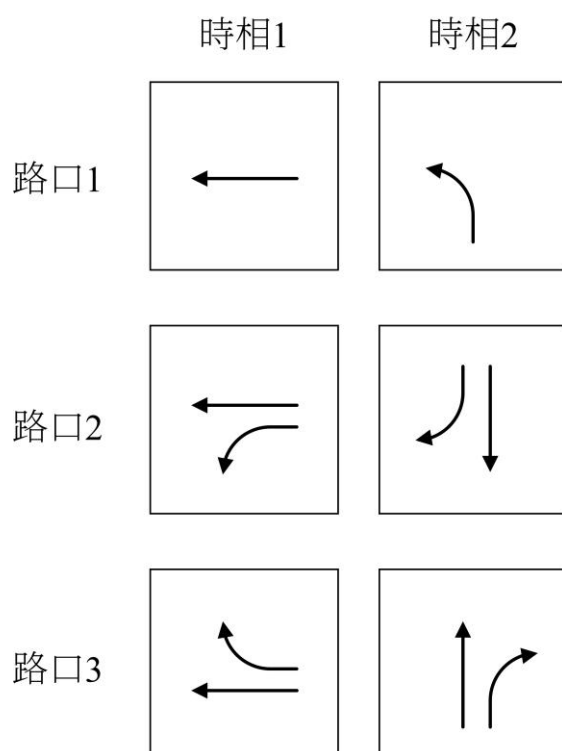


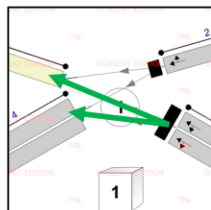
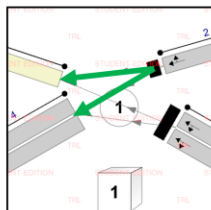
圖 4-13 旺角路網時相圖（可參考圖 4-1 配置）

¹³ 主要修正顯示綠燈時間與有效綠燈時間的差異，分起始位移（加速時間的損失）與結束位移（駕駛在黃燈時視為綠燈移動的情況）（TRL, 2024）。本研究所稱之損失時間為起始位移-結束位移。

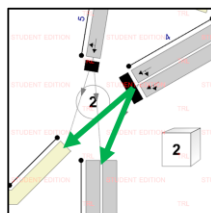
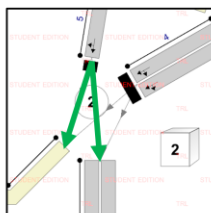
時相1 (環外綠燈) 時相2 (環內綠燈)



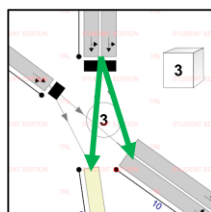
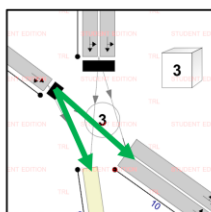
肢腳1



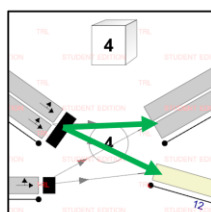
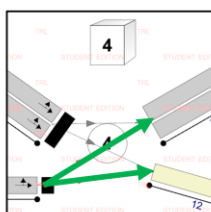
肢腳2



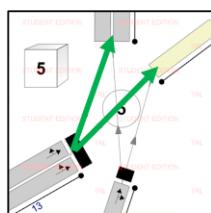
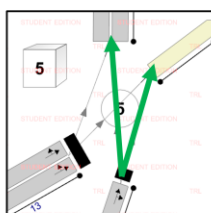
肢腳3



肢腳4



肢腳5



肢腳6

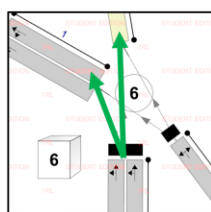
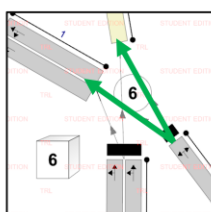


圖 4-14 現代圓環路網時相圖 (可參考圖 4-2 配置)

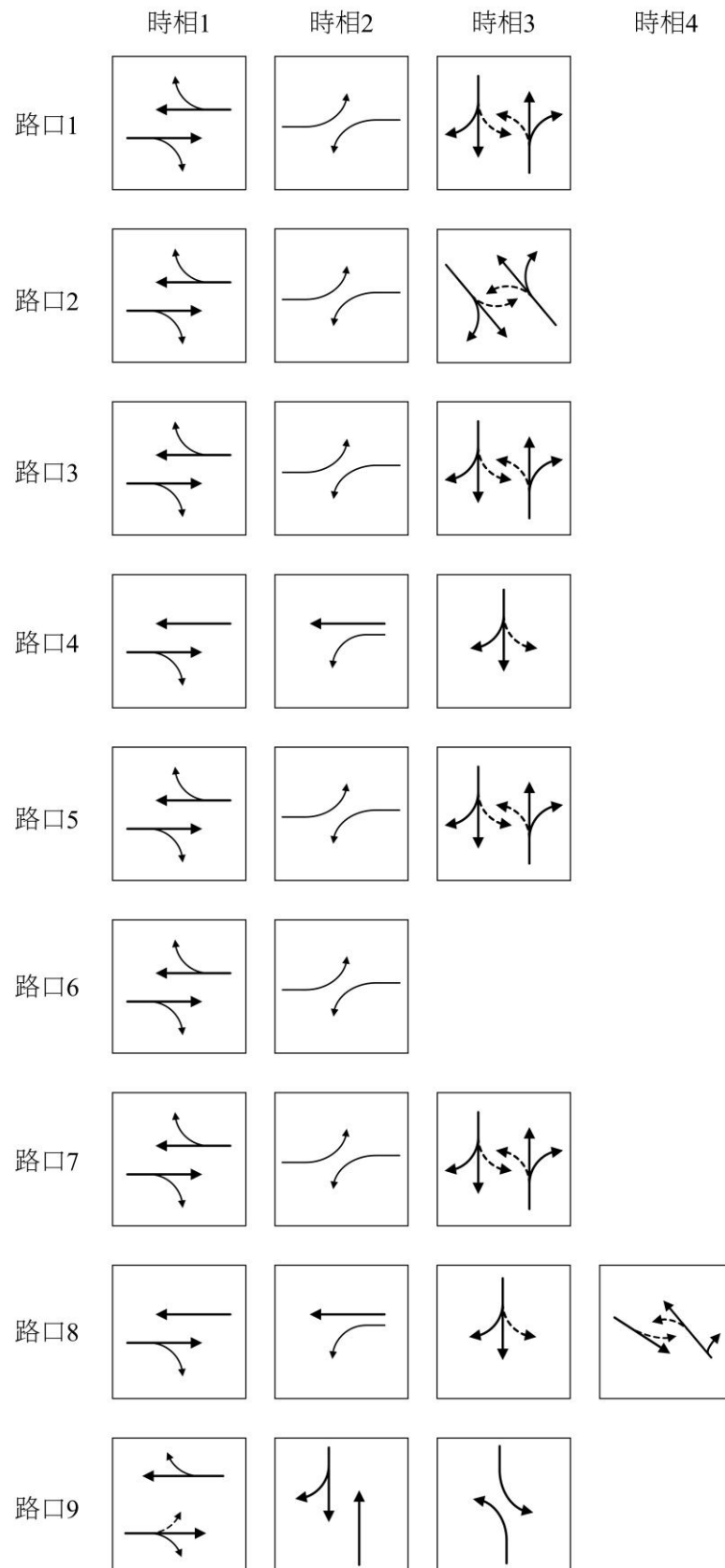


圖 4-15 雙幹道路網時相圖（可參考配置圖 4-3 及圖 4-4 至圖 4-12 配置）



4.2.2 演算法參數設定

本研究共通的演算法參數主要是迭代步長的部分參數與擾動幅度的部分。Gerencsér (1999) 驗證了當迭代步長與擾動幅度皆為常數時，求得的解是有界的。在本研究中，當擾動幅度收斂到一個較小值時，因為本研究的決策變數時差與綠燈長皆為整數，在梯度估計（如式(3.10)）時，可能導致 $D_k^+ = D_k^- = D(\mathbf{x}^k)$ 的情形發生，因此，本研究將所有擾動幅度皆固定為 1，即 $c_k = c = 1$ ，以預防以上問題。

在迭代步長部分則是使用式(3.11)，每次迭代的步長會形成一個數列，其中 A 依經驗設定為最大迭代次數的 10%， α 則定為 0.6 (Spall, 1998)。另外在區別變數型態 SPSA 的演算法中，所有路網的共通點即當連續 3 次迭代未改進超過 0.1% 時，可決定是否提前進入下一個階段（一個階段最多 20 次迭代）。一個時差階段與一個綠燈長階段為一個循環，當連續 2 個循環未改善超過 1% 時，可提前結束迭代。而每一次完整迭代之總最大迭代次數（所有時差與綠燈長迭代次數加總）限制為旺角路網 50 次、現代圓環及雙幹道路網為 100 次。

另外，本研究將三個路網的模擬時間設為 3600 秒（1 小時），即各路網 1 小時內所產生的總延滯。

4.2.3 初始解搜尋

由於總延滯一般來說非凸函數，當僅使用一個初始解做迭代時，並無法證明這個初始解的迭代最小值就會是全域的最小值。因此，為了找到更好的點，本研究根據路網大小，分別取了不同數量的初始解。在旺角路網中，取了 2 個隨機初始解做迭代；在現代圓環取 3 個隨機初始解；而考慮雙幹道路網的尺度較大，取 5 個隨機點作為初始解分別迭代。這些初始解僅有時差上的差異，在綠燈長部分如 3.5.6 小節，每個路網會有一組起始綠燈長，並在每個不同的初始解中保持固定，各路網的初始解如表 4-19、表 4-20 與表 4-21 所示。



表 4-19 旺角路網初始點列表（2 個點）

	點1	點2	生成方式
θ_2	28	44	隨機
θ_3	70	107	
$\phi_{1,1}$	64	64	Akçelik公式
$\phi_{2,1}$	40	40	
$\phi_{3,1}$	93	93	

表 4-20 現代圓環初始點列表（3 個點）

	點1	點2	點3	生成方式
θ_2	55	20	1	隨機
θ_3	91	30	43	
θ_4	98	8	2	
θ_5	77	83	15	
θ_6	46	93	50	
$\phi_{1,1}$	25	25	25	Akçelik公式
$\phi_{2,1}$	41	41	41	
$\phi_{3,1}$	55	55	55	
$\phi_{4,1}$	54	54	54	
$\phi_{5,1}$	50	50	50	
$\phi_{6,1}$	31	31	31	

表 4-21 雙幹道路網初始點列表 (5 個點)



	點1	點2	點3	點4	點5	生成方式
θ_2	36	101	69	45	64	隨機
θ_3	91	50	88	39	110	
θ_4	103	0	5	73	9	
θ_5	49	108	93	9	111	
θ_6	97	37	73	107	37	
θ_7	112	83	84	1	70	
θ_8	25	61	40	50	49	
θ_9	24	44	110	94	55	
$\phi_{1,1}$	46	46	46	46	46	Akçelik公式
$\phi_{1,2}$	11	11	11	11	11	
$\phi_{2,1}$	62	62	62	62	62	
$\phi_{2,2}$	13	13	13	13	13	
$\phi_{3,1}$	46	46	46	46	46	
$\phi_{3,2}$	20	20	20	20	20	
$\phi_{4,1}$	52	52	52	52	52	
$\phi_{4,2}$	5	5	5	5	5	
$\phi_{5,1}$	33	33	33	33	33	
$\phi_{5,2}$	8	8	8	8	8	
$\phi_{6,1}$	88	88	88	88	88	
$\phi_{7,1}$	29	29	29	29	29	
$\phi_{7,2}$	6	6	6	6	6	
$\phi_{8,1}$	40	40	40	40	40	
$\phi_{8,2}$	5	5	5	5	5	
$\phi_{8,3}$	26	26	26	26	26	
$\phi_{9,1}$	21	21	21	21	21	
$\phi_{9,2}$	66	66	66	66	66	



4.3 分析結果—旺角路網

旺角路網總共 3 個路口、5 個決策變數以及 2 個隨機初始解，分別為 θ_2 、 θ_3 、 $\phi_{1,1}$ 、 $\phi_{2,1}$ 、 $\phi_{3,1}$ 。在這個路網中，首先將比較在不區別變數型態下 SPSA 迭代步長的參數 a 與迭代結果。另外使用區別變數型態 SPSA 與結合不同變數型態（不區別變數型態）SPSA 做簡單的比較。另外也將使用最陡下降法（Steepest Descent Method）做迭代，並搭配黃金分割法（Golden Section Method）做迭代步長的線性搜尋¹⁴，比較 SPSA 與最陡下降法的最適性與收斂速度。

4.3.1 結合不同變數型態之 SPSA—不同迭代步長參數結果

在旺角路網中，首先取點 1 做了 3 個不同迭代步長參數 $a = 1.44, 1.8, 2.16$ 的結合不同變數型態 SPSA¹⁵，即 5 個變數以同一個擾動與梯度大小來移動，而不做其他處理。如同傳統 SPSA，在這裡僅考慮最大迭代次數限制，依照 SPSA 做迭代直到達最大迭代次數為止。本案例路網最大迭代次數設定為 50 次，各個 a 得到的結果如圖 4-16、圖 4-17 與圖 4-18 所示。

¹⁴ 最陡下降法在本節只列出結果，迭代過程詳見附錄 A。

¹⁵ 此處的 a 取值因之前使用其他模擬工具（SIGMIX）分析時的單位為 PCU-s/hr，與本研究使用的參數 PCU-hr/hr 不同，故使用過去經驗調整 a 至相似的尺度（SIGMIX 結果詳見附錄 B）。

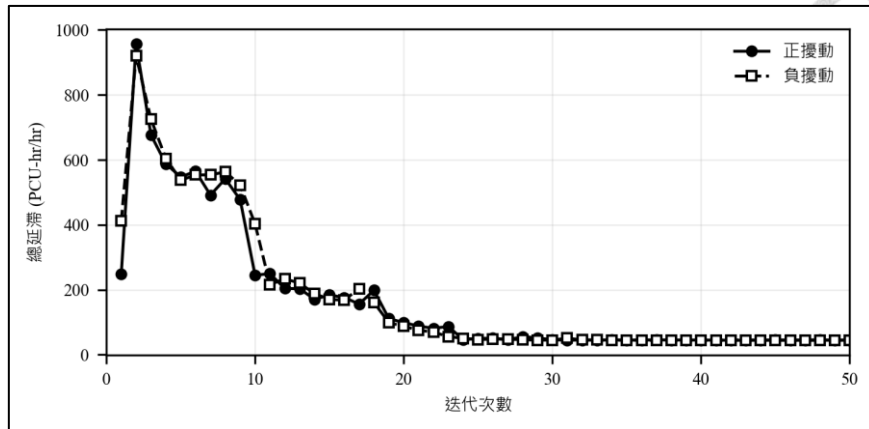


圖 4-16 旺角路網：結合不同變數型態之 SPSA（點 1） — $a = 1.44$

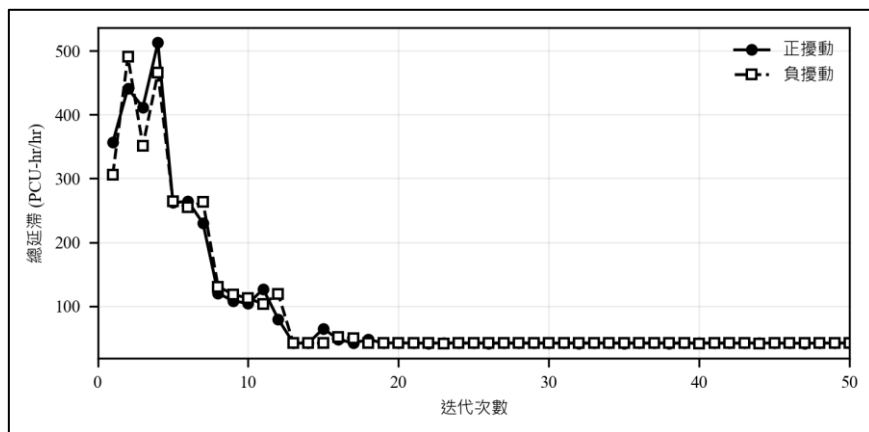


圖 4-17 旺角路網：結合不同變數型態之 SPSA（點 1） — $a = 1.8$

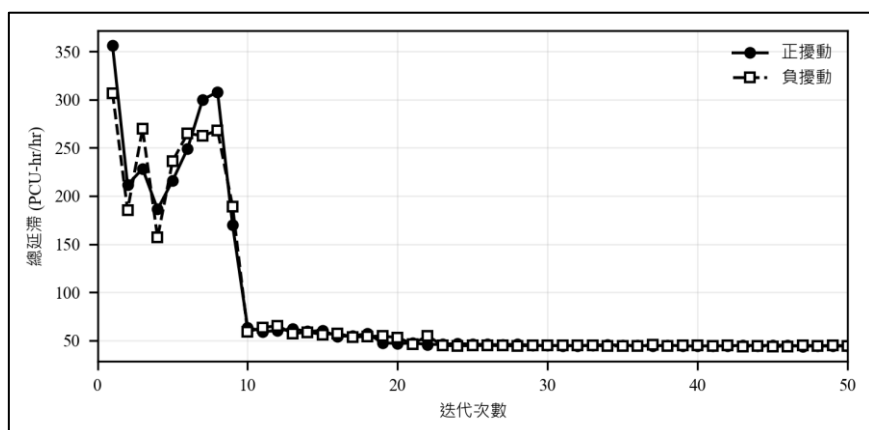


圖 4-18 旺角路網：結合不同變數型態之 SPSA（點 1） — $a = 2.16$



由以上 3 個參數的結果，可以找到每個結果包含正擾動與負擾動的整體最小值，將這 3 個最小值另外整理成長條圖來比較，如表 4-22 所示。

表 4-22 旺角路網：3 種不同的 a 值得到的最小值與求解時間表

	$a = 1.44$	$a = 1.8$	$a = 2.16$
迭代次數	50	50	50
迭代最小值 (總延滯 PCU-hr/hr)	43.99	41.63	43.92
求解時間 (sec)	63.77	69.06	64.54

從表 4-22 中可以看出，在迭代次數的部分，3 個不同的結果迭代次數都是 50 次，即 100 次模擬。而在迭代的最小值部分，當 $a = 1.8$ 時，整體的最小值會達到最小。在本情況中，因為只取了一個初始解及各取一個擾動作迭代，並未有完整統計證明 $a = 1.8$ 會是最好的結果。然而，因為本研究的时间有限，在後續旺角路網的分析中主要還是以 $a = 1.8$ 作為參考的迭代步長參數。¹⁶

4.3.2 區別變數型態之 SPSA

接續 4.3.1 節，使用 $a = 1.8$ 分別作點 1 與點 2 的區別變數型態之 SPSA，將時差與綠燈長分開做迭代。以下將使用 $a = 1.8$ 同時作為時差階段與綠燈長階段的迭代步長參數，其他參數則如 4.2.2 節所示。結果綜合整理如下表 4-23 所示。另外畫出結果 1 與結果 2 的函數改善過程，如圖 4-19 與圖 4-20 所示。

¹⁶ 後續研究可能要對 a 值的選擇做完整的統計分析。

表 4-23 旺角路網區別變數型態之 SPSA 結果整理

結果順序	初始解	時差 ^a	綠燈長 ^a	迭代最佳解 (總延滯 PCU-hr/hr)	迭代次數	求解時間 (sec)
1	點 1	1.8	1.8	40.75	35	43.93
2	點 2	1.8	1.8	43.44	43	50.74
3	點 1	1.8	2.88	40.32	30	37.13

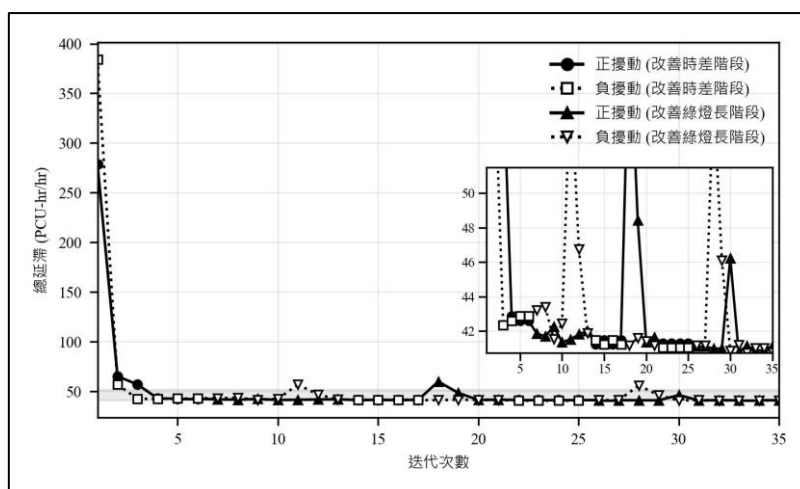


圖 4-19 旺角路網：區別變數型態之 SPSA 迭代過程（結果 1）

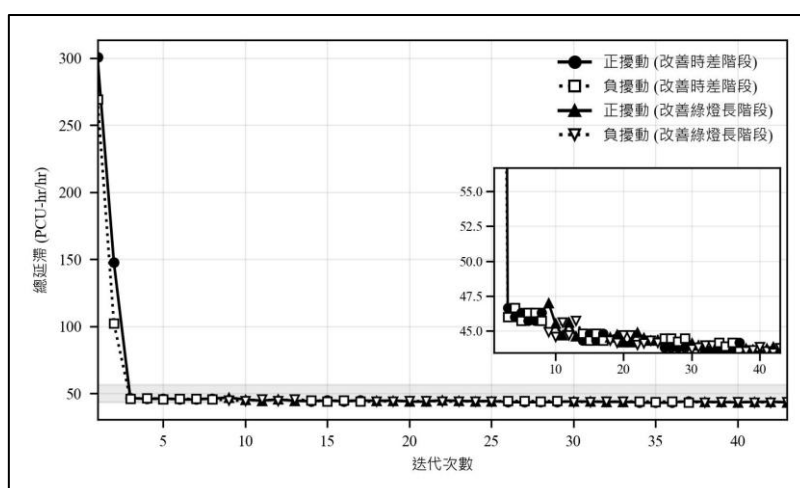


圖 4-20 旺角路網：區別變數型態之 SPSA 迭代過程（結果 2）

由以上圖 4-19 與圖 4-20 可以看出，比起結合不同變數型態之 SPSA，區別變數型態之 SPSA 可以更快速的到達較低點，並有利繼續往下搜尋找到一個更佳解。這是因為在決定初始變數時，已經一定程度了解綠燈長的大約秒數，如此一來，幾乎只要控制時差在最好的點，綠燈長僅需進行微調即可。從每次迭代的正擾動與負擾動的變數變化圖，以結果 1 為例，如圖 4-21、圖 4-22 所示，亦可得知。

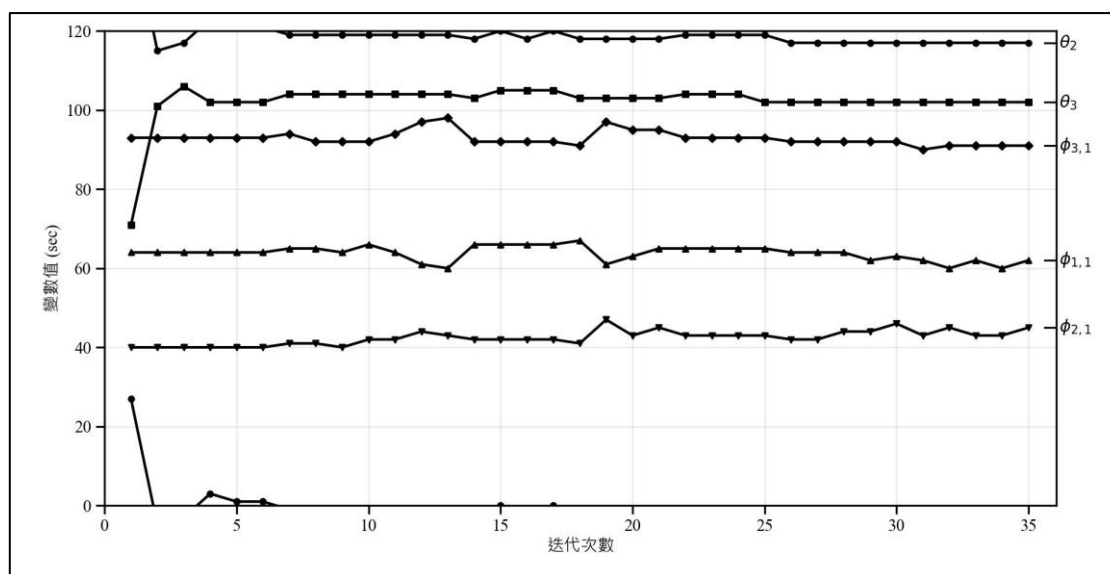


圖 4-21 旺角路網：結果 1 各變數值隨迭代變化（正擾動）

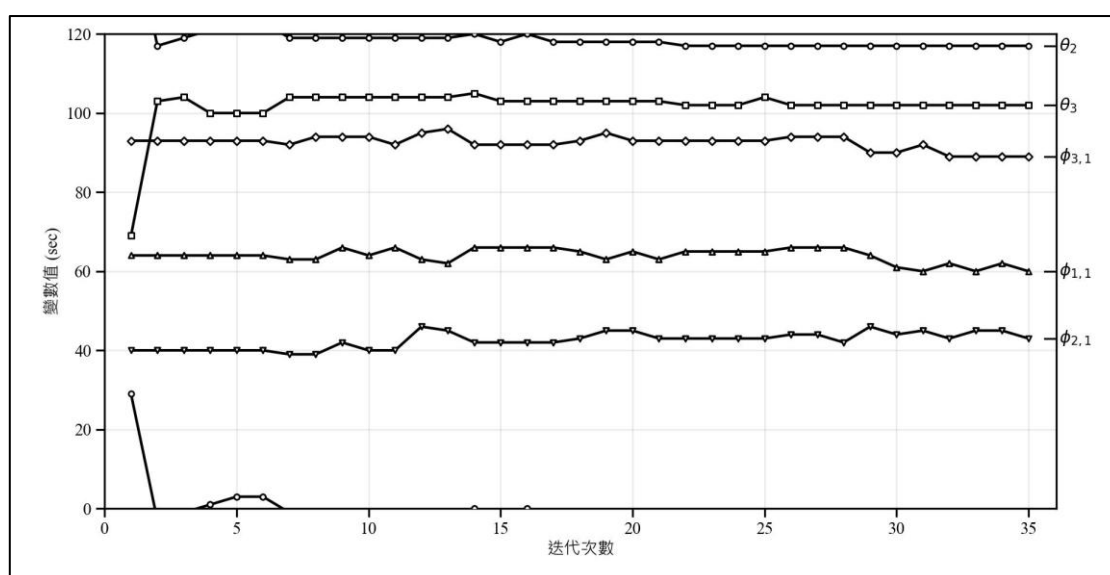


圖 4-22 旺角路網：結果 1 各變數值隨迭代變化（負擾動）

由於綠燈長變化量看起來沒有太大的改善，本研究另外調整綠燈長 $a = 2.88$ 對表現較好的初始解點 1 做一次新的迭代，即結果 3，其函數改善過程如圖 4-23 所示；其變數值隨迭代變化如圖 4-24、圖 4-25 所示。

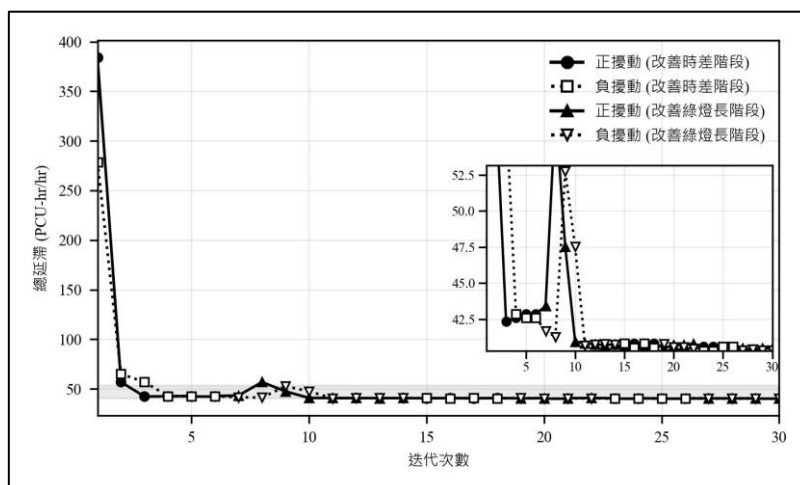


圖 4-23 旺角路網：區別變數型態之 SPSA 迭代過程（結果 3）

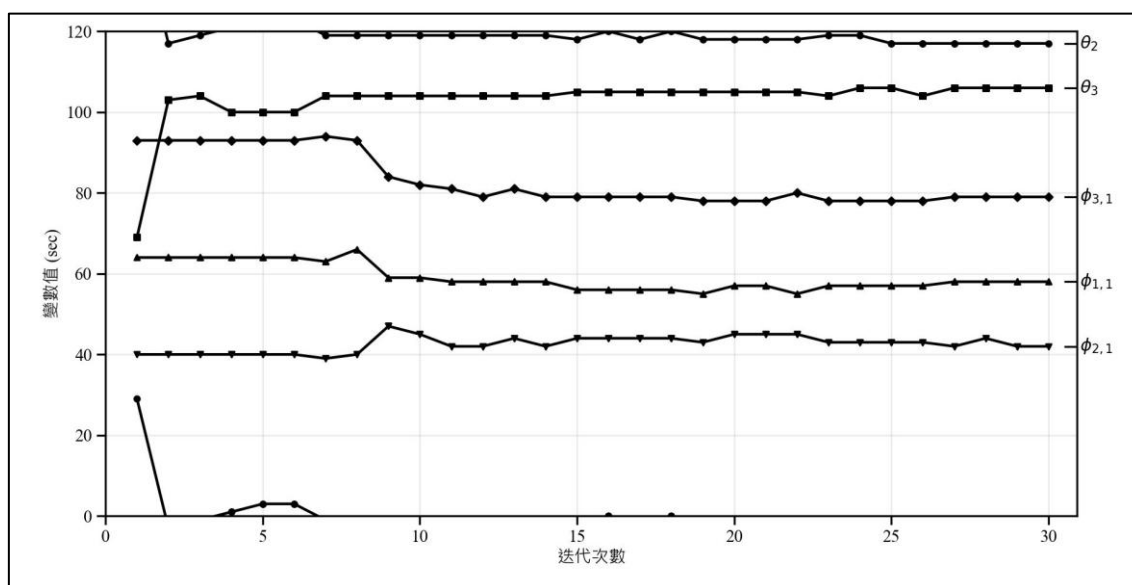


圖 4-24 旺角路網：結果 3 各變數值隨迭代變化（正擾動）

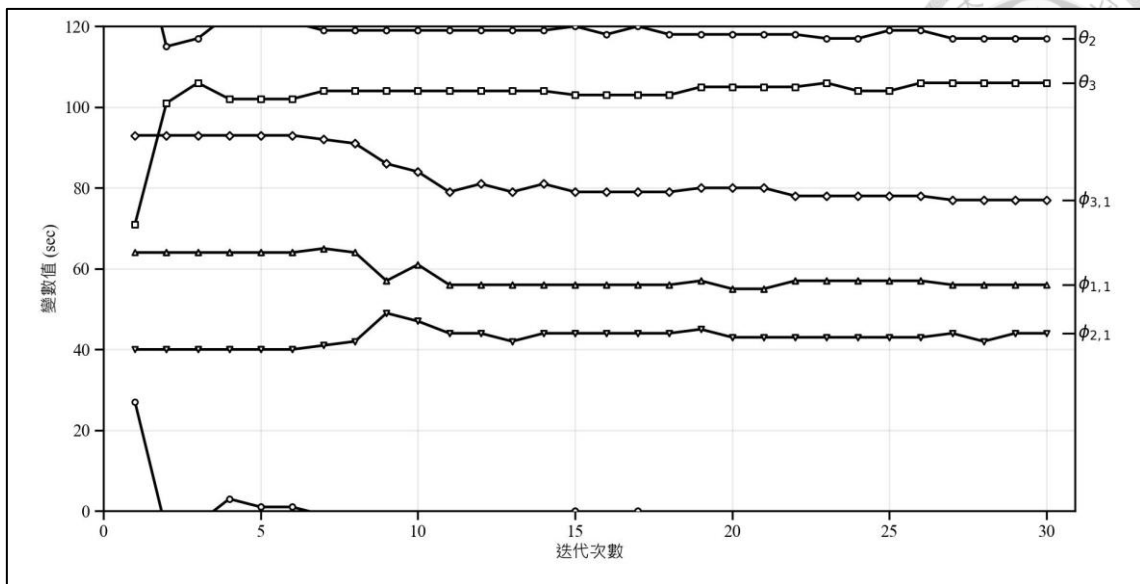


圖 4-25 旺角路網：結果 3 各變數值隨迭代變化（負擾動）

在以上圖 4-24 與圖 4-25 中，因為時差變數的週期性。如式(3.3)、式(3.4)所提到的，時差變數在 $[0, 120)$ 區間，當算出的值超出 120 時會自動除以 120 取餘數，因此時差變數會出現從 120 附近突然變到 0 附近（反之亦然）的現象。比較結果 1 與結果 3，結果 3 能夠找到更好的延滯，並且在改善綠燈長部分也有更相對明顯的改善，因此在本案例路網中，最終使用結果 3 作為區別變數型態之 SPSA 結果。其最佳時制 $\mathbf{x}^{\text{opt}} = [\boldsymbol{\theta}^{\text{opt}^T} \quad \boldsymbol{\phi}^{\text{opt}^T}]^T$, $\boldsymbol{\theta}^{\text{opt}} = [0 \quad 117]^T$, $\boldsymbol{\phi}^{\text{opt}} = [57 \quad 43 \quad 78]^T$ 。

4.3.3 不同演算法間的比較

在本案例中，為驗證本研究之演算法的最適性與收斂速度，比較結合不同變數型態 SPSA、區分變數型態 SPSA、最陡下降法、TRANSYT Hill Climbing 與本研究標竿方法 Net-GA (Lo 等人, 2001)，如表 4-24 所示。

在最陡下降法中，本研究亦選擇初始解 1 與初始解 2 分別執行，其主要是先利用差分法估計梯度後，使用該梯度的反方向（負梯度方向）作為方向，再應用黃金分割法做線性搜尋，以求得該方向上使函數值最小的步長，並重複該步驟直至收

斂。更詳細的演算法步驟與迭代過程詳見附錄 A。



表 4-24 旺角路網：不同演算法之比較

方法	迭代次數	模擬次數	迭代最小值 (總延滯 PCU-hr/hr)	求解時間 (sec)
結合不同變數型態 SPSA (初始解 1)	50	100	41.63	69.06
區分變數型態 SPSA (初始解 1)	30	60	40.32	37.13
最陡下降法 (初始解 1)	9	142	40.43	89.57
最陡下降法 (初始解 2)	15	244	40.44	146.63
TRANSYT Hill Climbing	--	--	39.33	6.44
Net-GA (10 min)	--	--	44.94	600
Net-GA (40 min)	--	--	37.31	2400

由上圖可以發現，若單純比較最陡下降法與本研究使用的 SPSA，可以更快地找到一個最佳解，並且相較於最陡下降也能找出一個更好的績效（總延滯）。而對於 TRANSYT Hill Climbing 與 Net-GA，本研究將其解出的最佳績效作為一個最適性指標，比較了與本研究結果的差異，並在求解時朝該指標改善。

4.4 分析結果—現代圓環

4.4.1 迭代步長參數搜尋

現代圓環路網以區別變數型態之 SPSA 作為演算法求解，所使用的共同參數亦如 4.2.2 節。然而，如同旺角路網，需要決定適當的迭代步長 α 以完成演算法。考慮式(2.1)，迭代步長的尺度要使梯度的估計值能夠有效地控制變化量。因為無法直接做到這點，且若像旺角路網直接使用結合不同變數型態 SPSA，可能會需要花

更多時間。因此本案例先試著用點 1 作為迭代初始解，選 4 組不同組合的 a 值，並使用績效最佳者作為後續 3 個初始解的 a 值來做迭代，4 組不同組合如所示。



表 4-25 現代圓環：不同組合 a 值的比較

組合順序	時差 a	綠燈長 a	迭代最小值 (總延滯 PCU-hr/hr)	迭代次數
1	0.5	0.5	209.85	16
2	1	1	171.15	99
3	2	1	168.36	40
4	5	1	178.46	38

表 4-25 描述了不同組的 a 值的迭代結果。其中組合 1 迭代次數明顯較少乃因過程中發現變數的更新速度需要使用更大一點的步長，即每次得出的 D_k^+ 與 D_k^- 有些差距卻無法反映在變數變化量上，因此提前取消迭代。而組合 3 能得到較好的結果，故後續使用組合 3（時差 $a = 2$ ，綠燈長 $a = 1$ ）來做迭代。

必須注意的是，上述結果僅各為某一組隨機擾動所產生的結果。也就是說，當使用不同的隨機擾動，是有可能發生與表 4-25 的截然不同的結果的。因此，若真的要決定最佳 a 值，可能須針對每種情況（甚至每個 a 值）使用大量不同的隨機擾動，並找到各個 a 值中何者的分布有較佳的績效，可能會比較嚴謹。然而，當需要快速求解時，此方法的效率必須被慎重考慮。



4.4.2 各初始解的迭代過程

決定好 a 值後，使用 3 個隨機初始解進行區別變數型態之 SPSA，並整理成表格如下表 4-26 所示。

表 4-26 現代圓環：三個不同初始解之迭代結果

初始解	時差 a	綠燈長 a	迭代最小值 (總延滯 PCU-hr/hr)	迭代次數	求解時間 (sec)
1	2	1	167.36	44	116.47
2	2	1	174.34	80	223.26
3	2	1	159.56	93	284.99

另外畫出各初始解之函數改善過程圖，如圖 4-26、圖 4-27、圖 4-28 所示。

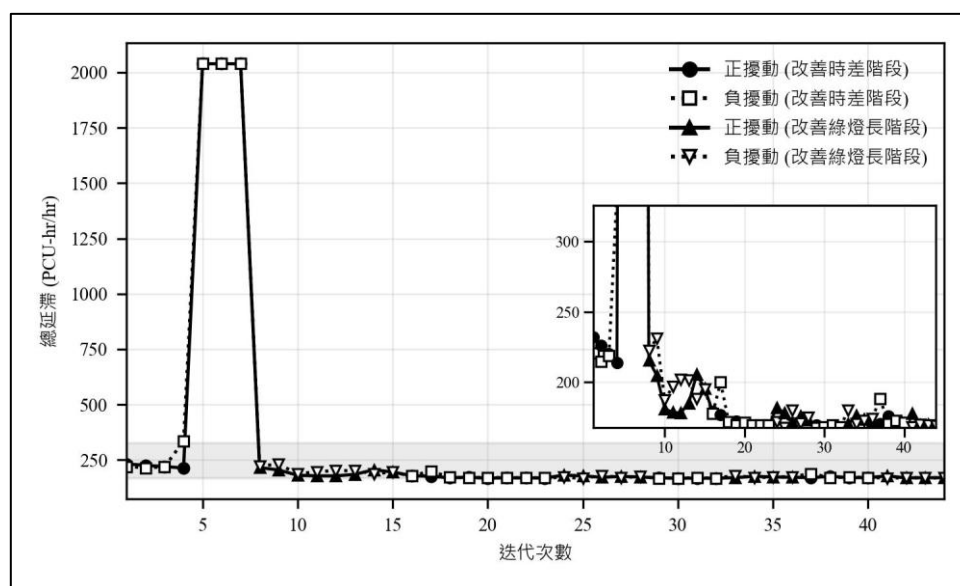


圖 4-26 現代圓環：初始解 1 迭代過程

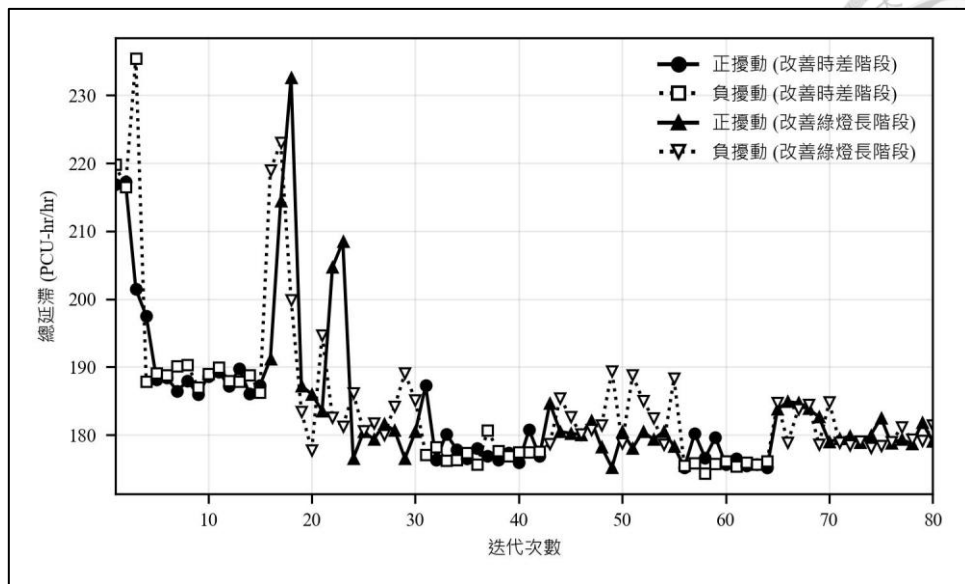


圖 4-27 現代圓環：初始解 2 迭代過程

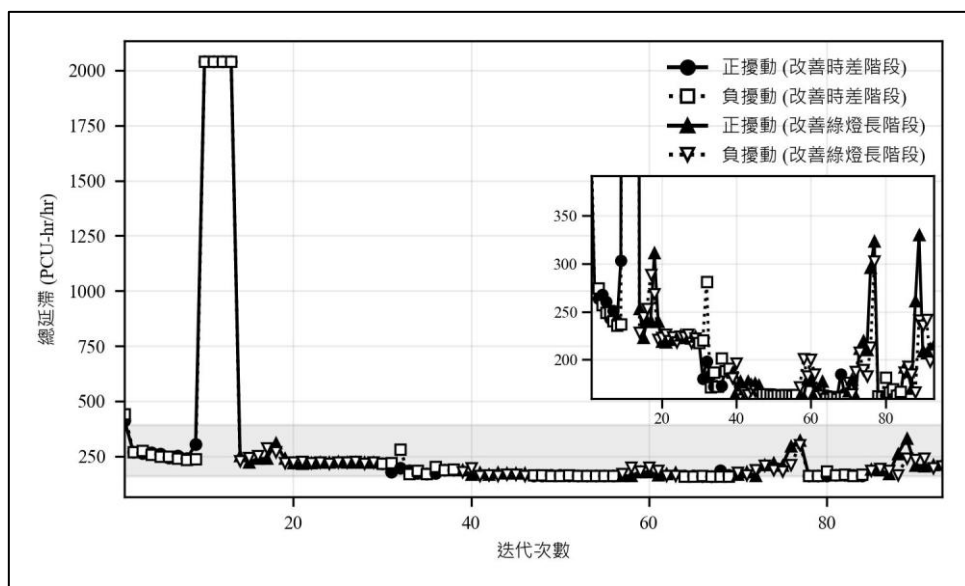


圖 4-28 現代圓環：初始解 3 迭代過程

因為有些初始解的結果被離群值影響刻度，故在旁邊加入一張放大圖以便識別較小範圍的總延滯變化。而造成這些離群值的原因則是當步長某一步跳太快時，可能會跳到一個非常差的解，但在本研究的演算法中，進入下一個階段前發現這些值並不是最佳解的時候就會自動忽略並進入下一階段，所以很快就會回到較合理

的總延滯值。

初始解 3 的最佳時制 $\mathbf{x}^k = [\boldsymbol{\theta}^{kT} \quad \boldsymbol{\phi}^{kT}]^T$, $\boldsymbol{\theta}^k = [30 \quad 39 \quad 6 \quad 88 \quad 83]^T$, $\boldsymbol{\phi}^k = [33 \quad 38 \quad 61 \quad 49 \quad 47 \quad 29]$ 。另外，也顯示出變數值隨迭代的變化，如圖 4-29、圖 4-30 所示。

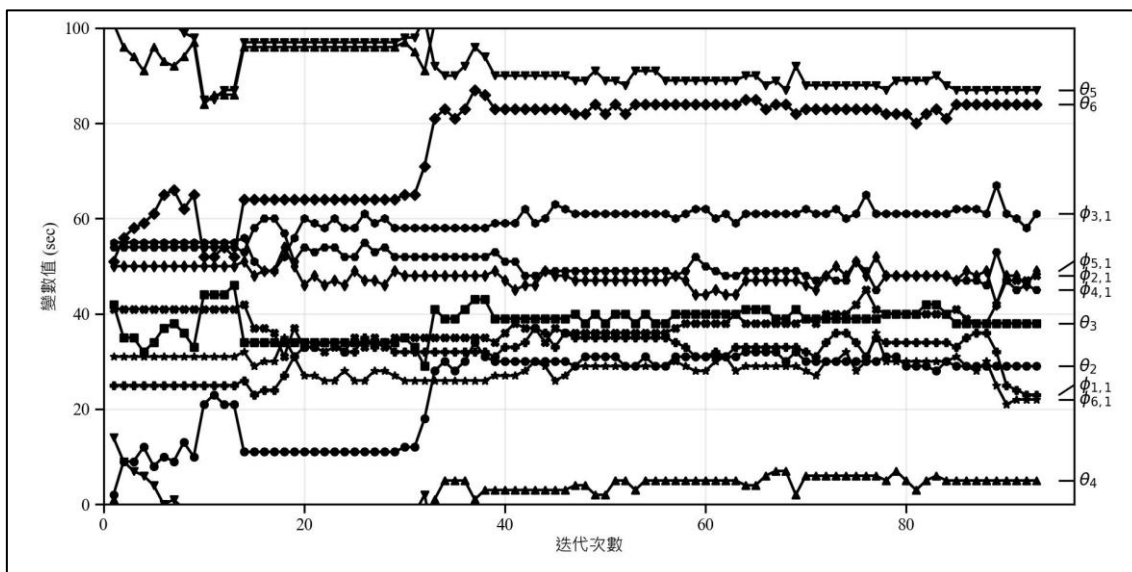


圖 4-29 現代圓環：初始解 3 各變數值隨迭代變化（正擾動）

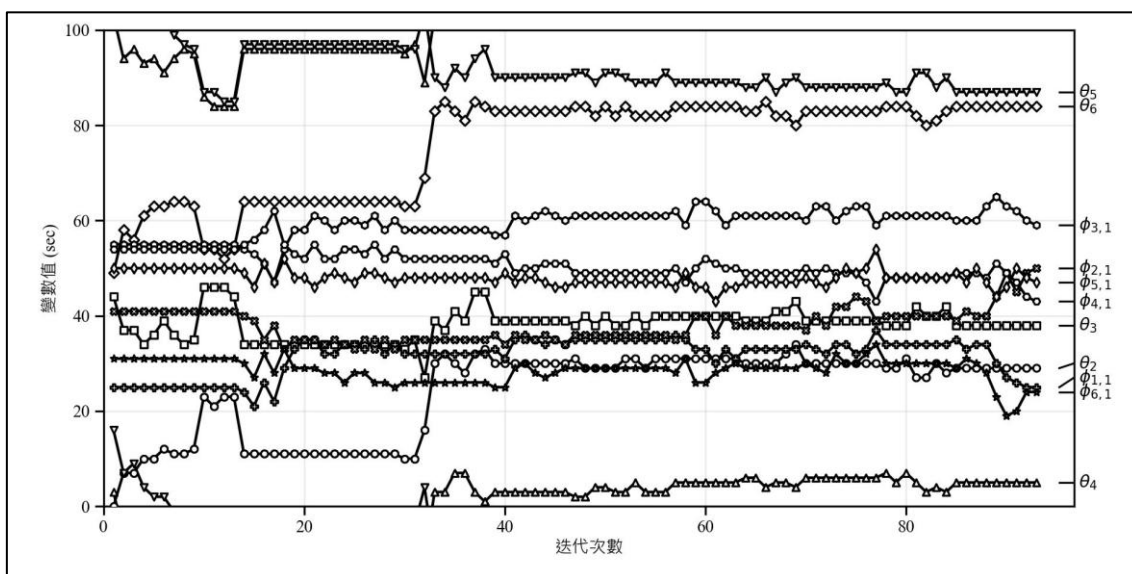


圖 4-30 現代圓環：初始解 3 各變數值隨迭代變化（負擾動）



4.4.3 不同演算法間的比較

本案例也使用初始解 3 之結果與最陡下降法以及 TRANSYT Hill Climbing 做比較。其中因為初始解 3 之最陡下降法結果較差，在本研究中多取了一個初始解 2 做最陡下降，有關最陡下降之步驟與迭代過程亦詳見附錄 A。表 4-27 為各演算法之結果比較。

表 4-27 現代圓環：不同演算法之比較

方法	迭代次數	模擬次數	求解時間 (sec)	迭代最小值 (總延滯 PCU-hr/hr)
區別變數型態之 SPSA (初始解 1)	44	88	116.47	167.36
區別變數型態之 SPSA (初始解 2)	80	160	223.26	174.34
區別變數型態之 SPSA (初始解 3)	93	186	284.99	159.36
最陡下降法 (初始解 1)	5	116	144.29	171.50
最陡下降法 (初始解 2)	5	112	157.70	173.87
最陡下降法 (初始解 3)	2	45	69.41	205.23
TRANSYT Hill Climbing	--	--	126.32	159.38

從表 4-27 可以得到 3 種方法的結果，其中 3 個初始解的樣本平均數為 167.02 PCU-hr/hr；樣本標準差為 7.50 PCU-hr/hr。最陡下降法不同於 SPSA 的估計梯度向量內所有值的絕對值皆相同，最陡下降因為梯度向量中的每個值都不同，且使用黃金分割法搜尋步長也需額外模擬，在變數變多時，每次迭代的模擬次數會增加。

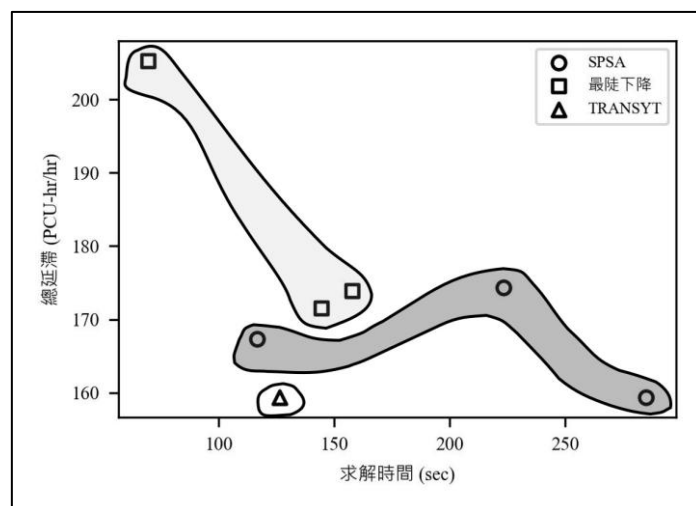


圖 4-31 現代圓環：不同演算法之求解時間與迭代最小值（總延滯）比較

如圖 4-31 所示，左下角代表各方法表現越佳，可以看出在這個案例中，雖然最陡下降的求解時間較短，但整體位於散佈圖的上方；反之雖區別變數型態 SPSA 需要較長的模擬時間，但整體位於散佈圖的下方，即能得到更小的總延滯值。

另外，在本案例中，因最陡下降法的參數固定¹⁷，因此當反覆重新迭代這 3 個初始解，若參數不調整，可以確定最佳績效（最小總延滯）是 171.50 PCU-hr/hr；然而，區別變數型態 SPSA 在不調整參數情況下，因每次隨機擾動會有不同結果，可能還有改善的空間。因此，在這個案例中，雖在收斂速度部分較無法與最陡下降來比較，但在最適性部分區別變數型態 SPSA 可以有更好的表現。比較兩者的迭代最小值，區別變數型態 SPSA 較最陡下降改善了 7.08%。

4.4.4 隨機擾動與演算法的收斂性

因為隨機擾動的隨機性，就算使用同一個初始解做迭代，也可能產生不一樣的

¹⁷ 僅有對於步長線性搜尋時的最大範圍與最小範圍，以及差分所需要的小變化量（因變數為整數，故都設為 1）。

結果。因此，在本案例路網中，嘗試了在同一個初始解的 20 個不同隨機擾動下的機率分布，一方面可以驗證演算法的收斂性，另外也可以得知其最適性。

表 4-28 取 20 個不同的初始解，列出其在第 20 次迭代時的最佳解與最終結果的最佳解，共 40 筆資料。再取其最大值與最小值，分成 10 段，計算每段長度為 17.11 (PCU-hr/hr)，並記錄各段的機率密度函數值，最終整理成直方圖如圖 4-32、圖 4-33 所示。

另外，亦使用核密度估計 (Kernel Density Estimation, KDE)¹⁸，估計平滑的機率密度函數，並繪製其與迭代次數的關係圖，第 20 次迭代之最佳與最終迭代結果比較如圖 4-34 所示。

¹⁸ 使用 KDE 需要決定核 (kernel) 與帶寬 (bandwidth)，其中核使用高斯核 (Gaussian Kernel)，帶寬 (bandwidth) 則以 Scott (1992) 方法決定，並且使用全部 40 筆資料決定共同帶寬，其值約為 14.21。



表 4-28 現代圓環：初始解 3 各不同隨機擾動與其迭代結果

隨機擾動	第 20 次迭代最佳 (總延滯 PCU-hr/hr)	最終迭代結果 (總延滯 PCU-hr/hr)
1	219.15	159.56
2	202.94	175.99
3	224.48	141.48
4	176.25	173.57
5	194.77	184.45
6	219.35	180.05
7	224.85	145.78
8	180.84	166.78
9	171.38	168.25
10	223.48	178.55
11	190.33	187.41
12	175.41	172.80
13	208.95	186.00
14	312.61	170.74
15	238.10	188.65
16	190.52	182.70
17	176.65	147.53
18	196.88	186.49
19	198.46	176.56
20	187.33	167.75

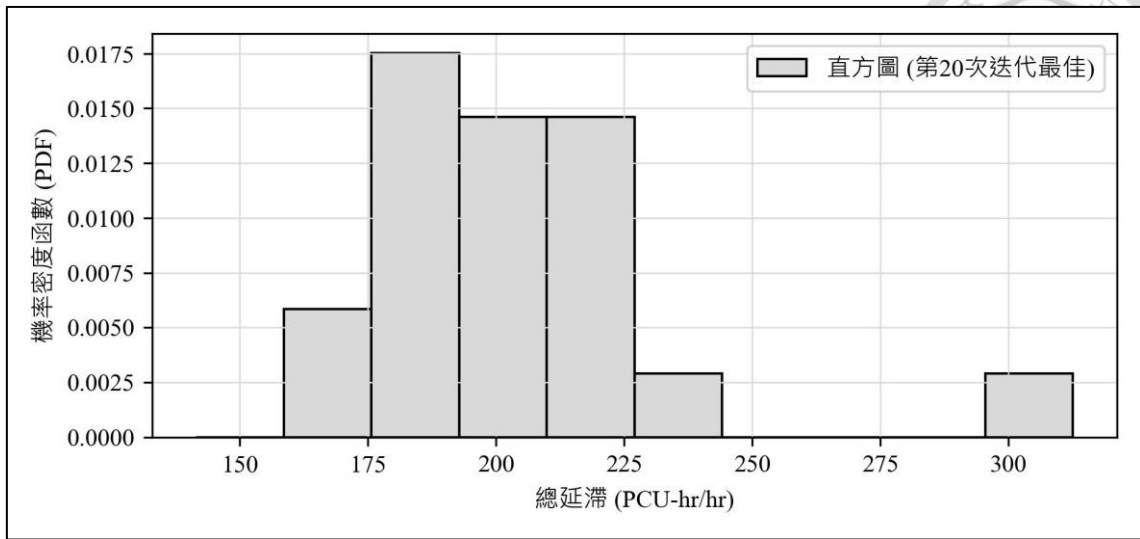


圖 4-32 現代圓環：初始解 3 不同隨機擾動第 20 次迭代後最佳結果直方圖

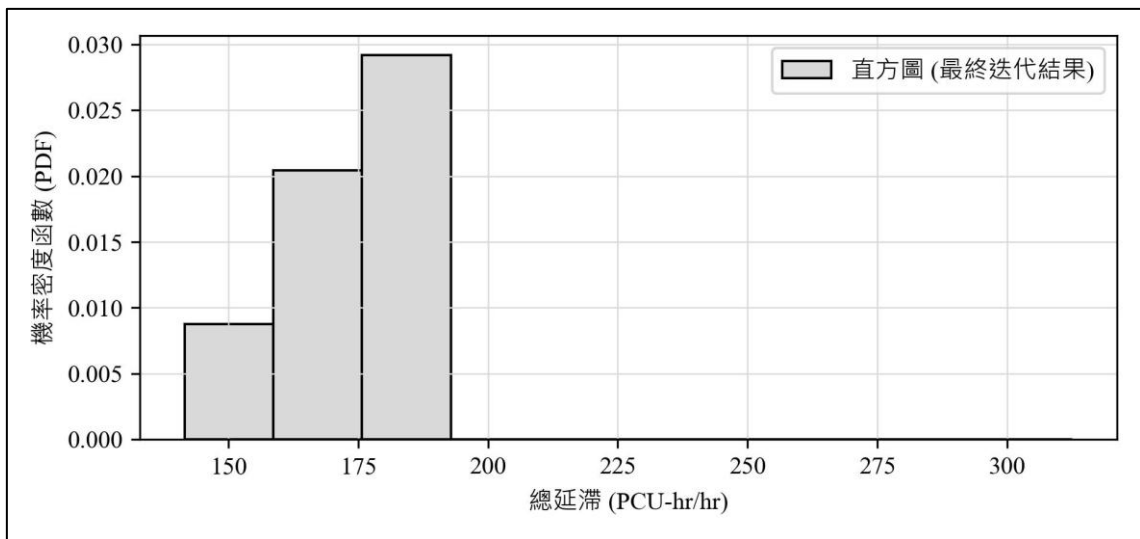


圖 4-33 現代圓環：初始解 3 不同隨機擾動最終迭代結果直方圖

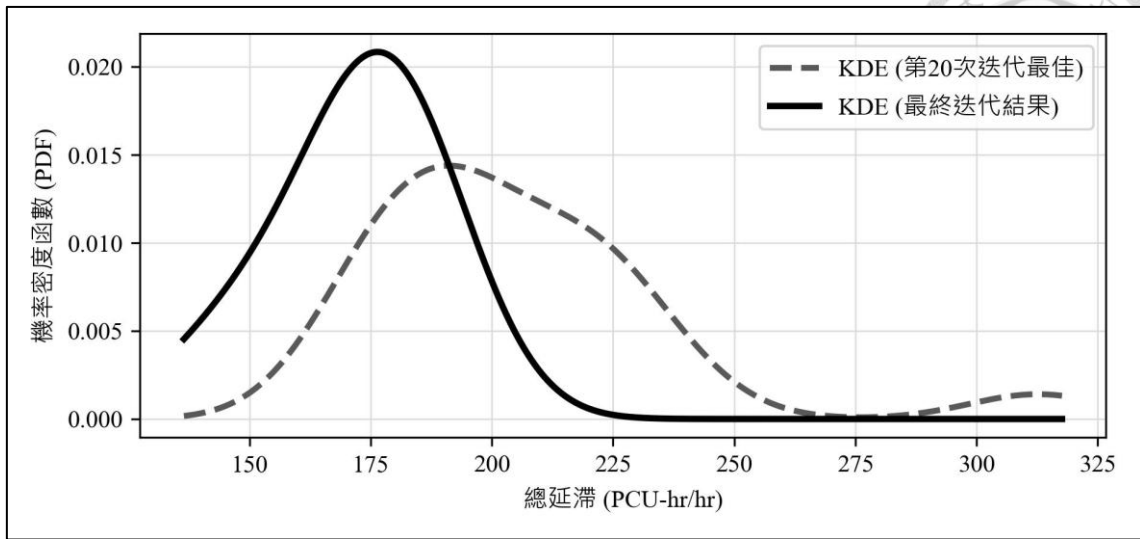


圖 4-34 現代圓環：初始解 3 不同隨機擾動之 KDE 結果比較

從圖 4-34 可以看出，最終迭代結果的最佳解相比第 20 次迭代結束時的最佳解，曲線往左上移動，亦即最終迭代結果的績效變佳，且得到好的績效的機率也變佳。這也驗證本演算法的收斂性，對於一個初始解，取不同的隨機擾動，會有進步的趨勢。

表 4-29 現代圓環：初始解 3 不同隨機擾動之平均數、標準差與變異係數

	樣本平均數	樣本標準差	變異係數
第 20 次迭代結果	205.64	31.94	15.53%
最終迭代結果	172.05	14.09	8.19%

表 4-30 現代圓環：初始解 3 不同隨機擾動之四分位數與四分位距

	第一四分位數 (Q1)	中位數	第三四分位數 (Q3)	四分位距 (Q3 - Q1)
第 20 次迭代結果	185.71	197.67	220.38	34.67
最終迭代結果	167.51	174.78	183.14	15.63



由表 4-29 與表 4-30 可看出，標準差從 31.94 降到 14.09；變異係數從 15.53% 降到 8.19%；四分位距從 34.67 降到 15.63。另外在圖 4-34 也能看出，最終迭代結果曲線的峰值較高且曲線較窄，而第 20 次迭代結果曲線較平坦。以上結果也代表在不同隨機擾動下，演算法在第 20 次迭代的結果分布較為分散，而隨迭代進行，最終結果的總延滯值會變得相對穩定。

4.5 分析結果—雙幹道路網

在雙幹道路網中，主要測試區別變數型態 SPSA 在多幹道的路網上是否能夠順利運作。本案例將取 5 個初始解做區別變數型態 SPSA，並在最後與 TRANSYT Hill Climbing 做比較。最大迭代次數為 100 次，迭代步長參數亦由簡易測試決定。與其他路網不同的是，本案例將使用 TRANSYT 中的 PDM（車隊擴散模型，Platoon Dispersion Model）（Robertson, 1969）來做模擬，而非旺角路網與現代圓環使用的 CTM（格位傳遞模型，Cell Transmission Model）（Daganzo, 1994; 1995）。主因是當使用 CTM 時，若路網出現過飽和情況，此時路口車道的車輛會往後溢流（Spillover），導致總延滯大幅增加。而當模擬大路網且若 D_k^+ 與 D_k^- 在未飽和與過飽和之間移動時，函數值可能變化很大，但卻不是函數變化的方向或誤判變數變化量的大小。因此，在本案例中，使用只考慮路口行為的 PDM，不將 CTM 的溢流效果納入考慮。

若同時以 TRANSYT Hill Climbing 解此路網，PDM 的求解時間為 133.58 秒，最小總延滯為 187.88 PCU-hr/hr，而 CTM 的求解時間為 4279.43 秒，最小總延滯為 273.04 PCU-hr/hr，兩者在最適性與收斂速度都有一定程度的差異，且若用 CTM 模擬 Hill Climbing 解出的 PDM 時制，其值會是 226.89 PCU-hr/hr¹⁹，還是比直接用

¹⁹ 該時制用 CTM 模擬之路口最大飽和度為 100.35%，可能有少數溢流導致總延滯相較 PDM（算出最大飽和度 88%）高上一些。

CTM 求解績效更佳。因此在後續本案例中，總延滯都將會是使用 PDM 得出的解。



4.5.1 迭代步長參數搜尋

在本案例中，迭代步長參數 a 僅憑模擬時梯度估計值與變數變化量的尺度差異來選擇。在時差迭代階段，因梯度估計值偏小，時差迭代需要較大的步長才能使變數有較明顯的變化，進而產生較佳的績效；反之在綠燈長迭代時若步長太大，可能會使梯度估計量變很大，為避免變數變化太大而讓績效變差，綠燈長階段的 a 值也相對較小，因為這部分參數決定並布嚴謹，可能需後續研究的進一步驗證。

本案例最終使用初始解 1 比較兩組不同的 a 值，中間經過一些零星測試，在發現使用的 a 值效果不佳，即立即更新新的 a 值，最終有完整結果的有兩個，如表 4-31 所示。由這兩者中選當次迭代較佳者，即組合 1，時差 $a = 25$ ，綠燈長 $a = 0.5$ 來做後續迭代。

表 4-31 雙幹道路網：不同組合 a 值的比較

組合順序	時差 a	綠燈長 a	迭代最小值 (總延滯 PCU-hr/hr)	迭代次數
1	25	0.5	193.45	59
2	25	1	196.04	53

4.5.2 各初始解的迭代過程

5 個初始解使用上節 a 值的結果如所示，另外因為表 4-31 使用初始解 1 搜尋迭代步長參數，該解直接作為初始解 1 之迭代結果。

表 4-32 雙幹道路網：五個不同初始解之迭代結果

初始解	時差 a	綠燈長 a	迭代次數	求解時間 (sec)	迭代最小值 (總延滯 PCU-hr/hr)
-----	--------	---------	------	---------------	--------------------------

1	25	0.5	59	1538.07	193.45
2	25	0.5	75	1919.93	193.21
3	25	0.5	66	1710.20	194.73
4	25	0.5	100	2620.15	185.49
5	25	0.5	47	1224.89	211.18

由於此路網規模較大，在模擬時需要花更多的時間，因此求解時間變得也比較長，其中 5 個初始解的樣本平均數為 195.61 PCU-hr/hr；樣本標準差為 9.43 PCU-hr/hr。函數值隨迭代改善過程如圖 4-35 至圖 4-39 所示。

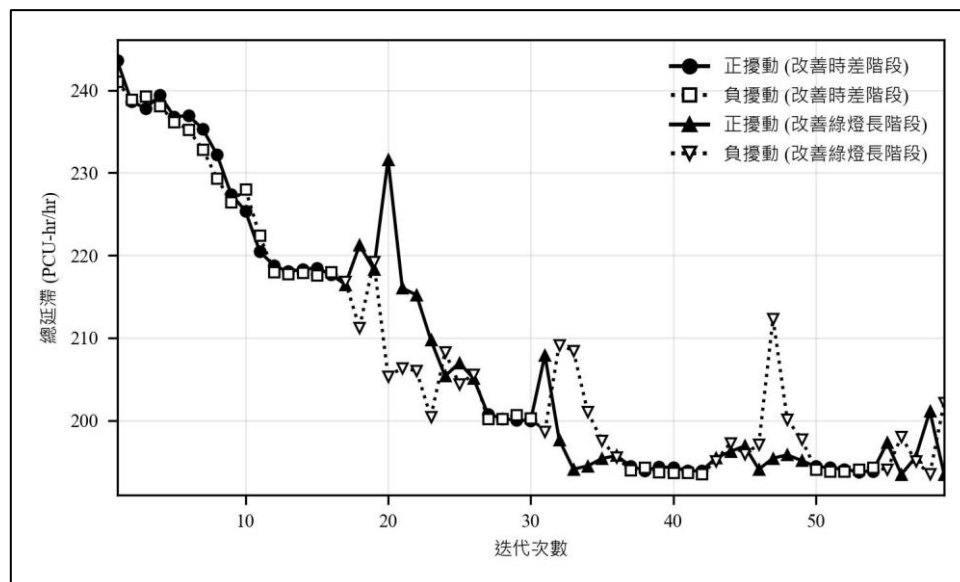


圖 4-35 雙幹道路網：初始解 1 迭代過程

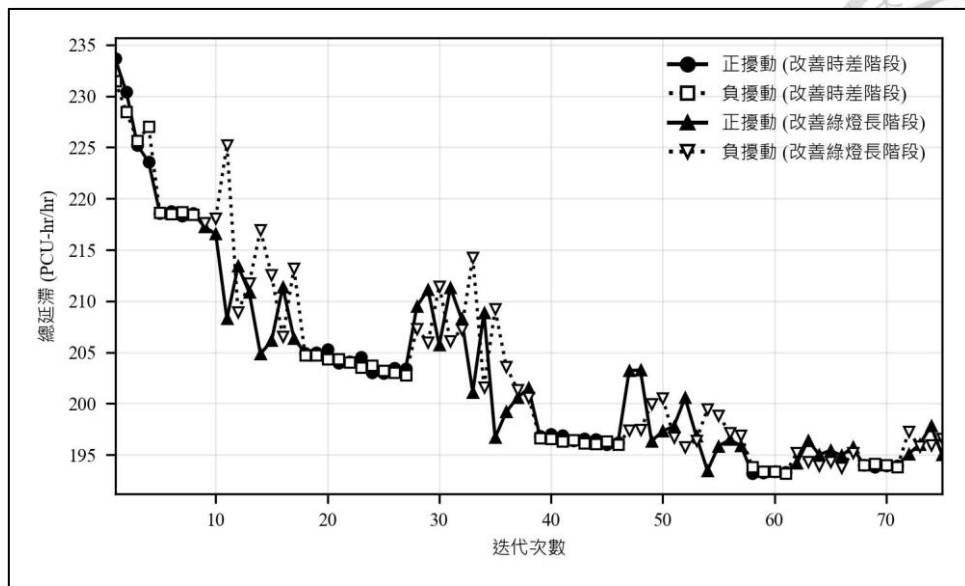


圖 4-36 雙幹道路網：初始解 2 迭代過程

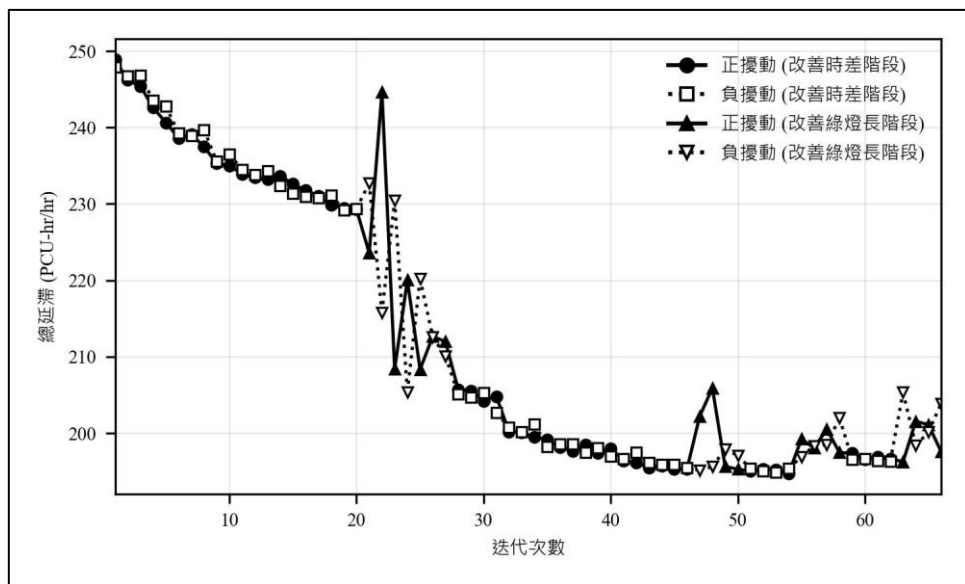


圖 4-37 雙幹道路網：初始解 3 迭代過程

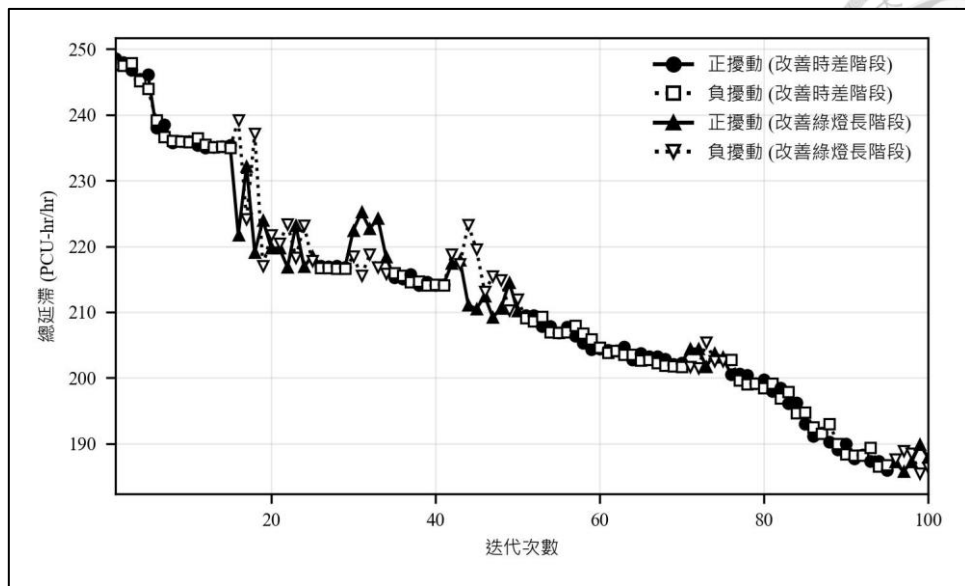


圖 4-38 雙幹道路網：初始解 4 迭代過程

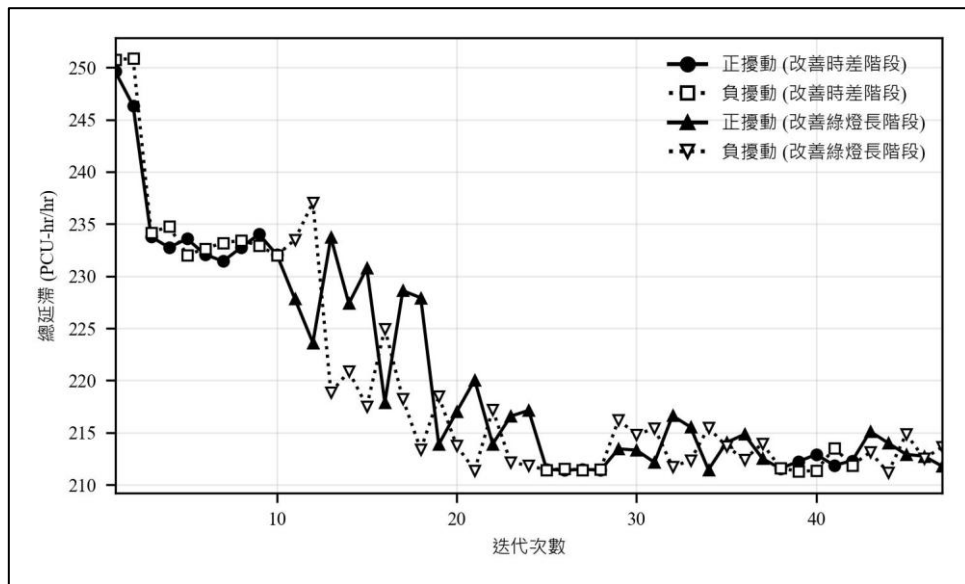


圖 4-39 雙幹道路網：初始解 5 迭代過程

由以上 5 張圖可以看出，在改善時差階段時，正擾動與負擾動函數值差距不大；而改善綠燈長階段時，正擾動與負擾動函數值差距擴大，也表示兩者的 α 值需要有一定尺度上的差異才能讓變數有適當的變化量。

其中，因初始解 4 找到的解最佳，且完成 100 次最大迭代限制，故後續將使用初始解 4 作為分析對象。另外，初始解 4 的求得之最佳時制如表 4-33 所示。



表 4-33 雙幹道路網：初始解 4 之最佳時制

	最佳時制 (起始點4)
θ_2	98
θ_3	60
θ_4	58
θ_5	18
θ_6	107
θ_7	102
θ_8	66
θ_9	108
$\phi_{1,1}$	48
$\phi_{1,2}$	15
$\phi_{2,1}$	62
$\phi_{2,2}$	14
$\phi_{3,1}$	46
$\phi_{3,2}$	22
$\phi_{4,1}$	54
$\phi_{4,2}$	10
$\phi_{5,1}$	31
$\phi_{5,2}$	10
$\phi_{6,1}$	89
$\phi_{7,1}$	28
$\phi_{7,2}$	11
$\phi_{8,1}$	36
$\phi_{8,2}$	10
$\phi_{8,3}$	23
$\phi_{9,1}$	24
$\phi_{9,2}$	64

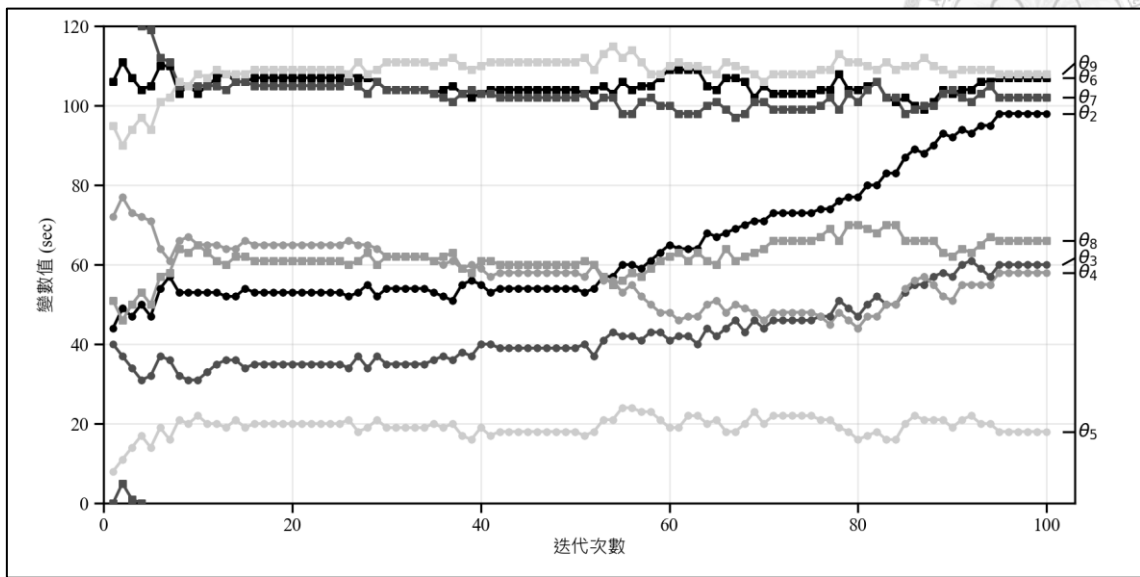


圖 4-40 雙幹道路網：初始解 4 時差變數值隨迭代變化（正擾動）

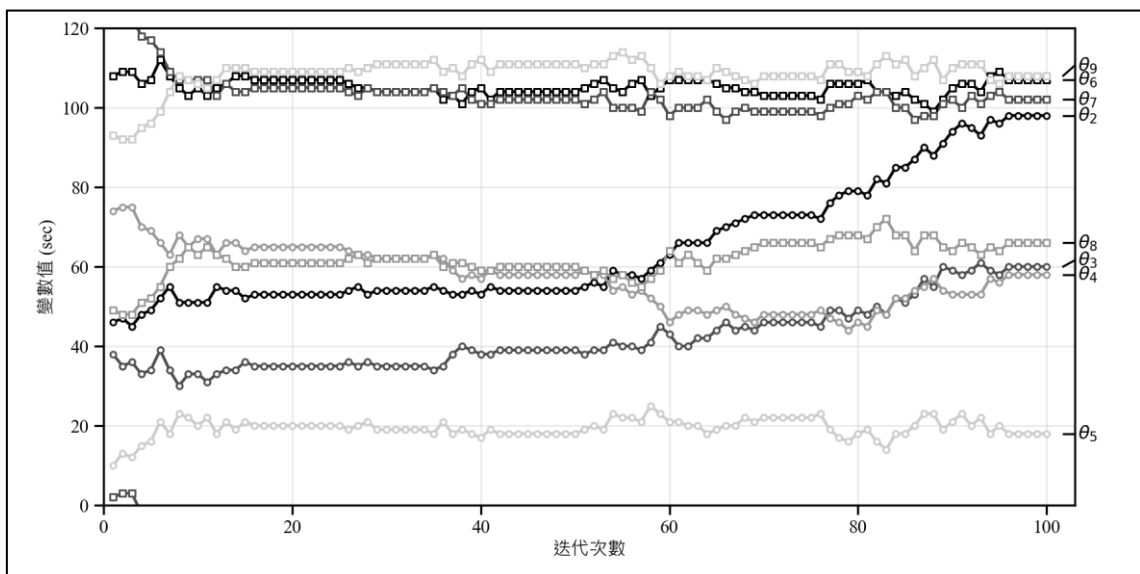


圖 4-41 雙幹道路網：初始解 4 時差變數值隨迭代變化（負擾動）

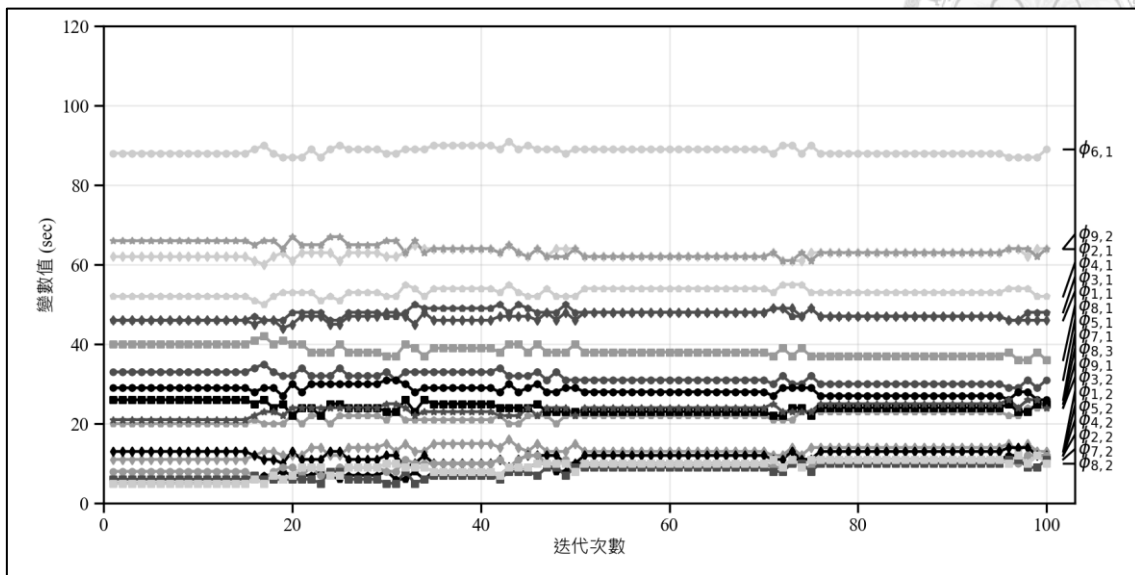
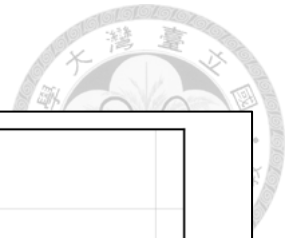


圖 4-42 雙幹道路網：初始解 4 綠燈長變數值隨迭代變化（正擾動）

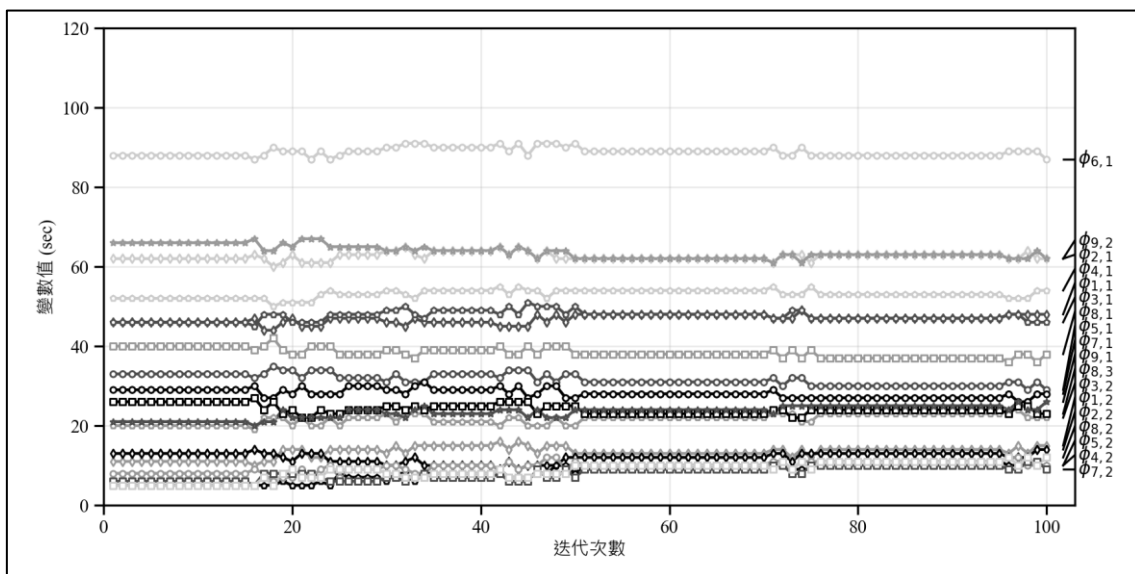


圖 4-43 雙幹道路網：初始解 4 綠燈長變數值隨迭代變化（正擾動）

圖 4-40 至圖 4-43 分別畫出初始解 4 時差與綠燈長的正擾動與負擾動變數隨迭代變化過程。可以得知的是，雖然改善時差階段的函數值變化較小，但時差變數仍有一定的變化；反而在函數值變化大的綠燈長部分，變化並未非常明顯。



4.5.3 不同演算法間的比較

因本路網主要測試能否應用在多幹道路網中，並未強調在與最陡下降的比較，因此，本路網之結果將僅與 TRANSYT Hill Climbing 做一個比較，如表 4-34 所示。

表 4-34 雙幹道路網：不同演算法之比較

方法	迭代次數	求解時間 (sec)	迭代最小值 (總延滯 PCU-hr/hr)
區別變數型態之 SPSA (最佳初始解/點 4)	100	2620.15	185.49
TRANSYT Hill Climbing	--	133.58	187.88

表 4-34 列出 5 個初始解中表現最好的初始解 4 做為比較對象。由於剛好達到迭代次數上限，可以得知迭代 100 次（模擬 200 次）約需要 2620.15 秒。雖然區別變數型態之 SPSA 能得到的績效較 TRANSYT Hill Climbing 佳，但區別變數型態之 SPSA 求解時間需要將近 Hill Climbing 的 20 倍，這也是未來需要進步的方向之一。

4.6 小結

本研究共分析了三個不同大小的路網。旺角路網主要比較了結合所有變數之 SPSA 與區別變數型態之 SPSA，另外也與最陡下降法進行了比較，發現其最適性與收斂速度都比最陡下降法來得佳；在現代圓環的部分，雖也比較了最陡下降法與區別變數型態之 SPSA，但整體求解時間略劣於最陡下降法，但也發現在最適性部分區別變數型態之 SPSA 明顯優於最陡下降法。另外也做了 20 個不同隨機擾動的統計分析，發現區別變數型態之 SPSA 的績效會隨迭代漸漸改善；最後是雙幹道路網，可以得知區別變數型態之 SPSA 可以用在多幹道的路網中，然而求解時間也還需要再做改善。

第五章 結論與建議



5.1 結論

本研究建立了一個區別變數型態的 **SPSA**，提供限制式處理方法與初始解搜尋，並使用了三個不同尺度的路網分析 **SPSA** 在號誌最佳化中的表現，可整理為以下結論：

- 一、相比於最陡下降法每次使用時同一個初始解會得到相同的值，本研究之號誌最佳化方法在同一個初始解但不同隨機擾動下結果可能較為相異。但也因在迭代結果上不如最陡下降法穩定，反而更有可能在非凸函數（**non-convex function**）之複雜問題下找到更佳之解。
- 二、在三個不同大小且性質各異的路網中，使用本研究之 **SPSA** 方法可以得到與 **TRANSYT Hill Climbing** 或旺角路網的 **Net-GA** 接近的結果。
- 三、以本研究提出之區別變數型態 **SPSA** 與最陡下降法相比，旺角路網可節省 **58.5%** 的求解時間，惟最適性相似（僅改善 **0.272%**）；號誌化現代圓環雖在求解時間較最陡下降法長，但可改善最佳解 **7.08%** 以上。
- 四、最陡下降法每次迭代需完成多次模擬，且隨著變數增多模擬次數也會變多，函數改善與否需要在完成一整次的模擬才能得知。而本研究之 **SPSA** 方法因每次迭代僅需 **2** 次模擬，每 **2** 次模擬就能得到一個新的改善方向，回饋較為及時。

另外並綜整本研究之貢獻如下：

- 一、本研究提供了一個基於梯度估計的號誌最佳化方法，透過不同的隨機擾動驗證此方法具有收斂性，可做為一個有效的號誌最佳化方法使用。
- 二、本研究的區別變數型態之 **SPSA** 考慮了時差與綠燈長的尺度可能有所差異，將其分開改善；且在迭代過程中藉由階段性的改善，當函數值太大且無法下降

回來時（例如到達模擬極限時績效可能停滯），下一階段的更新可以從函數值變大前重新開始，即可儲存一部分記憶。

三、本研究提供了一個改善的 **SPSA**，給予一些迭代次數或收斂閾值限制讓演算法有機制提早結束。另外，給予一組限制式處理的方法讓模擬結果始終有效。

四、多幹道的交通控制較難以以數學模型計算，而本研究提供之 **SPSA** 使用基於模擬之最佳化，並可套用在多條幹道的路網中，為號誌最佳化提供了一個選項。

5.2 建議

對於區別變數型態 **SPSA** 與案例分析結果，本研究提出以下建議供後續研究參考：

一、對於本研究提出之區別變數型態 **SPSA**，其在收斂速度上仍有進步空間。建議未來可針對求解速度進行改良，以提升效率。另外，也可比較最陡下降法與本研究之 **SPSA** 在相同求解時間限制的情況下的績效差異。

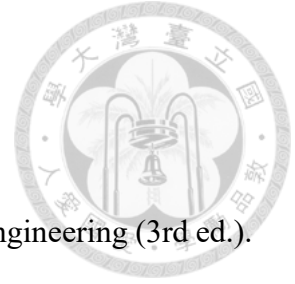
二、對於迭代步長參數（ α 值），由於本研究未能找到具確定性的數學公式，建議透過多次實驗累積經驗公式或進行敏感性分析，以獲得更穩健的設定依據。演算法內的迭代次數限制與收斂閾值亦同。

三、由於本研究在不同隨機擾動下所得結果還是有一些差異，建議後續研究可考慮引入減少隨機性的方法，以提升結果穩定性。另外也建議未來可加入二階資訊，利用函數曲率資訊改進搜尋方向，進而提升收斂效率。

四、本研究使用假想或簡易的路網案例來做分析，建議未來可加入國內案例分析。

五、本研究使用 **TRANSYT** 進行模擬，惟因其為確定性（**deterministic**）巨觀模擬軟體，僅對給定的道路參數進行分析，未考慮路網內部的隨機性，建議未來可再加入微觀模擬分析路網有隨機性時演算法是否適用。

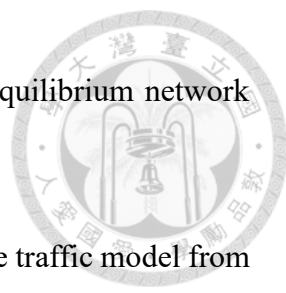
參考文獻

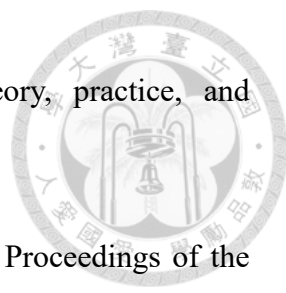


- (1) Roess, R. P., Prassas, E. S., & McShane, W. R. (2004). Traffic engineering (3rd ed.). Pearson Prentice Hall.
- (2) Transportation Research Board. (2010). Highway capacity manual (5th ed.). Washington, DC: National Research Council.
- (3) Fu, M. C. (2002). Optimization for simulation: Theory vs. practice. *INFORMS Journal on Computing*, 14(3), 192-215.
- (4) Ceylan, H., & Bell, M. G. (2004). Traffic signal timing optimisation based on genetic algorithm approach, including drivers' routing. *Transportation Research Part B: Methodological*, 38(4), 329-342.
- (5) Stevanovic, A., Martin, P. T., & Stevanovic, J. (2007). VisSim-based genetic algorithm optimization of signal timings. *Transportation Research Record*, 2035(1), 59-68.
- (6) Wallace, C. E., Courage, K. G., Reaves, D. P., Schoene, G. W., & Euler, G. W. (1984). *TRANSYT-7F user's manual* (No. UF-TRC-U32 FP-06/07).
- (7) Cauchy, A. (1847). Méthode générale pour la résolution des systemes d'équations simultanées. *Comp. Rend. Sci. Paris*, 25(1847), 536-538.
- (8) Wright, S., & Nocedal, J. (1999). Numerical optimization. *Springer Science*, 35(67-68), 7.
- (9) Spall, J. C. (1987, June). A stochastic approximation technique for generating maximum likelihood parameter estimates. In *1987 American control conference* (pp. 1161-1167). IEEE.

- 
- (10) TRL Software. (2024). TRANSYT 17 User Guide (Issue A). TRL Software.
- (11) Himmelblau, D. M. (1972). *Applied Nonlinear Programming*. McGraw-Hill.
- (12) Zoutendijk, G. (1970). Nonlinear programming, computational methods. *Integer and nonlinear programming*, 37-86.
- (13) Armijo, L. (1966). Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific Journal of mathematics*, 16(1), 1-3.
- (14) Wolfe, P. (1969). Convergence conditions for ascent methods. *SIAM review*, 11(2), 226-235.
- (15) Dennis Jr, J. E., & Schnabel, R. B. (1996). *Numerical methods for unconstrained optimization and nonlinear equations*. Society for Industrial and Applied Mathematics.
- (16) Robbins, H., & Monro, S. (1951). A stochastic approximation method. *The annals of mathematical statistics*, 400-407.
- (17) Kiefer, J., & Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function. *The Annals of Mathematical Statistics*, 462-466.
- (18) Spall, J. C. (1992). Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE transactions on automatic control*, 37(3), 332-341.
- (19) Spall, J. C. (1998). Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Transactions on aerospace and electronic systems*, 34(3), 817-823.

- 
- (20) Zhu, X., & Spall, J. C. (2002). A modified second - order SPSA optimization algorithm for finite samples. *International Journal of Adaptive Control and Signal Processing*, 16(5), 397-409.
- (21) Kushner, H. J., & Yin, G. G. (1997). *Stochastic approximation and recursive algorithms and applications*. New York, NY: Springer New York.
- (22) Gerencsér, L., Hill, S. D., & Vágó, Z. (1999, December). Optimization over discrete sets via SPSA. In *Proceedings of the 31st conference on Winter simulation: Simulation---a bridge to the future-Volume 1* (pp. 466-470).
- (23) Robertson, D. I. (1969). TRANSYT: a traffic network study tool.
- (24) Daganzo, C. F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation research part B: methodological*, 28(4), 269-287.
- (25) Daganzo, C. F. (1995). The cell transmission model, part II: network traffic. *Transportation Research Part B: Methodological*, 29(2), 79-93.
- (26) Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H., & Mishalani, R. (1998, February). DynaMIT: a simulation-based system for traffic prediction. In *DACCORD short term forecasting workshop* (Vol. 12). Delft The Netherlands.
- (27) Lo, H. K., Chang, E., & Chan, Y. C. (2001). Dynamic network traffic control. *Transportation Research Part A: Policy and Practice*, 35(8), 721-744.
- (28) Allsop, R. E. (1971). Delay-minimizing settings for fixed-time traffic signals at a single road junction. *IMA Journal of Applied Mathematics*, 8(2), 164-185.
- (29) Webster, F. V. (1958). Traffic signal settings (No. 39).

- 
- (30) Chiou, S. W. (1999). Optimization of area traffic control for equilibrium network flows. *Transportation Science*, 33(3), 279-289.
- (31) Wong, S. C. (1995). Derivatives of the performance index for the traffic model from TRANSYT. *Transportation Research Part B: Methodological*, 29(5), 303-327.
- (32) Wong, S. C. (1996). Group-based optimisation of signal timings using the TRANSYT traffic model. *Transportation Research Part B: Methodological*, 30(3), 217-244.
- (33) Lan, C. L., & Chang, G. L. (2016). Optimizing signals for arterials experiencing heavy mixed scooter-vehicle flows. *Transportation Research Part C: Emerging Technologies*, 72, 182-201.
- (34) Spall, J. C., & Chin, D. C. (1997). Traffic-responsive signal timing for system-wide traffic control. *Transportation Research Part C: Emerging Technologies*, 5(3-4), 153-163.
- (35) Chen, Y. H., & Hartono, M. F. (2023, December). Optimizing Arterial Traffic Signal Settings: Shotgun Version for Simultaneous Perturbation Stochastic Approximation Approach. In *2023 Winter Simulation Conference (WSC)* (pp. 1735-1746). IEEE.
- (36) Spall, J. C. (2003). *Introduction to stochastic search and optimization: estimation, simulation, and control*. John Wiley & Sons.
- (37) Akcelik, R. (1981). *Traffic signals: capacity and timing analysis*.
- (38) FHWA. (2000). *Roundabouts: An informational guide* (No. FHWA-RD-00-067; Project 2425). United States. Federal Highway Administration.
- (39) National Highways. (2023). *CD 116: Geometric design of roundabouts* (Version 2.1.0). Design Manual for Roads and Bridges.

- 
- (40) Scott, D. W. (1992). Multivariate density estimation: Theory, practice, and visualization. Wiley.
- (41) Kiefer, J. (1953). Sequential minimax search for a maximum. Proceedings of the American mathematical society, 4(3), 502-506.
- (42) Chen, Y. (2016). Sigmix User's Manual.
- (43) Flötteröd, G., & Rohde, J. (2011). Operational macroscopic modeling of complex urban road intersections. Transportation Research Part B: Methodological, 45(6), 903-922.



附錄 A 最陡下降法



本研究使用最陡下降法 (steepest descent, Cauchy, 1847) 與區別變數型態 SPSA 等演算法來作比較。最陡下降法的主要概念是即為在每次迭代時找出一個函數值下降最快的方向 (即負梯度方向)，並搭配適當步長改善變數值，逐步降低目標函數，直到找到最小值。當應用於本研究時，關鍵在於如何計算梯度以及如何決定步長。

在計算梯度時，理想上需獲得目標函數對各個變數的偏導數。然而，本研究的目標函數並無解析式，無法直接利用數學公式計算。因此，本研究加入一個小變化量，採用前向差分 (forward difference gradient approximation, FDGRAD) 的方式估計各次迭代的梯度，如式 (A.1) 所示 (Dennis & Schnabel, 1996)。

$$\hat{g}_n^k(\mathbf{x}^k) = \frac{D(\mathbf{x}^k + h_n^k \mathbf{e}_n) - D(\mathbf{x}^k)}{h_n^k} \quad (\text{A.1})$$

其中 $\hat{g}_i^k$ 為第 $k + 1$ 次迭代之估計梯度向量的第 n 個分量值； h_n^k 為一個小變化量，由於本研究中的決策變數 (時差、綠燈長) 皆為整數，這些小變化量皆設為最小單位的 1； $\mathbf{e}_n$ 為第 n 個單位向量，即除了第 n 個分量為 1 之外其餘分量皆為 0。

步長的搜尋使用黃金分割法 (Golden Section Method, Kiefer, 1953)。此方法為一種線性搜尋 (line search) 演算法，先給定一組參數決定步長搜尋範圍，接著依黃金比例 (約為 0.618) 將區間分割，反覆比較函數值並縮小區間，直到收斂於能使函數值下降最多的步長。此方法的特點在於，從第二次迭代開始 (步長搜尋的迭代，而非最陡下降迭代)，即可沿用一個既有模擬過的點，因此每次只需新增一次模擬，計算效率較高。本研究中步長搜尋的初始區間設定為 $(-100, 100)$ 。

A.1 最陡下降法之步驟

旺角路網使用表 4-19 提供的初始點做最陡下降。對於每個變數，必須分別對其做前向差分來估計梯度的每一個分量。例如在 $\mathbf{x}^0 = [\theta^{\text{optT}} \quad \phi^{\text{optT}}]^T$, $\theta^0 =$



$[28 \ 70]^T$, $\phi^0 = [64 \ 40 \ 93]^T$ (初始點 1) 的情況下，為了求梯度的估計值，必須跑 6 次模擬來得到結果，如表 A-1 所示。

表 A-1 旺角路網初始解 1 第一次迭代估計梯度需模擬之變數組合

需模擬之變數組合	θ_2	θ_3	$\phi_{1,1}$	$\phi_{2,1}$	$\phi_{3,1}$
初始解	28	70	64	40	93
第一分量+小變化	29	70	64	40	93
第二分量+小變化	28	71	64	40	93
第三分量+小變化	28	70	65	40	93
第四分量+小變化	28	70	64	41	93
第五分量+小變化	28	70	64	40	94

模擬完以上 6 組變數之後，即可估計出梯度，也就是變數移動的方向，在本研究中，將梯度轉為單位向量 ($\hat{\mathbf{g}}_k(\mathbf{x}^k)/\|\hat{\mathbf{g}}_k(\mathbf{x}^k)\|$) 使步長尺度在合理範圍。接下來需要尋找步長。圖 A-1 說明了黃金比例的特性，當區間 (a, b) 依黃金比例分割後，可得到兩個內點 λ 與 μ 。比較模擬出的績效後，若保留右區間 (λ, b) ，因黃金比例之特性，新的內點 λ' 會是原先的 μ ；反之當保留左區間 (a, λ) ，新的內點 μ' 會是原先的 μ ，由於內點可重複的特性，每次縮小區間後只需模擬一次即可，直至區間收斂。這樣能節省一部分的模擬次數。

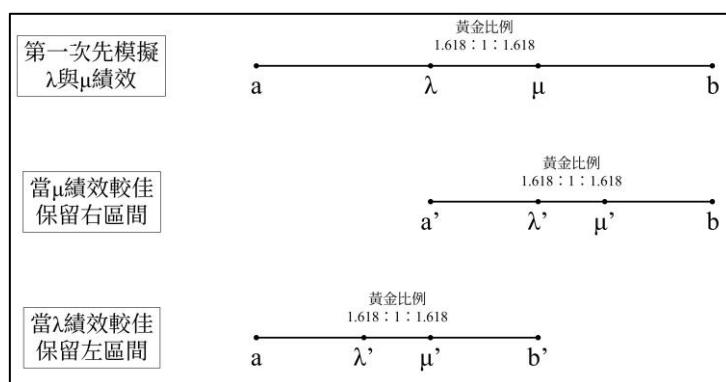


圖 A-1 黃金分割法說明



A.2 最陡下降法結果

旺角路網與現代圓環使用與區別變數型態 SPSA 相同的兩個點做最陡下降。以下舉旺角路網案例解釋，並類推至現代圓環。旺角路網初始解 1 的迭代過程如表 A-2 所示；其第一次迭代的梯度估計如表 A-3 所示；其第一次步長搜尋如表 A-4 所示。

表 A-2 旺角路網：初始解 1 最陡下降法迭代過程

iteration	θ_2	θ_3	$\phi_{1,1}$	$\phi_{2,1}$	$\phi_{3,1}$	Delay	a_k
1	28	70	64	40	93	331.29	45.58382
2	0	95	66	40	104	79.21	20.16474
3	0	95	46	41	102	46.58	19.15818
4	-1	95	46	43	83	41.2	3.12740
5	-4	95	47	42	83	40.78	1.63138
6	-5	95	47	43	83	40.74	1.62199
7	-4	95	48	42	83	40.73	1.63138
8	-5	95	48	43	83	40.58	1.62199
9	-5	95	49	42	83	40.54	0.00290
10	-5	95	49	42	83	40.54	
best	-5	95	49	43	83	40.43	

表 A-3 旺角路網：初始解 1 之第一次迭代梯度估計與單位向量

變數	θ_2	θ_3	$\phi_{1,1}$	$\phi_{2,1}$	$\phi_{3,1}$
$D(\mathbf{x}^k + h_i \mathbf{e}_i)$	359.06	306.57	328.91	331.76	306.41
梯度單位向量分量	0.61985	-0.55177	-0.053123	0.010491	-0.55534

表 A-4 旺角路網：初始解 1 之第一次迭代步長搜尋結果

ϕ_2	03	$\phi_{1,1}$	$\phi_{2,1}$	$\phi_{3,1}$
28	70	64	40	93

d_1	d_2	d_3	d_4	d_5
0.619849	-0.55177	-0.05312	0.010491	-0.55534

iteration	a	b	λ_t	μ_t	$\theta_2 - \lambda_t * d_1$	$\phi_{1,1} - \lambda_t * d_1$	$\phi_{2,1} - \lambda_t * d_1$	$\phi_{2,1} - \lambda_t * d_1$	$\theta_3 - \lambda_t * d_1$	$\phi_{1,1} - \mu_t * d_1$	$\phi_{2,1} - \mu_t * d_1$	$\phi_{2,1} - \mu_t * d_1$	Delay(x- λd)	$\theta_2 - \mu_t * d_1$	$\theta_3 - \mu_t * d_1$	Delay(x- μd)
1	-100	100	-23.6	23.6	43	57	63	40	80	83	65	40	286.24	13	83	82.35
2	-23.6	100	23.6	52.7848	13	83	65	40	105	99	67	39	82.35	-5	99	125.49
3	-23.6	52.7848	5.578994	23.6	25	73	64	40	96	83	65	40	97.88	13	83	82.35
4	5.578994	52.7848	23.6	34.75218	13	83	65	40	105	89	66	40	82.35	6	89	81.18
5	23.6	52.7848	34.75218	41.63621	6	89	66	40	105	93	66	40	81.18	2	93	80.59
6	34.75218	52.7848	41.63621	45.89634	2	93	66	40	105	95	66	40	80.59	0	95	80.3
7	41.63621	52.7848	45.89634	48.52604	0	95	66	40	105	97	67	39	80.3	-2	97	108.8
8	41.63621	48.52604	44.26812	45.89634	1	94	66	40	105	95	66	40	80.45	0	95	80.3
9	44.26812	48.52604	45.89634	46.89951	0	95	66	40	105	96	66	40	80.3	-1	96	125.4
10	44.26812	46.89951	45.27331	45.89634	0	95	66	40	105	95	66	40	80.3	0	95	80.3
45.58381766																



以上為估計梯度以及搜尋步長之示意表。按照此方式繼續迭代即可得到收斂的解。另外，在表 A-2 中，第二次迭代時變數 $\phi_{1,1}$ 原本應該是 105 秒，但由於此值恰好達到綠燈長上限，為了使該變數在梯度估計過程中仍保有改善空間，特別將其調整為 104 秒並重新進行模擬，後續結果也確實展現了更佳的表现。

值得注意的是，最佳績效（最小總延滯）並非必然出現在最終收斂點，而是在梯度估計過程中，變數加上變化量後所產生的模擬結果中績效表現最優者。以下附上初始解 2 的迭代過程如表 A-5 所示。另外初始解 1 與初始解 2 之總迭代與模擬次數如表 4-24 所示。其中每一次最陡下降迭代約需 9 至 13 次步長搜尋迭代。

表 A-5 旺角路網：初始解 2 最陡下降法迭代過程

iteration	θ_2	θ_3	$\phi_{1,1}$	$\phi_{2,1}$	$\phi_{3,1}$	Delay	a_k
1	44	107	64	40	93	285.03	15.40772
2	34	107	74	45	93	47.94	3.128705
3	33	107	74	42	93	47.47	10.64162
4	24	108	77	45	93	45.66	2.123576
5	24	108	77	43	93	45.18	5.878934
6	19	110	75	41	93	44.6	2.631653
7	18	110	75	43	92	43.89	13.9652
8	7	110	70	41	86	42.4	2.633709
9	6	110	69	43	86	41.77	1.121711
10	5	110	69	43	85	41.6	3.751894
11	2	110	69	42	84	41.32	3.879014
12	-1	110	68	44	83	40.79	1.432153
13	-1	110	68	43	83	40.76	3.133987
14	-1	110	65	43	83	40.72	1.813957
15	-3	110	65	43	82	40.44	-0.0029
16	-3	110	65	43	82	40.44	

另外，現代圓環的最陡下降法迭代過程如下頁表 A-6、表 A-7 與表 A-8 所示。因迭代過程與旺角路網相同，現代圓環部分僅列出迭代過程與結果，有關總迭代與模擬次數，見表 4-27。

表 A-6 現代圓環：初始解 1 最陡下降法迭代過程

iteration	θ_2	θ_3	θ_4	θ_5	θ_6	$\phi_{1,1}$	$\phi_{2,1}$	$\phi_{3,1}$	$\phi_{4,1}$	$\phi_{5,1}$	$\phi_{6,1}$	Delay	a_k
1	55	91	98	77	46	25	41	55	54	50	31	221.19	12.2704
2	55	91	93	80	42	25	41	55	46	47	26	183.18	2.012901
3	55	91	92	81	41	25	41	54	46	46	26	179.81	1.429997
4	54	91	92	81	41	25	41	54	46	46	27	176.28	1.813957
5	54	90	91	82	41	25	41	53	46	46	27	173.63	-0.18649
6	54	90	91	82	41	25	41	53	46	46	27	173.63	
best	54	90	91	82	41	25	41	53	47	46	27	171.5	

表 A-7 現代圓環：初始解 2 最陡下降法迭代過程

iteration	θ_2	θ_3	θ_4	θ_5	θ_6	$\phi_{1,1}$	$\phi_{2,1}$	$\phi_{3,1}$	$\phi_{4,1}$	$\phi_{5,1}$	$\phi_{6,1}$	Delay	a_k
1	20	30	8	83	93	25	41	55	54	50	31	232.32	6.545019
2	20	30	12	86	93	25	41	55	56	47	31	206.54	3.136828
3	20	29	13	85	93	25	41	54	54	47	31	195.18	4.253219
4	20	29	16	83	93	25	41	54	55	46	30	188.24	4.762829
5	20	28	17	84	92	25	41	53	51	46	31	174.93	0.002903
6	20	28	17	84	92	25	41	53	51	46	31	174.93	
best	20	28	17	84	92	25	41	53	51	47	31	173.87	

表 A-8 現代圓環：初始解 3 最陡下降法迭代過程

iteration	θ_2	θ_3	θ_4	θ_5	θ_6	$\phi_{1,1}$	$\phi_{2,1}$	$\phi_{3,1}$	$\phi_{4,1}$	$\phi_{5,1}$	$\phi_{6,1}$	Delay	a_k
1	1	43	2	15	50	25	41	55	54	50	31	464.96	-19.3257
2	4	41	7	7	43	25	44	47	47	40	28	206.43	0.504751
3	4	41	7	7	43	25	44	47	47	40	28	206.43	
best	4	41	7	8	43	25	44	47	47	40	28	205.23	



附錄 B SIGMIX 迭代步長測試



在使用 TRANSYT 進行車流模擬之前，曾使用過 SIGMIX 車流模擬軟體 (Chen, 2016) 作為模擬器測試，該車流模擬軟體在路段部分基於 Lo 等人 (2001) 的 DISCO 數學模型（該模型又基於 CTM (Daganzo, 1994)）；在節點部分基於 Flötteröd & Rhode (2011) 提出的 INMC。

該軟體所使用的總延滯單位為車輛-秒，因此與 TRANSYT 的 PCU-hr 差了 3600 倍。由於曾經使用過 SIGMIX 做結合不同變數型態 SPSA 來測試旺角路網迭代步長 a 值，因此在 4.3 節中，使用的 a 值數字比較特別一些。以下是初始解 1 使用 SIGMIX 模擬產生的結果，並附上 SIGMIX 內建最佳化器中使用的基因演算法 (Net-GA, 亦參考 Lo (2001)) 的結果做為比較。其中因梯度單位較大，事先將梯度除以 10000 再做步長搜尋。

表 B-1 旺角路網：SIGMIX a 值搜尋（初始解 1，結合不同變數型態 SPSA）

a 值	迭代次數	迭代最小值 (總延滯 veh-s/hr)
3	50	148,993.82
4	50	127,811.90
5	50	136,115.70
6	50	126,484.25
7	25	317,950.39
(SIGMIX GA 比較)	--	135,714.00

以上因為不是正式結果，數據僅供參考，其中 $a = 7$ 時因步長太大導致擾動太大，造成停滯故提前結束迭代。基本上 4.3 節所使用的步長就從以上 a 值乘回 10000 再除以 3600 的結果得來。另外，將此 6 個的迭代過程列出如圖 B-1、圖 B-2 所示。

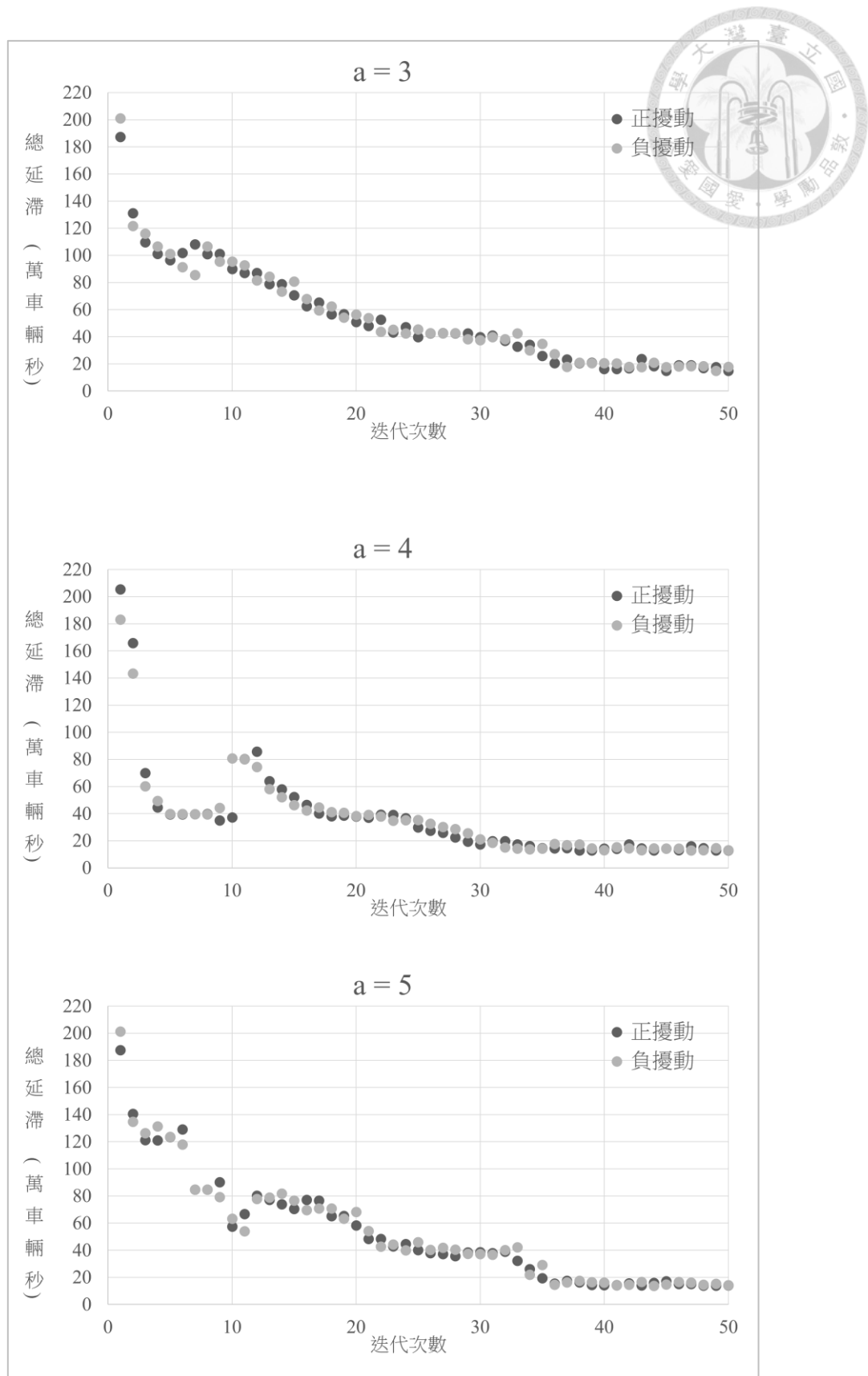


圖 B-1 旺角路網：表 B-1 迭代過程 ($a = 3, 4, 5$)

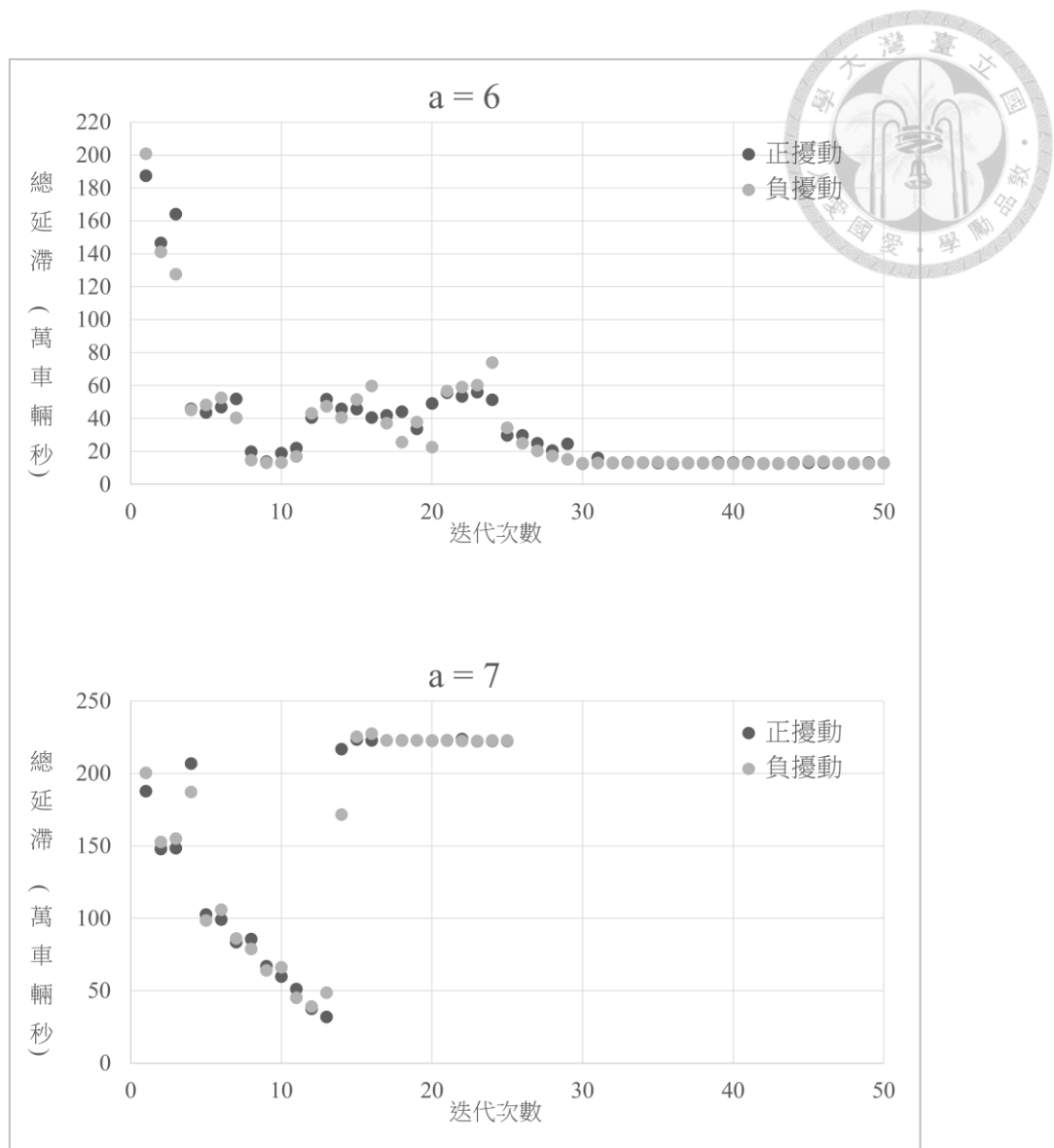


圖 B-2 旺角路網：表 B-1 迭代過程 ($a = 6, 7$)



附錄 C 口試委員意見回覆表



口試委員意見	回覆
論文跟 Chen (2023)那篇文章有何不同？除了多考慮綠燈時間長之外。	本研究多考慮了迭代步長參數的選擇；另外還有模擬器差異—2023 使用微觀，本研究則使用巨觀模擬。
近五年除了 Chen (2023)，還有相關研究嗎？	近年難以找到結合 SPSA 與號誌最佳化的文獻。
第 41 頁的其中一個飽和流率 1904 PCU/hr 是如何算出來的？	當時不慎輸入失誤，但對求解結果影響不大。
擾動幅度 c 假設為 1，不同的 c 值會對求解有影響嗎？	求解是希望 c 值越小越好，但因為收斂到後期會使演算法的正負擾動一樣，會使演算法無法計算，因此固定為 1。
第 82 頁的圖 4-32 與圖 4-33 有何不同？	圖 4-32 為第 20 次迭代結束時之最佳結果與最後一次迭代結束時的最佳結果。
為何在比較演算法時（最陡下降法、SPSA、Hill climbing），不同案例比較的演算法不一樣？	在雙幹道路網中，最陡下降法會花較長時間，未來研究可以加入討論。
隨機生成多個起始點後，該選擇哪一個起始點？	時差是隨機生成，綠燈時長是根據流率分配，最後選擇表現最好的起始點作為最佳結果。
解是否完全取決於時差，綠燈長沒有影響？先前提到論文與 Chen (2023)的差異是額外考慮「綠燈長」，這樣是否真正納入考量？	綠燈長是輔助角色，有些迭代過程綠燈長會稍有變化，但不會變化太大；另外還是有想辦法讓綠燈長變動，且有使用限制式去限制最小綠燈等，讓綠燈時長能夠納入考慮。
演算法設計中，起始點跟初始解（initial solution）是否為相同意義？	初始解與起始點在本研究是一樣的意思。

初始解選擇好壞是否會影響演算法收斂狀況跟速度？	初始解好壞也會影響收斂。
為何起始點是用隨機產生，而不是自己設定一個尚可接受的解？一般常見的都是研究者自己設定。	雖然本研究是隨機生成，但在選擇其中一個點之後仍會稍微調整，讓起始點代入模擬器後的目標值還可接受。
是否可能因固定綠燈長減弱其隨機性，因此演算法收斂狀況都是好的？	後續研究可以探討隨機的綠燈長影響，可直接使用本研究方法做處理。
最佳化問題的目標是控制延滯，那如果延滯沒控制好會發生什麼事？	由於延滯的定義是真實旅行時間與期望旅行時間的差值，若有延滯即代表用路人須要比預期的多等一段時間。
是否會增加其他想控制的目標？	目前常用的目標是排放廢氣、延滯等不同目標，在最小化延滯時也能同時最小化排放廢氣
若最小化延滯到極致，是否會有副作用？	綠燈長度可能會太長，可能會浪費一些時間；另外如指導教授補充，在 TRANSYT 中亦有使用停車次數等目標，最小化延滯可能導致用路人需要較多的停車次數等。而因本研究主要探討 SPSA 應用於號誌最佳化的可能性，因此暫不考慮多目標最佳化。
參考文獻中沒有 Chen (2023) 文章，請補上。	已修正參考文獻的順序與缺漏文獻。
為何 SPSA 模擬兩次就夠了？做更多次是否會更好？	因為兩次結果能找到一個方向，但確實也有看到其他使用 SPSA 的文獻是做兩次以上。
模擬時間跟求解時間何者較短？這兩個時間通常是設計演算法的權衡。	本研究的求解時間為每次模擬器運行的時間（即委員提到的模擬時間）之總和。由於最佳化器的求解過程未把人力輸入資料時間加入，因此難以計算真正的求解時間，後續研究可加入最佳化器的求解時間討論。



鼓勵使用英文進行論文寫作。	感謝委員建議。
若 SPSA 跟基因演算法或最陡下降結合，是否可行？	由於 SPSA 已有 Steepest descent 的概念，不是非常清楚如何做結合。
承上，是否可以部分做 SPSA 、部分做差分？是否這樣就可以減少模擬時間？	還是會多出很多個一次微分結果，可能會在收斂速度上產生影響。
啟發式(或是基因演算法)跟 SPSA 有辦法結合嗎？假設結合的話 SPSA 嵌入基因演算法還是基因演算法嵌入 SPSA 會比較好？為何文獻中沒有人使用類似方法？是否能得到比 TRANSYT 更好的解？	結合應該是可行的；個人會想將 SPSA 嵌入基因演算法。例如使用基因演算法找到一個適當的解後，對該解使用 SPSA 找到一個更好的局部最佳解，解出來有比 TRANSYT 最佳化結果更佳的可能。但也因基因演算法無法驗證其收斂性，而 SPSA 則可驗證其收斂性。兩者結合後無法確定其是否具有收斂性，且可能需要較長的求解時間。
請用兩句話解釋第 82 頁的圖 4-32 與 4-33。	第 20 次迭代結束的最佳解分布及最終迭代的最佳解分布。
承上，針對此演算法有讓解慢慢進步可以補充的清楚一點。	第 83 頁有說明圖 4-34 最終迭代結束與第 20 次迭代結束的曲線往左上移動，也代表解的整體趨勢是有變更佳的。