



國立臺灣大學共同教育中心統計碩士學位學程

碩士論文

Master Program in Statistics

Center for General Education

National Taiwan University

Master's Thesis

以 OPTICS 演算法識別

階層性密度差異的時空群聚結構

An OPTICS-based Algorithm for Identifying

Spatio-Temporal Density Faults

in Hierarchical Clustering Structures

游孟純

Meng-Chun You

指導教授: 溫在弘 博士

Advisor: Tzai-Hung Wen, Ph.D.

中華民國 113 年 8 月

August 2024



誌謝

本論文之完成，首先感謝溫在弘教授的悉心指引與教導。自學士班加入研究室起，老師的指導為我逐步開啟了資料分析領域的大門，始終不厭其煩的回覆著學生們的疑問，不吝於給予發展上的建議，也在碩士班推甄時，積極鼓勵我申請統計碩士學程。而在分析思維上的培養，老師總是深入探究所使用方法所隱含假設或意義，並對於量化分析結果給予完善的詮釋，這樣的培養也使我在分析上與他人具備著不同的觀點。此外，亦感謝口試老師余清祥教授與蔡政安教授提供許多寶貴的建議與指正。

接著感謝陪我走過碩士班兩年，在臺大資料分析與決策社相遇的夥伴們，不論是作為社員或作為幹部，都讓我認識了一群很棒的夥伴，彼此相互學習，我也透過這群夥伴獲得許多不同的視野。其中特別感謝第五屆課程長昱辰、家緯和瑾叡，除了在職涯上給了我很多啟發性的建議，也帶領我支持我去做了很多本來覺得很困難的事情，成為我在碩二時最重要的心靈支柱。

最後感謝我的家人，感謝我的父母自小始終以身作則，不論是對工作或家庭都盡心盡力並認真負責，這樣的態度成為我人格養成最重要的養分，也是我能克服困難走到今天的原因。此外，也感謝我的兩個哥哥，總是走在前方做我最好的榜樣，若沒有他們，我沒有機會可以走到這一天。衷心感謝家人的支持與付出。

游孟純 謹識于國立臺灣大學

2024 年 8 月 5 日



摘要

點事件之時空群聚代表事件之發生熱區，過去之時空群聚演算法僅能識別出密度值高於特定閾值之群聚範圍，無法識別出具備密度差異與階層性之時空群聚結構。本研究基於 OPTICS 演算法，發展 HST-OPTICS 演算法，此演算法可以用來識別點事件階層性時空群聚之密度斷層，進而獲取完整之時空群聚結構。密度斷層發生在具有密度明顯差異的時空邊界，邊界內的範圍可識別為群聚範圍，範圍內的時空密度與範圍外的時空度具有極劇差異。所得之時空群聚除了具排除雜訊點、總群聚數非經給定、群聚範圍明確以及形狀任意之特性，由群聚結構亦可識別出過去演算法無法得知的密度差異與群聚階層關係。本研究發展 HST-OPTICS 演算法時，放寬 OPTICS 演算法中陡度的定義，彈性地查找密度斷層範圍並進一步切分出時空群聚結構，並模擬群聚數、階層關係不同之多組群聚結構，以驗證演算法可以找出過去時空群聚演算法無法有效識別之階層性時空群聚結構。研究結果表示，HST-OPTICS 可以有效識別出重疊且具有高低密度差異之群聚結構、多個群聚隸屬於同一群聚的階層性時空群聚結構，以及時空群聚範圍變動且具備階層性關係的時空群聚結構。未來研究可以善用此演算法於各領域之實務應用中，探討合適之參數設定方式，並著重階層性時空群聚驗證指標設計以及演算法效能提高。

關鍵字：時空群聚演算法、階層群聚、密度斷層、OPTICS、空間分析



Abstract

The HST-OPTICS algorithm improves upon previous spatio-temporal clustering methods by identifying density faults in hierarchical clusters. This approach reveals complete clustering structures, including density differences and hierarchical relationships previously undetectable. The algorithm relaxes the OPTICS steepness definition, allowing for flexible identification of density fault ranges. It can detect overlapping clusters with varying density, hierarchical structures where multiple clusters belong to one cluster, and clustering structures with varying spatial ranges. HST-OPTICS produces clusters with noise exclusion, an undefined total count, clear boundaries, and arbitrary shapes. Simulations have verified its effectiveness in identifying complex hierarchical spatio-temporal clustering structures. Future work could explore practical applications, the design of verification metrics, and efficiency improvements.

Keywords: spatio-temporal clustering algorithm, hierarchical cluster, density faults, OPTICS, spatial analysis



目次

	Page
誌謝	i
摘要	ii
Abstract	iii
目次	iv
圖次	vii
表次	ix
第一章 緒論	1
1.1 研究背景與動機	1
1.2 研究目的	3
1.3 章節介紹	3
第二章 文獻回顧	5
2.1 時空群聚演算法	5
2.2 時空鄰近定義	6
2.3 基於密度之時空群聚演算法	8
2.4 綜合評析	9
2.4.1 密度各異群聚識別問題	9
2.4.2 階層性時空群聚識別問題	10

2.4.3 密度斷層識別彈性問題	11
第三章 研究方法	12
3.1 研究流程	12
3.2 基於密度之群聚演算法	12
3.2.1 DBSCAN	13
3.2.2 OPTICS	14
3.3 研提演算法	15
3.3.1 時空可及圖	16
3.3.2 時空密度斷層與群聚識別	20
3.4 演算法成效驗證	21
3.4.1 模擬資料驗證	22
3.4.2 密度斷層敏感度分析	25
第四章 研究結果	27
4.1 時空群聚結果比較	27
4.2 密度斷層敏感度分析結果	30
第五章 討論	31
5.1 時空群聚結構與密度斷層	31
5.2 參數設定	31
第六章 結論	34
參考文獻	36
附錄 A — 模擬資料生成規則	39
A.1 模擬群聚分布範圍與密度	39



附錄 B — 群聚查找相關圖表

B.1	HST-OPTICS 敏感度分析參數設定	40
B.2	HST-OPTICS 可及圖	42
B.3	HST-OPTICS 群聚圖	46
B.4	HST-OPTICS 群聚結構圖	46

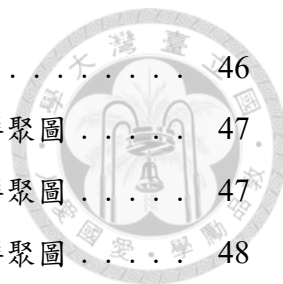




圖次

3.1	研究流程圖	13
3.2	DBSCAN 概念圖	14
3.3	可及圖示意圖	16
3.4	時空搜尋範圍示意圖	18
3.5	搜尋窗大小 window size 與可及分數差異 diff 參數示意圖	20
3.6	群聚切分位置示意圖	22
3.7	模擬資料時空群聚結構特性分類	23
3.8	時空模擬資料分布圖	24
4.1	第一組模擬資料時空群聚結果	28
4.2	第二組模擬資料時空群聚結果	28
4.3	第三組模擬資料時空群聚結果	28
4.4	第四組模擬資料時空群聚結果	29
4.5	第五組模擬資料時空群聚結果	29
4.6	第六組模擬資料時空群聚結果	29
4.7	第七組模擬資料時空群聚結果	29
B.1	最佳參數設定下時空模擬資料 HST-OPTICS 時空可及圖	42
B.2	第二組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖 . .	43
B.3	第三組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖 . .	43
B.4	第四組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖 . .	44
B.5	第五組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖 . .	44
B.6	第六組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖 . .	45
B.7	第七組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖 . .	45

B.8	最佳參數設定下時空模擬資料 HST-OPTICS 群聚圖	46
B.9	第二組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖	47
B.10	第三組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖	47
B.11	第四組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖	48
B.12	第五組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖	48
B.13	第六組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖	49
B.14	第七組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖	49
B.15	最佳參數設定下時空模擬資料 HST-OPTICS 時群聚結構圖	50





表次

2.1	既有時空群聚演算法比較	6
2.2	時空鄰近定義比較	8
2.3	基於密度之時空群聚演算法比較	10
A.1	各組群聚結構下群聚之分布範圍與密度	39
B.2	第二組時空群聚敏感度分析參數設定	40
B.3	第三組時空群聚敏感度分析參數設定	40
B.4	第四組時空群聚敏感度分析參數設定	40
B.5	第五組時空群聚敏感度分析參數設定	41
B.6	第六組時空群聚敏感度分析參數設定	41
B.7	第七組時空群聚敏感度分析參數設定	41



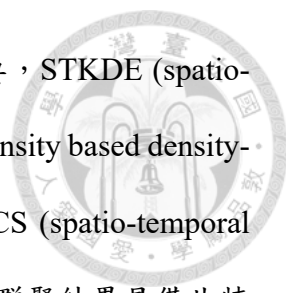
第一章 緒論

1.1 研究背景與動機

時空群聚所在位置為事件在時間和空間上密集發生的地區，亦即事件發生熱區 (hot spot) 之表示方式。傳染病、犯罪、交通事故等領域中，經常以含有時間和空間資訊之點資料，紀錄事件發生之時空位置。多個事件於相近地點發生，會被視為空間上的熱區，而多個事件於相近的時間發生，則會被視為時間上之熱區，僅有當多個事件密集發生於相近的時間和空間時，該區域才會被視為時空上之熱區。熱區在不同領域中有不同的意義，可能表示傳染病爆發地、犯罪高風險區或交通之易肇事路段。

時空群聚的研究中專注在兩大面向，一為全域分析，二為局部分析 (O'sullivan and Unwin, 2003)。全域分析中關注研究區內事件點是否有時空相依的現象，時空相依表示事件點出現之時間與空間並不獨立，事件點之發生可能集中於特定之時間與空間範圍。然而全域分析無法得知事件群聚發生之時間與空間，僅能得知整個研究區之時空相依情況有無。相對地，局部分析則欲找出研究區中事件在時間和空間上的發生熱區。若欲清楚描繪事件發生熱區的明確時間與空間位置，則應進行局部分析。

現有時空群聚局部分析之演算法在四個群聚結果特性上，已經達成成熟的方法研究與應用成果，四個群聚結果特性包含了總群聚數非經給定、排除雜訊點、群聚形狀任意以及群聚範圍明確。具有總群聚數非經給定特性之群聚結



果，因可基於點事件分布而自動查找適合數量的群聚而重要，STKDE (spatio-temporal kernel density estimate)、ST-DBSCAN (spatio-temporal density based density-based spatial clustering of applications with noise) 以及 ST-OPTICS (spatio-temporal ordering points to identify the clustering structure) 等多數演算法之群聚結果具備此特性。具有排除雜訊點特性之群聚結果，因可排除不位於高密度範圍內的事件點進而更關注於熱區而重要，ST-Hierarchical Clustering、ST-DBSCAN 以及 ST-OPTICS 等演算法之群聚結果具備此特性。具有群聚形狀任意特性之群聚結果，因可查找出非特定形狀的群聚進而更貼近更真實之群聚範圍而重要，STKDE、ST-DBSCAN 以及 ST-OPTICS 等演算法之群聚結果具備此特性。具有群聚範圍明確特性之群聚結果，因可以有助於回應群聚查找目的後解決問題的資源之配置以及管理策略之制定而重要，SaTScan、ST-DBSCAN 以及 ST-OPTICS 演算法之群聚結果具備此特性。這些演算法也已被廣泛於使用於應用領域，如 ST-KDE 被用於公共安全領域，查找犯罪之時空分布群聚位置 (Brunsdon et al., 2007)，SaTScan 被用於運輸領域，查找交通意外事故之群聚 (Song et al., 2018)，ST-DBSCAN 被用於海洋學和氣候學領域，查找海洋溫度變化相似之群聚範圍 (Birant and Kut, 2007)，ST-Hierarchical Clustering 被用於生態學領域，查找動物生活之熱區範圍 (Lamb et al., 2020)，ST-OPTICS 被用於遙測領域，將衛星影像資料進行分類以辨識地表不同的植物 (Agrawal et al., 2016)。

現有時空群聚局部分析之演算法在兩個群聚結果特性上之方法研究與應用，則相對處於發展中之狀態，此兩個群聚結果特性包含了具密度差異與具階層性。具有密度各異特性之群聚結果，因可了解各群聚的密度進而知悉不同群聚範圍內的嚴重程度而重要，在空間群聚上，KDE (Kernel density estimation) (Silverman, 2018) 以及 OPTICS (ordering points to identify the clustering structure) (Ankerst et al., 1999) 之群聚結果具備此特性。在時空群聚上，ST-KDE、ST-OPTICS 演算法之群

聚結果雖嘗試識別出具備此特性之群聚結果，但在 ST-KDE 使用時，需逐一設定欲用於切分群聚範圍的多個密度閾值，且閾值之選定缺乏依據。而 ST-OPTICS 演算法之群聚結果具備此特性，且已被使用於查找地表特性各異之群聚的應用領域 (Agrawal et al., 2016)，所查找之群聚結果含有密度差異之特性。

具有階層性特性之群聚結果，因可同時了解同一母群聚之下群聚的相似性，又可以了解子群聚的特殊性而重要，在空間群聚上，OPTICS 演算法之群聚結果具備此特性，但在使用時面臨陡度 (steepness) 切分過於嚴格之問題，導致子群聚的切分與否缺乏彈性。在時空群聚上，ST-OPTICS 使用時，也面臨與 OPTICS 相同之陡度 (steepness) 切分過於嚴格之問題，導致時空群聚的切分與否缺乏彈性。此外，ST-OPTICS 在 Agrawal et al. (2016) 研究中之目的較偏向分類而非查找熱區，因此該研究在最後一步驟中將階層性時空群聚結構中的小範圍群聚以特徵相似性合併，此步驟使得合併後所得之時空群聚失去階層性的特性。

1.2 研究目的

基於過去時空群聚局部分析之演算法及其群聚結果特性的發展，此篇研究基於 OPTICS 演算法，發展 HST-OPTICS 演算法，所建立之時空群聚演算法 HST-OPTICS 可達成三項目的：

1. 可識別出時空密度斷層位置，以作為不同密度群聚之劃分依據。
2. 可識別出群聚間的階層性關係，以描述時空群聚結構。
3. 可彈性調整時空密度斷層定義之嚴格程度，以廣泛考量斷層存在的可能性。

1.3 章節介紹

本文後續共分為 5 個章節，第 2 章為文獻回顧，總結了現有時空群聚演算法之識別概念與時空鄰近定義方式，以及基於密度之群聚演算法中的基礎概念與應

用，並綜合評析了現有演算法的限制與待改善問題；第 3 章為研究方法，細部介紹了本篇之研究流程、所提出之 HST-OPTICS 演算法以及演算法成效驗證方法；第 4 章為研究結果，呈現了演算法之群聚查找以及密度斷層敏感度分析之結果；第 5 章為討論，對比過去演算法，探討密度斷層之識別如何協助 HST-OPTICS 演算法查找到具密度差異與階層性之時空群聚結構，並提供參數設定之建議；最後的第 6 章中則統合整篇文章之結論，以及未來研究的方向與建議。



第二章 文獻回顧

2.1 時空群聚演算法

時空群聚演算法在識別群聚上的概念大致可分為四種類型，基於不同的識別概念所產生之群聚查找結果，具有不同的特性與適用情境。常見之群聚演算法可分為四種識別概念(表 2.1)，包含：查找中心點、鄰近整合、基於密度以及基於混合模型 (Rhys, 2020)。為了知曉研究區中事件發生時空熱區之明確範圍，需使用時空群聚之局部分析，群聚演算法適用於群聚之局部分析，可透過不同的識別概念，來查找出群聚。查找中心點之識別概念，透過最小化群聚內距離來查找群聚，常見之演算法如 k-means (MacQueen, 1967)。鄰近整合之識別概念，透過將距離相近之成對點進行合併來查找群聚，常見之演算法如 hierarchical clustering (Ward Jr, 1963)。基於密度之識別概念，透過識別資料點密度較高的區域來查找群聚，常見之演算法如 KDE (Silverman, 2018)、DBSCAN (Ester et al., 1996) 以及 OPTICS (Ankerst et al., 1999)。基於混合模型之識別概念，透過使用機率模型來描述點事件分佈來查找群聚，常見之演算法如 Gaussian mixture modeling (Stauffer and Grimson, 1999)。時空群聚演算法通常是將一般群聚演算法，透過進一步對於時空資料進行特定的考量，而延伸發展而來，因此時空群聚演算法之識別概念同樣可分為四種類型。基於查找中心點概念之時空群聚演算法包含 fuzzy C-means，依此概念所查找出的時空群聚具有聚中心明確的特色，但對於非球形的群聚查找效果不佳。在過去 Izakian et al. (2013) 曾用此類演算法來對於各地之降水量、溫

度和濕度進行群聚分析，以協助流行病學家更好地了解環境健康狀態與動物健康狀態。基於鄰近整合概念之時空群聚演算法包含 ST-Hierarchical Clustering，依此概念所查找出的時空群聚具有層次結構明確且可以顯示不同層級群聚的特色，但層級之選定缺乏依據。在過去 Lamb et al. (2020) 曾用此類演算法查找動物活動之熱區範圍。基於密度概念之時空群聚演算法包含 ST-DBSCAN 和 ST-OPTICS，依此概念所查找出的時空群聚具有形狀任意且可排除雜訊點的特色，但對於參數的設定較為敏感。在過去 Birant and Kut (2007) 曾用此類演算法查找海洋溫度變化相似之群聚範圍，Agrawal et al. (2016) 曾用此類演算法來將衛星影像資料進行分類以辨識地表不同的植物。基於混合模型概念之時空群聚演算法包含 STGMM (spatio-temporal gaussian mixture model)，依此概念所查找出的時空群聚具有能夠發現重疊群聚的特色，但計算複雜度較高。在過去 Kim et al. (2012) 曾用此類演算法進行影片之背景和前景分離。

表 2.1: 既有時空群聚演算法比較

	查找中心點	鄰近整合	基於密度	基於混合模型
方法	fuzzy C-means	ST-Hierarchical Clustering	ST-KDE ST-DBSCAN ST-OPTICS	STGMM (ST-gaussian mixture model)
時空群聚結果特性	群聚中心明確的特色	層次結構明確，可以顯示不同層級群聚	形狀任意且可排除雜訊點	能發現重疊階層性群聚
缺失	對於非球形的群聚查找效果不佳	層級之選定缺乏依據	對於參數的設定較為敏感	計算複雜度較高

2.2 時空鄰近定義

時空群聚分析中對於事件點間相似度的衡量，因時空資料的特性，相對於一般群聚分析需要更複雜或細緻的定義。時空群聚分析中對於時空鄰近的定義，與

一般群聚中對於點事件中的相似性判定不同，在一般群聚中，對於點事件間相似度的判定，大多是將不同特徵在執行群聚分析時直觀地以各種距離的判斷方式來衡量相似度。然而在時空群聚分析中是基於時空資料來進行分析，在此類的議題上，可能因為點事件間時間上的關聯性或空間上的關聯性有其不同的重要性或需特別考量與整合之處，因此在時空群聚分析中對於時空鄰近的定義，相對於一般群聚分析更為複雜與細緻。


時空群聚分析中對於時空鄰近的定義，代表事件點在時間與空間維度上相似程度與鄰近程度的衡量，也影響群聚存在與否之判定。時空群聚局部分分析中對於時空鄰近之定義，與時空群聚全域分析中的定義相似，參考全域分析中時空鄰近之定義，可作為發展時空群聚局部分分析演算法之參考。常見之時空群聚全域分析中，識別時空鄰近之定義可分為三個類型(表 2.2)，包含閾值切分、距離遞減以及前幾鄰近。使用特定閾值作為鄰近判斷依據來定義時空鄰近之全域分析方法包含 Knox's Test (Knox and Bartlett, 1964)、Baker's Max Test (Baker, 1996) 以及 Diggle's Test (Diggle et al., 1995)，其中 Baker's Max Test 在給定之參數搜尋範圍內測試出最佳之閾值參數，Diggle's Test 中透過時空 K-function 減去時間 K-function 與空間 K-function 的乘積，來衡量時空相依性之顯著性，其中的 K-function 透過不同之時空閾值設定找出時空群聚顯著之閾值參數，此種定義方式認為只有兩事件點時間與空間上皆小於給定之閾值時，彼此才具有鄰近關係。使用距離遞減概念來定義時空鄰近之全域分析方法包含 Mantel's Test (Mantel, 1967)，Mantel's Test 對於鄰近的定義則非二元分別，時間與空間距離之倒數被用以衡量鄰近與否，此種定義方式認為所有點間皆具有鄰近關係，鄰近程度則隨距離上升而遞減。使用前幾鄰近概念來定義時空鄰近之全域分析方法包含 Jacquez's k-NN Test (Jacquez, 1996)，該方法中會將離某一點前 k 個時空鄰近之事件點視為鄰近，此種定義方式中，各事件點對於鄰近與否的衡量會因為該點附近事件點密度的差異而有不同。

表 2.2: 時空鄰近定義比較

	閾值切分	距離遞減	前 k 鄰近
時空鄰近 定義	只有兩事件點時間與空間上皆小於給定之閾值時，彼此才具有鄰近關係	所有點間皆具有鄰近關係，鄰近程度則隨距離上升而遞減	各事件點對於鄰近與否的衡量會因為該點附近事件點密度的差異而有不同
時空相依 檢定方法	Knox's Test Baker's Max Test Diggle's Test	Mantel's Test	Jacquez's k-NN Test

2.3 基於密度之時空群聚演算法

基於密度之時空群聚演算法，為查找具密度差異以及具階層性的群聚結構帶來可能性，但仍因所參考而延伸發展的一般群聚演算法特性而有查找時空群聚結構的限制。此類演算法多為基於一般群聚演算法，再加上對於時空鄰近的考量設定，而延伸發展為時空群聚演算法。透過在執行演算法時設定不同的參數，基於密度之時空群聚演算法，可查找出不同密度的群聚結果，同時也對於所查找出之群聚的密度有所了解。而透過多次不同參數設定之群聚查找結果的整合，基於密度之時空群聚演算法，可嘗試獲取具備階層性的時空群聚結構。常見之基於密度之時空群聚演算法包含 ST-KDE (Brunsdon et al., 2007)、ST-DBSCAN (Birant and Kut, 2007) 以及 ST-OPTICS (Agrawal et al., 2016) (表 2.3)。ST-KDE 透過核密度估計 (kernel density estimation) 獲取各時空範圍內的時空密度值，在選取適當的多個切分閾值的情況下，可以查找出不同密度之群聚結果，並可同時了解整體之時空群聚結構與階層性關係，但此過程中需自行指定切分之閾值，是否需切分出新的子群聚缺乏密度變異程度的考量。在應用上，Nakaya and Yano (2010) 曾將其用於分析犯罪事件之分布，成功對於犯罪群聚的空間範圍和持續時間進行視覺化，但所獲得之群聚結果無法清晰界定每個熱區的確切邊界。ST-DBSCAN 透過點事件間可及與可連結的關係，識別出範圍明確且形狀任意特性之時空群聚，在設定不



同參數後，可查找出不同密度之群聚結果，並透過多次群聚結果的統整，嘗試獲取整體之時空群聚結構以及群聚間的階層性關係，但此過程中需設定適當的參數值，於統整時亦無法得知群聚間的階層性關係。在應用上，Kuo et al. (2018) 曾將其用於傳染病群聚分析，透過改良過後的 ST-DBSCAN 來查找時空群聚，並藉由時空群聚之中心點與面積大小來分析傳染病群聚隨時間之動態變化過程，成功描繪群聚之演化類型，然而該研究中指明，所獲得之群聚結果無法找出階層性之時空群聚，換言之，該研究中所查找出之時空群聚具有共同之密度，群聚中之密度更高的群聚無法被識別出來。ST-OPTICS 透過更新的點事件的排序以及判斷點事件之相似性，識別出具有不同密度層級特性之時空群聚，可透過可及圖與陡度 (steepness) 之參數設定，統整出整體之時空群聚結構以及群聚間的階層性關係，但其群聚查找結果受參數設定影響很大。在應用上，Agrawal et al. (2016) 曾將其用於辨識地表植物，將衛星影像資料進行分類，成功辨識地表不同的植物，但在該研究中之目的較偏向分類而非查找熱區，因此該研究在最後一步驟中將階層性時空群聚結構中的小範圍群聚以特徵相似性合併，此步驟使得合併後所得之時空群聚失去辨識階層性時空群聚結構的能力。

2.4 綜合評析

綜合過去研究成果，目前在時空群聚結構識別的領域中，尚面臨三大問題，包含密度各異群聚識別、階層性時空群聚結構識別以及密度斷層識別彈性。

2.4.1 密度各異群聚識別問題

不同時空密度的群聚代表事件在各個時空範圍內的嚴重程度，Kuo et al. (2018) 的研究中也提及，基於 ST-DBSCAN 改良過後之 MST-DBSCAN 方法無法知悉所辨識出的每個群聚的密度，因而無法呈現傳染病在不同群中的嚴重程度，

表 2.3: 基於密度之時空群聚演算法比較

	ST-KDE	ST-DBSCAN	ST-OPTICS
識別潛力	在選取適當的多個切分閾值後，可以查找出不同密度之群聚結果，並可同時了解整體之時空群聚結構與階層性關係	在不同參數設定下，可查找出不同密度之群聚結果。透過統整多次群聚結果，可嘗試獲取整體之時空群聚結構以及群聚間的階層性關係	可透過可及圖與陡度參數設定，查找有密度差異的群聚，並統整出整體之時空群聚結構以及群聚間的階層性關係
識別限制	需自行指定切分之閾值，閾值之選定缺乏依據，是否需切分出新的子群聚缺乏密度變異程度的考量	需設定適當的參數值，於統整時需自行歸納群聚間的階層性關係	群聚查找結果受參數設定影響很大，陡度參數切分過於嚴格，導致子群聚的切分與否缺乏彈性。Agrawal et al. (2016) 研究中最後一步驟階層性時空群聚結構中的小範圍群聚依相似性合併，失去階層性的特性

這反映不同密度群聚識別的重要性。ST-KDE 方法雖可透過多個密度閾值，切分識別出密度各異之群聚，但過程中仍面臨無法精確界定群聚範圍的限制。

2.4.2 階層性時空群聚識別問題

群聚間的階層性關係，代表子群聚具備母群聚類似的特性，但又同時進一步具備相對於母群聚更細部的特徵，Levin (1992) 的研究中也提及透過階層性之概念理論與詮釋，對於自然系統和生態學之研究非常有幫助。ST-OPTICS 方法在 Agrawal et al. (2016) 的研究中，雖嘗試識別出群聚間的階層性關係，但卻將階層性的群聚結果合併，未善加詮釋查找出的階層性時空群聚的意義。

2.4.3 密度斷層識別彈性問題

群聚結構中的密度斷層代表了密度的急遽變化處，其存在與否可作為是否應切分出新的子群聚之依據。在實務應用情境上，新的子群聚範圍可能有獨立或獨特的導致因素或特性，因而有被識別出的價值。然而時空密度的變化程度應在多劇烈的程度下才應被識別為密度斷層，亦即密度在怎麼樣的範圍內變化多大時，應多劃分出一個新的子群聚，此類密度斷層識別的彈性問題，在過去的研究中仍缺乏討論。ST-OPTICS 所參考之 OPTICS 演算法，雖嘗試識別出具密度差異以及階層關係的群聚，以建構整體之群聚結構，但 OPTICS 演算法原先在判定陡度 (steepness) 切分閾值時，僅考量前後排序之事件點間之間可及分數差異，僅在密度有顯著大之差異時才會識別出一個新的子群聚，此判斷法較為嚴苛與缺乏彈性。



第三章 研究方法

3.1 研究流程

本研究的流程包含三個階段(圖 3.1)，包含發展 HST-OPTICS 演算法，驗證演算法成效，以及提供演算法參數設定的建議。在第一階段中發展 HST-OPTICS 演算法，對 OPTICS 演算法進行調整，以發展 HST-OPTICS 演算法。在發展 HST-OPTICS 演算法時，定義時空資料點鄰近關係以及時空距離衡量方式，以獲取時空可及圖，接著基於 OPTICS 中陡度 (steepness) 之概念進行調整，定義出搜尋窗大小 window size 以及可及分數差異 diff 兩個參數，以識別出密度斷層範圍，並劃分出階層性時空群聚結構以及各群聚的範圍。在第二階段中模擬時空群聚結構來驗證演算法成效，以衡量不同演算法之群聚查找成效。七組給定之時空群聚結構資料點被模擬出來，其中包含群聚數、階層關係各異的群聚結構。在第三階段中進行 HST-OPTICS 演算法密度斷層識別之敏感度分析，以提供參數設定的建議。

3.2 基於密度之群聚演算法

所發展之 HST-OPTICS 演算法參考 DBSCAN (Density-Based Spatial Clustering of Application with Noise) 中之參數定義以及點事件間鄰近與否的概念，也參考 OPTICS (Ordering Points to Identify the Clustering Structure) 中將點排序繪製可及圖與透過陡度切分出群聚以識別階層性群聚之概念。

第一階段：發展 HST-OPTICS 演算法

時空可及圖：

1. 定義時空資料點鄰近關係
2. 定義時空距離
3. 獲取時空可及圖

時空密度斷層與群聚識別：

1. 基於 OPTICS 陡度概念，定義搜尋窗大小 window size 以及可及分數差異 diff 兩個參數
2. 識別出密度斷層範圍
3. 劃分階層性時空群聚結構

第二階段：驗證演算法成效

1. 模擬七組群聚數、階層關係各異之時空群聚結構
2. 衡量研提演算法與既有演算法之群聚查找成效

第三階段：提供演算法參數設定建議

1. 進行搜尋窗大小與可及分數差異兩個參數之敏感度分析
2. 提供演算法參數設定的建議

圖 3.1: 研究流程圖

3.2.1 DBSCAN

Ester et al. (1996) 發表 DBSCAN 演算法，DBSCAN 為基於密度之群聚查找演算法，此演算法定義了兩個參數，其一為空間搜尋範圍 Eps，其二為群聚最少應包含的點數 $MinPts$ ，藉由這兩個參數可以將事件點分成三個類別，分別為核心點 (core point)、邊緣點 (border point) 以及雜訊點 (noise)。某點之搜尋範圍內若包含比 $MinPts$ 還多的事件點數，則該點為核心點。若非核心點，但落在核心點之搜尋範圍內之事件點則為邊緣點。事件點若非核心點且非邊緣點，則為雜訊點。

在偵測群聚時，DBSCAN 定義兩個事件點間鄰近關係的概念，分別為可及 (reachable) 與可連結 (connected)。如圖 3.2 中，點 p 為核心點，且點 r 落在點 p 之搜尋半徑內，則 p 對 r 直接密度可及 (directly-density-reachable)。而點 p 直

接密度可及點 r ，點 r 又直接密度可及點 q_1 ，因此點 p 則對點 q_1 密度可及 (density-reachable)，另外，因為點 p 密度可及點 q_1 和點 q_2 ，點 q_1 和點 q_2 彼此關係為密度可連結 (density-connected)。

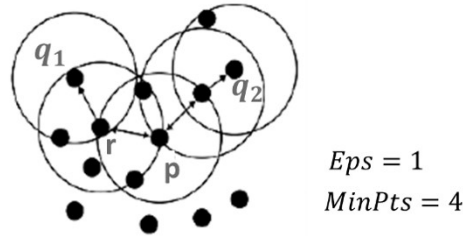


圖 3.2: DBSCAN 概念圖

DBSCAN 最後會將彼此密度可連結的多個點歸為同一群聚，並把不具有可連結關係的點視為雜訊點。在此判斷依據之下，所查找出之群聚密度會由起初對於 Eps 和 $MinPts$ 的設定所影響。

基於 DBSCAN 演算法，Birant and Kut (2007) 之 ST-DBSCAN 演算法引入了非空間屬性的距離參數，在衡量時空鄰近關係時考量了事件點間空間與非空間的相似性，使其可以在包含時間和非空間屬性的資料集上達成更有效的群聚分析成果。Kuo et al. (2018) 之 MST-DBSCAN 演算法則又進一步改進時間鄰近上的定義，設定新的參數，使得在判定事件點鄰近關係時，僅將具備潛在傳染關係的點視為潛在鄰近點。

3.2.2 OPTICS

Ankerst et al. (1999) 發表 OPTICS 演算法，OPTICS 與 DBSCAN 同為基於密度之群聚演算法，OPTICS 可在創建點的排序與對應之可及分數 (reachability score) 後，在可及圖 (reachability plot) 中藉由陡度 (steepness) 參數切分出群聚並識別階層性群聚結構。

OPTICS 中與 DBSCAN 定義了兩個相同參數，分別為 $MinPts$ 與 Eps 。 $MinPts$ 表示每個群聚最少應包含的點數， Eps 則表示空間搜尋半徑。以某一事件點為圓

心，Eps 為半徑所畫出的圓中，若包含比 MinPts 還多的事件點數，則該點為核心點。此外，OPTICS 中定義了兩種距離的概念，分別為核心距離 (core distance)，以及可及距離 (reachability distance)。某點之核心距離則定義為該點能涵蓋 MinPts 個事件點的最小搜尋半徑。可及距離定義為某一核心點至其未造訪之鄰居事件點之空間距離，最小不會小於核心點之核心距離。

OPTICS 演算法會由編號 1 的點開始逐點造訪與運作，其運作流程如下：

1. 紀錄該點既有可及距離中最小的數值當作該點之可及分數 (reachability score)，其中編號 1 的可及分數不存在。
2. 計算核心距離，若該點為核心點，則計算它對空間搜尋範圍內所有鄰近點的可及距離，若鄰近點之新的可及距離較小，則取代原有之可及距離數值。
3. 按現有的可及距離由小到大排列，更新點的造訪順序。

在所有點皆造訪過後，即可繪製可及圖，該圖以各點之造訪順序為橫軸，各點之可及分數為縱軸 (圖 3.3)。接著 OPTICS 透過陡度 (steepness) 參數，當圖中排序前後的兩點之可及分數差異大於所設定之陡度值，則兩點間為群聚切分位置，當成對之群聚切分起始與終止位置形成，則可識別出群聚。圖下方以不同色的橫條表示不同群聚，由呈現的趨勢可發現階層性之空間群聚，舉例而言，圖中灰色群聚中亦包含黃色以及紅色等較密集的群聚。

基於 OPTICS 演算法，Agrawal et al. (2016) 之 ST-OPTICS 演算法新增非空間距離之參數，並且額外計算最小可及距離 (minimum of reachability distance) 與最大核心距離 (maximum of core distances) 來取代了原先 OPTICS 演算法中的 Eps 參數。

3.3 研提演算法

HST-OPTICS 演算法是一種改良的群聚演算法，旨在處理時空資料，查找階層性時空群聚。本研究在發展 HST-OPTICS 演算法的過程，修正時空鄰近關係與

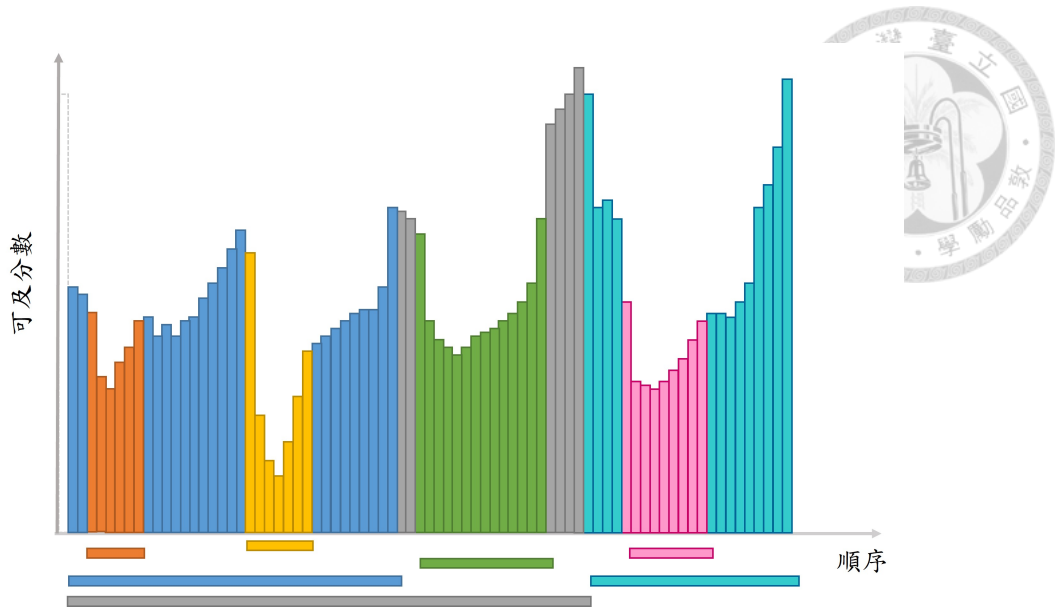


圖 3.3: 可及圖示意圖

時空距離衡量方式，獲取點的排序與對應之時空可及分數，以繪製時空可及圖。接著基於時空可及圖，建立彈性地識別群聚切分位置之方式，以獲取最終之時空群聚結構。

HST-OPTICS 演算法所需參數共有七個，包含建立時空可及圖所使用的群聚最少應包含的點數 $MinPts$ 、空間搜尋範圍 $EpsS$ 、時間搜尋範圍 $EpsT$ 、空間距離權重 w_S 與時間距離權重 w_T ，以及彈性地識別密度斷層與群聚結構所使用的搜尋窗大小以及 window size 和可及分數差異 $diff$ 。

3.3.1 時空可及圖

在 HST-OPTICS 中，基於 OPTICS 中空間鄰近關係與可及距離之概念進行調整，以將 OPTICS 中之可及圖擴展至 HST-OPTICS 中時空資料適用之時空可及圖。HST-OPTICS 中定義時間搜尋範圍參數，衡量點事件是否在時空上鄰近，並定義時間距離權重與空間距離權重兩個參數，衡量事件點間的時空可及距離，將 OPTICS 中由點的排序與對應可及距離所組成之可及圖，擴展至 HST-OPTICS 中，由點的排序與對應時空可及距離所組成之時空可及圖。

除了 OPTICS 既有之群聚最少應包含的點數 $MinPts$ 以及空間搜尋範圍 $EpsS$

外，在 HST-OPTICS 中還定義時間搜尋範圍 EpsT 參數，以衡量點事件間時空鄰近關係，並定義事件點是否為核心點。在既有 OPTICS 之下，事件點間是否鄰近，以及事件點是否為核心點之判定，是基於空間搜尋範圍與群聚最少應包含的點數來決定。為同時考量事件點時間和空間上之相近性，本研究將原先 OPTICS 中的空間搜尋範圍 Eps 標示為 EpsS，並在 HST-OPTICS 新增時間搜尋範圍 EpsT，式 3.1 與式 3.2 定義事件點 i 與事件點 j 是否具有鄰近關係，其中 S 是任意兩個事件點組合中，具有時空鄰近關係事件點組合的集合。當兩事件點間之空間距離小於等於空間搜尋範圍，且兩事件點間之時間差異小於等於空間搜尋範圍，則兩事件點具有時空鄰近關係。再者，式 3.3 定義事件點 i 是否為核心點，其中 C 為核心點的集合，當一事件點在其時間與空間搜尋範圍內，包含自己後，所涵蓋之事件點數大於等於群聚最少應包含的點數 MinPts，則該事件點為核心點。舉例而言，圖 3.4 中，藍色點為 i 事件點之時空鄰近點，而若 i 事件點之時空鄰近點數達到 MinPts，則 i 事件點為核心點。

$$\forall i, j \in S \iff \text{cond} \quad (3.1)$$

其中

$$\text{cond} = \left\{ \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \leq \text{EpsS} \right\} \wedge \{|t_j - t_i| \leq \text{EpsT}\} \quad (3.2)$$

$$p \in C \iff \text{Count}(j) \geq \text{MinPts} \quad \forall i, j \in S \quad (3.3)$$

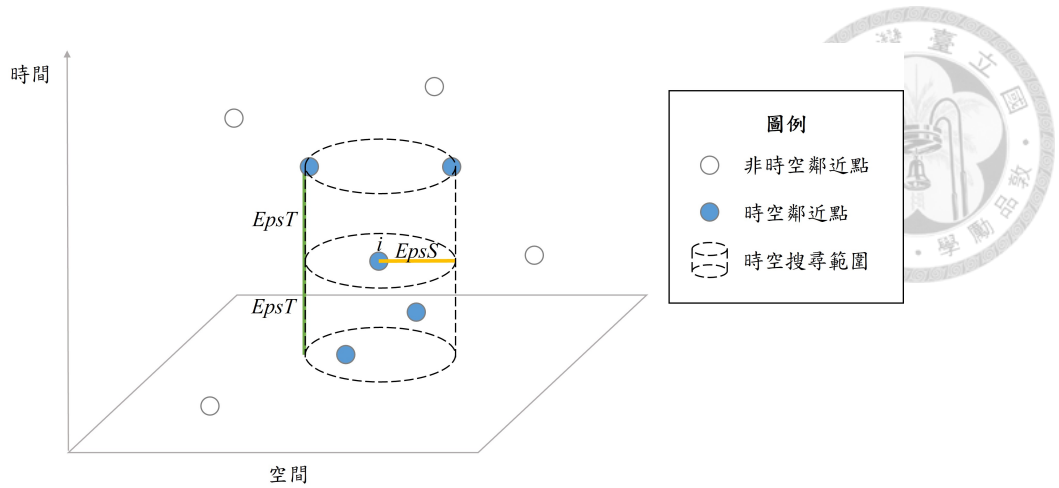


圖 3.4: 時空搜尋範圍示意圖

在 OPTICS 既有之空間距離衡量方式上，HST-OPTICS 新增空間距離權重 w_S 與時間距離權重 w_T 兩參數，衡量點事件間時空距離，以獲取時空可及圖所需之時空可及分數，若無特別指定，兩權重之預設值皆為 1。HST-OPTICS 演算法中，擴展 OPTICS 中空間距離之衡量為時空距離的衡量，兩事件點之時空距離計算方式 (式 3.4)，定義為兩事件點間空間距離與時間差之組合，權重即為 HST-OPTICS 所設定之空間距離權重 w_S 與時間距離權重 w_T 。在既有 OPTICS 中之核心距離與可及距離兩種距離概念，也被擴展為 HST-OPTICS 中的時空核心距離以及時空可及距離。在 HST-OPTICS 中，某點之時空核心距離 (式 3.5)，定義為該點能涵蓋 MinPts 個事件點的最小時空距離，而時空可及距離 (式 3.6)，則定義為某核心點至其未造訪之時空鄰近事件點之時空距離，最小不會小於核心點之時空核心距離。舉例而言，圖 3.4 中，若 i 核心點至 j 事件點間的時空距離小於 i 核心點之時空核心距離，則由 i 核心點至 j 事件點之時空可及距離為 i 核心點之時空核心距離。 j 事件點會記錄多個核心點至其的時空可及距離，而在演算法中會將最小的時空可及距離紀錄為 j 事件點之時空可及分數。



$$d(i, j) = \left\{ \left[w_S * \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \right]^2 + (w_T \times |t_j - t_i|)^2 \right\}^{\frac{1}{2}} \quad (3.4)$$

s.t. $w_S, w_T \geq 0$

$$\text{STCD}(i) = \begin{cases} \text{Undefined, if } i \notin C \\ d(i, j^M), \text{ if } i \in C \end{cases} \quad (3.5)$$

$$\text{STRD}(j, i) = \begin{cases} \text{Undefined, if } i \notin C \\ \max \{ \text{STCD}(i), d(i, j) \}, \text{ if } i \in C \end{cases} \quad (3.6)$$

其中 M 為 MinPts， j 已依據與 i 之時空距離由小到大排序， d 為時空距離，STCD 為時空核心距離， j^M 為與 i 第 MinPts 時空鄰近的時空鄰近點，STRD 為時空可及距離。

HST-OPTICS 在逐點造訪運作後，會記錄下各點之排序與對應之時空可及分數，並可繪製成時空可及圖。HST-OPTICS 演算法會由編號 1 的點開始逐點造訪與運作，其運作流程如下：

1. 紀錄該點既有時空可及距離中最小的數值當作時空可及分數，其中編號為 1 之事件點的時空可及分數不存在。
2. 計算時空核心距離，若該點為核心點，則計算它對時空搜尋範圍內所有點的時空可及距離。
3. 按現有的時空可及距離由小到大排列，更新點的造訪順序。

在既有 OPTICS 中，在所有點皆造訪過後，可繪製可及圖，HST-OPTICS 中則可繪製時空可及圖，時空可及圖中之橫軸不變，同樣為各點之順序，而縱軸則

由 OPTICS 中各點對應之可及分數，改為 HST-OPTICS 中各點對應之時空可及分數。



3.3.2 時空密度斷層與群聚識別

HST-OPTICS 基於 OPTICS 中陡度 (steepness) 之概念進行調整，定義出搜尋窗大小 window size 以及可及分數差異 diff 兩個參數 (圖 3.5)，以更彈性地由時空可及圖中，識別出群聚中密度斷層之範圍與群聚切分位置，以及最終之整體階層性時空群聚結構。

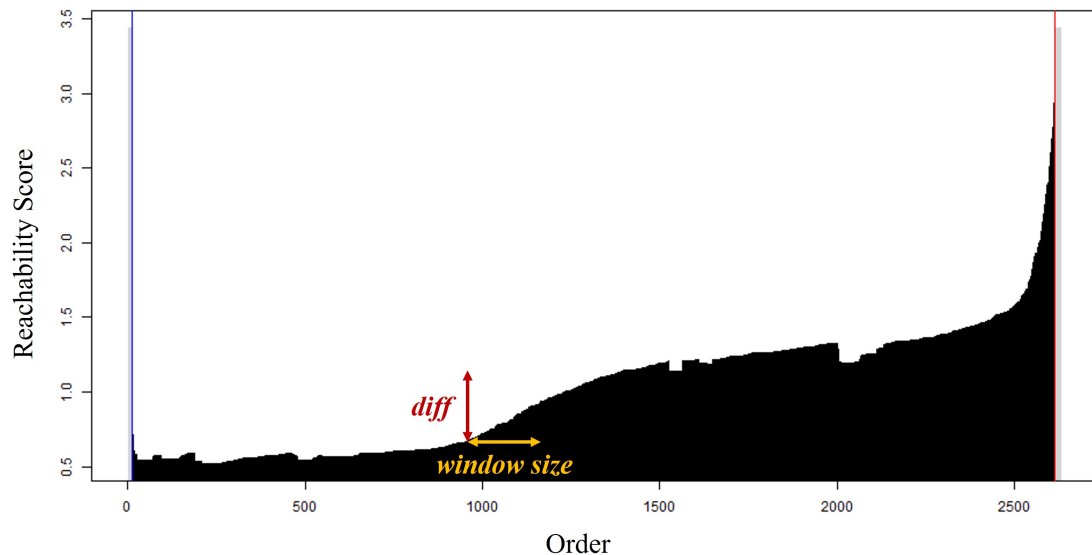



圖 3.5: 搜尋窗大小 window size 與可及分數差異 diff 參數示意圖

OPTICS 將可及圖中排序前後點間陡度夠大的位置識別為群聚切分位置，此識別依據過於嚴格缺乏彈性。在既有 OPTICS 之下，唯有當排序前後的兩點間之可及分數差異大於指定之陡度值時，才會識別出群聚的切分位置，此識別依據過於嚴格且缺乏彈性，即使前後數個排序的點間有足夠大的陡度差異存在，仍不會被識別為群聚切分位置。

HST-OPTICS 定義搜尋窗大小 window size 以及可及分數差異 diff 兩個參數，在時空可及圖中，更彈性地識別出密度斷層發生範圍。HST-OPTICS 執行時，會



搜索各點排序前後搜尋窗範圍內的點，是否有可及分數差異夠大的點，若有達成條件，則該點被視為位於密度斷層範圍內。連續多個達成條件且可及分數變化正負方向相同的點會被視為位在同一個密度斷層範圍內，密度斷層即代表密度有急遽變化的範圍，為潛在的群聚切分位置。其中，若可及分數差異參數固定，當搜尋窗大小參數越大時，代表對於急遽變化的需求嚴格度較低，搜尋窗大小參數越小時，則代表嚴格程度較高，當搜尋窗為 1 時，則與原始 OPTICS 中陡度的判斷概念相同。如此一來，透過搜尋窗大小 window size 以及可及分數差異 diff 兩個參數的設定，即可彈性地調整密度斷層識別的嚴格程度。

在 HST-OPTICS 中，密度斷層範圍中最前或最後排序的點被設定為群聚切分位置，以劃分出各群聚的範圍，並獲得整體之階層性時空群聚結構。在同一密度斷層範圍內，若可及分數變化方向為負，則取排序最後的點作為群聚切分開始位置，如圖 3.6 中藍色標註之密度斷層範圍與切分點，若可及分數變化方向為正，則取排序最前的點作為群聚切分結束位置，如圖 3.6 中紅色標註之密度斷層範圍與切分點。如此一來，被劃分在同一群聚範圍內的點具備相似的可及分數，亦即範圍內的點附近還具備相似高的時空密度。最終，一組開始與結束群聚切分位置會形成一個群聚。若群聚切分開始位置數量大於群聚切分結束位置，則重複使用最後一個群聚切分結束位置，若群聚切分開始位置數量小於群聚切分結束位置，則重複使用第一個群聚切分開始位置，以使得最終群聚切分開始與結束位置可以成對。透過多組成對群聚切分位置，可劃分出多個群聚範圍，並得出最終的階層性時空群聚結構。

3.4 演算法成效驗證

本研究生成七組具備不同群聚結構的時空模擬資料，以驗證 HST-OPTICS 演算法查找時空群聚結構之成效，此外，本研究設定九組不同的 HST-OPTICS 演算

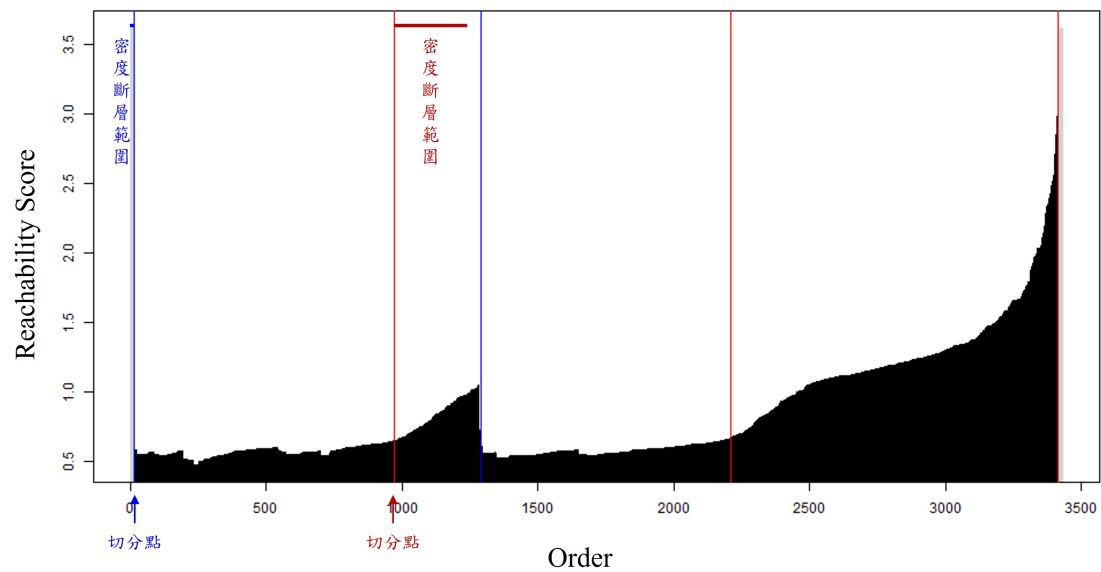


圖 3.6: 群聚切分位置示意圖

法參數，以分析不同參數設定對於時空密度斷層與階層性時空群聚查找結果影響的敏感度。

3.4.1 模擬資料驗證

本研究模擬七組群聚結構特性不同的時空資料(圖 3.7)，以驗證 HST-OPTICS 相對於 ST-DBSCAN 演算法，在指定群聚數以及群聚階層關係上的群聚查找成效。圖 3.7 中呈現模擬資料時空群聚結構特性之分類，上半之時空群聚結構中，群聚間彼此不重疊且不具密度差異，所模擬之第一組至第三組的時空模擬資料屬於此類型的時空群聚結構，下半之時空群聚結構中，群聚間彼此具有密度差異，且互相重疊，所模擬之第四組至第七組的時空模擬資料屬於此類型的時空群聚結構。

由不具備任何子群聚的模擬資料開始，第一組模擬資料僅包含雜訊點，檢驗演算法是否能排除雜訊點，僅將聚集密度夠高之點事件識別為群聚，其後之第二至七組資料也都包含這些雜訊點。接著模擬有子群聚的群聚結構，第二組模擬資料包含一個子群聚(圖 3.8a)，檢驗演算法是否能識別密度夠高之群聚範圍。接著擴展至存在多個子群聚的情況，第三組模擬資料包含兩個群聚(圖 3.8b)，檢驗演

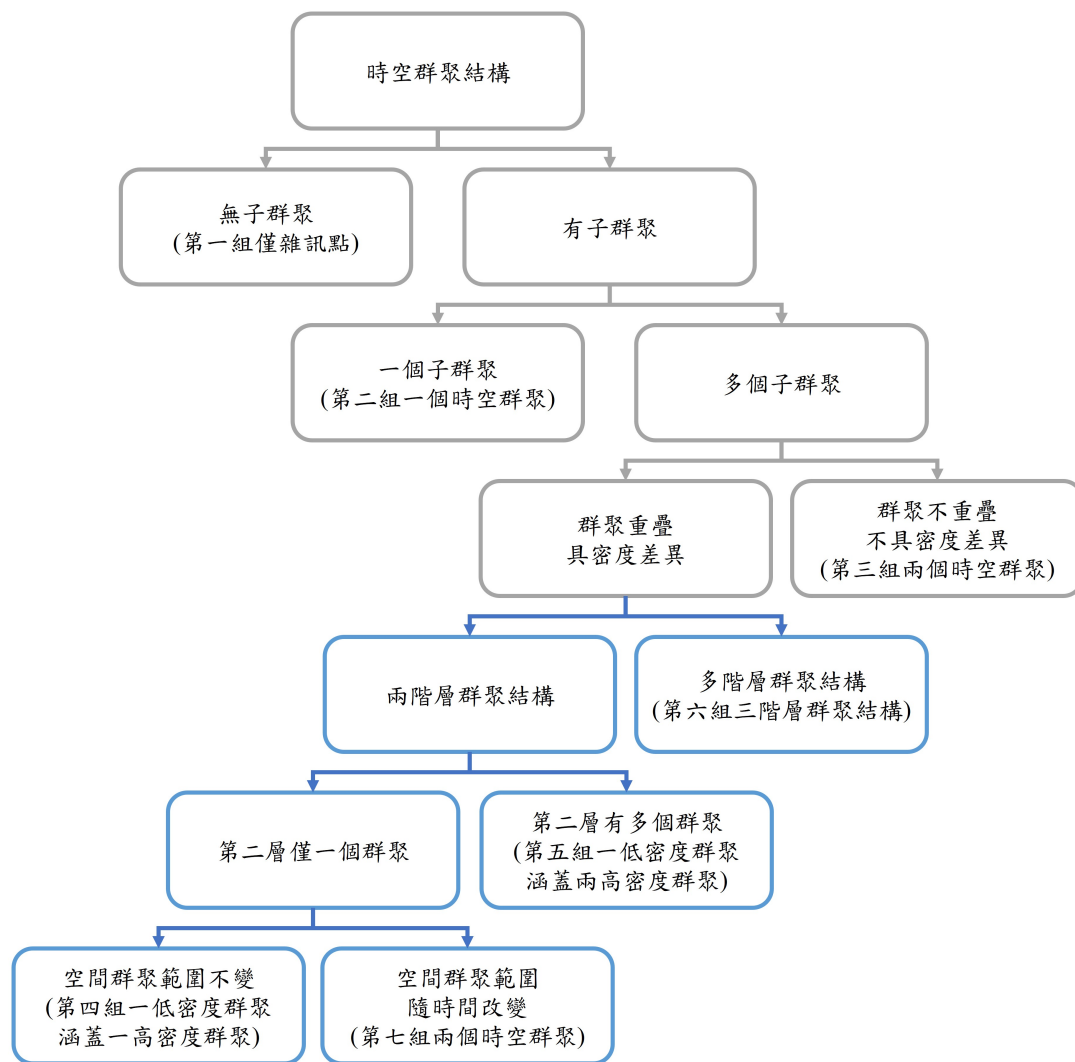
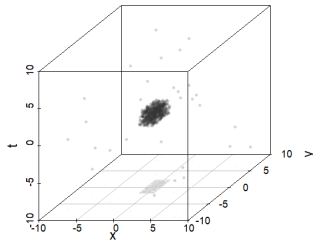


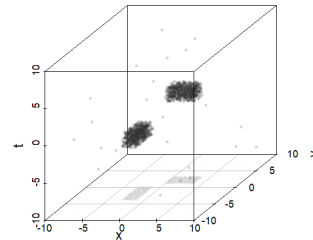
圖 3.7: 模擬資料時空群聚結構特性分類

算法是否能識別多個不重疊，密度夠高之群聚範圍。透過第一組至第三組資料，可檢驗演算法是否能識別出與過去基於密度之時空群聚演算法所能識別出的相同時空群聚結果。接著開始模擬子群聚具備密度差異，且子群聚有重疊關係的群聚結構。在僅具有兩階層群聚結構，且第二層僅包含一個群聚的情況下，第四組模擬資料包含兩個群聚(圖 3.8c)，兩者間具有階層性的包含關係，檢驗演算法是否能識別有重疊，密度有高低差異之兩個群聚範圍。接著考量空間群聚範圍可能隨時間變化的情形，第七組模擬資料也包含兩個群聚(圖 3.8f)，兩者間具有階層性的包含關係，且兩群聚接隨著時間推移而變化空間群聚範圍，此組模擬資料檢驗演算法是否能識別時空群聚範圍變動，且具備階層性關係的時空群聚結構。另外

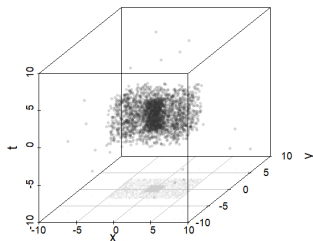
模擬當第二層有多個群聚的情況，第五組模擬資料包含三個群聚(圖 3.8d)，三者間的階層性關係為，密度較高之兩個群聚被包含於一個密度較低之群聚，此組模擬資料檢驗演算法是否能識別多個群聚隸屬於同一群聚的階層性時空群聚結構。最後模擬多階層群聚結構的情況，第六組模擬資料包含三個群聚(圖 3.8e)，三者間的階層性關係為，密度最高的群聚被包含於密度次之的群聚，密度次之的群聚又被包含於密度最低之群聚，此組模擬資料檢驗演算法是否能識別具有多層階層性關係的群聚結構。



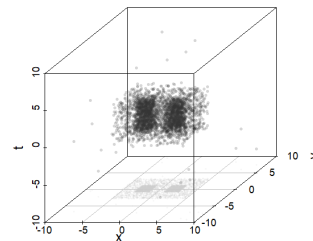
(a) 第二組模擬資料



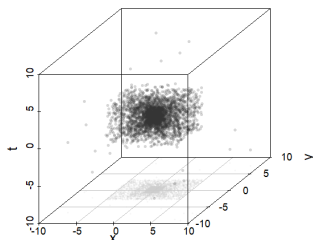
(b) 第三組模擬資料



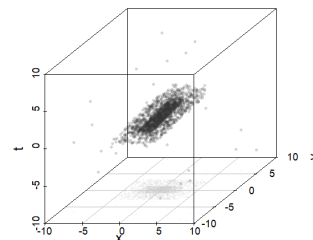
(c) 第四組模擬資料



(d) 第五組模擬資料



(e) 第六組模擬資料



(f) 第七組模擬資料

圖 3.8: 時空模擬資料分布圖

模擬資料中各群聚的分布範圍與密度詳細記錄於附錄 A.1 中，模擬資料中的每個點皆具備 x 與 y 兩軸座標代表空間分布位置， t 軸座標代表時間分布，三軸之

座標值皆各自服從均一分布 (uniform distribution)。在模擬一子群聚中的資料點時，會先透過 x 上界 (upper_x)、x 下界 (lower_x)、y 上界 (upper_y)、y 下界 (lower_y)、t 上界 (upper_t) 以及 t 下界 (lower_t) 來計算三軸之全距，並將三軸之全距相乘得出該時空子群聚之時空體積，時空體積之單位為立方單位，在給定之時空密度之下，將指定之時空密度乘以時空體積可得出該時空群聚範圍內應生成的點數。依據上述模擬資料規則，各群聚之 x、y 與 t 服從聯合機率密度函數 (joint probability density function)：


$$f(x, y, t) = \left(\frac{1}{\text{upper}_x - \text{lower}_x} \right) \times \left(\frac{1}{\text{upper}_y - \text{lower}_y} \right) \times \left(\frac{1}{\text{upper}_t - \text{lower}_t} \right) \quad (3.7)$$

其中

$$x \in [\text{lower}_x, \text{upper}_x], y \in [\text{lower}_y, \text{upper}_y], t \in [\text{lower}_t, \text{upper}_t] \quad (3.8)$$

3.4.2 密度斷層敏感度分析

在 HST-OPTICS 演算法所需的七個參數中，搜尋窗大小 window size 和可及分數差異 diff 兩個參數，對識別密度斷層與群聚結構的結果具有決定性影響。搜尋窗大小代表在多大的排序前後範圍內計算時空可及分數的差異，當可及分數差異固定，搜尋窗越大，越容易將對應位置納入在密度斷層的範圍內，密度斷層分布的範圍越寬廣，識別出的群聚切分位置也會越少，群聚結構層次傾向較少，相對地，搜尋窗越大，則密度斷層的分布範圍越窄，越容易識別出群聚切分位置，群聚結構層次傾向較多且複雜。可及分數差異代表在多大的差距下，兩事件點間應視為存在密度斷層，當搜尋窗大小固定，可及分數差異需求越大，越不容易將對應位置視為有密度斷層存在，最後之查找出的群聚間，傾向有較大的時空密度差異，相對地，可及分數差異需求越小，則越容易將對應位置視為有密度斷層存在，最後之查找出的群聚間，傾向具有較小的時空密度差異。



本研究針對搜尋窗大小 window size 和可及分數差異 diff 兩個參數進行敏感度分析，個別對兩參數設定高中低三種數值進行比較分析，共形成九組參數設定方式，以探討兩個參數設定對不同群聚結構識別結果的影響。針對搜尋窗大小 window size，最大可設定為非雜訊點的總數量，最小可設定為 1，在進行敏感度分析時，高值設定為四成的非雜訊點數量，中值設定為三成的非雜訊點數量，低值則設定為兩成的非雜訊點數量。可及分數差異 diff，最大可設定為非雜訊點時空可及分數的全距，最小可設定為任兩雜訊點間時空可及距離差異的最小值，在進行敏感度分析時，高值設定為非雜訊點可及分數標準差的兩倍，中值設定為非雜訊點可及分數標準差，低值則設定為非雜訊點可及分數標準差的一半。其中，可及分數之標準差，代表任意非雜訊點與鄰近其中一個核心點的距離相對於平均值的變異量。在兩個參數皆有高中低三種設定值之下，共形成九組參數設定方式，九組參數設定方式被使用到七組模擬資料上，以獲取不同參數設定方式下，HST-OPTICS 識別不同群聚結構的成效。

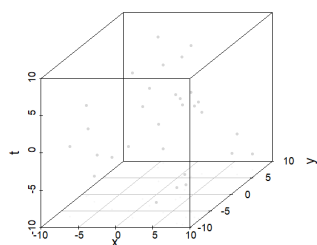


第四章 研究結果

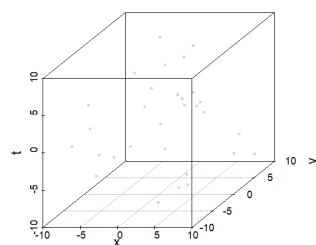
4.1 時空群聚結果比較

HST-OPTICS 相對於 ST-DBSCAN，除了可以排除雜訊點、查找出不同範圍的群聚外，亦可查找出具有階層關係之時空群聚結構。ST-DBSCAN 在第一組至第三組模擬資料上，雖達成預期之群聚識別結果，但在第四組至第七組模擬資料上，在具有重疊關係與密度差異的群聚間，ST-DBSCAN 僅能識別將所有群聚識別為單一群聚。圖 4.1 呈現了第一組時空模擬資料之群聚查找成果，可發現 HST-OPTICS 與 ST-DBSCAN 皆可排除雜訊點。圖 4.2 呈現了第二組時空模擬資料之群聚查找成果，HST-OPTICS 與 ST-DBSCAN 皆可識別出中央單一密度夠高之群聚範圍。圖 4.3 呈現了第三組時空模擬資料之群聚查找成果，HST-OPTICS 與 ST-DBSCAN 皆可識別多個不重疊，密度夠高之群聚範圍。圖 4.4 呈現了第四組時空模擬資料之群聚查找成果，ST-DBSCAN 僅識別出單一群聚範圍，HST-OPTICS 則將具有階層性包含關係且具有密度差異的兩個群聚範圍。圖 4.5 呈現了第五組時空模擬資料之群聚查找成果，ST-DBSCAN 僅識別出單一群聚範圍，HST-OPTICS 則識別出一個密度較低的群聚，又識別出該群聚下又識別出兩個密度較高的群聚，完成多個群聚隸屬於同一群聚的階層性時空群聚結構查找。圖 4.6 呈現了第六組時空模擬資料之群聚查找成果，ST-DBSCAN 僅識別出單一群聚範圍，HST-OPTICS 則識別出高中低密度三個時空群聚，成功識別出多層階層性關係的群聚結構。圖 4.7 呈現了第七組時空模擬資料之群聚查找成果，

ST-DBSCAN 僅識別出單一群聚範圍，HST-OPTICS 則識別出兩個密度不同的群聚，群聚之範圍亦與預期中的相同，兩個群聚的空間群聚範圍皆隨時間推移而改變。

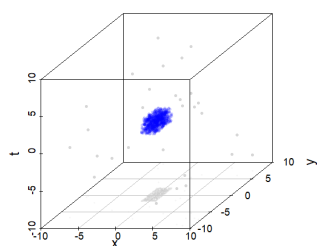


(a) ST-DBSCAN

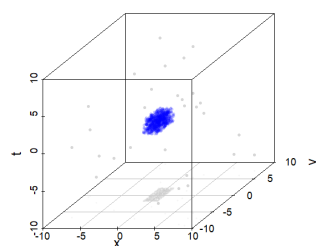


(b) HST-OPTICS

圖 4.1: 第一組模擬資料時空群聚結果

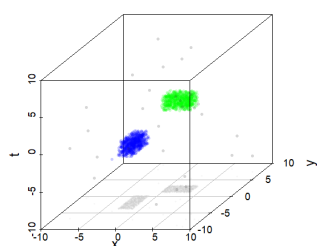


(a) ST-DBSCAN

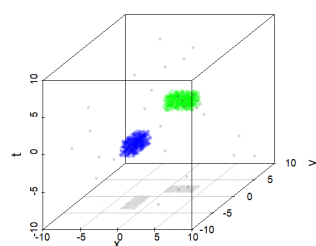


(b) HST-OPTICS

圖 4.2: 第二組模擬資料時空群聚結果

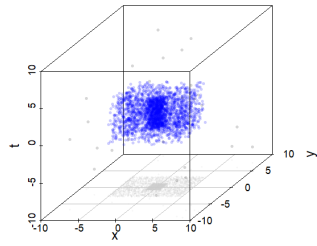


(a) ST-DBSCAN

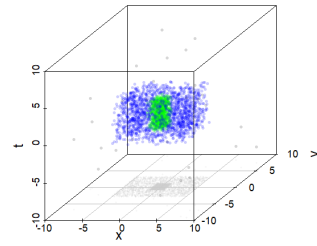


(b) HST-OPTICS

圖 4.3: 第三組模擬資料時空群聚結果

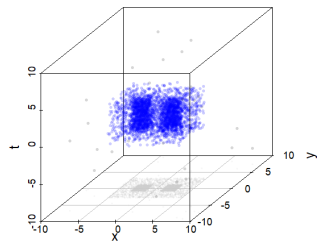


(a) ST-DBSCAN

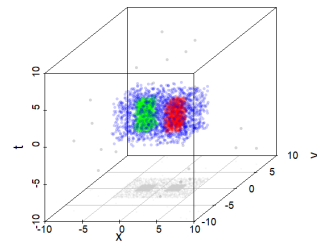


(b) HST-OPTICS

圖 4.4: 第四組模擬資料時空群聚結果

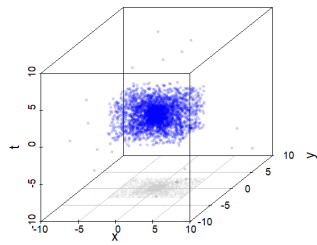


(a) ST-DBSCAN

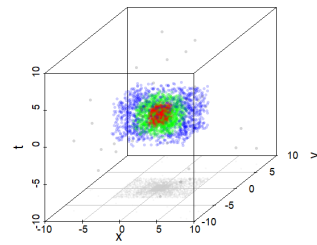


(b) HST-OPTICS

圖 4.5: 第五組模擬資料時空群聚結果

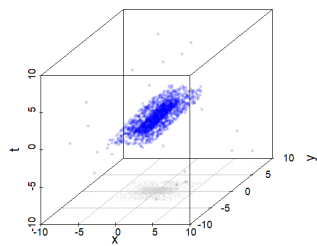


(a) ST-DBSCAN

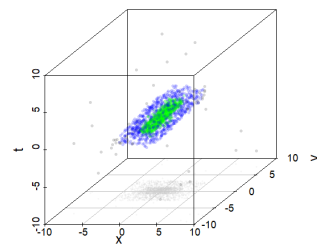


(b) HST-OPTICS

圖 4.6: 第六組模擬資料時空群聚結果



(a) ST-DBSCAN



(b) HST-OPTICS

圖 4.7: 第七組模擬資料時空群聚結果



4.2 密度斷層敏感度分析結果

圖 B.9 呈現了不同參數設定在第二組時空模擬資料之群聚查找成果，其中在子圖 (a) 和 (e) 中的參數設定導致查找出過多群聚。圖 B.10 呈現了不同參數設定在第三組時空模擬資料之群聚查找成果，其中在子圖 (d)、(g)、(h) 和 (i) 中的參數設定導致查找出過多群聚。圖 B.11 呈現了不同參數設定在第四組時空模擬資料之群聚查找成果，其中在子圖 (a)、(d) 和 (f) 中的參數設定導致查找出過多群聚，在子圖 (c)、(g)、(h) 和 (i) 中的參數設定導致查找出過少群聚。圖 B.12 呈現了不同參數設定在第五組時空模擬資料之群聚查找成果，所有試驗的參數組合皆沒有找出預期的群聚結果，所查找出的群聚數量皆小於預期，其中當可及分數差異 diff 過大時，甚至連僅識別整體為一個群聚，而無法查找出子群聚範圍。圖 B.13 呈現了不同參數設定在第六組時空模擬資料之群聚查找成果，其中在子圖 (a) 和 (f) 中的參數設定導致查找出過多群聚，在子圖 (b)、(c)、(d)、(e)、(g)、(h) 和 (i) 中的參數設定導致查找出過少群聚。圖 B.14 呈現了不同參數設定在第七組時空模擬資料之群聚查找成果，所有試驗的參數組合的群聚識別結果皆查找出過少群聚。

各組模擬資料詳細之敏感度分析的依據以及結果圖紀錄於附錄 B 中，包含敏感度分析所使用的參數 (附錄 B.1)、不同參數設定所得之時空可及圖 (附錄 B.2) 以及群聚識別結果圖 (附錄 B.3)。



第五章 討論


5.1 時空群聚結構與密度斷層

既有之 ST-DBSCAN 演算法，不適用於具備密度差異與階層性時空群聚的識別。ST-DBSCAN 演算法僅能在特定的參數設定下，將密度大於指定程度的範圍劃分為群聚，但當群聚結構中存在密度差異時，ST-DBSCAN 無法有效地識別具有不同密度群聚之群聚結構。

本研究所研提之 HST-OPTICS 演算法因能夠識別時空密度斷層，有助於點事件的階層性時空群聚結構之建立。HST-OPTICS 在建立時空可及圖後，可以透過點的排序與點所對應的可及分數變化，識別出在指定搜尋窗大小 window size 範圍內，達成條件的點，條件為該點與範圍內的點之可及分數差異 diff 足夠大，接著對應出多個達成條件且排序連續的點之位置，將這些點的位置分布範圍視為密度斷層所在的範圍，在密度斷層範圍內，依演算法規則選取一個點作為切分點，因點的排序具有特定規則，該切分點可作為群聚劃分的依據，切分點前後的範圍分別為兩個群聚，兩個群聚間具有階層性關係以及時空密度差異。

5.2 參數設定


搜尋窗大小 window size 以及可及分數差異 diff 兩個參數的設定主要依據演算法使用者對群聚寬限範圍和密度差異的期望。搜尋窗大小的參數設定對應於寬限範圍的期望，而可及分數差異的參數設定則對應於密度差異的期望。敏感度分析



的成果為兩個參數的設定提供了建議，然而，兩參數的設定尚無統一的指引，這與過去基於密度演算法面臨的限制相似，這些演算法的群聚查找結果受到參數設定的影響很大。當搜尋窗大小參數的設定小且可及分數差異參數的設定大時，演算法對於密度差異的要求較為嚴格，因此所識別出的時空群聚切分點，可能代表了較為區域化的密度差異發生處。相對地，當搜尋窗大小參數的設定大且可及分數差異參數的設定小時，演算法對於密度差異的要求較為寬鬆，因此所識別出的時空群聚切分點，可能反映出較大範圍中密度斷層存在的情況。

在未來應用中，使用者應根據資料本身的可能群聚結構來設定參數。當研究的點事件傾向出現小範圍內的密度變化，且需要識別出細微的時空群聚，則應使用嚴格度較大的參數設定。在查找交通事故時空群聚，亦即交通之易肇事區時，採用嚴格度較大的參數設定，有助於查找出區域化的時空群聚密度斷層，在密度斷層之發生處，交通事故發生的風險有急遽變化，則高風險群聚範圍則可能是潛在交通改善位置，該亦肇事區之產生可能是源於範圍內號誌設計或道路設計等小範圍的缺失，此為過去 ST-DBSCAN 或 ST-KDE 等方法無法自動查找出群聚範圍。相對地，若研究的點事件傾向在大範圍內才具有密度變化，且需要識別時空群聚的細緻度較低，則應使用較為寬鬆的參數設定。在查找傳染病病例點時空群聚，亦即傳染病嚴重區時，採用嚴格度較小的參數設定，有助於查找出較大範圍中受傳染風險有急遽變化的時空密度斷層發生位置，斷層發生位置為較為平滑化、趨勢化的密度變化位置，斷層範圍內之群聚傳染風險較高的原因，可以連結到更趨勢化的因素，如人流間的密集交流，或是範圍內之環境更有助於病源生長等，管制單位可針對這些造成傳染高風險的因素制定較一般化，適用於大範圍中的管制或改善措施。

計算時空距離時所使用的時間與空間距離權重，代表對於時空差異的重視度，在實務上應考量時間與空間的紀錄單位。空間距離權重 w_S 與時間距離權重



w_T 兩參數用於計算兩點間的時空距離，時空距離越接近的兩點，越容易被歸於通一群聚中。當時間距離權重大於空間距離權重，代表時空距離的衡量受到時間差異變化的影響程度較大，在實務意義上則表示時間上相近的程度會更主導性地影響兩點是否會被視作屬於同一群聚。相對地，當空間距離權重大於時間距離權重，則代表時空距離的衡量受到空間差異變化的影響程度較大。在實務意義上則表示空間上相近的程度會更主導性地影響兩點是否會被視作屬於同一群聚。兩參數的設定若與切分群聚所使用的可及分數差異 diff 參數等比例放大縮小，可以獲得相同的群聚結構查找結果。舉例而言當 $w_S : w_T : \text{diff}$ 由 $1 : 2 : 0.5$ 改變為 $2 : 4 : 1$ ，任兩點間的時空距離會變為原先的 2 倍，所得之時空可及圖，會具有相同的樣態，僅在縱軸上各點對應之可及分數上變為原先的兩倍，此時若也將可及分數差異 diff 參數設定為原先的兩倍，則可以獲得與原先相同之群聚查找結果。

在實務應用上，空間距離與時間差異的衡量都有不同的單位，空間距離可能使用公尺、公里等單位，時間差異可能使用分鐘、秒鐘、週數、天數等單位，在設定空間距離權重 w_S 與時間距離權重 w_T 兩參數時，應考量任兩點間空間差異與時間差異的變異程度，以選定適切的權重參數來決定時間距離與空間距離對於兩點間時空距離的主導程度。



第六章 結論

本文中提出了一種新的基於密度群聚演算法 HST-OPTICS，旨在處理時空群聚之局部分析問題。此演算法基於 OPTICS 演算法調整，能夠識別出包含不同密度群聚之階層性時空群聚結構。為了克服 OPTICS 在群聚識別上過度嚴格的限制，方法中引入了搜尋窗大小 window size 的概念，旨在放寬 OPTICS 中陡度之識別嚴格度，讓搜尋窗內達一定可及分數差異 diff 時，即可認定出對應之密度斷層，以劃分出群聚範圍。驗證結果表示，HST-OPTICS 可以識別過去演算法亦可以識別，具備四大特性的群聚，四大特性包含總群聚數非經給定、排除雜訊點且群聚形狀任意，除此以外，HST-OPTICS 也可以識別過去演算法無法識別，具備兩大特性的群聚結構，兩大特性為具密度差異與具階層性，群聚結構如重疊且具有高低密度差異之多個群聚結構，多個群聚隸屬於同一群聚的階層性時空群聚結構，以及時空群聚範圍變動且具備階層性關係的時空群聚結構。

儘管有上述之研究貢獻，本文在時空距離衡量以及時空群聚成效量化驗證上有研究之限制。在時空距離衡量上，由於本文使用的是模擬資料，資料在三維度上沒有意義差異，因此在時空距離之衡量公式上，僅考慮使用歐式距離 (euclidean distance) 加上空間距離權重 w_S 與時間距離權重 w_T 來衡量時空距離。在時空群聚成效量化驗證上，本文旨在查找出子群聚具備密度差異與階層性的時空群聚結構，然而現有之時空群聚衡量指標，多以測量群聚內相似度或計算群間相異度來設計驗證指標，不適用於衡量階層性時空群聚查找結果之好壞，因此本研究僅以三維分布圖標註顏色的方式來呈現時空群聚之查找結果，並未使用量化指標來呈

現群聚之查找成效。

在後續研究上，可以著重在實務應用、階層性時空群聚驗證指標設計以及演算法效能提高上。首先，本文已展示了演算法在識別出過去無法查找之包含密度差異時空群聚結構的有效成果，未來研究可以善用此方法於各領域之實務應用中，查找適用於各式實際資料之參數設定值，參數包含計算時空距離所需的空間距離權重 w_S 與時間距離權重 w_T ，以及劃分群聚所需的搜尋窗大小 window size 與可及分數差異 diff，以達成階層性時空群聚的查找。此外，現有之時空群聚成效衡量指標並不適用於階層性時空群聚查找的成效比較，未來研究可嘗試設計適合用以評斷階層性時空群聚查找成效的指標。最後，面對實務上龐大的資料量，未來研究可著重在提高演算法之運行效能，以及優化演算法的運行速度上。



參考文獻

- Agrawal, K. P., Garg, S., Sharma, S., and Patel, P. (2016). Development and validation of optics based spatio-temporal clustering technique. Information Sciences, 369:388–401.
- Ankerst, M., Breunig, M. M., Kriegel, H.-P., and Sander, J. (1999). Optics: ordering points to identify the clustering structure. SIGMOD Rec., 28(2):49–60.
- Baker, R. D. (1996). Testing for space-time clusters of unknown size. Journal of Applied Statistics, 23(5):543–554. doi: 10.1080/02664769624080.
- Birant, D. and Kut, A. (2007). St-dbscan: An algorithm for clustering spatial-temporal data. Data Knowledge Engineering, 60(1):208–221.
- Brunsdon, C., Corcoran, J., and Higgs, G. (2007). Visualising space and time in crime patterns: A comparison of methods. Computers, Environment and Urban Systems, 31(1):52–75.
- Diggle, P., Chetwynd, A., Häggkvist, R., and Morris, S. (1995). Second-order analysis of space-time clustering. Statistical Methods in Medical Research, 4(2):124–136.
- Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, KDD'96, page 226–231. AAAI Press.

Izakian, H., Pedrycz, W., and Jamal, I. (2013). Clustering spatiotemporal data: An augmented fuzzy c-means. IEEE Transactions on Fuzzy Systems, 21(5):855–868.

Jacquez, G. M. (1996). A k nearest neighbour test for space–time interaction. Statistics in Medicine, 15(18):1935–1949.

Kim, Y. S., Hae, Y., and Intaek (2012). Spatio-temporal gaussian mixture model for background modeling. In 2012 IEEE International Symposium on Multimedia, pages 360–363.

Knox, E. G. and Bartlett, M. S. (1964). The detection of space-time interactions. Journal of the Royal Statistical Society. Series C (Applied Statistics), 13(1):25–30.

Kuo, F.-Y., Wen, T.-H., and Sabel, C. E. (2018). Characterizing diffusion dynamics of disease clustering: A modified space–time dbscan (mst-dbscan) algorithm. Annals of the American Association of Geographers, 108(4):1168–1186.

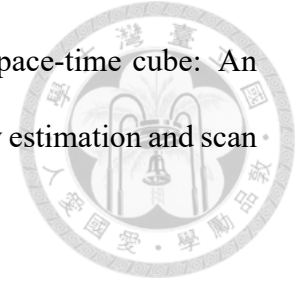
Lamb, D. S., Downs, J., and Reader, S. (2020). Space-time hierarchical clustering for identifying clusters in spatiotemporal point data. ISPRS International Journal of Geo-Information, 9(2):85.

Levin, S. A. (1992). The problem of pattern and scale in ecology: The robert h. macarthur award lecture. Ecology, 73(6):1943–1967.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, volume 1, pages 281–297. Oakland, CA, USA.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. Cancer Research, 27(2_Part_1):209–220.

Nakaya, T. and Yano, K. (2010). Visualising crime clusters in a space-time cube: An exploratory data-analysis approach using space-time kernel density estimation and scan statistics. Transactions in GIS, 14(3):223–239.



O'sullivan, D. and Unwin, D. (2003). Geographic information analysis. John Wiley Sons.

Rhys, H. (2020). Machine Learning with R, the tidyverse, and mlr. Simon and Schuster.

Silverman, B. W. (2018). Density estimation for statistics and data analysis. Routledge.

Song, J., Wen, R., and Yan, W. (2018). Identification of traffic accident clusters using kulldorff's space-time scan statistics. In 2018 IEEE International Conference on Big Data (Big Data), pages 3162–3167.

Stauffer, C. and Grimson, W. E. L. (1999). Adaptive background mixture models for real-time tracking. In Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149), volume 2, pages 246–252 Vol. 2.

Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58(301):236–244.



附錄 A — 模擬資料生成規則

A.1 模擬群聚分布範圍與密度

表 A.1: 各組群聚結構下群聚之分布範圍與密度

群聚結構	子群聚	時空密度	x 下界	x 上界	y 下界	y 上界	t 下界	t 上界	點數
第一組	僅雜訊點		-9.9	9.9	-9.9	9.9	-9.9	9.9	30
第二組	1	50	-1	1	-2	2	-1	1	800
第三組	1	50	-3	-1	-4	0	-3	-1	800
第三組	2	50	0	4	1	3	1	3	800
第四組	1	5	-5	5	-3	3	-3	3	1800
第四組	2	50	-1	1	-1	1	-2	2	800
第五組	1	5	-5	5	-3	3	-3	3	1800
第五組	2	50	-3	-1	-1	1	-2	2	800
第五組	3	50	1	3	-1	1	-2	2	800
第六組	1	5	-5	5	-3	3	-3	3	1800
第六組	2	10	-2.5	2.5	-2	2	-2	2	800
第六組	3	100	-1	1	-1	1	-1	1	800
第七組	1	5	-5*	5*	-3	3	-3	3	1058
第七組	2	50	-3*	3*	-1	1	-2	2	791

* x 上下界隨 t 而變動



附錄 B — 群聚查找相關圖表

B.1 HST-OPTICS 敏感度分析參數設定

表 B.2: 第二組時空群聚敏感度分析參數設定

參數大小	<i>window size</i>	<i>diff</i>
低	160	0.014778905
中	240	0.029557809
高	320	0.059115619
最佳參數	500	0.4

表 B.3: 第三組時空群聚敏感度分析參數設定

參數大小	<i>window size</i>	<i>diff</i>
低	320	0.016370116
中	480	0.032740231
高	640	0.065480462
最佳參數	500	0.4

表 B.4: 第四組時空群聚敏感度分析參數設定

參數大小	<i>window size</i>	<i>diff</i>
低	521	0.201339798
中	781	0.402679597
高	1041	0.805359194
最佳參數	500	0.4



表 B.5: 第五組時空群聚敏感度分析參數設定

參數大小	<i>window size</i>	<i>diff</i>
低	681	0.20361365
中	1021	0.4072273
高	1361	0.8144546
最佳參數	400	0.4

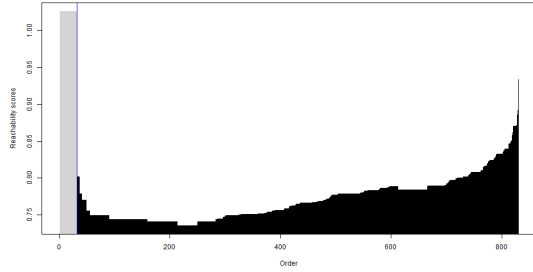
表 B.6: 第六組時空群聚敏感度分析參數設定

參數大小	<i>window size</i>	<i>diff</i>
低	681	0.226263133
中	1021	0.452526266
高	1361	0.905052531
最佳參數	340	0.13

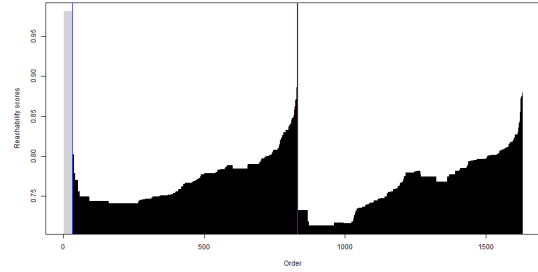
表 B.7: 第七組時空群聚敏感度分析參數設定

參數大小	<i>window size</i>	<i>diff</i>
低	364	0.202875388
中	545	0.405750776
高	727	0.811501552
最佳參數	500	0.45

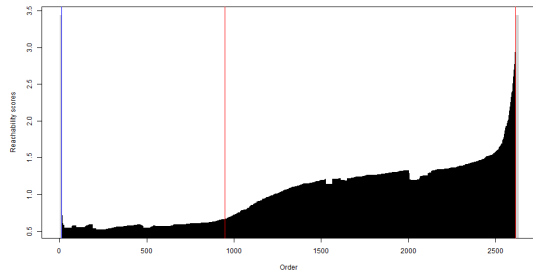
B.2 HST-OPTICS 可及圖



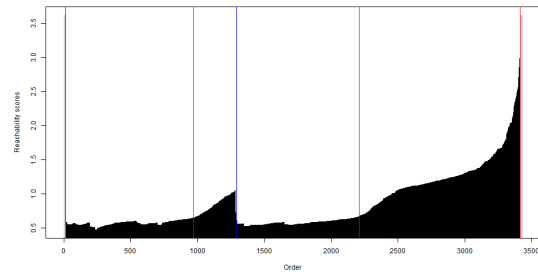
(a) 第二組模擬資料最佳參數設定



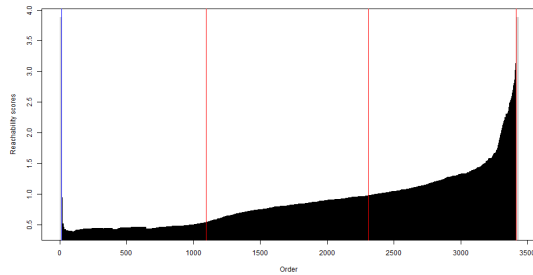
(b) 第三組模擬資料最佳參數設定



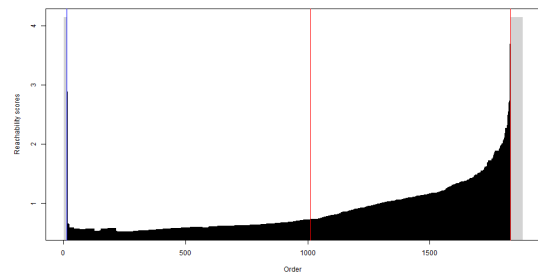
(c) 第四組模擬資料最佳參數設定



(d) 第五組模擬資料最佳參數設定



(e) 第六組模擬資料最佳參數設定



(f) 第七組模擬資料最佳參數設定

圖 B.1: 最佳參數設定下時空模擬資料 HST-OPTICS 時空可及圖

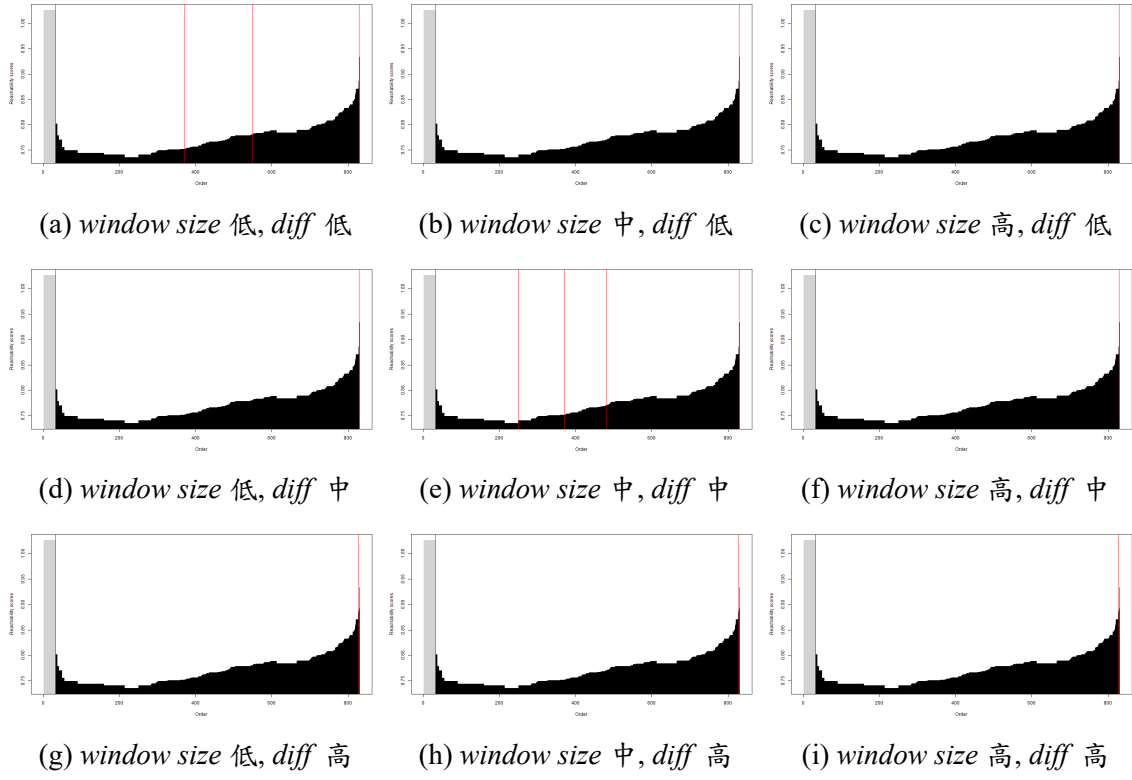
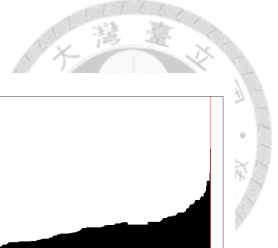


圖 B.2: 第二組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖

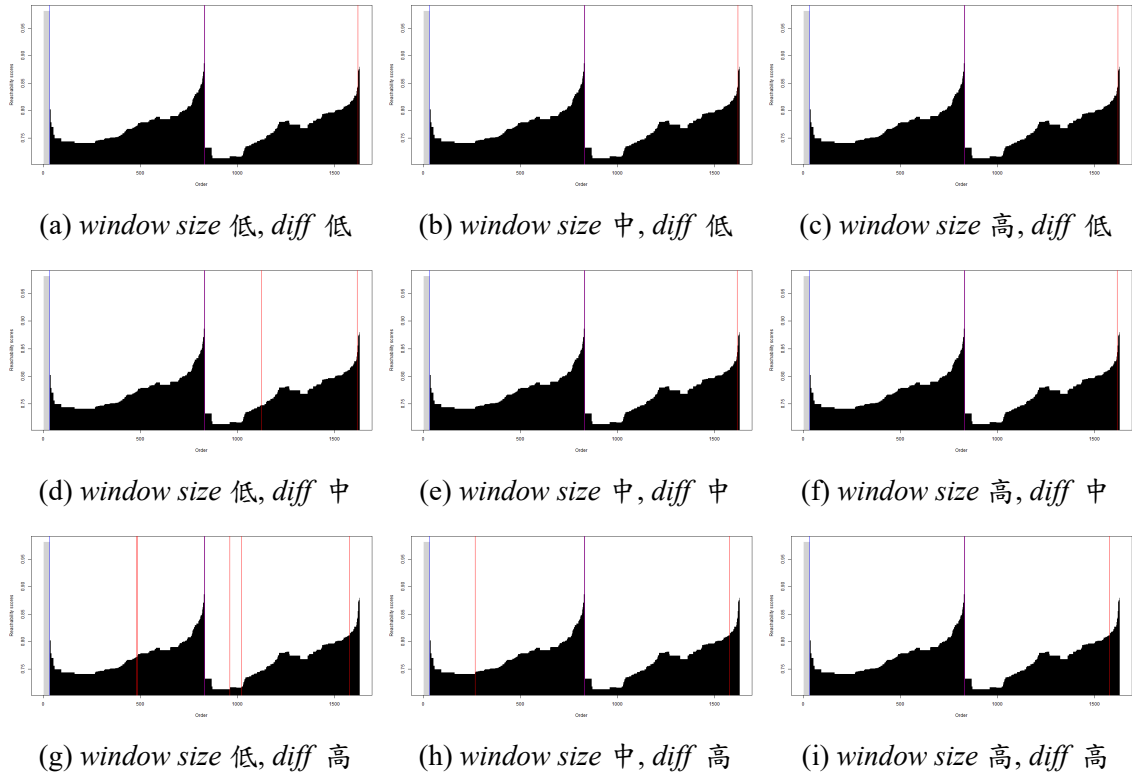


圖 B.3: 第三組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖

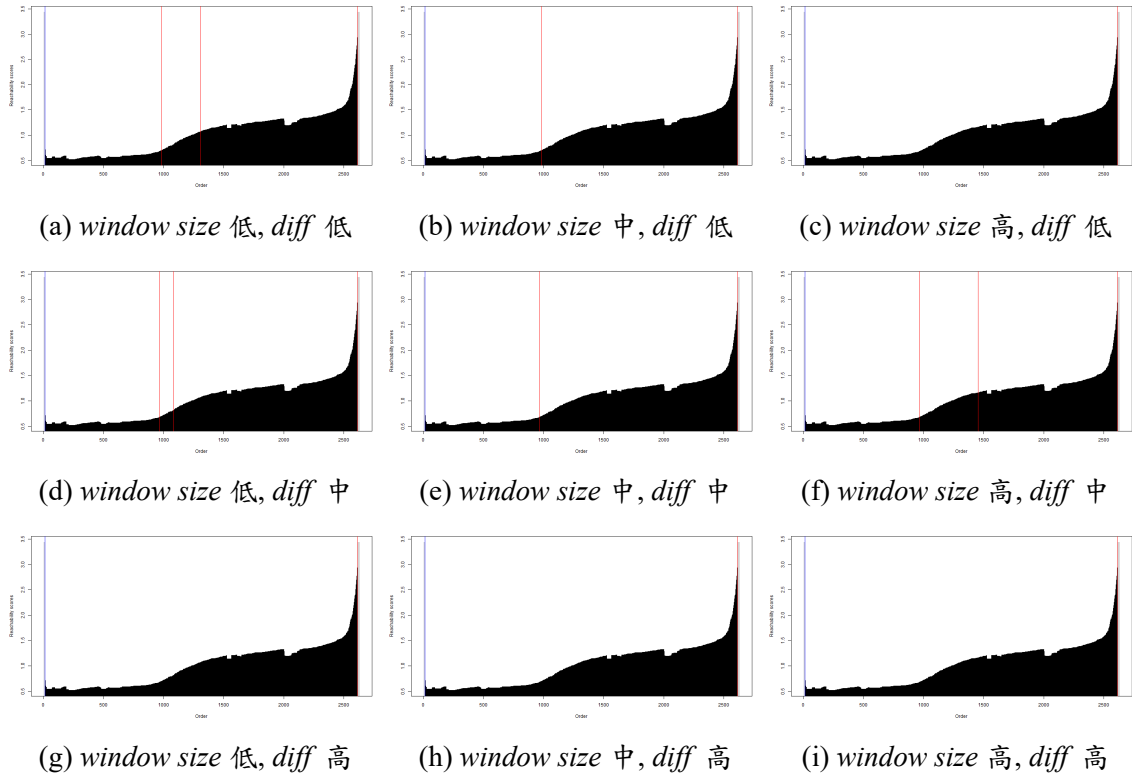
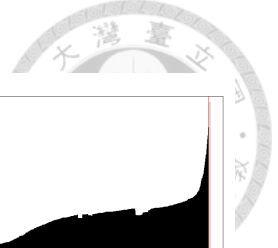


圖 B.4: 第四組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖

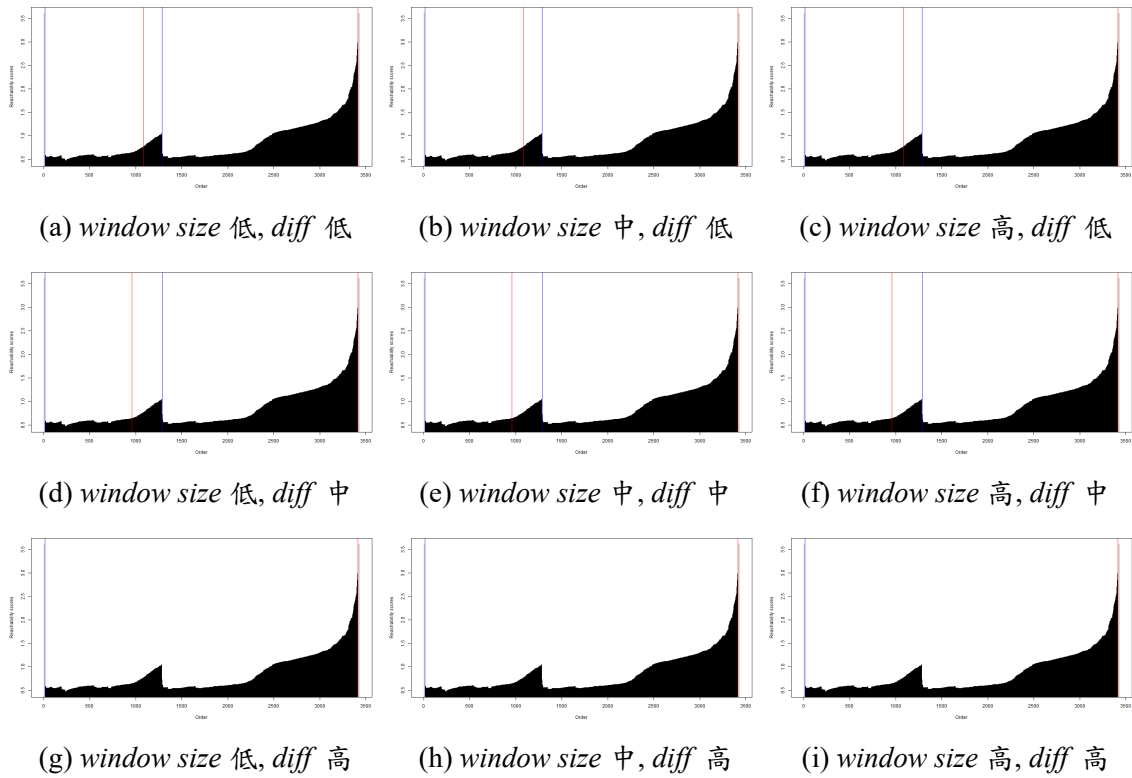


圖 B.5: 第五組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖

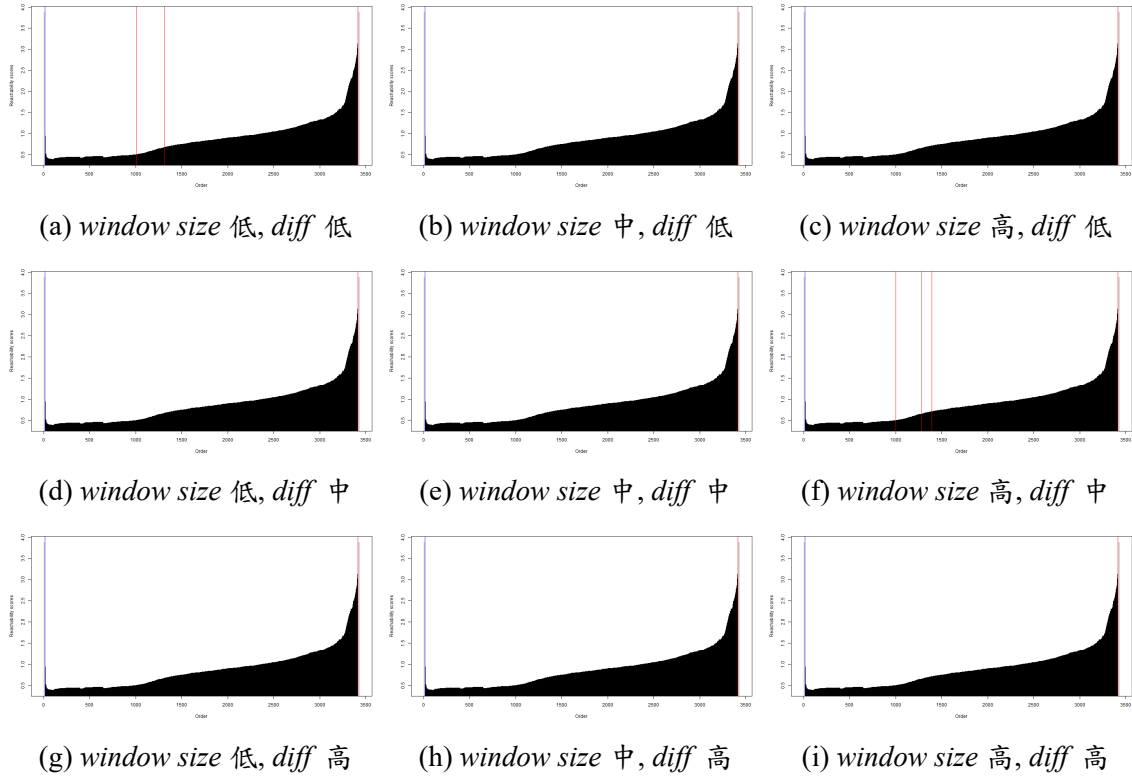
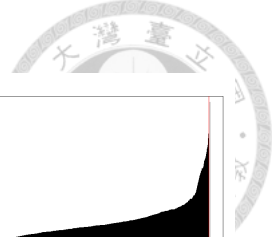


圖 B.6: 第六組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖

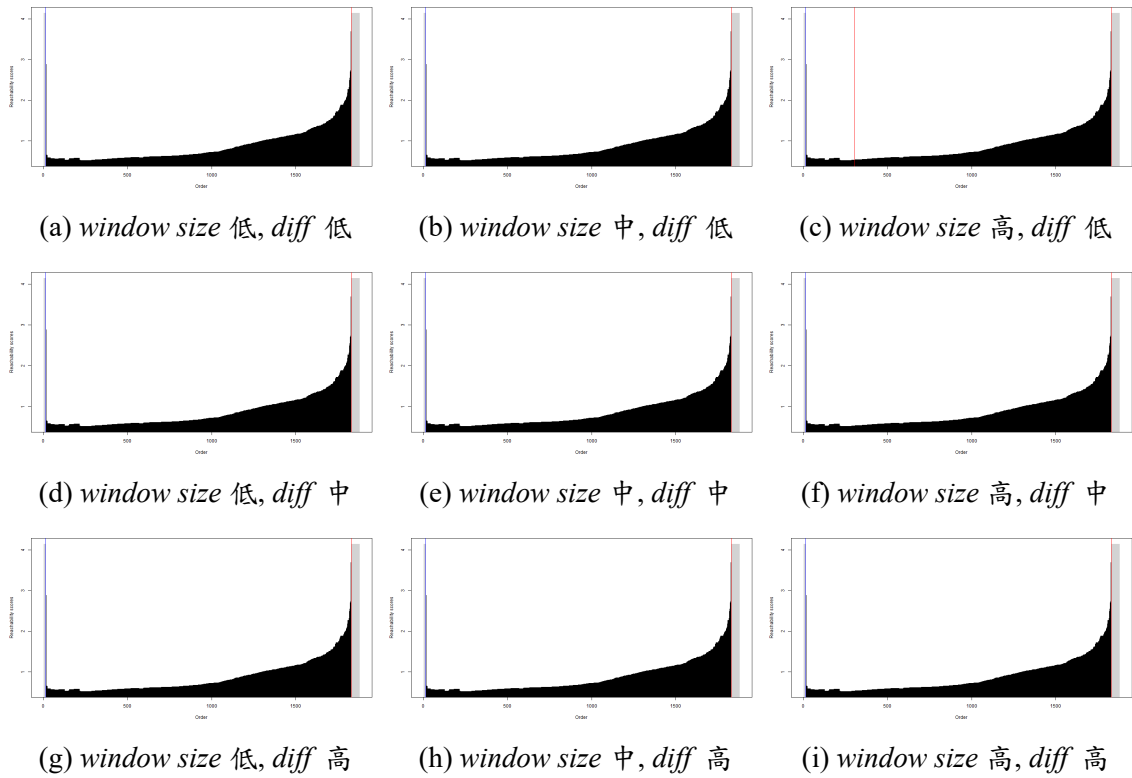
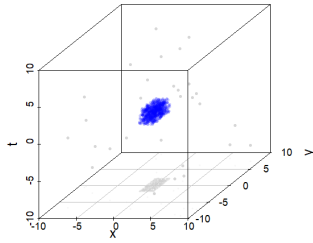
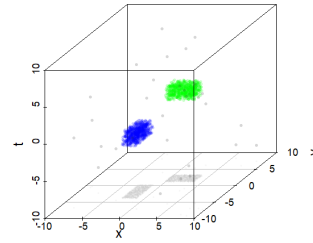


圖 B.7: 第七組時空模擬資料在各參數設定下之 HST-OPTICS 時空可及圖

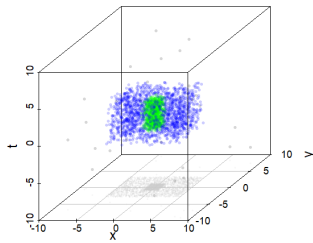
B.3 HST-OPTICS 群聚圖



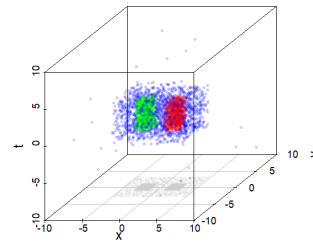
(a) 第二組模擬資料最佳參數設定



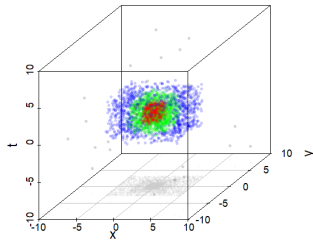
(b) 第三組模擬資料最佳參數設定



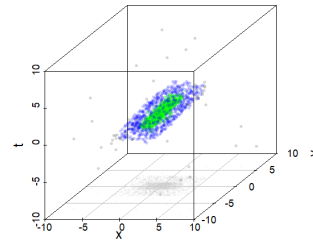
(c) 第四組模擬資料最佳參數設定



(d) 第五組模擬資料最佳參數設定



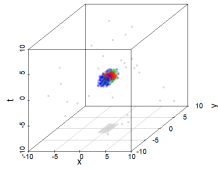
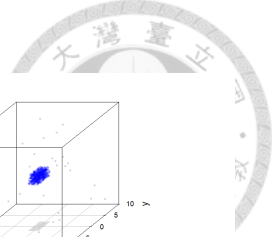
(e) 第六組模擬資料最佳參數設定



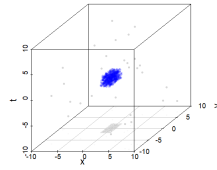
(f) 第七組模擬資料最佳參數設定

圖 B.8: 最佳參數設定下時空模擬資料 HST-OPTICS 群聚圖

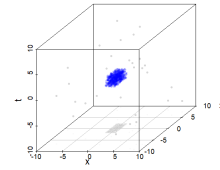
B.4 HST-OPTICS 群聚結構圖



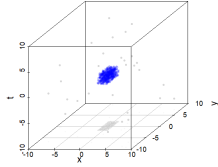
(a) window size 低, diff 低



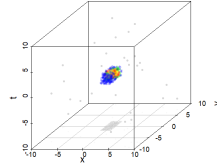
(b) window size 中, diff 低



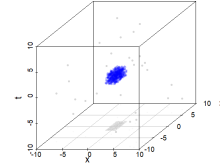
(c) window size 高, diff 低



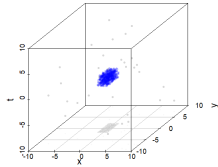
(d) window size 低, diff 中



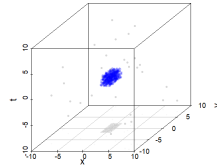
(e) window size 中, diff 中



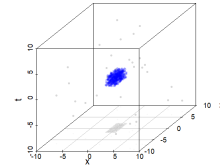
(f) window size 高, diff 中



(g) window size 低, diff 高

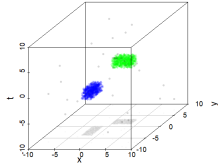


(h) window size 中, diff 高

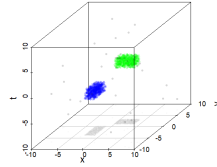


(i) window size 高, diff 高

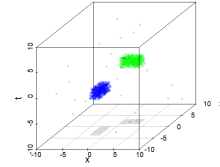
圖 B.9: 第二組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖



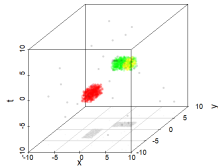
(a) window size 低, diff 低



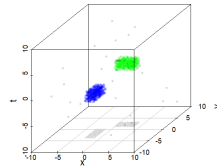
(b) window size 中, diff 低



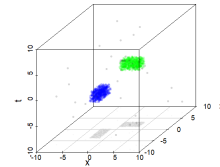
(c) window size 高, diff 低



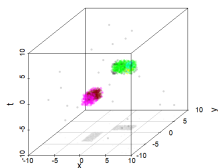
(d) window size 低, diff 中



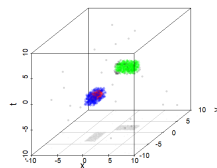
(e) window size 中, diff 中



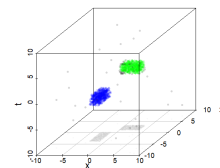
(f) window size 高, diff 中



(g) window size 低, diff 高

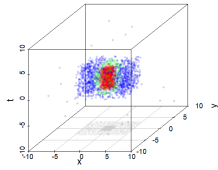
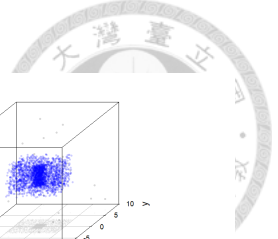


(h) window size 中, diff 高

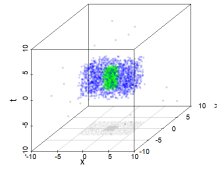


(i) window size 高, diff 高

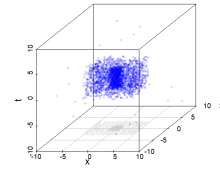
圖 B.10: 第三組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖



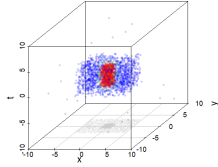
(a) window size 低, diff 低



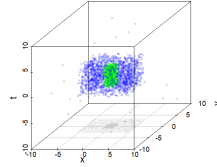
(b) window size 中, diff 低



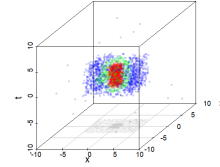
(c) window size 高, diff 低



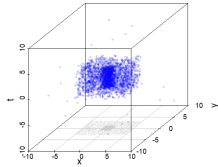
(d) window size 低, diff 中



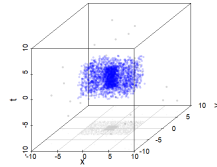
(e) window size 中, diff 中



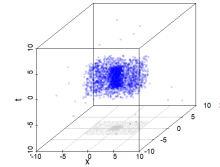
(f) window size 高, diff 中



(g) window size 低, diff 高

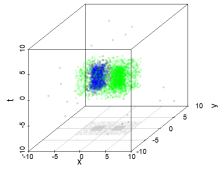


(h) window size 中, diff 高

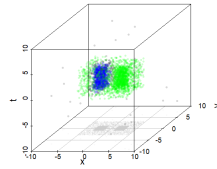


(i) window size 高, diff 高

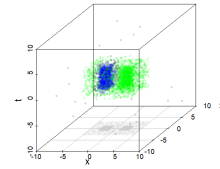
圖 B.11: 第四組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖



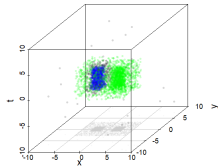
(a) window size 低, diff 低



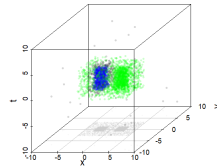
(b) window size 中, diff 低



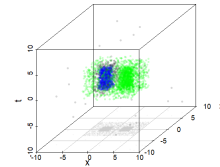
(c) window size 高, diff 低



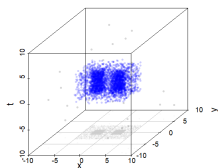
(d) window size 低, diff 中



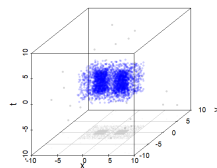
(e) window size 中, diff 中



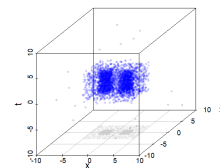
(f) window size 高, diff 中



(g) window size 低, diff 高

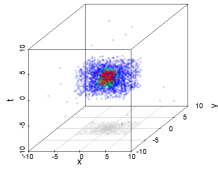
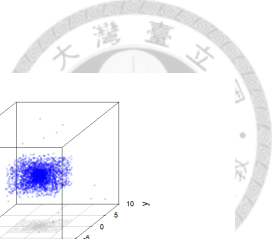


(h) window size 中, diff 高

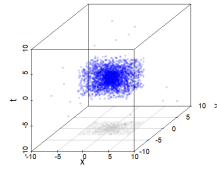


(i) window size 高, diff 高

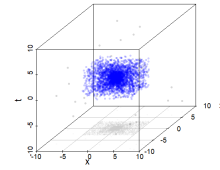
圖 B.12: 第五組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖



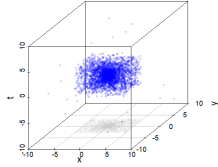
(a) window size 低, diff 低



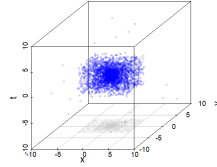
(b) window size 中, diff 低



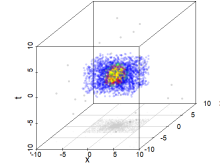
(c) window size 高, diff 低



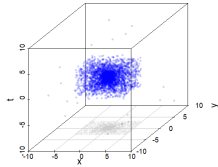
(d) window size 低, diff 中



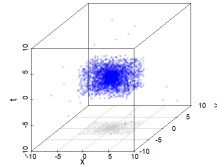
(e) window size 中, diff 中



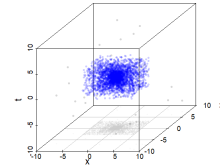
(f) window size 高, diff 中



(g) window size 低, diff 高

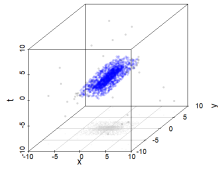


(h) window size 中, diff 高

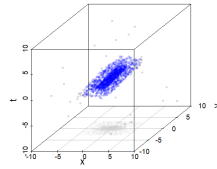


(i) window size 高, diff 高

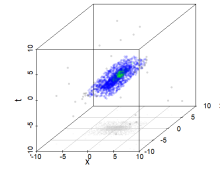
圖 B.13: 第六組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖



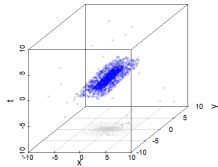
(a) window size 低, diff 低



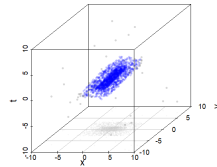
(b) window size 中, diff 低



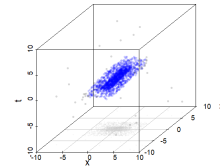
(c) window size 高, diff 低



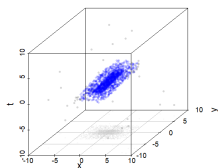
(d) window size 低, diff 中



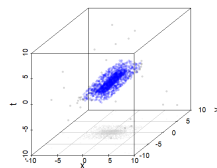
(e) window size 中, diff 中



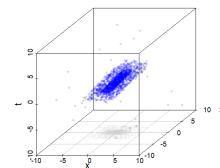
(f) window size 高, diff 中



(g) window size 低, diff 高



(h) window size 中, diff 高

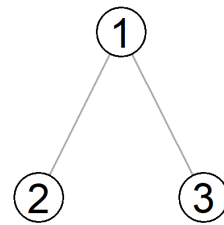


(i) window size 高, diff 高

圖 B.14: 第七組時空模擬資料在各參數設定下之 HST-OPTICS 群聚圖



(a) 第四組模擬資料群聚結構圖



(b) 第五組模擬資料群聚結構圖



(c) 第六組模擬資料群聚結構圖



(d) 第七組模擬資料群聚結構圖

圖 B.15: 最佳參數設定下時空模擬資料 HST-OPTICS 時群聚結構圖