國立臺灣大學理學院地理環境資源學系

博士論文

Department of Geography
College of Science
National Taiwan University
Doctoral Dissertation

理解傳染病的空間擴散：地理計算的方法

Understanding Spatial Diffusion of Infectious Diseases: Geo-computational Approaches

郭飛鷹

Fei-Ying Kuo

指導教授：溫在弘 博士

Advisor: Tzai-Hung Wen, Ph.D.

中華民國 110 年 7 月

July, 2021

doi:10.6342/NTU202101357

誌謝

碩士班一年加上逕讀博士班五年，這六年是我成長幅度最多的時光。不管是從事學術研究、撰寫程式語言、資料量化分析、問題邏輯思考等能力，幾乎都是在這段時間所培養出來的，也正是這些能力使我能順利完成這本論文，並取得博士學位。

能有如此豐碩的成果，首要感謝的自然是我的指導教授溫在弘老師。在研究所的生涯中，溫老師的鞭策雖然有時候會帶來莫大的壓力，但事後回想時才發覺這也是推動自己向前邁進的助力。另一方面，在研究陷入低潮時，溫老師也時常透過自身的經驗分享來激勵人心，並適時地指引可以調整、嘗試的方向，讓研究能夠順利進行下去。總而言之，溫老師的從旁協助與引導是我能夠完成這本論文的重要因素之一。

除了指導教授，也要感謝口試委員在資格考試和學位考試的兩次口試階段提出發人深省的問題，協助我釐清研究的主軸和方向。同時，也要感謝同是甘苦人的威全和竣翔，時常與我討論研究，協助釐清自己沒有發現的研究盲點，並分享研究上創新的想法或訊息；特別感謝威全當初手把手地教我 Python，讓我能在段時間內奠定基礎，達到能夠上網自學的程度。感謝曾經一起待在研究室的助理同仁和學長姐弟妹們，不管是研究上的討論或是生活趣事的分享都調劑了單調的學術研究生活。最後，感謝家人、朋友的支持，讓我在龐大的壓力下可以有心靈和精神上的支柱。

這本論文，各位或多或少都參與其中，也終於到了我們一起品嘗豐收成果的時刻。這六年的時光說長不長，說短也不短，感謝在這段人生最黃金的歲月中，有各位的協助讓我完成了一件意義非凡的事情！

doi:10.6342/NTU202101357

中文摘要

對抗傳染性疾病有兩大面向：防止傳染病擴散以及準確評估各地區對於醫療資源可及性的高低。前者是避免疾病的地理影響範圍繼續擴大；後者則是力求保護疫區居民的健康、減少生命財產損失。而要有效達成這兩個面向的防疫工作，關鍵前提就是要對於傳染病擴散的時空特徵有足夠瞭解。從理解到防止再到保護，這三個面向的整合涉及了防疫過程中的大部分重要決策。然而，文獻上針對這三個面向的研究皆存在方法學上的不足。首先，過往已有文獻透過量性描述群聚的擴散型態來理解傳染病的擴散特徵，但是僅止於學理性的討論，缺乏系統性的分類來定義各種可能的群聚擴散型態；同時，也缺乏相對應的方式來自動剖析各種型態出現的時間和位置。第二，移動管制分區是最為嚴厲的防止疾病擴散的手段，但在面臨疾病威脅的極端狀況時（例如新冠肺炎），也是普遍被多數國家採用。實務上的移動管制分區大多以現行的行政區層級做為管制單位，忽略了人口移動行為的特徵，導致容易劃定地理範圍過大或過小的管制分區，致使防止擴散的效益不如預期。最後，在傳染病疫情中，醫療資源需求會隨著疫情嚴重程度起伏而變動，導致各地區的可及性出現時序性的變化。然而，目前並未有文獻考量到如此的變動特性，所以也缺乏合適的模型進行動態評估。這些方法學上的不足可能會產生有偏誤的分析結果或是無法揭露重要的疫情訊息，進而影響到防疫決策的判斷。因此，本論文旨在針對這三個面向的分析方法分別進行改良。研究成果包含了能自動判斷各種群聚演化擴散型態的 MST-DBSCAN 演算法、考量人口移動規律性進行區域劃分的 HuMoRZ 演算法、以及 Epi-RA 模式利用傳染病擴散模擬將資源需求的動態變化整合至空間可及性評估模型當中。本論文以實際案例應用的方式來證實這三個方法的實用性，相關成果也已透過國際學術期刊文章或專書專章的形式進行發表，進一步證實本論文的價值和貢獻。

關鍵字：傳染病擴散、時空過程、MST-DBSCAN、HuMoRZ、Epi-RA

doi:10.6342/NTU202101357

英文摘要

Confronting infectious diseases contains two major aspects: preventing epidemic diffusion and evaluating spatial accessibility to medical resources. The former is to control the influenced area; the latter is to protect residents' health and lives. Moreover, understanding space-time process of epidemic diffusion is critical for these two aspects. The three aspects are related to almost all decision makings during an epidemic; however, some methodological shortages exist. First, to understand epidemic diffusion, a systematical classification of evolution types of disease clusters and a method to automatically profile location and time of every type is still lack. Second, to prevent epidemic diffusion, containment zone is the most severe approach used in an extreme situation. A containment zone is usually formed by administrative districts, yet it is not effective due to the neglect of human movement properties. Finally, demand for medical resource may vary over time based on severity of an epidemic, which result in spatiotemporally dynamic accessibility, yet such dynamic is still neglected in the literature These methodological shortages may generate inappropriate results and thus negatively affect decision makings. Therefore, this thesis developed the following three new methods: (1) MST-DBSCAN, which can automatically profile various cluster evolution types, (2) HuMoRZ, which considers the regularity of human mobility to delineate containment zones, and (3) Epi-RA, which integrates evaluation of spatial accessibility with simulation of epidemic diffusion. These methods' feasibilities are demonstrated through real case studies, and they have also been published in international academic journal or book, which further proves their contributions.

Keywords: Epidemic diffusion, Space-time process, MST-DBSCAN, HuMoRZ, Epi-RA

doi:10.6342/NTU202101357

目錄

doi:10.6342/NTU202101357

圖目錄

doi:10.6342/NTU202101357

表目錄

doi:10.6342/NTU202101357

第一章 緒論

1.1 動機背景

近年來，韌性城市的概念逐漸興起。因為當防災工程無法完全阻隔天然災害於人類社會之外，學習如何與災害共處、調適，才能夠降低災害帶來的衝擊和損失 (Evans 2011)。而除了天然災害以外，人類社會近年來也不斷遭受到新興傳染病的威脅，包括：2003 年的 SARS、2009 年的 A 型流感、2014 年的伊波拉病毒、以及從 2019 年底持續到現在的新冠肺炎。當這些傳染病演變成季節性流行並永久存在於人類社會中，便如同天然災害一樣隨時會有大規模發生的可能性。因此，為了確保社會經濟秩序得以在不受傳染病嚴重影響的情況下維持正常的運作，人們必須學習如何防止傳染病大規模擴散並找出醫療防疫資源可能較缺乏的地區以便進行補強。而要使這兩項措施發揮其最大效益，關鍵前提便是要對於傳染病擴散的時空特性具備足夠的瞭解 (Sabel, Pringle, and Schoerstrom 2009)。從理解到防止再到保護，這三個面向的整合涉及了防疫過程中的大多數重要決策；不過目前文獻上關於這三個面向的量性分析方法仍舊存在學理上的瑕疵或疏漏，致使分析結果無法有效揭露重要的疫情擴散或是醫療資源缺乏等訊息，降低了其本身對於決策的支援能力，底下將詳述之。

傳染病擴散指涉到疾病從發源地向外蔓延至其他地區的現象 (Meade and Emch 2010)。這樣的擴散過程往往是透過新病例不斷出現於新地點所揭露，顯示出傳染病疫情在時間和空間維度上的演化狀況 (Cliff and Haggett 2006)。為了量性描述擴散的時空過程，過往有文獻利用傳染病單一群聚的中心點和群聚範圍等兩項特徵提出學理性的描述 (Cliff et al. 1981)；也有文獻發現疫情群聚之間可能存在交互作用，例如兩地的疫情群聚融合、或是單一群聚分裂開來等現象 (Barreto et al. 2008; Kalnis, Mamoulis, and Bakiras 2005; Kan et al. 2008; Lian et al. 2007)。然而，這些觀察或發現皆僅止於概念層次的討論，尚未有一個系統性的整理和分類來考

doi:10.6342/NTU202101357

量各種可能的擴散樣態。同時，也缺乏了相對應的方法來剖析一場疫情當中，每個群聚在各個時間點下的擴散樣態。究竟是群聚的中心點移動還是範圍擴大？又或者是與其他群聚發生交互作用而改變了疫情的發展？缺乏系統性的整理分類和相對應的分析方法，導致無法在此脈絡下清楚地量性描述傳染病的擴散過程，進而造成我們對於擴散的時空過程無法完整理解。

以此次的新冠肺炎疫情為契機，人們認識到了多種防止傳染病擴散的措施，包含了配戴口罩、維持社交距離、國家邊境管制、甚至是封城這種較為嚴厲的手段。在這些措施當中，封城管制會造成最嚴重的經濟損失和人身自由限制，而如此巨大的犧牲是否能換來等價的防疫效果，仍舊是文獻上經常討論的議題。其主要問題有二：首先，大規模的封城管制可能會涵蓋到許多沒有顯著風險的民眾 (Rubin and Wessely 2020)，所引起的社會恐慌容易迫使區域內的居民逃離 (Coltart et al. 2017)，進而造成防疫的漏洞。第二，維持大規模封城需要龐大的人力物力，還得搭配如同中國那樣的政治體制，才有可能像中國政府一樣成功以封城措施來控制武漢的疫情。因此，並非所有國家或政府都能夠輕易實施 (Kupferschmidt and Cohen 2020)。另一方面，在北京或是上海等疫情相對較不嚴重的地區，中國政府實施了以住宅區為單位的小規模封城措施，稱為「小區管制」(Shao 2020)。由於區域過小，區域內的無症狀感染者可能在被管制之前，就已經將病毒攜帶到區域之外，傳播給其他民眾，使得防疫單位忽略了潛在的疫情風險。因此，防疫管制分區的地理範圍既不能太大也不能太小，必須要在一個合理的規模達到最佳的防疫效果，符合過往研究的觀點 (Barbera et al. 2001)。然而，文獻上依舊缺乏完善的方法來劃分適當大小的防疫管制分區，導致實務上的防疫仍舊是以既有的大尺度行政區為管制單位。基於人或動物的移動行為是造成傳染病擴散的一大主因 (Lima et al. 2015)，本論文認為移動行為特徵或許能協助劃定更有效的防疫管制分區。

最後，一個地區獲取醫療資源的容易程度通常被定義為資源可及性

doi:10.6342/NTU202101357

(Guagliardo 2004)。民眾位於可及性程度較高的地區，能夠在疫情來襲時，及時獲得足夠的醫療或防疫資源，故疫情較不容易在此處蔓延擴散。相反地，可及性程度較低的地區因為不易取得防疫醫療資源，導致疾病較容易入侵該區域且快速蔓延。在此次的新冠肺炎疫情中，醫療資源可及性較低的地區也確實都伴隨著較多的染病人數和死亡人數 (Ji et al. 2020)。因此，準確評估防疫醫療資源的空間可及性將能協助後續的物資佈署規劃，優先在較脆弱的地區投注額外籌措來的資源，讓資源發揮最大的效益。在量性評估的層面上，醫療資源的空間可及性主要被三項因素決定 (Ma et al. 2018a)：需求的空間分布、供給的空間分布、以及需求地區和供給設施之間的移動阻抗（旅行成本）。實務上，各個地區的可及性無法獨立評估，因為一間供給設施的資源可能需要同時服務好幾個地區，造成鄰近地區之間存在資源競爭的狀況；同樣地，民眾可能會有跨區就醫的行為，造成供給設施之間也存在互相競爭的情形 (Luo and Qi 2009)。因此，文獻上常以二階段浮動服務範圍模型（Two-step Floating Catchment Area, 2SFCA）這項考量了供給和需求之間競合關係的模型或其改良版本來評估研究區內每個地點的可及性 (McGrail 2012; Wang 2012)。過往研究大多假設需求分布是固定不變的；然而，傳染病疫情所引起的需求是會隨著時間變動的。好比 2021 年初發生在台灣的部立桃園醫院新冠肺炎群聚感染事件，就在短時間內增加了大量的採檢需求，直到狀況排除。因此，忽略疫情引起的需求變動將使得模式無法準確評估每個地區對於資源的可及性，進而影響到後續防疫資源的分配規劃。

1.2 研究目的

由前述可知，目前文獻上關於傳染病擴散的三個重要面向皆存在方法學上的不足，可能導致實務上的防疫策略和防疫工作無法有效回應傳染病擴散所帶來的威脅，進而導致人類社會失去對於傳染病的調適能力。因此，本論文旨在以地理

doi:10.6342/NTU202101357

學或空間資訊分析的觀點針對上述三個面向進行方法學上的改良。具體的研究目的包含以下三點：

1. 系統性定義疫情群聚擴散的類型，並發展相對應的分析方法來自動判斷群聚於不同時期下的擴散演化類型。
2. 以能夠劃分合適地理範圍大小的防疫管制分區為目標，考量人口移動行為的相關特徵，發展有效的區域劃分方法。
3. 發展一套整合模型將醫療資源需求受疫情影響而產生的動態變化納入 2SFCA 模型當中，藉此來細緻剖析隨時序變化的空間可及性。

1.3 大綱

延續第一章的緒論，本論文的第二章將呈現文獻的回顧與整理，藉此更仔細地顯示本論文的研究缺口是基於何種學術脈絡背景下所推演出來。本論文規劃將三個研究目的以三個獨立研究的形式完成，故第三章的書寫架構係分別介紹每個獨立研究的具體目標、研究方法以及成果摘要，讓讀者清楚理解各個獨立研究的內容。最後，第四章將討論三個獨立研究的成果如何回應本論文的研究目的，以及對於學術理論上和實務應用上的可能貢獻。全文以第五章結論做收尾。

doi:10.6342/NTU202101357

第二章 文獻回顧

2.1 傳染病擴散之空間型態

傳染病擴散的空間型態可分為階層型（Hierarchy）、接觸型（Contagion）和位移型（Relocation）三種 (Meade and Emch 2010)；而後兩者又可合併稱為擴張型（Expansion）(Cliff et al. 1981)，每一種型態各自隱含著不同的流行病學意義和空間傳播機制。首先，階層型擴散關注疾病如何在網絡結構中的不同等級之城市或節點傳播，例如 Grenfell, Bjørnstad, and Kappey (2001) 發現英格蘭和威爾士地區的麻疹是從大城市擴散到環繞其周圍的小鎮，顯現出疾病在階層網絡中的傳播情形。位移型擴散強調傳染病會因為宿主移動或是交通流動的關係，被帶至新的地點然後在該地蔓延開來，例如 Kuo and Fukui (2007) 發現霍亂在 1882 到 1895 年間於日本福島縣西南邊的病例紀錄是到處移動的，他們推測這是居民在不同地區間往來互動導致的現象，所以將其認定為位移型擴散。最後，接觸型擴散強調疫情擴散受限於地理空間的距離限制，近距離的接觸或互動是疫情傳播的要件 (Haggett, Cliff, and Frey 1977)。例如 Lai et al. (2004) 透過追蹤香港地區每個 SARS 病例之間的接觸關係，將所有病例依照接觸關係區分成 3 組並各自進行群聚分析，結果發現病例之間的距離若是愈靠近，則疫情的群聚就愈強，顯示近距離的接觸確實會加劇疫情的傳播。

Cliff et al. (1981) 試圖以疫情中心點和疫情範圍大小這兩項特徵的變化來量性描述傳染病的擴散過程。作者認為擴張型的擴散過程是有個固定的疫情中心點，並且逐漸地向外擴大疫情範圍（圖 1A）；而位移型則是疫情中心點朝向某個方向移動，但疫情範圍大小是沒有劇烈變化的（圖 1B）。除此之外，作者也認為這兩種過程可能會同時發生，產生混合型的擴散型態（圖 1C）。後續也有文獻呼應了此兩項特徵（疫情中心點與範圍大小）的重要性。Barreto et al. (2008)認為雖然疫情最後擴散至整個城市，但疫情中心點並沒有太大的變化，所以建議政府的防治和醫護

doi:10.6342/NTU202101357

資源應該優先投入疫情中心點，因為該處一直都是病例最集中的地方；而 Fernandez et al. (2012)透過分析 2008 年到 2009 年非洲哈拉雷地區的霍亂疫情，發現疫情中心點的位置之土地利用狀態如果不同，會影響到疫情的規模大小。由此可知，疫情中心點可能是防疫的關鍵地點；同時，若疫情中心點處於不同位置，則可能會導致疫情的嚴重程度出現變化。因此，掌握疫情中心點在不同時間點的位置是有助於理解疫情的擴散過程的。另一方面，Kan et al. (2008) 表示，當疫情群聚的範圍如果愈來愈大時，可能反映出防疫措施有漏洞或是沒有徹底消滅感染源，導致疫情愈來愈嚴重；不過 Watve and Jog (1997) 則指出當群聚規模愈大時，反而會降低疾病傳播至該群聚外的機率。這顯示了群聚規模的改變，確實會影響到疫情的發展。

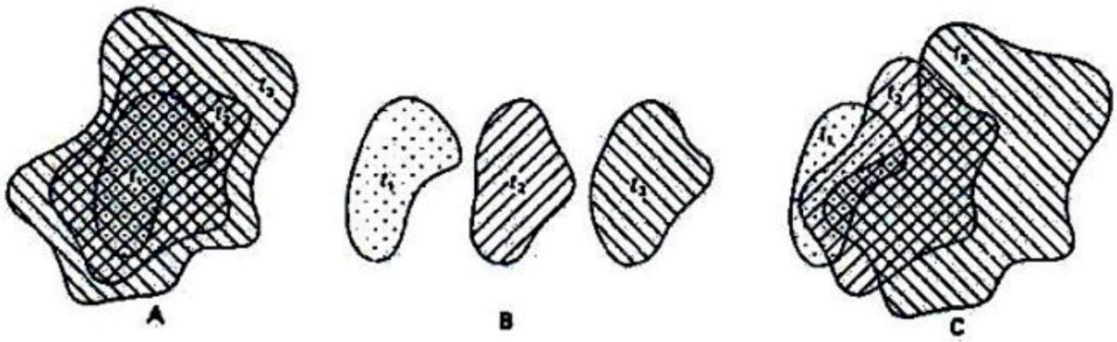


圖 1. 疫情擴散之三種空間型態示意
來源：Cliff et al. (1981)

上述的空間型態描述皆聚焦在「單一」疫情群聚的演化；然而，文獻上也發現不同群聚之間可能存在交互作用的影響關係。Lian et al. (2007) 指出不同地區的疫情群聚可能會發生融合的現象，加劇了後續疫情的嚴重程度；Kalnis, Mamoulis, and Bakiras (2005) 認為一個群聚可能會在後續時間點分裂成若干個群聚，在不同地區各自發展演化。除此之外，疾病常常因為人或動物的移動而被帶往一個新的地點，出現長距離的位移型擴散 (Kan et al. 2008)，所以新地點的疫情群聚本質上是舊有的群聚所造成。換言之，新舊地區的疫情群聚之間也存在著交互作用關係。

doi:10.6342/NTU202101357

2.2 防疫管制分區劃分

傳染病擴散往往是因為人或動物的移動行為而加劇，因為移動會增加感染者和健康者之間的接觸頻率 (Fèvre et al. 2006; Mao 2014)；透過移動管制來降低兩者之間的接觸頻率，可以協助控制疫情的擴散 (Bajardi et al. 2011; Velthuis and Mourits 2007)。在實務上，當一個新興的未知傳染病出現時，移動管制也是爭取時間來開發疫苗或解藥的手段 (Gensini, Yacoub, and Conti 2004)。移動管制的效果也在過去幾次的全球性傳染病疫情中，包括：SARS (Day et al. 2006; Fraser et al. 2004)、A 型流感 (Cooper et al. 2006; Epstein et al. 2007)、伊波拉病毒 (Peak et al. 2018) 等，皆被證實是有效的。移動管制經常被區分常兩種不同的形式，分別為感染者隔離（Isolation）和接觸者隔離（Quarantine）(Barbera et al. 2001; Day et al. 2006; Fraser et al. 2004)。兩種模式都會管制感染者的移動，但接觸者隔離的模式更強調將與感染者接觸過的人一併納入管制的範圍，進一步降低傳染病擴散的風險。然而，當確診人數快速增加時，要完整找出所有與確診者接觸過的民眾是非常困難的 (Fong et al. 2020; Peak et al. 2020)。在這種情況下，接觸者隔離的模式傾向鎖定一個地理區域來涵蓋所有可能的接觸；如此大規模的管制，又被稱為防疫管制分區（Containment zone），是最嚴厲的防止傳染病擴散的措施 (Wilder-Smith, Chiew, and Lee 2020)。

在實務的防疫工作上，防疫管制分區多半是以既有的行政區層級進行劃分。以台北市為例，2009 年因為擔憂 H5N1 禽流感病毒會演化出人傳人的傳播途徑，市政府即以村里為最小的空間單位，制定了台北市的防疫分區劃分（表 1）。以行政區層級為劃分的優勢在於能夠迅速指認出權責的地方單位，加快政府不同單位之間的溝通和防疫準備工作。然而，這樣的劃分方式忽略了人或是家禽的移動樣態，容易忽略疾病已經擴散至管制區域外部的風險。基於此原因，Lee et al. (2019) 提出一個名為移動基準分區（Movement-based zoning）的劃定管制區之方法（圖 2），

doi:10.6342/NTU202101357

表 1. 台北市 2009 年防疫分區

區名	次分區	次分區代碼	涵蓋村里
松山	三民	1	莊敬、東榮、三民、新益、富錦、新東、富泰、介壽
	東社	2	精忠、東光、龍田、東昌、東勢、中華、民有、民福、松基
	本鎮	3	慈祐、安平、鵬程、自強、吉祥、新聚、復盛
	中崙	4	中正、中崙、美仁、吉仁、敦化、復源、復建、復勢、福成
信義	三張犁	1	西村、正和、興隆、中興、新仁、興雅、敦厚、廣居、安康
	五分埔	2	六藝、雅祥、五常、五全、永吉、長春、四育、四維、永春、富台
	福德	3	國業、松隆、松友、松光、中坡、中行、大道、大仁
	吳興	4	景新、惠安、三張、三犁、六合、泰和
	六張犁	5	景新、惠安、三張、三犁、六合、泰和
大安	和平	1	錦華、古莊、龍坡、福住、龍安、龍泉、永康、光明、錦安、錦泰
	學府	2	古風、龍門、大學、龍淵、學府
	瑞安	3	仁慈、德安、住安、龍雲、龍圖、龍陣、新龍、龍生
	新生	4	誠安、昌隆、義村、和安、民輝、民炤
	敦南	5	華聲、正聲、車層、建安、建倫、光武、仁愛
	安和	6	法治、臨江、敦煌、通安、全安、光信、敦安、通化、義安
	臥龍	7	群英、臥龍、虎嘯、芳和、黎孝、黎元、黎和、群賢
中山	林森	1	正守、正義、正得、民安、康樂、中山
	圓山	2	聚盛、聚葉、集英、恆安、圓山、晴光
	新庄	3	新庄、新生、新福、新喜、中庄
	大直	4	劍潭、大直、成功、永安、北安、金泰
	下埤頭	5	大佳、松江、行政、行仁、行孝、下埤、江寧、江山
	長春	6	中吉、中原、中央、龍洲、朱馥、復華
	朱厝崙	7	興亞、朱園、埤頭、朱崙、力行
中正	古亭	1	河堤、螢圃、網溪、板溪、頂東、螢雪
	公館	2	水源、富水、文盛、林興
	南門	3	南門、新營、龍福、南福、愛國
	崁頂	4	廈安、忠勤、永功、永昌、龍興、龍光
	城內	5	黎明、光復、建國
	東門	6	東門、幸福、梅花、幸市、文北、文祥、三愛

doi:10.6342/NTU202101357

表 1. 台北市 2009 年防疫分區（續）

區名	次分區	次分區代碼	涵蓋村里
大同	大龍	1	揚雅、斯文、至聖、重慶、保安、蓬萊
	蘭州	2	隆和、景星、國順、國慶、鄰江、老師
	建城	3	雙連、星明、光能、建功、建泰、建明、民權
	延平	4	玉泉、永樂、朝陽、大有、南芳、延平
萬華	西門	1	西門、菜園、福星、萬壽、新起
	龍山	2	福音、柳鄉、仁德、富民、青山、富福
	大理	3	綠堤、糖部、雙園、頂碩、和平、華江
	西園	4	和德、日善、忠德、全德、孝德、錦德
	東園	5	銘德、壽德、保德、興德、榮德、華中
	青年	6	新安、新忠、新和、日祥、忠貞、騰雲、凌霄
文山	景美	1	景美、景行、景東、景慶、景仁、景華、萬有、萬祥、萬隆、萬年、萬和、萬盛
	興隆	2	興豐、興光、興家、興得、興業、興安、興福、興旺、興泰、興昌
	木柵	3	木柵、木新、明義、明興、順興、忠順、樟新、樟腳、試院、華興、樟林
	二格山	4	萬興、指南、老泉
	萬芳	5	萬芳、萬美、博嘉
南港	三重埔	1	南港、三重、中南、新富
	新庄仔	2	東新、西新、東明、新光
	後山坡	3	玉成、聯成、合成、成福、仁福、百福、鴻福、萬福
	中研	4	中研、九如、舊莊
內湖	東湖	1	五分、東湖、安湖、葫洲、內溝、金湖、康寧、明湖、蘆洲、安泰、樂康
	西湖	2	港漧、港富、港都、港華、西康、西湖、西安、麗山
	紫陽	3	紫星、紫雲、清白、紫陽、瑞陽、瑞光
	金龍	4	湖濱、內湖、金龍、金瑞、碧山、大湖、秀湖
	灣仔	5	湖興、湖元、石潭
	洲尾	6	週美、行善

doi:10.6342/NTU202101357

表 1. 台北市 2009 年防疫分區（續）

區名	次分區	次分區代碼	涵蓋村里
士林	街上	1	仁勇、義信、福林、福德、福志、舊佳、福佳
	後港	2	後港、福中、前港、百齡、承德、福華、明勝
	社子	3	福順、富光、葫蘆、葫東、社子、社新、社園、永倫、福安、富洲
	芝山岩	4	岩山、名山、聖山、芝山、東山
	蘭雅	5	德行、德華、忠誠、蘭雅、蘭興
	天母	6	三玉、天母、天福、天祿、天壽、天和、天山、天玉
	陽明山	7	永福、公館、新安、陽明、菁山、平等、溪山、翠山、臨溪
北投	石牌	1	建民、文林、石牌、福興、榮光、榮華、裕民、振華、永欣、永和、洲美
	唭哩岸	2	永明、東華、吉利、吉慶、尊賢、立賢、立農
	關渡	3	八仙、豐年、稻香、桃源、一德、關渡
	舊北投	4	奇岩、清江、中央、大同
	新北投	5	長安、溫泉、林泉、中心
	陽明山	6	泉源、湖山、湖田
	大屯	7	中庸、開明、中和、大屯、智仁、秀山、文化

用以防止口蹄疫的疫情擴散。首先，利用豬舍的位置和豬隻的載運紀錄建構出一個豬隻的移動網絡；然後利用既有的網絡分群方法將建構出的網絡區分成數個區塊，區塊內部會有強烈且緊密的豬隻流動、區塊與區塊之間的流動強度則相對低弱；最後，再將區塊映射到地理空間上，形塑出疫情管制分區。作者強調由於分區內的設施具有強烈的流動關係，所以當疫情出現於某個設施時，區塊內的其他設施也很有可能已經受到感染，只是因為潛伏期的關係所以豬隻尚未出現明顯病症而已；而跨區之間因為豬隻流動量少，所以疫情較不可能跨區傳播。因此，針對出現疫情的分區域進行移動管制，能夠有效地將疫情控制於該區域內，避免其他地區受到影響。

doi:10.6342/NTU202101357

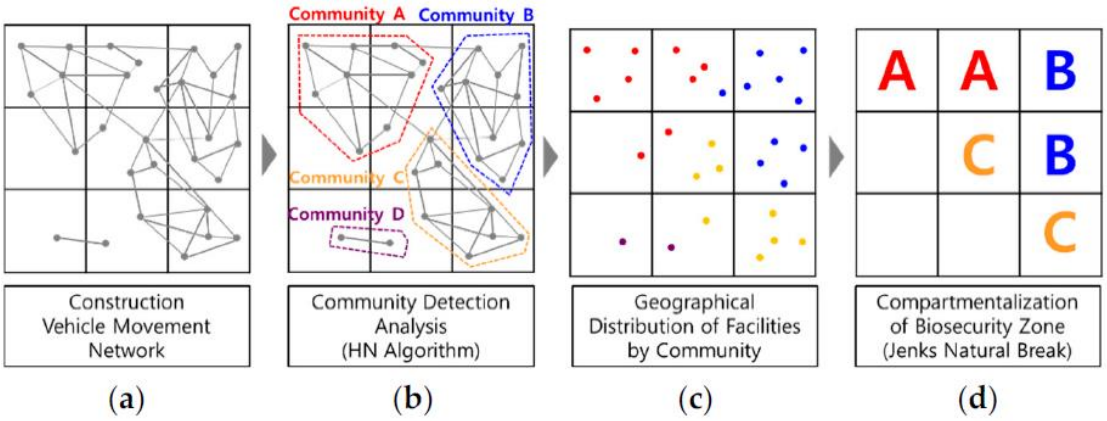


圖 2. 移動基準分概念圖

來源：Lee et al. (2019)

2.3 空間可及性評估模型

關於空間可及性的量測與分析，文獻上最簡單的方法是以需求地點和鄰近供給地點之間的移動阻抗之反比來反映該地點的可及性高低程度。但是這種方法忽略了每個供給點所能提供的資源數量，以及每個地區對於資源的需求量 (McGrail and Humphreys 2014)。引力模型能夠同時考量到移動阻抗和供給量與需求量等影響因素，但此方法必須事先決定距離遞減的效果強度 (Joseph and Bantock 1982)。過往文獻認為該效果強度是因地制宜的，且難以客觀決定，再加上分析結果難以詮釋，故此方法並未被廣泛使用 (Luo and Whippo 2012)。Khan (1992) 以每個行政區內的供給需求比，作為反映該行政區的資源可及性程度高低之量化指標。雖然此方法簡單且結果容易詮釋，但其最大的缺點在於忽略了民眾跨區尋求資源的移動行為 (Luo and Qi 2009)。

有鑑於上述三項方法的缺點，Luo and Wang (2003) 提出了改良版本的二階段浮動服務範圍模型（Two-step Floating Catchment Area, 2SFCA），此模型首先計算每個供給設施的供給需求比：

$$R_j = \frac{S_j}{\sum_{k\in\{d_{kj}\le d_0\}} P_k} \quad (1)$$

doi:10.6342/NTU202101357

其中 S_j 代表 j 供給設施的資源數量、P_k 代表 k 地區的資源需求量（大多以人口數來代表）、d_{kj} 代表 k 地區和 j 供給設施之間的移動阻抗或旅行成本（大多以距離遠近或是旅行時間長短來代表）、d_0 代表可容忍的旅行成本上限，反映出服務範圍大小的概念。因此，R_j 即代表 j 供給設施的供給需求比，反映了平均一個人能夠在此處獲得多少單位的資源。接著，再針對每個地區計算其供給需求比的總量：

$$A_i = \sum_{j\in\{d_{ij}\le d_0\}} R_j = \sum_{j\in\{d_{ij}\le d_0\}} \frac{S_j}{\sum_{k\in\{d_{kj}\le d_0\}} P_k} \qquad (2)$$

A_i 即代表地區 i 的空間可及性。多數學者認為 2SFCA 打破了區域邊界的限制，讓供給需求比的計算能夠考量到跨區移動的行為；同時，此方法也被視為是引力模型的特殊版本，以更簡單的公式反映出移動阻抗的效果，讓計算結果可以更直觀地被詮釋和理解 (McGrail 2012)，Wang (2020) 更以數學公式推演的方式詮釋了 2SFCA 模型本質上就是霍夫模型（Huff model）的延伸應用。因此，文獻上大多以 2SFCA 為評估空間可及性的主要方法，或是以 2SFCA 為原型進行模型的發展改良。

雖然 2SFCA 奠定了空間可及性的基本分析架構，但此方法被認為存在兩項缺點 (McGrail 2012; Wang 2012)。首先，此方法僅以二元區分（$d_{kj} \le d_0$）的方式來反映移動阻抗的效果，而這樣的形式只適用於小範圍地區的計算。當服務範圍擴大時，範圍內的需求地點與供給設施之間的移動阻抗便會出現明顯的距離遞減效果，忽略此效果將會導致計算出的數值無法反映實際的可及性程度高低 (McGrail 2012)。第二，每個供給設施之間或是每個需求地點之間都採用固定的旅行成本容忍上限（d_0）。然而，居住於郊區或是偏遠地區的民眾，因為已有心理準備居住於資源較缺乏的地區，故可能會願意付出更大的旅行成本來獲取資源 (Buzza et al. 2011; Tanser, Gijsbertsen, and Herbst 2006)。因此，學者認為當研究區同時涵蓋市區與郊區時，必須在每個需求地點或是供給設施採用符合當地狀況的旅行成本上限 (McGrail and Humphreys 2014)。

doi:10.6342/NTU202101357

基於這兩項缺點，近十年來陸續有文獻提出改良的模型來計算更合理的空間可及性。Luo and Qi (2009) 提出 Enhanced 2SFCA （E2SFCA）模型並採用「階段型」的距離遞減函式；此模型將服務範圍依照移動成本（距離或時間）由小至大區分成類似同心圓的若干個區塊，並分別給予由大至小的權重來反映距離遞減的效果。然而，這種階段型的距離遞減函式仍然被認為過於簡化。因為其無法反映出單一個區塊內的距離遞減效果；同時，兩個區塊的邊界上會有不符合現實狀況的權重落差 (Delamater 2013)。因此，Dai and Wang (2011)在服務範圍內採用核密度估計（Kernel Density Estimation）的函式來捕捉「連續型」的距離遞減效果，並將模型命名為 KD2SFCA。圖 3 展示了不同距離遞減函式的概念。

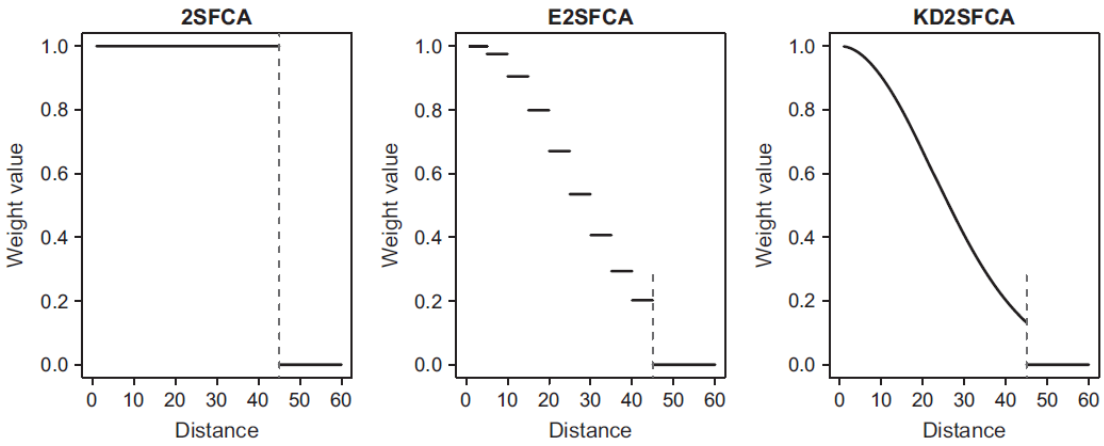


圖 3. 不同距離遞減函式之概念示意
備註：虛線為移動阻抗（旅行成本）的容忍上限
來源：Delamater (2013)

另一方面，關於固定旅行成本上限的改良，Luo and Whippo (2012)為 2SFCA 模型加上了兩項條件來分別浮動調整每個供給點和每個需求點的範圍大小，並將模型命名為 Variable 2SFCA（V2SFCA）。第一個條件是要求每個供給點的服務範圍都要涵蓋一定數量的需求人口數（公式 1 中的分母），所以在郊區人口稀疏的地方，供給點的服務範圍將能夠自動擴大，直到包含足夠多的需求量。第二個條件則

doi:10.6342/NTU202101357

是反過來要求每個需求點的供給需求比總量（公式 2）必須達到基本門檻，所以資源稀少地區的資源尋求範圍也會自動擴大。Dony, Delmelle, and Delmelle (2015) 認為一個設施所能提供的資源數量愈多，代表該設施愈有能力服務較大的範圍，便依照此條件來浮動調整每個供給設施的服務範圍。同時，整合了四種不同運輸模式下的旅行成本，形成兩地之間的綜合移動阻抗。作者將新模型命名為 Variable-width FCA（VFCA）。

除了距離遞減函式與服務範圍大小的改良之外，也有文獻提出其他 2SFCA 模型的問題並加以改良。首先，Delamater (2013) 認為 2SFCA 的分析結果是過度理想化，因為在一個複雜的都市系統內，可能存在一些絕對弱勢的族群或地區無法順利獲取醫療資源。因此，作者提出了 Modified 2SFCA（M2SFCA）模型，在 2SFCA 的模型架構下，多考量了一次移動阻抗的效果，藉此來突顯出可及性非常弱勢的地區。Jamtsho, Corner, and Dewan (2015) 進一步將 M2SFCA 的服務範圍大小設定改成浮動形式，要求每個供給設施的服務範圍只能包含兩個最鄰近的需求地區，反之亦然。新模型被命名為 Nearest-neighbor M2SFCA（NN-M2SFCA）。除此之外，Jin et al. (2019) 指出多數文獻忽略了不同層級的醫療資源供給設施之間的差異：通常診所或小醫院只能服務鄰近地區的居民，但是大醫院的服務範圍則相對較大，甚至能涵蓋整個城市。因此，作者提出了 Hierarchical 2SFCA（H2SFCA）模型，賦予不同層級的供給設施有不同的服務範圍大小和距離遞減程度。除了分別計算各地區對於不同層級醫療資源的可及性之外，此模型也使用加權總合的方式來計算每個地區的綜合資源可及性。

另一方面，Wan, Zou, and Sternberg (2012) 認為 2SFCA 模型存在需求膨脹的問題，因為若是一個地區鄰近若干個資源供給點，則該地區的資源需求量可以分散給不同的供給點來服務，而不必集中於某個特定的供給站。換言之，作者認為公式 1 在計算一個供給地點的供給需求比時，會高估了需求量（公式 1 中的分母）。為

doi:10.6342/NTU202101357

了解決此問題，作者提出了 Three-step FCA（3SFCA）模型。此模型在使用 2SFCA 的模型架構進行計算之前多加了一個步驟：依序將每個地點的需求量，按照其與鄰近供給設施之間的旅行成本反比作為權重，分配給其鄰近的供給設施。這樣的作法就可以確保每個單位的需求量不會被重複計算，避免需求膨脹的問題。Paez, Higgins, and Vivona (2019) 進一步指出供給層面也存在類似的膨脹問題，並認為能夠借助 3SFCA 的概念反過來將每個供給設施的供給需求比分配給鄰近的需求地點。因此，作者提出了一個分析架構來同時處理需求膨脹和供給膨脹的問題，並且允許使用者自行選擇如何定義服務範圍大小（固定或浮動）和距離遞減函式（階段型或連續型），再利用額外的參數來達到 M2SFCA 的效果。這樣普遍性的分析架構被 Pereira et al. (2020) 稱為 Balanced FCA（BFCA）模型。然而，Wang (2020) 認為並不需要特別處理需求膨脹和供給膨脹的問題，因為空間可及性的意義在於每個地點對於資源索取的容易程度，而非資源分配的均勻程度。

上述研究或模型改良發展研究大多關注於基礎醫療照護（Primary Health Care）的空間可及性，例如：診所醫師數量 (Luo 2014)、醫院病床數量 (Delamater 2013)。部分文獻則關注於餐廳 (Dai and Wang 2011)或是公園綠地 (Dony, Delmelle, and Delmelle 2015)等其他生活資源的空間可及性。由於這些資源屬於民眾日常生活所需的類型，所以一個地區對於這種類型的資源需求不太會在短期內有大幅度的變動。再加上供給的設施位置和資源數量也不會大幅度變動的前提下，過往文獻所評估出的空間可及性結果幾乎都是靜態不隨時間變動的。近年來，開始有學者提出空間可及性應該是會隨著時間演化的，因為三項用於評估空間可及性的重要因素（需求分布、供給分布、需求和供給之間的移動阻抗）都會隨著都市擴張或是都市化等建成環境的改變而出現變化 (Yang and Mao 2018)。因此，Xing et al. (2018) 使用 2SFCA 剖析了武漢地區公園綠地之空間可及性從 2000 年到 2014 年的變化。Hu et al. (2020) 使用 E2SFCA 剖析杭州地區食物市場之空間可及性 2016 年到 2018 年的

doi:10.6342/NTU202101357

變化。

表 2. FCA 系列模型整理

模型名稱	服務範圍大小	距離遞減函式	其他特點	文獻
2SFCA	固定	二元	無	W. Luo and Wang (2003)
E2SFCA	固定	階段型	無	W. Luo and Qi (2009)
KD2SFCA	固定	連續型	無	Dai and Wang (2011)
V2SFCA	浮動	階段型	無	W. Luo and Whippo (2012)
VFCA	浮動	連續型	整合不同運具模式	Dony, et al. (2015)
3SFCA	固定	階段型	處理需求膨脹問題	Wan, et al. (2012)
BFCA	固定或浮動	階段型或連續型	同時處理需求膨脹和供給膨脹問題	Paez, et al. (2019)
M2SFCA	固定	連續型	重複考量移動阻抗的效果，突顯弱勢地區	Delamater (2013)
NN-M2SFCA	浮動	連續型	固定鄰近供給點或需求點的數量	Jamtsho, et al. (2015)
CB2SFCA	固定	階段型	考量通勤移動行為	Fransen, et al. (2015)
H2SFCA	固定	連續型	考量醫療設施的等級階層	Jin, et al. (2019)
ST-E2SFCA	固定	階段型	以 GPS 資料考量實際的人口分布	Xia, et al. (2019)

除了長期的轉變之外，也有學者認為空間可及性會受到人口移動或是通勤行為的影響而在短期內出現動態變化。Fransen et al. (2015) 認為每個地區對於資源的需求應該在白天和晚上有明顯差異，因為人們白天時經常會移動到其他地區工作上學，晚上才回到居住地；而關於 2FSCA 模型的研究大多使用居住人口數代表需求量，就只能反映出需求在夜晚的空間分布。因此，作者提出 Commuter-based 2SFCA（CB2SFCA）模型，將通勤移動行為整合至 2SFCA 模型當中，藉此突顯出

doi:10.6342/NTU202101357

白天的需求分佈對於空間可及性的影響。然而，其分析結果與 2SFCA 相比，只是綜合考量了白天和夜晚的資源需求，依舊是一個靜態的空間分布。Ma et al. (2018b) 認為需求地點到供給設施之間的旅行成本，會在一天內隨著城市路網擁擠程度的改變而出現動態變化，進而造成空間可及性在時序上的變化。因此，作者以 3SFCA 模型為基底，整合了即時的路網擁擠程度，剖析出武漢地區一天內的醫療資源空間可及性之變化。Yun et al. (2020) 使用手機定位資料得知首爾民眾在一天之內的空間分布變化，搭配 2SFCA 模型分析出不同時間點下的空間可及性。Xia et al. (2019) 開發了 Spatio-temporal E2SFCA（ST-E2SFCA）模型，搭配 GPS 訂位資料，藉此準確評估每間供給設施在每個時間點下可能需要服務的人數。表 2 彙整了上述各個模型的概念和特點。

doi:10.6342/NTU202101357

第三章 實證研究摘要

3.1 群聚擴散演化時空樣態剖析

3.1.1 研究背景

從第二章的文獻回顧可知，群聚中心點和範圍大小的改變可以用來描述單一群聚擴散演化的樣態，而不同群聚之間的融合、分裂或是位移等現象，則代表了交互作用的發生。而雖然過去的文獻已經發現這些現象，仍舊無法實際勾勒出每種型態。原因在於文獻上多半是以核密度估計法（Kernel Density Estimation, KDE）搭配時序性的地圖來判斷疫情熱區的位置和連續變化 (Bailey and Gatrell 1995; Gatrell et al. 1996; Sabel et al. 2000)。不過 KDE 實際上只能產生風險的空間分布圖，搭配時序性地圖也只能揭露風險分布在不同時間點之間的差異。因此，這樣的方法並沒有辦法清楚界定每個群聚的邊界所在，也難以量測中心點的位置和實際的範圍大小；也無法判斷群聚之間是否出現融合或分裂等交互作用。如此一來，便無法細緻地追蹤每個群聚的發展演化過程，例如擴散方向或是群聚範圍等特徵的變化。再者，多數傳染病皆具備潛伏期的特性，表示一個人被病毒感染後，必須經過一段時間才會發病，進而被醫療系統確診並通報 (Chan and Johansson 2012)。換言之，新舊病例之間就算空間上非常鄰近，也必須考量兩者之間的時間差距是否超過潛伏期的長度，才有可能是傳播關係，也才有可能是擴散現象。

綜合上述，本論文於此部分的具體研究目的有二：其一，系統性地定義和分類可能的群聚擴散演化過程。包含以傳染病群聚中心點和範圍大小等兩項特徵的變化來定義單一群聚可能的擴散演化型態，以及不同群聚之間的融合、分裂、位移等交互作用。其二，發展一套演算法來自動偵測判斷所定義的各種群聚演化型態。本論文採用 Density-based Spatial Clustering Application with Noise（簡稱 DBSCAN）此分群方法作為基礎並加以修改，新的演算法命名為 Modified Space-Time DBSCAN（簡稱 MST-DBSCAN）。DBSCAN 擁有可偵測不規則形群聚、不需事先

doi:10.6342/NTU202101357

決定分群數目等優點 (Birant and Kut 2007; Ester et al. 1996)，也已有研究利用 DBSCAN 對於地表的點型態事件進行分群分析 (Birant and Kut 2007; Wang, Wang, and Li 2006)，證明此方法應用於空間資料分析的適用性。

3.1.2 研究方法與資料

傳染病群聚演化型態分類定義

本論文將可能的擴散型態大致區分為兩類：獨自型態（Single pattern）和互動型態（Interaction patter）。前者利用群聚中心點和範圍大小的變化來描述單一群聚自行演化擴散的可能型態；後者則聚焦於不同群聚之間的互動作用所衍伸出來的可能擴散型態。以下將分別介紹這兩種不同類別的各種型態之定義和其所反映出的疫情狀況；關於 MST-DBSCAN 演算法的詳細介紹，請參閱附錄 A。

首先，由於中心點只有是否移動的區別，而範圍則有變大、變小或不變 3 種可能，故若是將這兩項特徵的改變狀況進行交叉排列組合，則會得到 6 種不同的獨自型態，分別說明如下：

- 穩定（Steady，圖 4A）：中心點不動，範圍不變。表示群聚與上個時間點的狀態相比，並沒有明顯的差異，維持原本的疫情嚴重程度。
- 移動（Move，圖 4B）：中心點移動，範圍不變。表示原本地區的疫情緩和下來了，但疫情卻朝向特定方向蔓延到鄰近地區造成新的感染，所以群聚範圍保持不變。
- 成長（Growth，圖 4C）：中心點不動，範圍變大。表示疫情是均勻地向四面八方擴散出去。
- 方向性成長（Directional growth，圖 4D）：中心點移動，範圍變大。表示除了原本地區的疫情沒有減緩，疫情又朝向特定方向蔓延到鄰近地區，造成中心改變的同時，範圍也愈來愈大。

doi:10.6342/NTU202101357

- 衰退（Reduction，圖 4E）：中心點不動，範圍變小。表示疫情是很均勻地內縮而減緩。
- 方向性衰退（Directional reduction，圖 4F）：中心點移動，範圍變小。表示雖然原本地區的疫情逐漸消失了，但疫情還是擴散到鄰近地區。

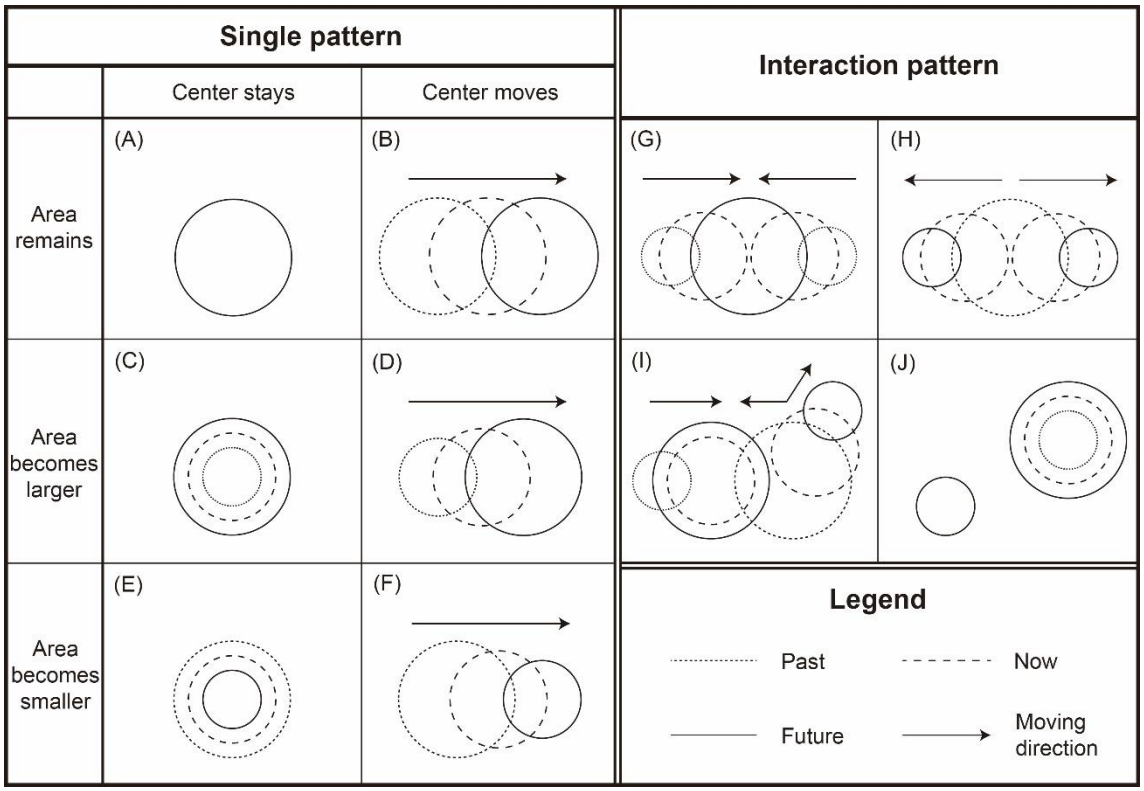


圖 4. 本論文所定義之 10 種群聚演化擴散型態

另一方面，針對本論文所定義之各種互動型態，分別說明如下：

- 融合（Merge，圖 4G）：當兩個群聚逐漸擴散時，可能會因為愈來愈靠近，或是範圍愈來愈大而重疊，形成一個更大的疫區，加劇疫情的嚴重性。
- 分裂（Split，圖 4H）：當疫情從中心開始減緩時，可能會分離成 2 個或數個小地區的疫情群聚，各自再經歷不同的擴散過程。
- 分裂融合（Split-Merge，圖 4I）：此型態表示一個群聚可能只吸收了另一個群聚的一部份，而非完整的融合。

doi:10.6342/NTU202101357

- 浮現（Emerge，圖 4J）：如同在前言的部分所提到，位移型擴散多半是因為宿主的移動而將病毒帶往其他地區，造成疫情在新的地區出現，就像圖 2.(c)一樣，原本的群聚在擴散的過程中，旁邊突然新冒出一個群聚。

MST-DBSCAN 演算法介紹

既有的 DBSCAN 演算法只需要給定 2 項參數，分別為空間搜尋半徑（EpsS）和最小點數（Minimum Points，以下簡稱 MinPts），前者是用以判斷兩個點之間是否能因為距離足夠接近而成為鄰居，後者在概念上則是用以界定一個群聚必須要包含多少點才能成為鄰居。然而，誠如前述，擴散現象是一個時空的過程，且必須考量到傳染病潛伏期特性所造成的時間差（Time lag）；故兩兩樣本點之間除了空間上的距離外，還有時間維度上的距離遠近，以及時間差等特徵必須被納入群聚分析當中。因此，本論文加上了兩項參數：EpsT1 和 EpsT2，前者是用於判斷時間距離是否足夠接近而讓兩兩樣本點成為鄰居關係，後者則是用以排除時間距離過近的鄰居關係。除此之外，為了使 DBSCAN 確實能追蹤每個群聚的發展演化和擴散過程，本論文針對 DBSCAN 的三個部分進行修改，分別為：資料集合、鄰居關係定義，以及鄰近關係定義，並將新演算法命名為 Modified Space-time DBSCAN（以下簡稱 MST-DBSCAN）。

首先，針對資料集合進行修改的部分，本論文在每個時間點皆進行一次分群分析，然後透過時間前後的比較來得知每個群聚的中心點和範圍改變的情形，故關鍵之處在於要能夠追蹤每個群聚的發展過程。為了達成此目的，本論文認為重點在於在每個時間點進行分群時，不應該只有當下時間點出現的病例要被考慮到，也應該納入不久前出現的病例，因為這些病例雖然不是當下出現的，但他們可能仍然處於生病的狀態，依舊有傳播病毒給他人的能力，會對於疫情的發展產生影響，所以必須一併考量。透過這樣的方式，便能夠達到追蹤群聚發展的目的，因為這些舊病例

doi:10.6342/NTU202101357

可能早就已經形成群聚了，所以當新發病例與他們一起進行分群分析時，就能夠利用舊病例來連接上個時間點和現在的群聚。另一方面，由於病例會康復或是死亡，不論哪種狀況，都表示該病例無法再傳播病毒，意即一個病例出現後，並不會一直對疫情造成影響，所以在每個時間點進行分群時，不能把距離現在太早的病例也放進來一起進行分群。綜合這兩點，本論文利用公式 3 來表達此資料集合的概念。

$$D_t = \{p | \, t - EpsT1 \leq T(p) \leq t\} \tag{3}$$

公式 3 中，t 代表第 t 個時間點，D_t代表在 t 時的資料集合，$T(p)$代表病例 p 的出現時間，而 EpsT1 在此則代表了一個病例的影響時間長度，所以$t - EpsT1$即可淘汰掉太早出現的病例，同時保留剛出現不久的病例。類似的資料集合之概念也已在過往的文獻出現 (Ester et al. 1998)，差別在於，Ester 等人所提出的方法只針對本身就有「持續時間」這種特性的資料進行分析，例如手機通話的起始時間點、或是人們進入和離開某個特定網站的時間點。然而，疾病或是其他點型態的地理資料大多只記錄了出現時間而沒有結束時間，所以無法使用 Ester et al. (1998) 的方法。

第二個修改 DBSCAN 的部分是針對鄰居關係定義進行修改，在 DBSCAN 中，一個點擁有的鄰居數目多寡決定了該點的密度，若數目超過最小點數（MinPts）這項參數的設定，就會被視為是核心點（表 3），並成為一個群聚形成的起始處，由此可知，鄰居關係的判斷方式對於分群分析的結果佔有一定的影響。過往的文獻在考量鄰居關係時，都將之視為雙向的關係，也就是兩兩樣本點之間的距離若小於 EpsS，彼此就會是鄰居。然而，當考量到病例點會隨著時間的推進而出現和消失時，原本的鄰居判斷方式可能會出現問題，因為一個病例通常都是被比它早或是同時間出現的病例傳染，不太可能是被較晚出現的病例傳染，所以若是採用原本的鄰居判斷方式進行計算，則很有可能將較早出現病例視為較晚出現病例的鄰居，高估了

doi:10.6342/NTU202101357

晚出現病例的重要性，使其變成核心點，進而導致分群分析的結果出現偏誤。因此，本研究認為在鄰居關係的判斷上，應該採用單向性的而非雙向性（公式 4、公式 5 和圖 4），並將此關係定義為「時空鄰居」。

$$\Delta T(i,j) = T(j) - T(i) \tag{4}$$

$$STNB_{i,t} = \{ j \in D_t \mid dist(i,j) \leq EpsS \cap EpsT2 \leq \Delta T(i,j) \leq EpsT1 \cap j \neq i \} \tag{5}$$

在公式 4 中，$\Delta T(i,j)$表示 i 點和 j 點的時間距離。在公式 5 中，$dist(i,j)$表示 i 點和 j 點的空間距離，$STNB_{i,t}$代表 i 點在 t 時間下所擁有的鄰居形成的集合。由於$\Delta T(i,j)$必須大於等於 0 和小於等於 EpsT1，i 病例點和 j 病例點之間的鄰居關係才會成立，又$\Delta T(i,j)$是用 j 的出現時間減去 i 的出現時間，所以唯有出現時間點比 i 晚或同時的病例，才有可能是 i 的鄰居。透過這樣的修改方式，確保當把 DBSCNA 轉化成時間維度上的分析模式時，不會高估每個點的重要性。

表 3. 樣本點在 DBSCAN 和 MST-DBSCAN 中的角色分類

角色	鄰居數目	是否屬於某分群
核心點 (Core)	≧ MinPts	是
邊緣點 (Border)	＜ MinPts	是
雜訊點 (Noise)	＜ MinPts	否

doi:10.6342/NTU202101357

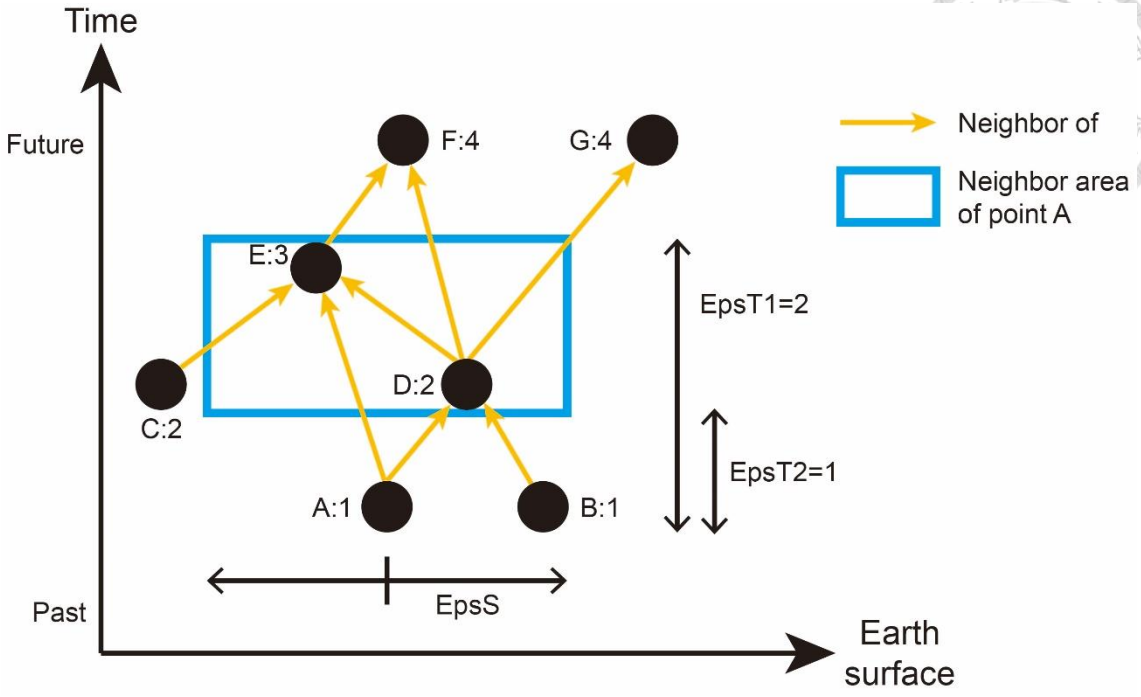


圖 5. MST-DBSCAN 的時空鄰居示意圖

備註：B 點雖然與 A 點在空間上鄰近，但因為兩者的出現時間一樣，不符合時空鄰居的定義，所以 B 點不是 A 點的鄰居。

最後一個修改 DBSCAN 的部分是針對鄰近關係定義進行修改。鄰近關係是用以判斷兩個點之間的接近程度，這些程度也是 DBSCAN 進行分群時的判斷依據，兩個點之間的接近程度如果改變可能就會影響到整個分群的結果。而因為 DBSCAN 的鄰近關係是根基於上述的鄰居關係發展出來的，所以當鄰居關係從原本只考慮空間維度的狀態變成同時考慮時間和空間上的鄰近時，就有必要重新定義鄰近關係。因此，本論文修改了 DBSCAN 原本的 3 種鄰近關係，並且為了能夠捕捉到疫情群聚的現象而新增了第 4 種鄰近關係，此 4 種關係定義為「時空鄰近」，分別說明如下（圖 5）：

1. 直接時空可及（*Directly spatiotemporally reachable*，圖 6A）：給定兩個點 p 和 q，如果 q 是核心點且 p 是 q 的時空鄰居，則 p 是直接時空可及於 q。
2. 時空可及（*Spatiotemporally reachable*，圖 6B）：給定兩個點 p 和 q，如果存在一條鍊狀的連結 $k_1, k_2, \ldots, k_n$，其中 $k_1 = q$ 且 $k_n = p$，並使得 k_{i+1} 是直接

doi:10.6342/NTU202101357

時空可及於 k_i，$\forall 1 \leq i < n$，$k_i \in D_t$，則 p 是時空可及於 q。

3. 時空相連（*Spatiotemporally connected*，圖 6C）：給定兩個點 p 和 q，若存在第 3 點 k 使得 p 和 q 皆時空可及於 k，則 p 和 q 之間的關係為時空相連。

4. 間接時空相連（*Indirectly spatiotemporally connected*，圖 6D）：給定兩個點 p 和 q，若存在第三個點 k 使得 p 和 q 皆與 k 時空相連，則 p 和 q 之間的關係為間接時空相連。此定義為本研究新增設的，目的在於捕捉群聚融合的現象。

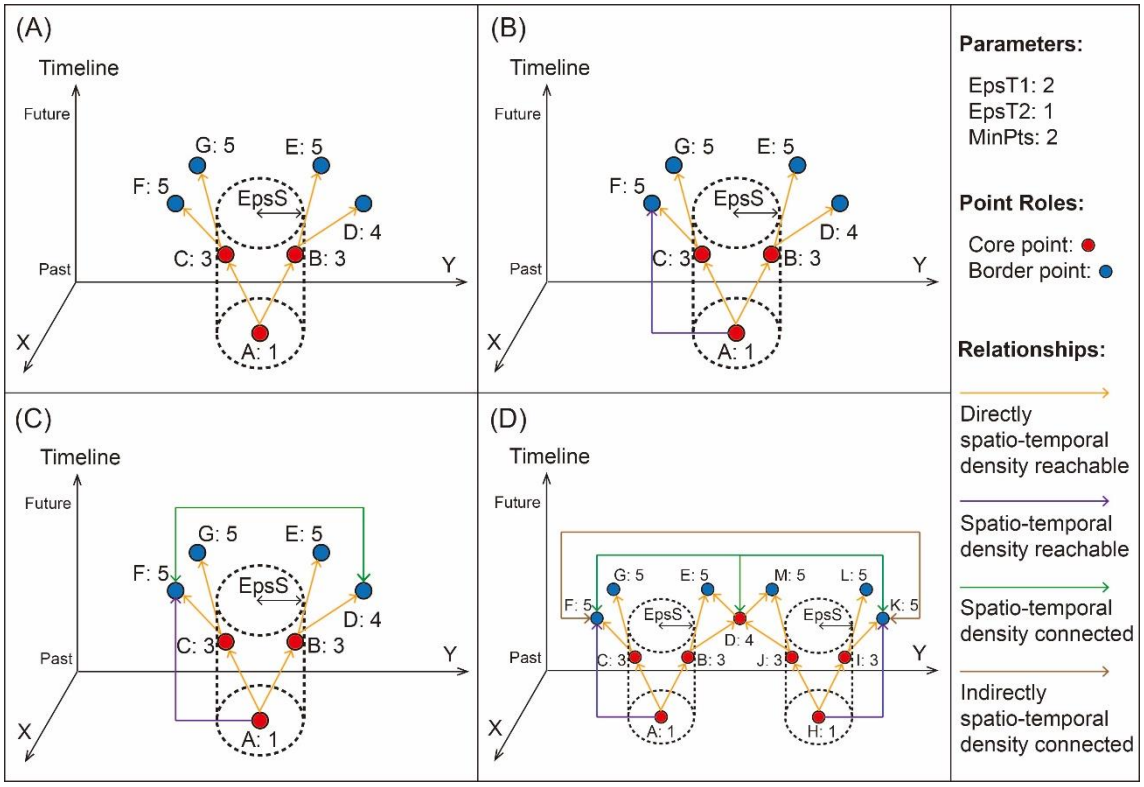


圖 6. MST-DBSCAN 的四種時空鄰近定義

備註：（A）直接時空可及、（B）時空可及、（C）時空相連、（D）間接時空相連

搭配以上 4 種時空鄰近關係的修改，本論文也針對 DBSCAN 本身的群聚定義進行修改，以便能確實達到追蹤群聚發展的目的，新的群聚定義如下：

群聚（C）：一個群聚必須滿足以下兩條件：

1. $\forall p, q \in D_t$，如果 $q \in C$ 且 p 是時空可及於 q，則 $p \in C$。

doi:10.6342/NTU202101357

2. $\forall p,q \in C$，p 必須與 q 間接時空相連。

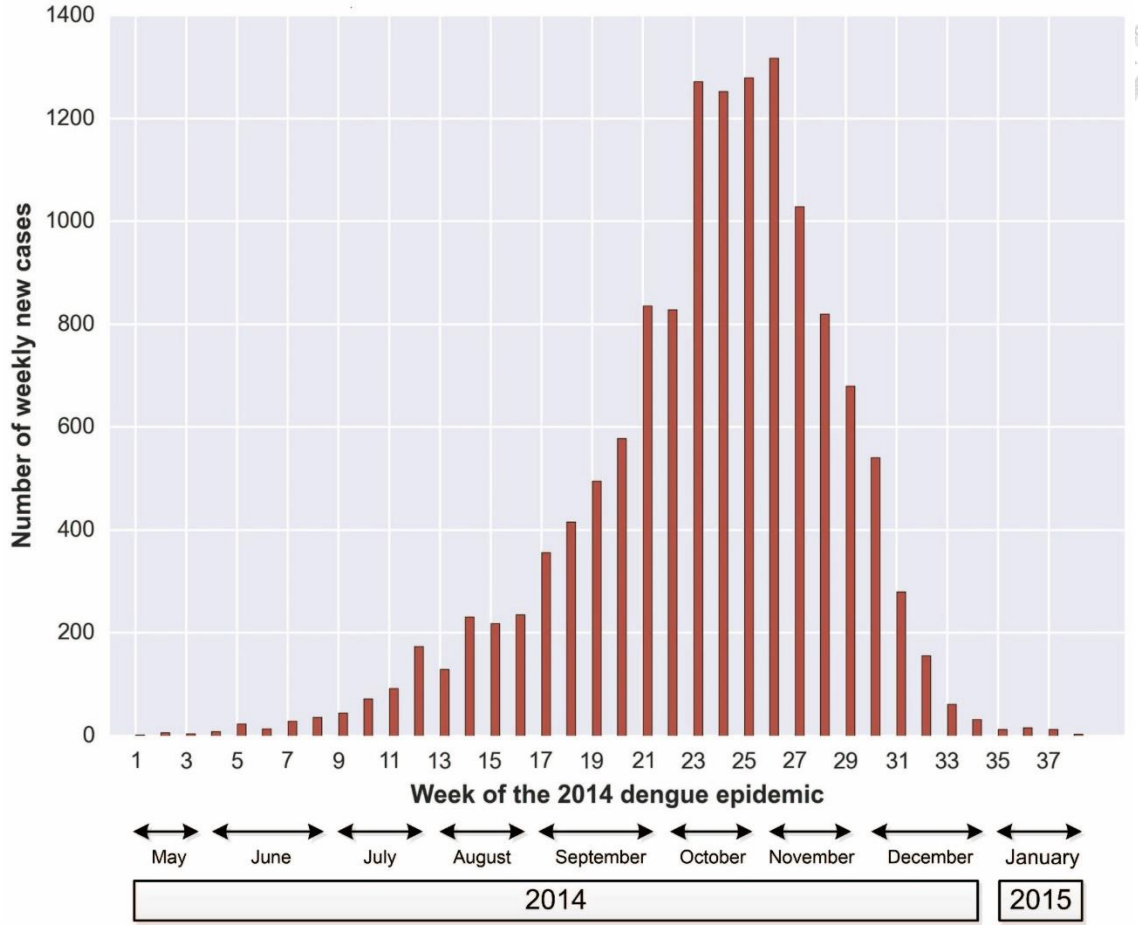


圖 7. 高雄市 2014 年登革熱疫情每周新發病例數

本論文以高雄市 2014 年的登革熱疫情做為案例應用，來檢視 MST-DBSCAN 演算法是否能剖析出疫情的擴散趨勢，並且判斷出每個群聚在每個時間點下的擴散型態。實際的研究區只包含了高雄市的 12 個行政區，因為這些區域才是登革熱疫情主要發生的地區，過往研究也都挑選一樣的研究範圍 (Kan et al. 2008; Wen and Tsai 2015)。登革熱病例的資料來自衛生福利部疾病管制署的開放資料，總計有 13606 個病例紀錄出現在本論文所選取的研究範圍和研究時間段。圖 7 以長條圖的方式呈現了這筆資料的每周新發病例數；與其他年份的疫情類似，疫情大約從六月份開始逐漸升溫，一直到十月份出現高峰，之後便迅速衰退。

doi:10.6342/NTU202101357

3.1.3 研究成果摘要

圖 8 同時展示了 MST-DBSCAN 和 KDE 在第 17 周到第 20 周的分析結果。KDE 的結果顯示在這幾周的時間內，有個熱區一直持續存在於高雄市中心，同時熱區有逐漸往北方擴散的現象。而 MST-DBSCAN 的結果除了能揭露這些訊息之外，更能夠進一步顯示每個群聚的擴散演化型態。例如：持續存在於市中心的熱區係因為有個群聚不斷與周遭的群聚發生融合。導致疫情升溫；此外，從第 19 周開始，許多在市中心靠近北方的群聚發生分裂的現象，同時也有許多新興的群聚出現於此，故這些群聚很有可能就是導致疫情往北方擴散的原因。因此，比起 KDE，MST-DBSCAN 能夠呈現更細緻的傳染病擴散訊息，讓防疫單位清楚瞭解擴散的時空過程。

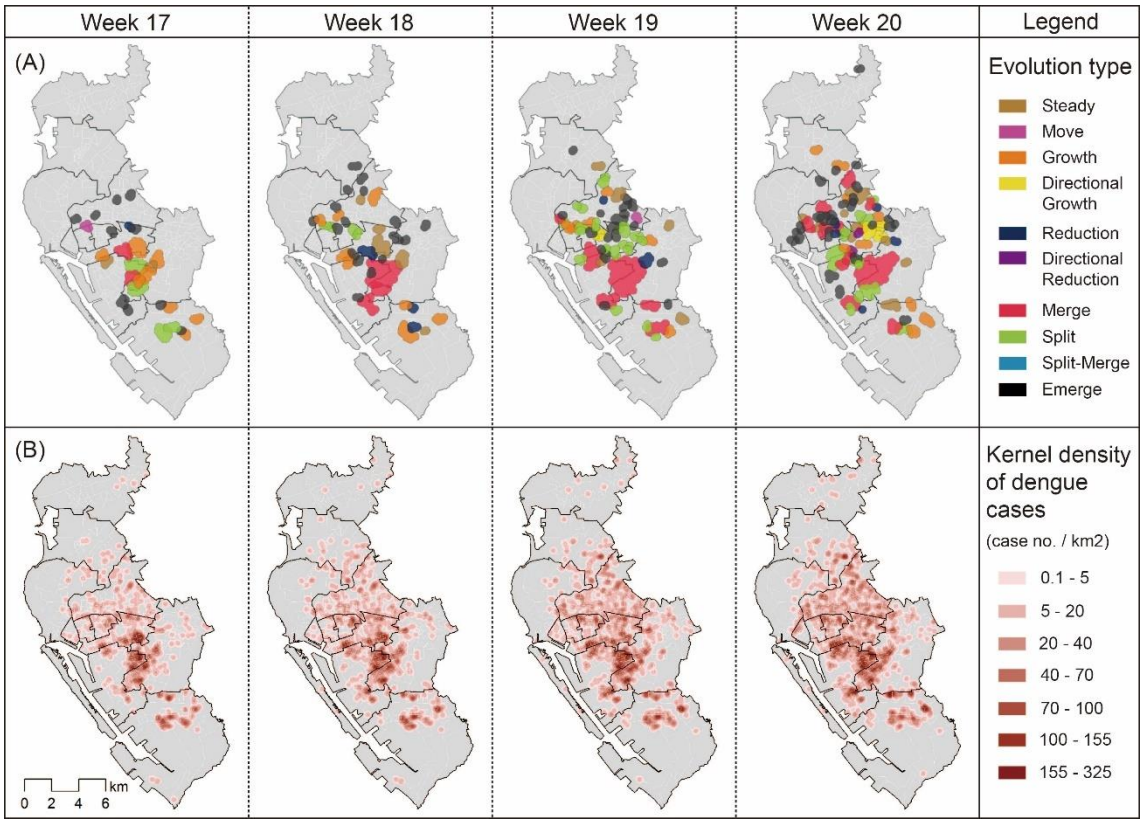


圖 8. MST-DBSCAN 與 KDE 的分析結果比較

doi:10.6342/NTU202101357

在高雄市 2014 年登革熱案例分析中，疫情群聚剛開始大多是在市中心以獨自型態（Single pattern）的方式擴散出去，直到互動型態（Interaction pattern）出現才發生大規模的疫情爆發。此項觀察與當地衛生局的長期經驗一致，因為高雄市每年的登革熱疫情幾乎都是源自於市中心的地帶。而過往研究也在 2001 到 2003 年的登革熱疫情 (Kan et al. 2008)，以及 2015 年的登革熱疫情 (Wen and Tsai 2015)發現了類似的擴散樣態。這兩篇文獻皆以接觸型擴散來描述疫情初期的擴散樣態，並認為此階段的擴散速度是較為緩慢的，且主要是受到環境因子（垃圾堆積或是積水等）的影響。另一方面，文獻以位移型擴散來描述疫情開始升溫時的擴散樣態，並認為此階段的擴散速度是相對較迅速的，且主要是受到人的通勤通學等跨區移動行為影響。因此，本論文所定義的獨自型態可比擬為接觸型擴散，而互動型態可比擬為位移型擴散。

上述比較顯示互動型態的群聚演化比起獨自型態更容易造成疫情大規模爆發，本論文認為這是因為其中的浮現型態（Emerge）。因為當一個新的小群聚浮現時，當地的大多數民眾都沒有抗體；這個小群聚便能以各種獨自型態快速擴張，直到該區域的多數民眾都被感染。雖然擴散速度可能因此緩慢下來，該群聚卻也早已發展成了一個地理範圍較大的群聚，並且有更多機會與周遭的其他群聚發生互動關係，特別是融合型態。當兩個群聚融合後，代表這些地方的傳播擴散關係趨向複雜，難以再區分每個病例的感染源頭，也增加了區域內民眾的感染機率。此時，受感染民眾透過移動將疾病帶往其他乾淨的地區，即有可能促成浮現型態的發生，再重複上述的擴散連動過程。本論文推測這樣的惡性循環便是病例數快速增加，疫情在中後期加劇的一大原因。因此，若是能有效降低浮現型態的發生頻率，即針對人口移動進行一些管制措施，或是針對新興地區的病媒蚊即早進行處理，應該會是有效防止疫情擴散的措施。

有別於過往文獻只能粗略地使用接觸型擴散或是位移型擴散來概括多種不同

doi:10.6342/NTU202101357

的擴散型態，本論文的定義和分類可以將疫情群聚的中心點和範圍變化等特徵，以及不同群聚之間的互動，視為擴散型態改變的線索。這些線索除了能協助理解每個疫情群聚的詳細擴散過程之外，也能夠協助防疫單位來探討疫情的發展是否與這些線索有關。面對不同線索特徵所造成的疫情轉變，防疫單位便能夠制定不同的防疫措施來避免傳染病進一步擴散，因地制宜達到有效的防疫目標。

doi:10.6342/NTU202101357

3.2 移動管制分區劃分

3.2.1 研究背景

城市規模的移動管制容易造成社會恐慌，引起人民的反抗和逃跑，導致推行防疫措施受到阻礙；另一方面，若是劃設的管制區太小，可能會因為疾病早已傳播至他處而忽略了潛在的疫情風險。因此，疫情管制區的地理範圍既不能太大也不能太小，必須要在一個合理的規模達到最佳的防疫效果 (Barbera et al. 2001)。目前已有研究利用移動基準分區的概念針對家畜之間的傳染病進行防疫分區劃分 (Lee et al. 2019)；然而，若是要應用此概念來劃分人類傳染病的防疫分區，本論文認為其忽略了都市人口移動規律性這項重要的人類移動行為特徵。過往研究已經發現人們的移動行為會在時間和空間維度上出現高度的規律性 (Sun et al. 2013)；同時，人們會時常返回至少數幾個經常訪問的地點 (González, Hidalgo, and Barabási 2008)。因此，學者推論多數人在一天的日常生活中只會訪問一定數量的地點，並推測其原因來自於人們的既定行程，例如通勤、採買生活用品等 (Schneider et al. 2013)。Belik, Geisel, and Brockmann (2011) 證明了若是忽略了人口移動的規律性，將會高估疫情擴散的速度；換言之，若是上述的移動基準分區概念不考量移動規律性，將有可能劃定過大的分區，進而引起不必要的恐慌和不便，阻礙防疫工作的進行並降低防疫措施的效果。

基於上述原因，本論文於此部分的具體研究目的係整合移動基準分區的概念和人口移動規律性，並發展了一個新的網絡分群演算法，名為 Human Mobility Regularity-based Zoning（簡稱 HuMoRZ），用以劃定地理規模大小更合適的防疫管制分區。HuMoRZ 演算法係以 Rosvall and Bergstrom (2008) 所發明的 Map equation 演算法作為發展的基礎，因為該方法可以有效地針對人口移動網絡這種有方向、有權重的網絡進行區塊劃分。

doi:10.6342/NTU202101357

3.2.2 研究方法與資料

由於 HuMoRZ 是基於 Map equation 進行演算法的發展，所以必須先理解 Map equation 的基本概念，才能夠瞭解本論文的創新性。Map equation 是一個最佳化演算法，給定一個分群結果（分群數量和每個節點所屬的分群），Map equation 即會根據這樣的樣態進行評分；接著，透過找解策略來不斷調整分群結果然後進行評分，直到評分結果無法再改善為止。針對每一次分群結果的評分，Map equation 認為一個網絡是隨機漫步者（Random walker）在節點之間不斷移動長時間後所產生的穩定態，所以演算法使用了漫步者的移動特性來描述一個給定的分群結果。

接著，本論文進一步說明如何在 Map equation 中考量人口移動規律性，形成 HuMoRZ。前述文獻提到移動規律性是來自人們日常生活中的既定行程，換言之，給定一個旅次目的，人們的目的地應該會集中在少數幾個地點，而非整個都市。例如人們通常都只有一個固定的工作地點，或是人們往往傾向於前往熟悉的餐廳或店家消費。因此，HuMoRZ 針對 Map equation 的創新有兩個階段：（1）藉由土地利用的資訊使得人口流動網絡可以反映出旅次目的；（2）針對每個節點的每種旅次目的量測規律性，以下將詳述之。

首先，公式 6 量測了當隨機漫步者從某一個節點移動到另一個節點是基於某種特定旅次目的之機率。

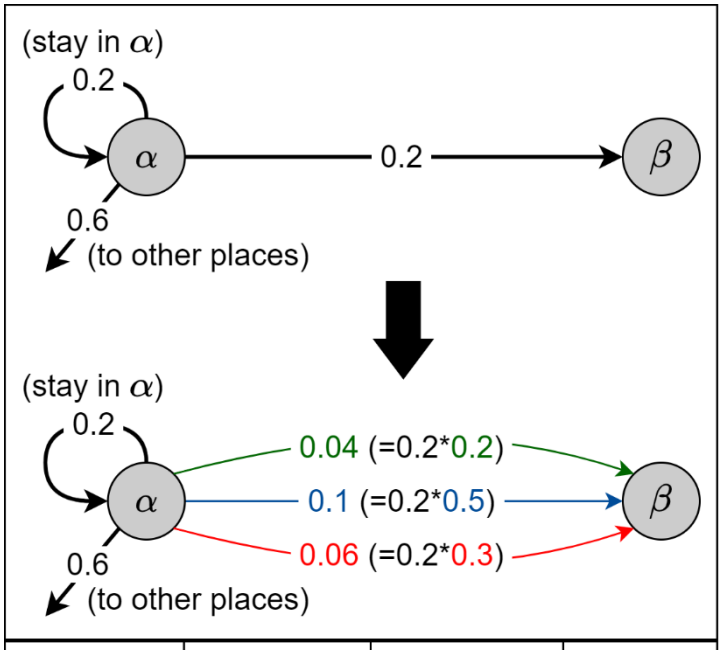
$$p_{\alpha\beta}^{k} = p_{\alpha}\omega_{\alpha\beta}z_{\beta}^{k} \tag{6}$$

$$\sum_{k=1}^{K} z_{\beta}^{k} = 1 \tag{7}$$

其中，α 是移動起點、β 是移動目的地，而 p_{α} 是隨機漫步者停留在 α 的機率；$\omega_{\alpha\beta}$ 是給定漫步者抵達 α 之後，移動到 β 的條件機率；z_{β}^{k} 代表第 k 種土地利用型態於β的面積比例。因此，$p_{\alpha}\omega_{\alpha\beta}$ 即代表從 α 移動到 β 的總機率，而

doi:10.6342/NTU202101357

HuMoRZ 利用了 β 的土地利用面積比例來切分這個總機率，讓每個部份的機率各自代表一種旅次目的之機率。圖 9 使用一個簡單的例子來展示公式 6 的概念。



Land use type	Residential	Commercial	Industrial
Area proportion in β	0.2	0.5	0.3
Arrow mark	→	→	→

圖 9. 土地利用反映旅次目的示意圖

接著，由於移動規律性表示給定一個旅次目的，目的地會集中在少數幾個節點而非遍布整個網絡中，所以 HuMoRZ 進一步利用熵值（Entropy）來針對上一步驟得到的旅次目的之機率進行量測，以便能量測集中程度並藉此反映移動規律性。公式 8 展示了 HuMoRZ 實際上使用熵值的方式。

doi:10.6342/NTU202101357

$$\mathrm{H}(\mathcal{Z}_{\alpha}^{k}) = -\sum_{\beta} \frac{p_{\alpha\beta}^{k}}{p_{\alpha}^{k}} \log_2(\frac{p_{\alpha\beta}^{k}}{p_{\alpha}^{k}}), \qquad \alpha, \beta \in \text{the same communiy} \tag{8}$$

$$p_{\alpha}^{k} = \sum_{\beta} p_{\alpha\beta}^{k}, \qquad \alpha, \beta \in \text{the same communiy} \tag{9}$$

其中，p_{α}^{k} 代表隨機漫步者因為第 k 種旅次目的離開 α，但仍然留在 α 所屬的分區內之機率。可以發現公式 8 和公式 9 都強調 α 和 β 必須屬於同一個分區，意即我們在使用熵值量測某個節點的移動規律性時，只考慮群內的移動狀況。原因在於原本的 Map equation 所產生的分區已經確保分區跟分區之間不會有大量的流動存在，故本論文認為可以忽略此部分的影響性而聚焦於區內流動的狀況來量測規律性。

現在，每個節點會有 K 個規律性分數，為了方便進行後續計算和比較，必須有一個指標來代表這 K 個分數的影響力。針對每個節點，HuMoRZ 使用每一種旅次目的之群內移動機率（p_{α}^{k}）來針對規律性（$\mathrm{H}(\mathcal{Z}_{\alpha}^{k})$）進行加權；接著，再把所有旅次目的加權後的規律性加總，來代表每個節點的綜合規律性。如公式 10 所示。

$$\mathrm{IR}(\alpha) = \sum_{\mathrm{k}=1}^{\mathrm{K}} \mathrm{p}_{\alpha}^{\mathrm{k}} \mathrm{H}(\mathcal{Z}_{\mathrm{a}}^{\mathrm{k}}) \tag{10}$$

基於熵值本身的特性，當 IR(α) 愈大時，代表從 α 離開的移動行為愈不規律；而當 IR(α) 愈小時，代表移動行為愈規律。因此，本論文將 IR(α) 稱呼為反規律性（Inverse-regularity），而圖 10 利用簡單的情境對比展示了公式 8 到公式 10 的概念。

doi:10.6342/NTU202101357

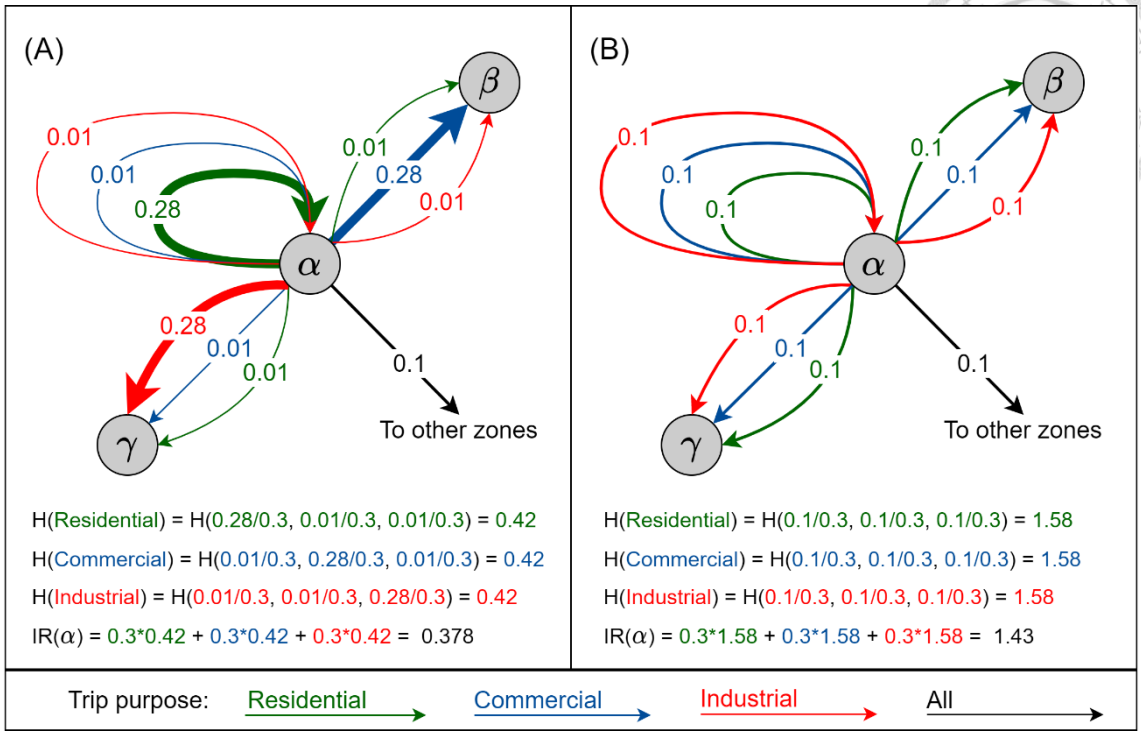


圖 10. 規律性量測示意圖

備註：在（A）中，可以看到某個人基於某種旅次目的而移動離開 α 時，有很明確的目的地，所以相對於（B），（A）的 IR(α) 具有較小的數值。

最後，將每個節點的非規律性分數加總，然後再與 Map equation 原本的目標式做結合，即得到 HuMoRZ 的目標式，如公式 11。

$$L_R(\mathrm{M}) = \mathrm{L}(\mathrm{M}) + \sum_{\alpha} H(\alpha) \tag{11}$$

誠如上述，HuMoRZ 遵循 Map equation 找解的策略，差別只在於使用不同的目標式來評估每個分群結果的好壞。因此，HuMoRZ 的目標是要最小化公式 11。

doi:10.6342/NTU202101357

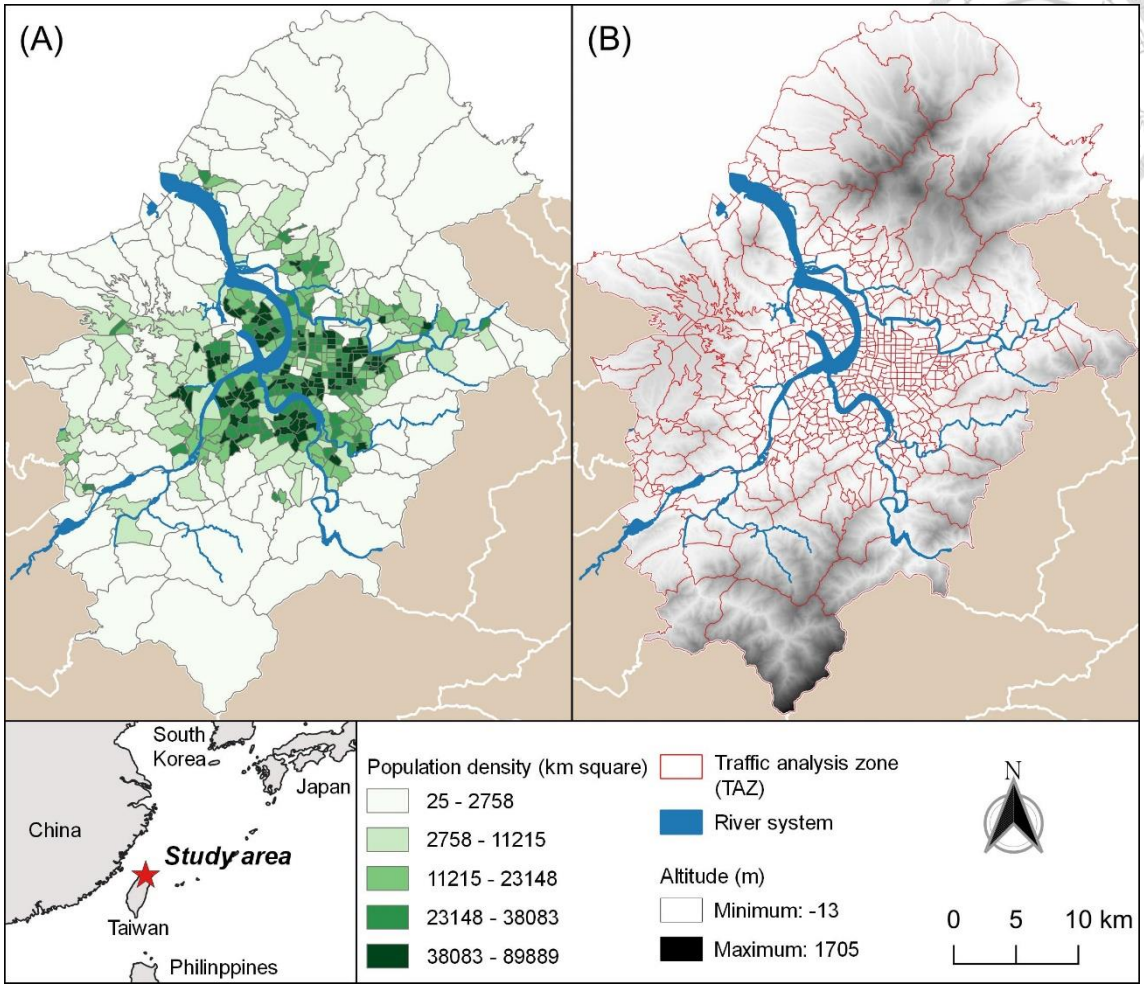


圖 11. 大台北都會區人口密度分布

本論文於此部分的研究選擇以大台北都會區作為研究區（詳見圖 11），涵蓋了大約 1362.57 平方公里的面積以及大約 634 萬的人口數。同時，本論文蒐集了土地利用和人口移動的相關資料，以利 HuMoRZ 演算法進行管制分區劃分。首先，土地利用資料是來自於國土測繪中心於 2012 年提供的國土使用現況調查資料。這份資料利用三個階層的分類系統來區分不同的土地使用類別，而我們使用了最小階層的分類尺度，因為這樣的尺度最能夠反映出細緻的人類活動行為和旅次目的。在我們的研究區當中，共計有 79 種不同的土地使用類別。另一方面，人口移動資料來自於台北市政府捷運局於 2009 進行的調查報告：第四版的「臺北都會區整體運輸需求模式建立與應用」。除了 2009 之外，政府也提供了 2012 年的推估移動樣態，故此資料能夠在時間和空間的維度上與土地使用資料一致。這份人口移動的資

doi:10.6342/NTU202101357

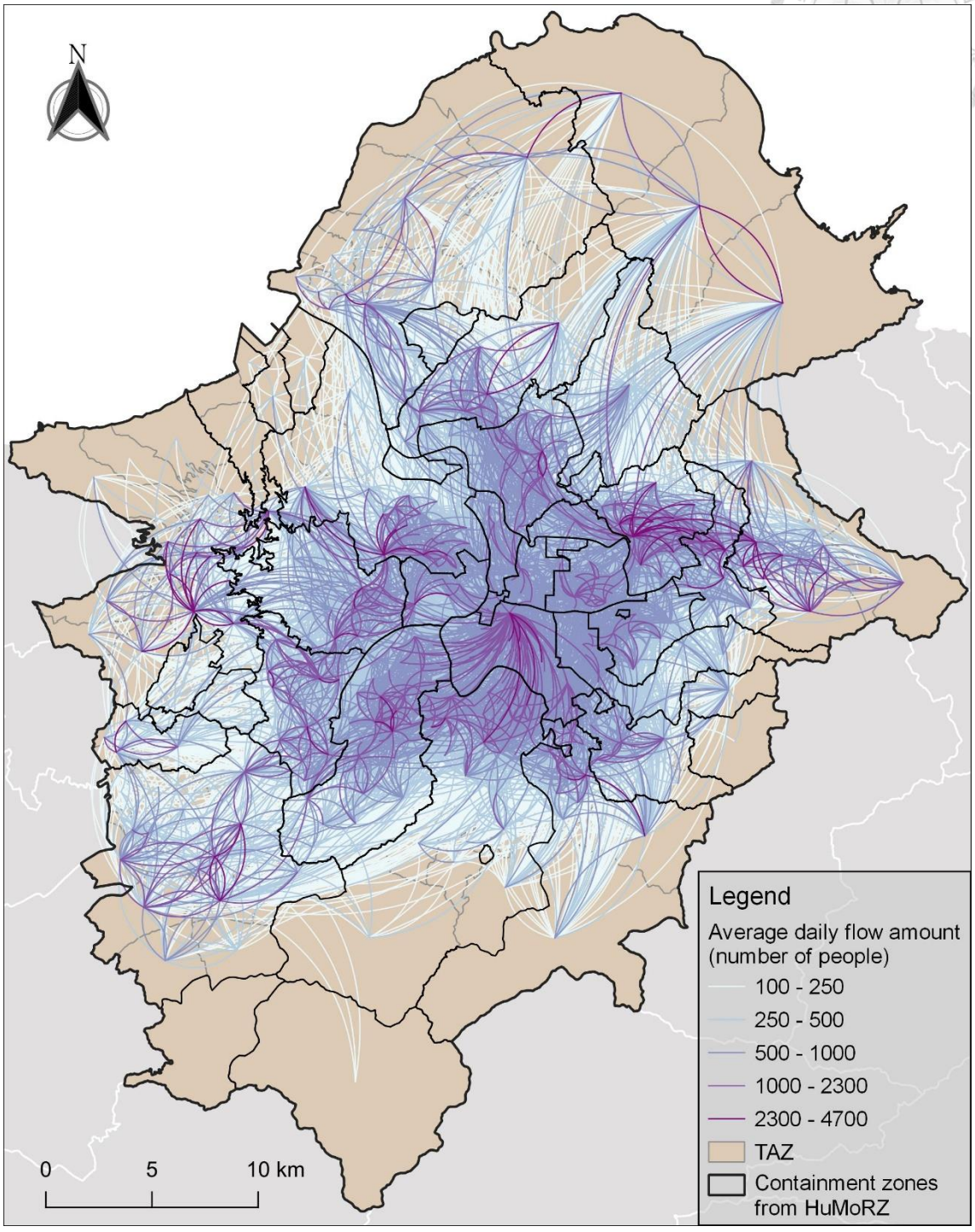


圖 12. 人口移動旅運需求日平均值之空間分布

料將整個大台北都會區劃分成 508 個不同的交通分析區（Traffic analysis zone, TAZ），並利用抽樣調查搭配模式推估的方式計算出從任一個 TAZ 到其他 TAZ 的日常移動需求量（圖 12 呈現了移動的空間分布）。TAZ 的主要劃分準則是盡

doi:10.6342/NTU202101357

可能讓每個 TAZ 的人口數不要差異太大，而在市中心和郊區的人口密度有顯著差異的情形下，TAZ 之間就會有明顯的面積大小差異。本論文的研究區包含了 498 個 TAZ，而圖 11A 呈現了它們的空間分布以及各自的人口密度。

3.2.3 研究成果摘要

基於上述的人口移動和土地使用資料，本團隊的演算法將 498 個 TAZ 劃分成 26 個不同的分區（圖 12）。從圖 12 可以發現，大多數高強度的移動路線還是維持在一個分區內，表示本團隊的分區結果仍然能夠盡可能地降低跨區移動的數量。

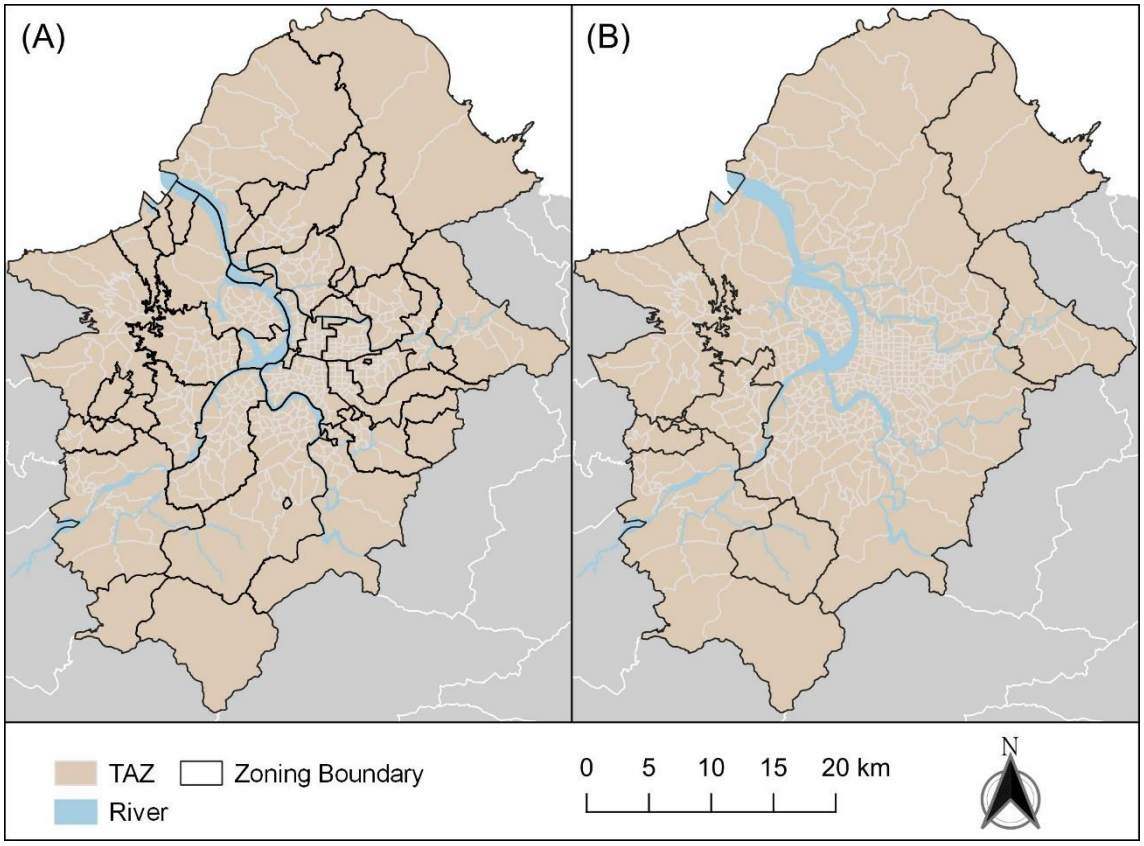


圖 13. HuMoRZ 與 Map equation 分析結果比較

圖 13 比較了我們的 HuMoRZ 演算法和既有的方法：Map equation；這兩者的主要差別在於演算法本身是否有考量移動規律性。由於 Map equation 只考慮流

doi:10.6342/NTU202101357

動的強度並最大化群內流動量而已，故它的分區結果顯示整個市中心地帶會組成一個巨大的分區來涵蓋大多數的 TAZ；而邊陲地帶的 TAZ 則組成若干個小規模的分區，反映出邊陲地帶的自我生活範圍。相較之下，我們的 HuMoRZ 演算法在邊陲地帶型塑出類似樣態的分區，但卻將市中心劃分成若干個緊湊的分區，而非一個單一的巨大分區。

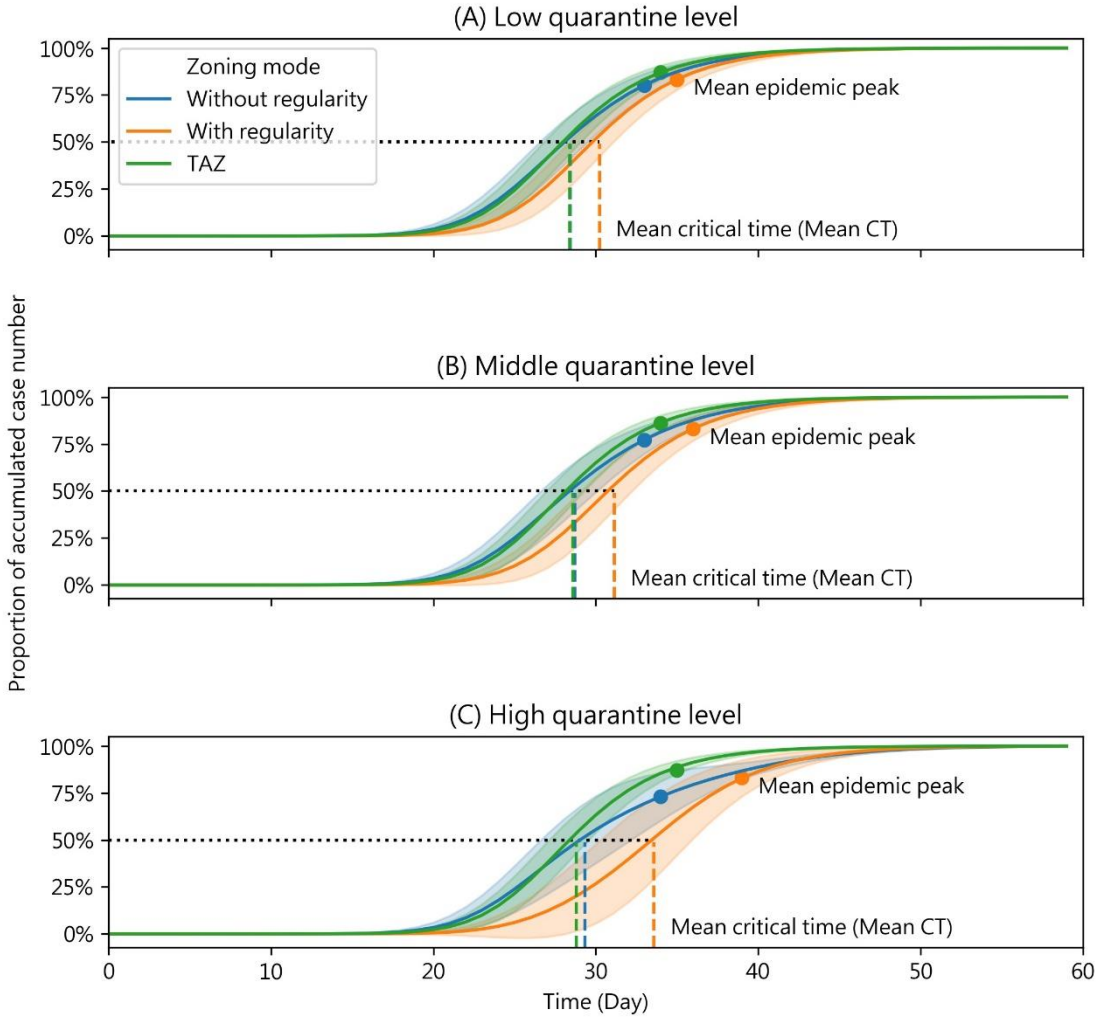


圖 14. 不同管制程度之疫情模擬比較

為了證明 HuMoRZ 演算法所產生的分區能夠防止疫情擴散，本論文以上呼吸道疾病為模擬案例，並且以 SEIR 傳染病流行病學數理模式進行疫情擴散模擬（詳細模擬模型請參見附錄 B），比較三種不同分區模式的效力。由於 Map equation 的

doi:10.6342/NTU202101357

分區結果將整個市中心劃分成同一分區，所以可以比擬為城市封鎖等級的疫情管控(圖 13)；而每個 TAZ 因為本身的特性，可以比擬為小區封鎖等級的疫情管控；至於 HuMoRZ 演算法的分區結果之每個分區規模大小介於上述兩種分區模式之間。比較結果(圖 14)顯示 HuMoRZ 的分區可以更有效地減緩疫情擴散的速度，同時，當管制等級愈嚴格時，HuMoRZ 分區的表現會比其他兩種分區模式的表現愈來愈好。在最嚴格的分區模式下（圖 14C），HuMoRZ 分區可以延緩疫情高峰將近一周。圖 15 進一步比較了三種分區模式的疫情嚴重程度(一個月內的累積病例人數，以 AN 為標記)以及擴散速度(研究區內一半人數被感染的時間點，以 CT 標記)。結果顯示 HuMoRZ 分區具有最低的 AN 數值和最晚的 CT 時間點，表示此管制分區劃分可以顯著地在疫情初期降低疫情嚴重性，同時有效地延緩疫情高峰出現的時間點。

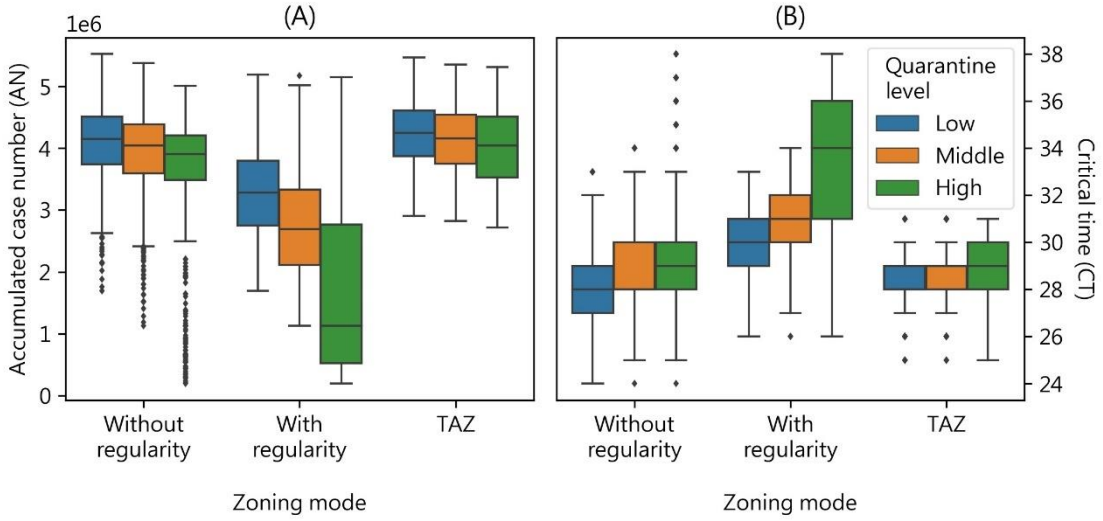


圖 15. AN 和 CT 兩項指標之數值分布比較

本論文的 HuMoRZ 演算法比起 Map equation 多考量了人口移動的規律特性，所以演算法可以產生許多地理規模介於城市封鎖和小區封鎖之間的分區，這也符合過往學者對於疾病管制區域大小的論述 (Barbera et al. 2001)。城市封鎖需要大量

doi:10.6342/NTU202101357

的人力物力進行管控，且容易造成社會恐慌。雖然中國看似成功執行了城市封鎖來對抗 COVID-19 疫情，但學者認為其他國家因為政治體系和制度與中國的明顯不同，故難以仿效中國這種較為極端的城市封鎖措施 (Kupferschmidt and Cohen 2020)。另一方面，小區封鎖其實難以確實將病原體控制在該範圍內，因為病毒很有可能早在實行封鎖之前就已經擴散到區域之外，導致傳播不會被斷絕；而類似的現象已出現在美國的 COVID-19 疫情當中 (Parodi and Liu 2020)。因此，我們的分區大小是一個比較適當的規模來執行移動管制措施。

在歐洲各國的 COVID-19 疫情中，義大利是最早淪陷的國家。學者認為主因是義大利政府太晚施行相關的防疫措施 (Indolfi and Spaccarotella 2020)，而政府的顧慮可能是擔心移動管制或封城等措施會影響國內的經濟、觀光產業等 (Auzan 2020)。在這種情形下，HuMoRZ 的分區模式便可提供有效的解方。由於比較分析的結果顯示了 HuMoRZ 的分區可以在疫情初期明顯降低感染人數，同時又可以延緩疫情擴散的速度和高峰時間點。這些改善都能協助政府和防疫人員爭取更多時間來提前佈署防疫資源、擬定防疫策略來對抗疫情的後續發展 (Desjardins, Hohl, and Delmelle 2020)。因此，若是義大利政府採取 HuMoRZ 演算法進行防疫分區規劃，也許就能夠在疫情初期只針對疫情較嚴峻的地區進行管制，讓國家的整體經濟和觀光產業不會受到劇烈的衝擊。

這次的 COVID-19 也在各國傳出有許多無症狀的感染者，能夠在無形之中將病原體傳播給他人，造成民眾間的惶恐。為了找出這些隱形的傳播者，包括南韓、德國等國家都進行了大規模的篩檢。雖然有一定的成效，但是其所耗費的人力、物資並非所有國家或是地方上的衛生單位都能夠負擔。基於此原因，本論文認為 HuMoRZ 的分區能夠協助解決問題。由於演算法本身考量了人口移動的規律性，所以與確診者頻繁接觸的他人基本上會被網羅在同一個分區內。在這樣的情況下，篩檢就可以聚焦於若干個高風險區域，而不必將整個城市或甚至是整個國家的人

doi:10.6342/NTU202101357

口都進行篩檢。除了可以避免篩檢的偽陰性、偽陽性等問題之外，也可以更有效率地運用篩檢的人力和試劑等資源。因此，搭配移動管制分區等措施可以讓防疫工作更有效益。

doi:10.6342/NTU202101357

3.3 防疫醫療資源之空間可及性

3.3.1 研究背景

過往關於空間可及性的研究大多是在靜態穩定的狀態下評估每個地區的醫療資源可及性；換言之，每個地區的需求量皆被假設為固定不變的。以此次的新冠肺炎為例，有文獻利用 2SFCA 系列模型分析採檢站 (Tao et al. 2020)、負壓隔離病房 (Kim et al. 2020) 以及加護病房 (Pereira et al. 2020) 等醫療資源的空間可及性，或是城市內各地區老年人口對於基礎醫療照護資源的空間可及性 (Guida and Carpentieri 2021)。然而，這樣的假設並不適用於一個城市面對急性傳染病疫情威脅的情況。

首先，急性傳染病具備突發且迅速擴散的特性，能夠在短時間內感染大量的民眾，產生額外的醫療資源需求。這種由傳染病疫情所引起的需求有兩種特性：（1）主要的需求來自於已感染或是高風險的族群，而非所有民眾都需要前往醫療院所索取醫療資源，例如篩檢試劑或是重症加護病房等。因此，必須準確地推估一個地區可能存在的已感染或是高風險的族群人數，才能夠合理地評估該地區的醫療資源需求量，並進一步推算可及性。（2）由於需求是來自於已感染或是高風險的族群，而這兩個族群的人數又會隨著疫情的發展而不斷改變，故同一個地區對於醫療資源的需求量也將會隨著時間而變動 (Weissman et al. 2020)，連帶造成可及性的動態變化。

雖然近年來也有文獻嘗試剖析空間可及性在時序上的變化，但這些研究是關注於長時間的建成環境改變所引起的可及性變化 (Hu et al. 2020; Xing et al. 2018)；或者通勤移動行為所帶來的需求分布變化並導致可及性在一天內的規律變化 (Xia et al. 2019; Yun et al. 2020)。這些文獻的分析架構或是模型並沒有考量到由疫情引起的需求變動及相關特性；再加上，空間可及性的變動又會反過來影響疫情的後續發展，所以兩者之間存在交互作用的關係，無法單獨評估。因此，現階段並沒有合

doi:10.6342/NTU202101357

適的模型架構能夠審慎評估疫情期間的醫療資源空間可及性之時序變化。

基於上述之研究缺口，本論文發展了一套命名為 Epi-RA 的整合模式用以整合醫療資源可及性評估（2SFCA）與疫情擴散模擬（SEIR）的時空風險分析。疫情擴散模擬能夠推演出疫情的可能發展狀況，協助模式評估醫療資源需求的動態變化；在模擬過程中，搭配 2SFCA 系列模型架構進行即時的空間可及性推算，即可捕捉到疫情發展和可及性變動之間的交互作用關係，更合理地剖析出傳染病疫情期間的醫療資源可及性在時間與空間上的變化。除此之外，本論文以新冠肺炎採檢試劑代表醫療資源，並以中央流行疫情指揮中心所公布的採檢站假設為資源供給地點。選擇採檢試劑為研究對象是因為在缺乏足夠數量疫苗的防疫中，特別像是此次新冠肺炎疫情含有許多無症狀感染者，篩檢是有效來防止傳染病疫情擴散的措施 (He et al. 2020)。同時，在美國的防疫研究也發現，篩檢資源可及性過低也是造成疫情大規模爆發的原因之一 (Sharfstein, Becker, and Mello 2020)。

3.3.2 研究方法與資料

本論文之 Epi-RA 模型的架構圖如圖 16 所示。首先，參考 Gatto et al. (2020) 所提出的 SEPIA 模型，本團隊將疫情的動態過程區分成八個階段：易感（S）、暴露（E）、發病前（P）、重症（I）、輕症或無症（A）、隔離（Q）、康復（R）、死亡（D）。由於只有 P、I、A 這三個階段的感染者具備傳播疾病給予他人的能力，故本論文假設只有這三個階段的感染者才會使用新冠肺炎採檢試劑。換言之，這三個階段的感染者人數代表了疫情在每個時間點所引起的需求量；而這三個階段的人數變化即反映出需求隨著疫情發展而改變的狀況。基於這樣的變化資訊，2SFCA 模型即可動態地評估每個地點在每個時間點下的篩檢資源可及性。篩檢資源可及性的評估狀況，影響著下一個時間點的疫情模擬結果，受影響的 P、I、A 三個階段的人數又反過來影響該時間點的可及性評估結果，兩者之間形成交互循環影響的

doi:10.6342/NTU202101357

關係。因此，Epi-RA 模式展示了空間可及性評估和疫情模擬兩種模型之間的整合。

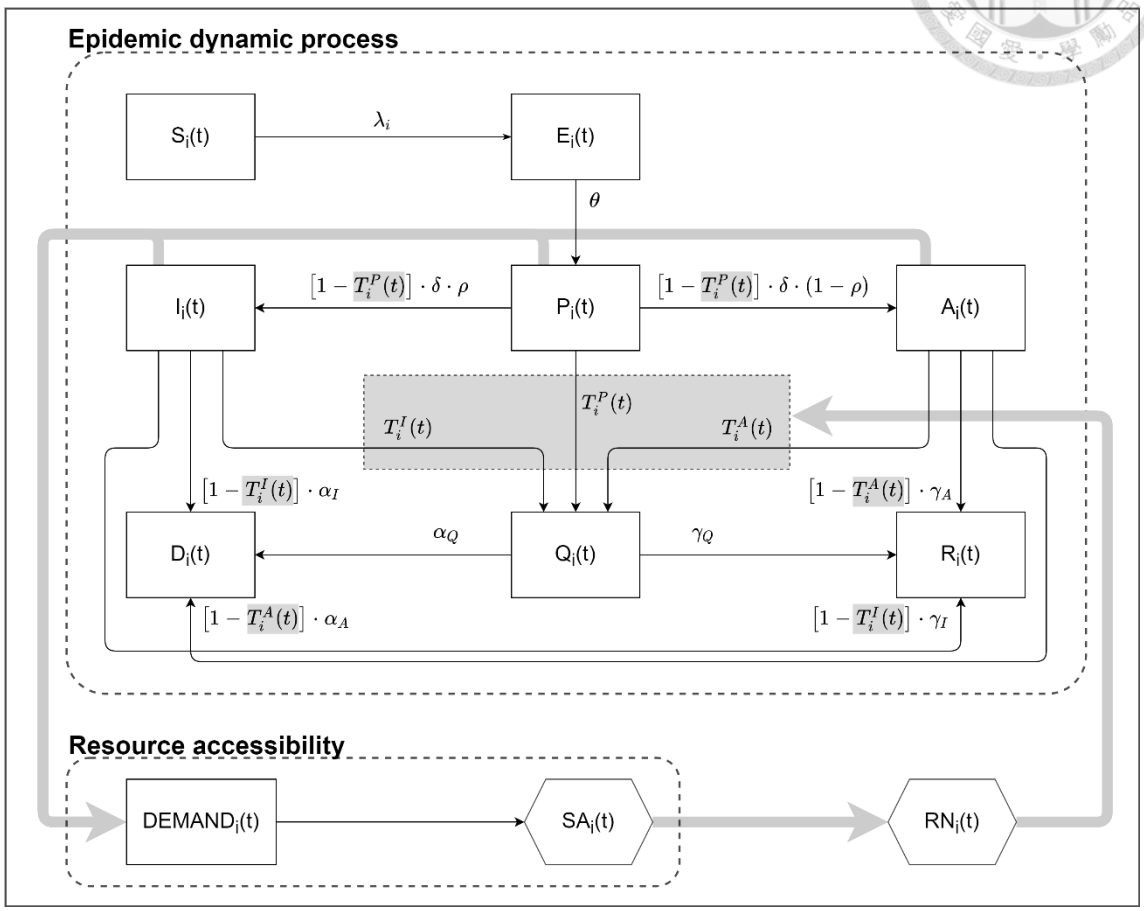


圖 16. Epi-RA 整合模式架構圖

A. 可及性評估模式

誠如上述，本論文的初步假設只會使用發病前（P）、重症（I）、輕症或無症（A）等三個階段的感染者來代表每個地區在每個時間點下的需求量（公式 12）：

$$DEMAND_i(t) = v_P P_i(t) + v_I I_i(t) + v_A A_i(t) \quad (12)$$

其中，$P_i(t)$、$I_i(t)$ 和 $A_i(t)$ 分別代表 i 地區在 t 時間點下的 P、I、A 階段感染者數量；v_P、v_I 和 v_A 則分別代表三個階段各自有多少比例的人數會索取資源。基於這種需求的變動，本論文使用 2SFCA 模型來量測每個地點的可及性時序動態（公

doi:10.6342/NTU202101357

式 13）：

$$SA_i(t) = \sum_{j=1}^{n} \frac{S_j f(d_{ij})}{\sum_{k=1}^{m} DEMAND_k(t) f(d_{kj})} \quad (13)$$

其中，S_j 代表 j 供給設施所能提供的資源數量；d_{ij} 代表 i 地區和 j 供給設施之間的距離；$f(d_{ij})$ 是轉換函數用以將距離換算成旅行成本。本論文採取負冪次形式的轉換函數（$f(d_{ij}) = d_{ij}^{-2}$），因為它是文獻上常用的一種轉換函數(Kwan 1998)。透過公式 12 和 13，本論文將疫情動態整合到 2SFCA 模型當中，讓評估出的可及性會隨著空間和時間變化。

B. 疫情擴散模擬模式

疫情擴散模擬是用以推估每個各地區的疫情發展過程，Epi-RA 的疫情擴散模擬是基於 Gatto et al. (2020)所提出的 SEPIA 模式進行開發（公式 14）：

$$\frac{dS_i}{dt} = -\lambda_i(t) \cdot S_i$$

$$\frac{dE_i}{dt} = \lambda_i(t) \cdot S_i - \theta E_i$$

$$\frac{dP_i}{dt} = \theta \cdot E_i - \{T_i^P(t) + [1 - T_i^P(t)] \cdot \delta\} P_i$$

$$\frac{dI_i}{dt} = [1 - T_i^P(t)] \cdot \rho \cdot \delta \cdot P_i - \{T_i^I(t) + [1 - T_i^I(t)](\alpha_I + \gamma_I)\} I_i$$

$$\frac{dH_i}{dt} = T_i^I(t) \cdot I_i - (\alpha_H + \gamma_H) \cdot H_i \quad (14)$$

$$\frac{dD_i}{dt} = [1 - T_i^I(t)] \cdot \alpha_I I_i + \alpha_H H_i$$

$$\frac{dA_i}{dt} = [1 - T_i^P(t)] \cdot (1 - \rho) \cdot \delta \cdot P_i - \{T_i^A(t) + [1 - T_i^A(t)] \cdot \gamma_A\} A_i$$

$$\frac{dQ_i}{dt} = T_i^P(t) \cdot P_i + T_i^A(t) \cdot A_i - \gamma_Q \cdot Q_i$$

$$\frac{dR_i}{dt} = [1 - T_i^I(t)] \cdot \gamma_I I_i + [1 - T_i^A(t)] \cdot \gamma_A A_i + \gamma_H H_i + \gamma_Q Q_i$$

doi:10.6342/NTU202101357

在此模式中，當易感（S）族群與具備傳播力的感染者（發病前（P）、重症（I）、輕症或無症（A）等三個族群）接觸後，會轉變成暴露（E）族群。各地區在不同時間點下的傳播力之定義如公式 15：

$$\lambda_i(t) = \sum_j \frac{\beta_P \cdot P_j(t) + \beta_I \cdot I_j(t) + \beta_A \cdot A_j(t)}{S_j(t) + E_j(t) + P_j(t) + I_j(t) + A_j(t) + R_j(t)} \cdot \frac{w(j,i)}{N_j} \qquad (15)$$

其中，β_P、β_I 和 β_A 分別代表 P、I、A 三個具有感染力的族群之傳播率；$w(j,i)$ 代表從 j 地區到 i 地區的日平均通勤移動人數；N_j 代表 j 地區的總人口數；$\frac{w(j,i)}{N_j}$ 即代表 j 地區的疫情如何透過人口移動的關係影響 i 地區的疫情。另外，當 j 與 i 相等時，公式 15 即反映出在地的疫情傳播情形。

在 Epi-RA 的模式中，暴露者（E）並沒有傳播疾病的能力。必須經過潛伏期（速率θ）之後，轉變成感染前（P）的階段才具備傳播疾病的能力。對於沒有被隔離的 P 階段感染者，在經過一段時間後（速率δ），一部分的人（比例ρ）會出現重度症狀（I），其餘的人（比例$1-\rho$）則是輕症或是無症狀（A）。這兩個階段的感染者除了被隔離以外，最終會進入康復狀態（速率γ_I或是 γ_A）；此外，重症患者還有可能進入死亡狀態（速率α_I）。同樣地，重症隔離（H）和輕症隔離（Q）的患者最終也會進入康復狀態（速率γ_H或是 γ_Q）；而重症隔離（H）的患者也有死亡的可能（速率α_H）。

從 P、I 和 A 這三個階段分別進入重症隔離（H）或是輕症隔離（Q）階段的人數，取決於每個時間點下空間可及性的評估結果。因為高可及性代表了該地區能夠獲取較多的篩檢試劑來檢測更多民眾是否被感染。公式 13 所計算出的空間可及性代表一個人在 i 地區在 t 時間點下平均能獲得多少資源，當該地充滿資源或是需求

doi:10.6342/NTU202101357

量不大時，這個數值可能會遠大於 1。然而，大於 1 的數值對於 Epi-RA 並沒有意義，因為以篩檢試劑來講，一個人一天只會需要用到一劑。因此，本論文利用公式 16 將空間可及性分數轉換成配給率：

$$RN_i(t) = \min\big(1, SA_i(t)\big) \tag{16}$$

此公式確保了配給率（$RN_i(t)$）的數值介於 0~1 之間，1 代表所有對於篩檢試劑的需求都能夠被滿足。而公式 17 進一步定義了每個地區的民眾在不同時間點下分別從 P、I、A 三個階段轉移到隔離階段的速率：

$$\begin{aligned} T_i^P(\mathrm{t}) &= v_P \cdot RN_i(\mathrm{t}) \cdot \mathrm{tp} \\ T_i^I(t) &= v_I \cdot RN_i(\mathrm{t}) \cdot \mathrm{tp} \\ T_i^A(t) &= v_A \cdot RN_i(\mathrm{t}) \cdot \mathrm{tp} \end{aligned} \tag{17}$$

其中，v_P、v_I 和 v_A 與公式 12 中的定義一樣，分別代表三個階段各自有多少比例的人數會索取資源；tp 代表採檢試劑的真陽性率，反映出有多少接受採檢的感染者是能夠確實被捕捉到的。基於 $RN_i(t)$ 的特性，$T_i^P(\mathrm{t})$、$T_i^I(t)$、$T_i^A(t)$也會在時間與空間維度上變化；因此，公式 17 呈現了疫情擴散模擬如何整合空間及性的評估結果。表 4 呈現了模型中所有參數的定義與數值設定。

doi:10.6342/NTU202101357

表 4. Epi-RA 模型參數之定義與數值

參數	定義	數值	參考文獻
β_P	發病前患者的傳播率	4.8014	*
β_I	重症患者的傳播率	0.1632	*
β_A	輕症患者的傳播率	0.1585	*
$1/\theta$	發病潛伏期	3.32 天	*
$1/\delta$	傳染潛伏期	0.75 天	*
ρ	成為重症患者的比例	0.25	*
γ_I	重症患者康復速率	0.0698	*
γ_H	重症隔離患者康復速率	0.0698	*
γ_A	輕症患者康復速率	0.1396	*
γ_Q	輕症隔離患者康復速率	0.0698	*
α_I	重症患者死亡速率	0.0413	*
α_H	重症隔離患者死亡速率	0.0413	*
v_P	發病前患者的資源需求比例	0.1	-
v_I	重症患者的資源需求比例	0.5	-
v_A	輕症患者的資源需求比例	0.3	-
tp	採檢試劑真陽性率	0.84	Bastos et al. (2020)

備註：*標記代表該參數的數值是參考自 Gatto et al. (2020)，-標記代表該參數的數值為本論文自行定義。

本論文使用大台北都會區作為應用案例進行研究分析，藉此來檢驗 Epi-RA 整合模型的效用性，詳細的研究範圍如圖 17 所示。Epi-RA 整合模型需要三項輸入資料：地理人口分布、人口移動流量以及採檢站點位與資源數量。首先，台北市政府捷運局於 2009 年時，為了調查旅運需求量，將本計畫所匡列的研究範圍劃分成 543 個交通分析區（Traffic Analysis Zones, TAZ）。這些 TAZ 的平均面積大小約為 2.52 平方公里，平均人口數為 12,500，圖 17A 的面量圖呈現了 TAZ 人口密度的空間分布。在本計畫中，每個 TAZ 都將被視為單一的需求地區。第二，台北市政府捷運局於 2015 年時，依照 2009 年的研究架構，重新調查了一次旅運需求量之日平均值。因此，本團隊規劃使用 2015 年的調查結果作為本計畫的人口移動資料，圖 17B 呈現了 543 個 TAZ 之間的人口移動之日平均量。最後，本團隊規劃使用中

doi:10.6342/NTU202101357

央流行疫情指揮中心所公布的社區採檢站作為供給設施的位置，並挑選出位於本計畫之研究範圍內的共 22 處站點（圖 17A 的紅點）。

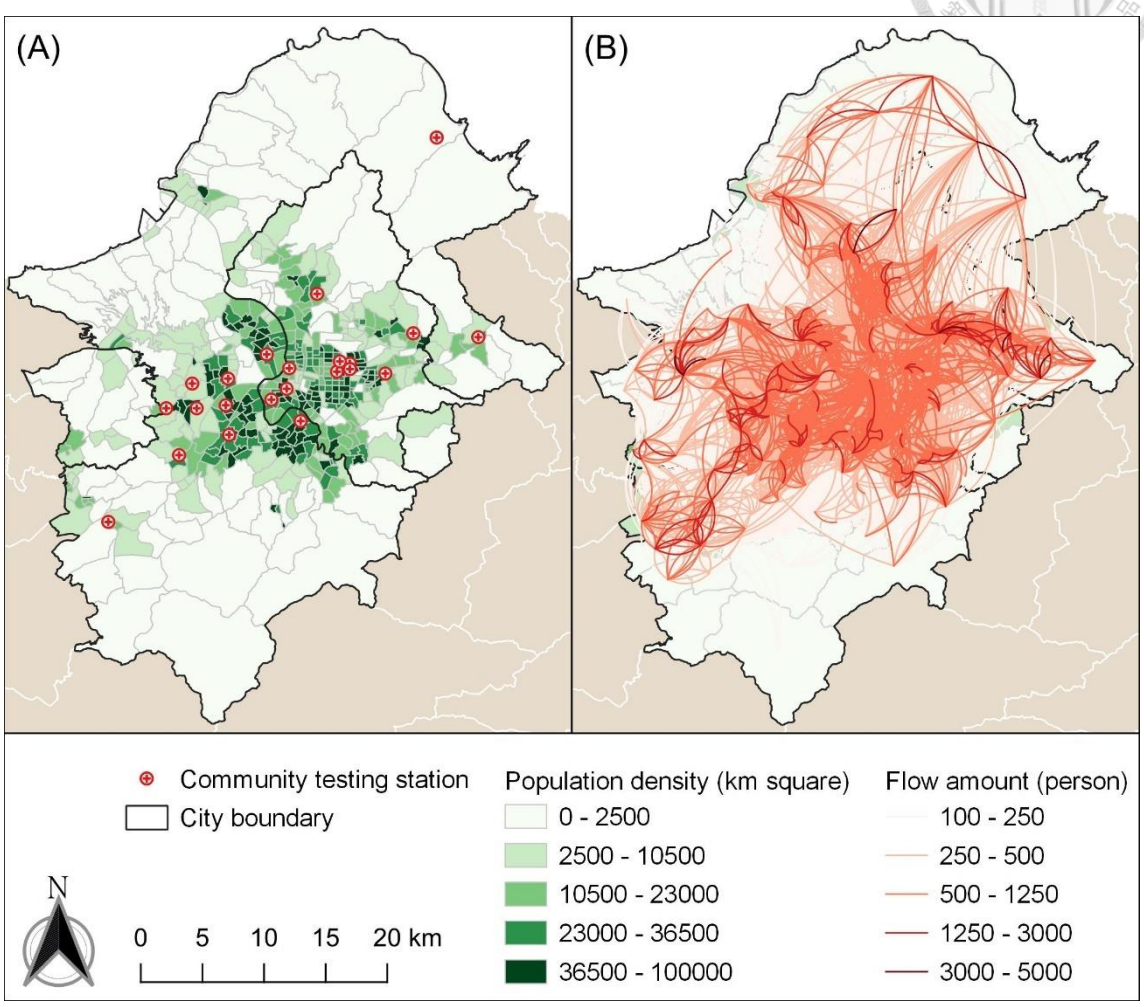


圖 17. 研究區與研究資料
（A）人口密度分布與採檢站分布、（B）2015 年人口移動旅運需求日平均分布

3.3.3 研究成果摘要

剖析疫情發展引起的需求變動，如圖 18 所示。圖 18A 顯示了整個大台北都會區，發病前（P）、重症（I）以及輕症或無症（A）等三個階段的人數（採檢需求來源）隨著時間而變化的過程。圖 18B、C、D 則分別呈現了空間可及性在第 20 天（指數上升期）、第 26 天（疫情巔峰期）以及第 32 天（衰退遞減期）的地理分布。在指數上升期間，對於採檢資源的需求多半集中在市中心地區；在疫情顛峰期間，包括市中心和外圍地區都有大量的需求；而到了衰退遞減期間，市中心的需求

doi:10.6342/NTU202101357

開始減少，不過外圍地區依然有高度的需求量。因此，圖 18 清楚地呈現了需求量由市中心擴散至外圍地區的時空動態變化過程。圖 19 進一步呈現了可及性在三個不同時間點下的空間分布。在疫情初期，外圍地區相較於市中心馬上出現資源不足的問題（圖 16A）；到了疫情高峰期，幾乎所有地方都面臨資源匱乏的狀況（圖 16B）；而當疫情開始消退時，卻是市中心那些鄰近採檢站的地區率先脫離資源缺少的困境（圖 16C）。圖 18 和圖 19 之間的對比顯示了在需求小的地區，也有可能因為資源過度缺乏而出現供需失衡（研究區的西北角）；而在市中心資源較充足的地方，也有可能因為疫情巔峰期間的資源需求量過大而導致供不應求。

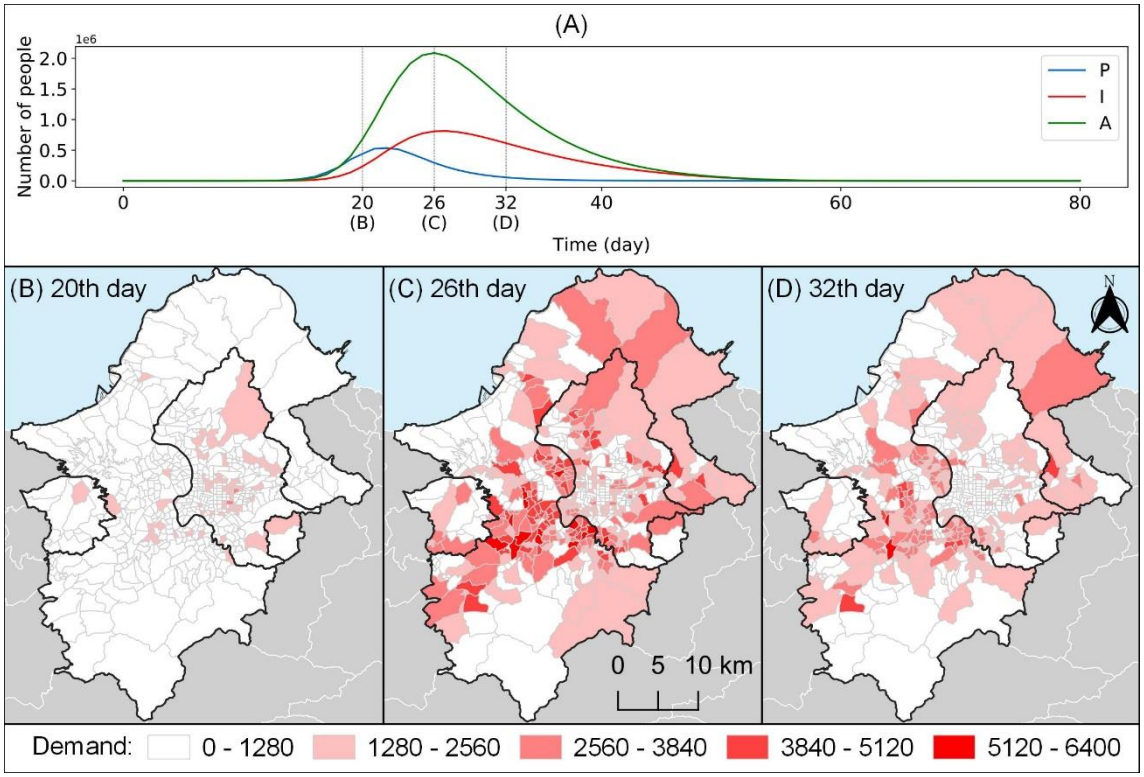


圖 18. 各地點之需求動態變化

doi:10.6342/NTU202101357

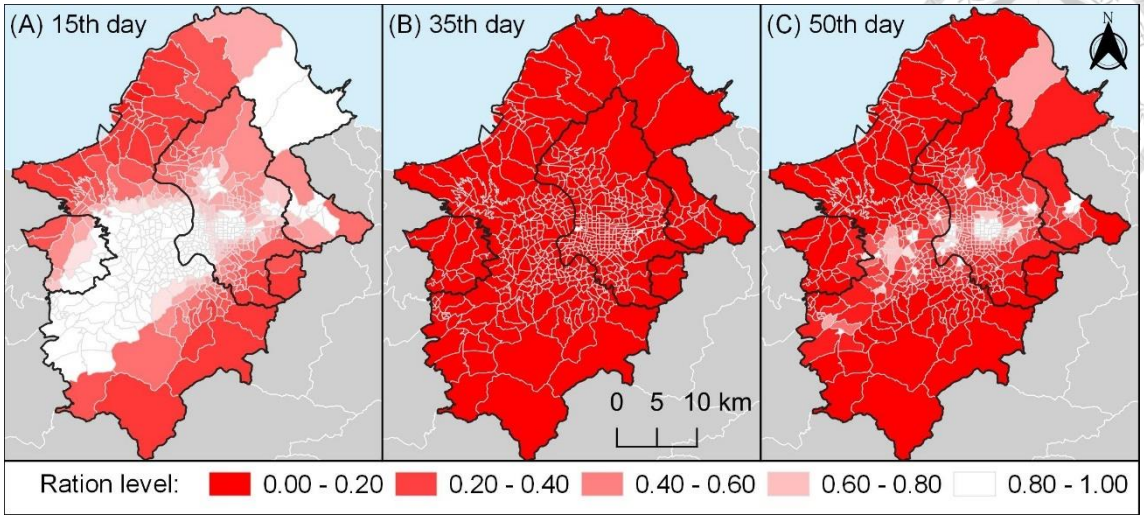


圖 19. 各地區之資源可及性（配給率）動態變化

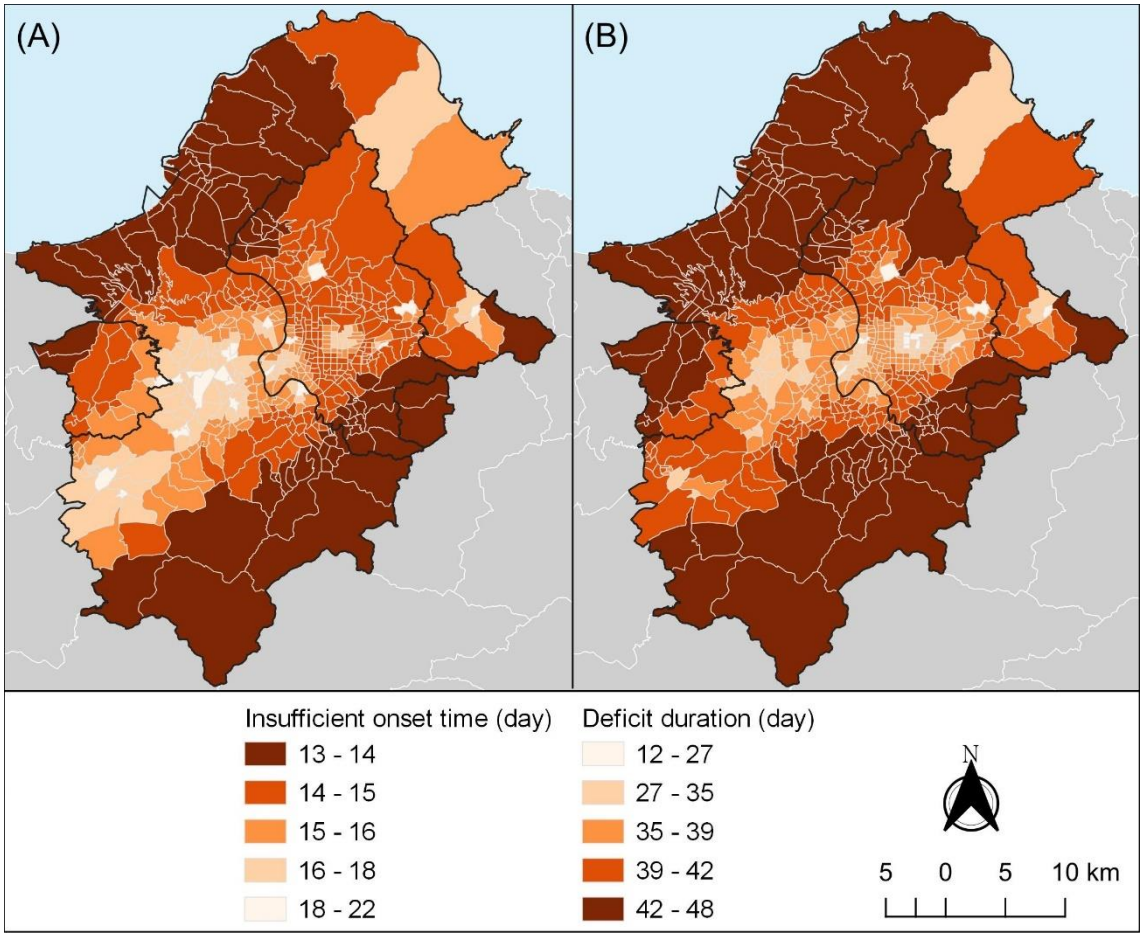


圖 20. 資源不足指標之空間分布

備註：（A）發生時間點、（B）持續時間的空間分布

doi:10.6342/NTU202101357

本論文進一步利用了兩項指標來總空間可及性時序變化的特性，分別為資源不足的發生時間點（Insufficient onset time）以及持續時間長度（Deficit duration）。圖 20 呈現了這兩項指標的空間分布，可以發現兩者呈現高度的負相關（皮爾森相關係數為-0.82）。因此，一個地區愈早出現資源不足的情形，通常也會經歷較久的資源不足時間，而且這些地區大多分布在研究區的東南角和西北角。雖然這些地區都屬於外圍地帶，但有些地區其實是人口密集區，例如淡水或是新店地區，凸顯了這些特別需要注意的幾個地區。因此，Epi-RA 模型可以反映出在疫情的每個階段下，資源不足和資源過剩的地區；這些分析結果將可提供給防疫單位作為未來醫療資源配置的參考，更有效率地使用和分配相關的醫療資源。這些資訊都能夠協助政府更理解傳染病疫情當中的資源供需問題，以及防疫措施是否會影響到民眾對於資源可及的程度，讓政府能夠擬定更合適的資源運用和防疫策略。

本論文的分析結果顯示在一個都會地區內，有些地方因為資源可及性過低，所以對於傳染病疫情是更加脆弱的。而這樣的地區有可能是人口稀疏區，也有可能是人口密集區。在人口稀疏區，因為人數較少所以對於醫療防疫資源的需求量也相對較少，但因為該區域的資源供給數量也相對較少，所以只要需求有稍微增加，就容易出現資源不足的狀況。另一方面，位於人口密集區的居民因為有較多的機會在公共場合與他人接觸，導致傳染病也比較容易在該地區蔓延 (Leung et al. 2017)，在短時間內造成大量的資源需求，超出了防疫單位的預期。因此，儘管人口密集區雖然比較容易獲取資源，仍然有可能出現資源不足的情形。這樣的發現也呼應了南韓針對此次新冠肺炎疫情的研究發現 (Yun et al. 2020)。

時空動態可及性的評估結果，同時顯示了資源不足的地區和資源較充沛的地區，這些資訊可以協助防疫單位進行資源配置的決策，盡可能地確保各個地區在疫情期間都能達到資源供需平衡的狀態。以本論文的分析結果為例，在疫情初期雖然市中心對於醫療資源的需求量較多，但並沒有出現資源不足的情形。相反地，外圍

doi:10.6342/NTU202101357

地區雖然需求量少卻很早就出現資源不足的情形。因此，防疫單位即能夠在疫情初期從市中心調配一些醫療防疫資源到外圍地區，對抗當地的疫情。類似的建議也曾在此次的新冠肺炎疫情中被提出來 (Ranney, Griffeth, and Jha 2020)。因此，Epi-RA 模式確實能夠協助找出資源不足和資源充沛的地區，協助資源調配任務的進行。另一方面，方艙醫院或是野戰醫院也是此次新冠肺炎疫情期間，經常被用來提升一個地區之醫療資源供給量的措施 (Arango 2020; Chen et al. 2020; Danguy des Déserts et al. 2020; Wallis et al. 2020)。透過 Epi-RA 模式，能夠協助防疫單位判斷以模擬的方式來事先判斷哪些地區在哪些時間點可能需要投放額外的醫療防疫資源，讓防疫單位能盡早準備相關的資源。

doi:10.6342/NTU202101357

第四章 討論

從理解到防止再到保護，針對這三個傳染病擴散的重要層面，本論文以地理計算的觀點開發了三項量性研究方法，藉此提供更為有效的疫情擴散資訊和防疫決策支援工具，強化人們對於傳染病防治的能力。實際案例應用分析的成果證實了三項方法的效用性。其中，MST-DBSCAN 能夠考量傳染病潛伏期的特性，自動判斷出傳染病群聚的各種擴散演化型態；HuMoRZ 利用人口移動的規律特性劃分移動管制分區，有效達到延緩傳染病擴散的目標；Epi-RA 整合疫情引起的需求變化，動態評估各地區對於醫療資源的可及性及其時空變化。本章節將分別討論三項方法各自的適用性，以及三者彼此之間的整合性，最後以研究限制的小節來提醒讀者這三項方法在使用上可能會遇到的限制。

4.1 方法適用性

本論文的三個獨立研究，採用了不同型態的資料和疾病對象作為案例應用來分別證明所開發之演算法或數理模式的效用性。此舉乃是基於三種方法本身的特性來選擇各自適用的分析對象。首先，MST-DBSCAN 的計算仰賴兩兩資料點之間的時間差距和地理距離等資訊，故其只適用於具備這兩項資訊的分析資料；以空間分析的學科領域而言，通常是點型態或是面型態的資料形式。例如本論文中的登革熱資料是以點型態呈現，故可以根據每個病例的居住地和發病時間來推算出兩兩病例之間的發病時間差距和地理距離，再進一步利用 MST-DBSCAN 來剖析疫情群聚的擴散演化型態。而除了登革熱這種屬於蚊媒傳染的疾病之外，目前也已有研究採用 MST-DBSCAN 來剖析瑞士日內瓦的新冠肺炎疫情擴散之時空特徵 (De Ridder et al. 2020)，證明此方法也適用於空氣或飛沫傳染病。如此多樣的適用性乃係因為 MST-DBSCAN 僅考量資料點之間的時間差距和地理距離這兩項因素，使得不同疾病的流行病學特性差異並不會造成方法適用性上的影響。除此之外，針對

doi:10.6342/NTU202101357

犯罪擴散 (Anselin et al. 2000) 或是創新擴散 (Feldman and Florida 1994) 等同樣具備上述兩項因素的社會現象，MST-DBSCAN 或許也能協助剖析這些現象的擴散時空特徵，擴大此方法的適用領域和貢獻。

第二，HuMoRZ 演算法並未直接針對疾病資料進行分析，而是利用人口移動的規律性進行防疫分區劃分。因此，在資料型態上，必須是以網絡（Network）的形式來呈現不同地區之間的人口移動往來數量，而這也是常見的人口移動資料的形式之一 (Gallotti and Barthelemy 2015)。同時，因為移動規律性為 HuMoRZ 的核心概念，故此方法較適用於會透過人與人之間的近距離接觸所傳播的疾病（空氣或飛沫傳染病），例如新冠肺炎、A 型流感等；較不適用於性傳播的疾病或是人畜共通傳染病（包含蚊媒傳染病），如愛滋病、登革熱、瘧疾等。基於本論文在分析實作之際，台灣本土的新冠肺炎疫情尚未爆發，故採用 A 型流感做為模擬分析對象，藉此驗證 HuMoRZ 所劃定之防疫分區是否能有效延緩傳染性疾病的擴散。再者，過往研究發現動物的移動行為也同樣具備規律性 (Fagan et al. 2013)，故 HuMoRZ 或許也能夠用於預防動物之間的傳染性疾病。

而除了用以劃定防疫分區以外，本論文認為 HuMoRZ 的分區結果也回應了人文地理學中的機能區域（Functional region）理論。一個機能區域通常被認定為類似一個生活圈並由具備不同土地利用型態的地區所組成，居民於此區域內便可獲取各種生活所需的資源 (Philbrick 1957)。如此自給自足的特性降低了跨區域移動的必要性，並在區域內部形塑出強烈的移動凝聚性 (Klapka and Halás 2016)。這樣的移動樣態與多樣化的土地利用組成使得一個機能區域在本質上與防疫分區的性質雷同。此外，過往研究也發現過大的機能區域會喪失內部移動的凝聚性；而過小的區域可能會導致無法達成自給自足 (Halás, Klapka, and Erlebach 2019)。如此不能過大也不能太小的特性，呼應了前述文獻對於防疫分區大小的建議 (Barbera et al. 2001)，故本論文認為 HuMoRZ 應該也可用於協助劃定機能區域或是生活圈的地理

doi:10.6342/NTU202101357

範圍，增加此方法的價值和論文的貢獻。

最後，Epi-RA 的主要目的在於協助評估隨著疫情發展而出現時空變化的醫療資源可及性，故此模式較適用於醫療資源數量有限、無法充分地供應所有民眾的情況。本論文係以新冠肺炎篩檢為疾病標的和資源對象進行案例應用分析，不過 Epi-RA 的核心概念在於由疫情引起的醫療資源需求變化，模式並未限定傳染病的類型或是必須具備特定的傳染病特性，所以其模型架構經過些微調整後，理當也能應用於登革熱、A 型流感等疾病，或者是未來的新興傳染病。另一方面，傳染病疫情的醫療資源並非只有篩檢試劑，諸如疫苗、重症加護病床、呼吸器、甚至是醫師和護理師人數等，都屬於醫療資源的範疇，故 Epi-RA 的模型架構也能協助評估各地區在一場傳染病疫情中對於這些醫療資源的可及性。

除了需求面會因為疫情發展而變化之外，供給層面以及移動阻抗層面也具備同樣的特性，進而影響到各地區對於醫療資源的可及性。面對新興傳染病突如其來的威脅，政府或防疫單位多半會調整平日的資源佈署模式來因應疫情中的各種突發狀況。例如將其他地區的物資調派到疫情較嚴重的區域 (Ranney, Griffeth, and Jha 2020)、增設篩檢站找出潛在感染源 (Cohen and Kupferschmidt 2020)、或是興建野戰醫院來緊急收容更多需要住院隔離的患者 (Chen et al. 2020)。因此，除了供給點的資源數量可能會隨著時間變動，供給點本身的位置也有可能改變遷移、甚至是增加供給地點的數量，這些狀況皆反映了醫療資源的供給隨著疫情的發展而變動的特性。另一方面，在過往的伊波拉病毒或是此次的新冠肺炎疫情中，多數國家皆採取封城或是宵禁等強烈的手段來防堵疫情擴散 (Peak et al. 2018; Wu, Leung, and Leung 2020)。在這種情形下，街區或是社區邊界上的移動管制措施便有可能阻斷了某個地區的民眾前往鄰近供給點索取資源的可能性，大幅增加了移動阻抗。換言之，即使資源供給點非常靠近該區域，也無法顯著提升該區域對於醫療資源的可及性。因此，一個地區對於醫療資源的可及性確實有可能受到政府因應疫情發展所採

doi:10.6342/NTU202101357

取的防疫措施而影響。基於上述原因，若是要更細緻地量測疫情中各地區對於醫療資源的動態可及性，必須將需求、供給和移動阻抗等三個層面的影響效果同時納入考量。目前的 Epi-RA 已經考量了需求面變化的影響，故本論文認為其模型架構經過適當的擴充和調整後應該可以達成此任務，反映出 Epi-RA 的潛在價值。

本論文的三個實際案例應用分析皆是以單一城市為尺度來分別展示三項方法的適用性，其中，若是要將 HuMoRZ 和 Epi-RA 應用於跨縣市或是國家等空間尺度，則最大的阻礙即在於跨縣市或是全國人口移動資料的取得。另一方面，HuMoRZ 的核心精神係人口移動於日常生活中的規律性，目前文獻上也僅有發現城市尺度的移動規律性。故若是要將 HuMoRZ 應用到跨縣市或是全國的尺度，可能必須先證實這種空間尺度下是否還存在日常生活的移動規律性，才能確保使用移動規律性進行移動管制分區劃定的合理性。

4.2 方法整合性

本論文所開發的三項方法雖然各自有不同的適用狀況，但是當這三者的適用狀況同時出現時，便能夠整合起來形成一套完整的防疫決策支援分析模式。以此次的新冠肺炎疫情為例，不同病例之間即具備了發病時間差和地理空間鄰近的特性，適合使用 MST-DBSCAN 來剖析疫情擴散的時空過程。而新冠肺炎的傳播途徑是人與人近距離接觸的飛沫傳染，多數國家皆採取移動管制的方式來降低人群接觸群聚的機會，故 HuMoRZ 可以協助劃定移動管制的分區。最後，由於新冠肺炎是一種新興的傳染病，在疫情爆發之初，尚未有足夠的疫苗讓大眾擁有抗體或是足夠的醫療資源來治療已經染疫的民眾；而 Epi-RA 即是協助評估各地區在這種醫療資源不足的狀況下對於資源之可及性的變化。因此，若是一種傳染病類似於新冠肺炎的疾病特徵和狀況，即可整合本論文的三項方法來支援防疫決策。然而，整合此三項方法必須同時具備細緻的病例點位和發病日期、研究區域內的人口移動和土地

doi:10.6342/NTU202101357

使用狀況、以及醫療資源供給的數量分布等資料。以台灣目前相關的開放資料而言，尚未有合適的場域和疾病來展示這三項方法實際上的整合形式以及具體的分析成果，所以本論文以討論的形式來說明可能的整合方式。

此三項方法的整合應用可以分別從時間和空間的維度來理解。就時間維度而言，當確診病例出現時，防疫單位可以率先使用 HuMoRZ 來劃定移動管制分區，協助判斷那些高風險地區需要與已經出現疫情的地區一併實施一定程度的移動管制，在第一時間盡可能地阻止疫情向外擴散。接著，利用 MST-DBSCAN 根據現有的病例資料來剖析疫情群聚的位置及其擴散演化型態，理解疫情的時空擴散過程，協助掌握疫情的發展現況和判斷未來的發展趨勢，並從中研判 Epi-RA 模式中與傳染病擴散特徵相關之參數的可能數值。最後，利用 Epi-RA 進行事先模擬來判斷哪些地區可能會在疫情大規模爆發之際最容易面臨到醫療資源不足的狀況，作為物資調度和分配的參考依據。

不僅在疫情初期而已，此三項方法在疫情的中後期依然能繼續發揮防疫決策支援的功用，協助滾動式調整已經實施的防疫措施來因應疫情的最新發展。當確診病例的數量隨疫情發展而累積，MST-DBSCAN 便有更豐富的資料來細緻剖析疫情的擴散演化，協助調整防疫單位對於疫情趨勢的認知和研判。除此之外，施行移動管制之後勢必會改變當地人口的移動特性，而人口移動特性的轉變又會改變疫情擴散的樣態 (Chan and Wen 2021)。因此，防疫單位或許能在疫情的不同階段使用 HuMoRZ 分析當下的人口移動資料，滾動式調整移動管制分區的劃分。另一方面，在疫情初期所獲得的 Epi-RA 模擬結果，可以在疫情中期與實際狀況進行比對，藉此來滾動調整模式中的各項參數，使得模擬結果能夠更接近實際狀況，並讓防疫單位對於各地區未來的醫療資源可及性有更精確的評估。因此，此三項方法在時間維度上的整合有助於防疫單位在疫情的各個階段因時制宜調整防疫策略。

而在空間維度上，此三項方法所關注的地理範圍大小呈現類似同心圓狀的向

doi:10.6342/NTU202101357

外擴張關係。首先，由於 MST-DBSCNA 是針對現有病例資料進行分析，故其所關注的地區都是已經被感染的地區，屬於同心圓的最內層。接著，HuMoRZ 旨在找出尚未有病例出現的高風險地區，意即與感染地區有密切人流往來的地區，並將這些地區劃定成一個分區，故其所關注的地理空間範圍比 MST-DBSCAN 更往外擴張了一些。最後，Epi-RA 是全盤檢視研究區內每個地區對於醫療資源的可及性在疫情當中的時序變化，並判斷出最脆弱的地區，也就是供不應求最嚴重的地區。由於脆弱地區不見得只會出現在已感染區或是高風險區，故 Epi-RA 所關注的範圍擴及了整個研究區。此三項方法在空間維度上分別關注不同性質和不同地理空間範圍大小的地區，這三者的整合有助於防疫單位因地制宜每個地區合適的防疫策略。

上述的整合方式在本質上還是將三項方法分開來單獨進行分析，只是將分析結果彙整起來形成完整的防疫決策支援的相關資訊；因此，後續研究或許可以朝向直接整合此三項方法的方向進行發展。首先，Epi-RA 模式單純假設每個時間點下的醫療資源需求是等比例於當下的染病人數，然而，Jun, Yoo, and Lee (2021)指出民眾尋求醫療資源的積極程度取決於民眾對於疫情狀況的認知以及危機意識感。換言之，當疫情群聚開始增強或衰退，皆會改變當地民眾對於資源尋求的積極性，而 MST-DBSCAN 針對疫情群聚的型態分類正好可以區辨每個疫情群聚的擴散演化強度是逐漸增強還是逐漸衰退，故其分析結果或許可以用以反映更實際的由疫情所引起的需求變化。另一方面，前述提到移動管制會影響到跨地區之間的移動成本，進而影響到醫療資源可及性的評估結果，而 HuMoRZ 正好提供了移動管制分區劃定的基礎。綜合上述，本論文建議後續研究可嘗試將 MST-DBSCAN 和 HuMoRZ 的分析結果整合到 Epi-RA 模式當中，模擬在 HuMoRZ 的分區劃定之情境下，由 MST-DBSCAN 所反映出的需求變化對於醫療資源可及性的時空變動狀況。

doi:10.6342/NTU202101357

4.3 研究限制

本論文所發展之演算法和數理模式，雖然能夠有效協助判斷疫情群聚擴散的型態、劃定移動管制分區並動態評估各地區在疫情期間對於醫療資源可及性的變化，依舊存在幾項缺點有待後續研究精進。首先，MST-DBSCAN 演算法僅關注群聚擴散的型態，並未區分不同群聚之間的強弱差異。例如：一個地理範圍大的群聚可能內部的病例數分布很稀疏，反而是一個地理範圍小的群聚可能存在較高的病例集中程度。這種特性其實是 DBSCAN 這種以密度為基準（Density-based）的演算法之本質，過往也有研究嘗試進行改良 (Liu, Zhou, and Wu 2007; Ram et al. 2009)。然而，本論文係以疫情群聚的中心點和範圍大小為擴散型態的判斷線索，群聚內部的病例密集程度並非用以判斷的線索之一，故 MST-DBSCAN 仍然能夠有效分辨出本論文所區分的各種擴散型態。不過，後續研究若是能夠進一步將病例密集程度也納入考量，相信能夠更細緻地理解不同擴散型態的意義。

第二，本論文的 HuMoRZ 演算法雖然考量了人口移動的規律性進行分區劃定，卻僅是以地區到地區之間這種總體人口樣態的形式來量測移動行為的規律性，並未考量到不同社會經濟族群可能具備不同的移動行為和特徵。例如過往研究發現女性通常因為肩負照顧家庭和小孩的角色，所以日常的生活移動範圍會比男性來得小 (Verhetsel, Beckers, and De Meyere 2018)；收入程度較低的族群可能因為無法負擔長途的移動成本或費用，所以其生活移動範圍也多半集中在住家附近 (Casado-Díaz 2000)。因此，後續研究可以嘗試朝此方向進行 HuMoRZ 演算法的改良，藉此來進一步劃定更有效的移動管制分區，協助防疫。

再者，不同社會經地位也影響著人們對於資源的可及程度 (Kim et al. 2020)，所以將此影響效果整合至本論文的 Epi-RA 模式也是值得後續研究發展的方向之一。同時，不同年齡層也存在不同的感染機率和疫情風險，對於資源的需求程度也

doi:10.6342/NTU202101357

不盡相同 (El Bcheraoui et al. 2018; Horner et al. 2015)，若是將其整合至 Epi-RA 模式中，將能夠剖析出更細緻的醫療資源可及性之時空變化過程。目前已有文獻開發出考量年齡結構的 SEIR 模式 (Tsai et al. 2011) 和 2SFCA 模式 (Hashtarkhani et al. 2020)，故將年齡結構的特性整合至 Epi-RA 模式中應當是可行的。

最後，人際關係網絡往往與疫情傳播有著密不可分的關聯性，因為其能夠反映出不同個體之間接觸互動的頻率和強度 (Keeling and Eames 2005)。本論文所開發的三項方法並未納入人際關係網絡，主要原因係本論文的主軸在於檢視文獻上關於傳染病擴散的理解、防止和保護等三個層面的量性研究方法，並從地理計算的觀點加以改良。有鑑於若要以地理空間的形式來呈現人際關係網絡，必須進行一定程度的加總和概略化；而在加總的過程中勢必會犧牲掉部分有關於個人的資訊，進而犧牲掉接觸關係與疫情傳播之間最直接的關聯性，導致失去了採用人際關係網絡進行分析的初衷。雖然沒有直接納入人際關係網絡，但本論文的 HuMoRZ 和 Epi-RA 在分析上皆必須使用到人口移動網絡資料來反映移動規律性或是個體之間的平均接觸頻率等現象，所以在某種程度上應該也是反映了一部份的人際關係網絡所造成的效果。而若是後續研究實際將人際關係因素整合到本論文的三項方法之中，必定能夠獲得更貼近現實狀況的分析結果。

doi:10.6342/NTU202101357

第五章 結論

理解傳染病擴散的時空樣態、防止傳染病擴散的措施、以及定位出資源可及性較低的地區以便保護該區域的民眾，這三者都是傳染病防疫決策中的重要面向。透過文獻爬梳，本論文發現這三個面向都存在方法學上的不足。因此，本論文分別發展了 MST-DBSCAN 和 HuMoRZ 演算法、以及 Epi-RA 模式，嘗試在方法學上做出貢獻。實際案例應用的分析結果顯示，MST-DBSCAN 確實能夠考量潛伏期的特性，剖析出傳染病群聚的各種擴散型態；HuMoRZ 能夠考量人口移動的規律特性來劃分出有效的移動管制分區，達到延緩傳染病擴散的目的；Epi-RA 能夠考量到疫情引起的需求變化，評估出資源可及性的時空動態。這三項方法能夠協助防疫單位在疫情初期更徹底理解傳染病擴散的時空樣態，判斷哪些人口特徵或是環境特徵可能會是造成疫情加劇的原因，做為後續防疫工作的參考資訊。同時，搭配移動管制分區進行疫情管控，雖然必須犧牲一定程度的人身自由和經濟損失，卻能夠有效地將疫情控制在一定的地理範圍內，延緩擴散的速度。而在延緩疫情擴散的同時，搭配模擬的方式來預先判斷哪些地區可能會在未來出現資源不足的情形，利用移動管制分區所爭取到的時間，讓防疫單位籌措額外的醫療防疫物資及佈署防疫資源於這些高風險地區，並執行相對應的防疫措施來對抗疫情。除了應用於傳染病擴散的領域之外，本論文所開發之方法也極可能適用於其他相關領域。MST-DBSCAN 可用於分析同樣具備時空擴散特徵的社會現象，例如犯罪或是資訊流通等；HuMoRZ 所劃定的移動管制分區與地理學中的機能區域有異曲同工之妙，都是反映出生活圈特性的地理範圍；Epi-RA 可以協助評估各地區其他醫療資源甚或是民生物資的可及性，使政府能全面整備各種物資的籌措和配置。因此，本論文的研究成果除了具有學術上的貢獻之外，相信也具備實務的應用價值來對抗傳染病擴散，以及協助分析相關的地理現象。

doi:10.6342/NTU202101357

參考文獻

Anselin, L., J. Cohen, D. Cook, W. Gorr, and G. Tita. 2000. Spatial Analyses of Crime. *Criminal Justice* 4 (2):213-262.

Arango, C. 2020. Lessons learned from the coronavirus health crisis in Madrid, Spain: How COVID-19 has changed our lives in the last 2 weeks. *Biological Psychiatry* 88 (7):e33-e34.

Auzan, A. A. 2020. The economy under the pandemic and afterwards. *Population and Economics* 4 (2):4-12.

Bailey, T. C., and A. C. Gatrell. 1995. *Interactive spatial data analysis*. Essex: Longman Scientific & Technical.

Bajardi, P., C. Poletto, J. J. Ramasco, M. Tizzoni, V. Colizza, and A. Vespignani. 2011. Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. *PloS one* 6 (1):e16591.

Barbera, J., A. Macintyre, L. Gostin, T. Inglesby, T. O'Toole, C. DeAtley, K. Tonat, and M. Layton. 2001. Large-scale quarantine following biological terrorism in the United States: Scientific examination, logistic and legal limits, and possible consequences. *The Journal of the American Medical Association* 286 (21):2711-2717.

Barreto, F. R., M. G. Teixeira, N. Costa Maria da Conceição, M. S. Carvalho, and M. L. Barreto. 2008. Spread pattern of the first dengue epidemic in the city of Salvador, Brazil. *BioMed Central Public Health* 8:51.

Bastos, M. L., G. Tavaziva, S. K. Abidi, J. R. Campbell, L.-P. Haraoui, J. C. Johnston, Z. Lan, S. Law, E. MacLean, A. Trajman, D. Menzies, A. Benedetti, and F. A. Khan. 2020. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. *BMJ* 370: m2516.

doi:10.6342/NTU202101357

Belik, V., T. Geisel, and D. Brockmann. 2011. Natural human mobility patterns and spatial spread of infectious diseases. *Physical Review X* 1:011001.

Birant, D., and A. Kut. 2007. ST-DBSCAN: An algorithm for clustering spatial–temporal data. *Data & Knowledge Engineering* 60 (1):208-221.

Buzza, C., S. S. Ono, C. Turvey, S. Wittrock, M. Noble, G. Reddy, P. J. Kaboli, and H. S. Reisinger. 2011. Distance is relative: Unpacking a principal barrier in rural healthcare. *Journal of General Internal Medicine* 26 (2):648-654.

Casado-Díaz, J. M. 2000. Local labour market areas in Spain: A case study. *Regional Studies* 34 (9):843-856.

Chan, C.-H., and T.-H. Wen. 2021. Revisiting the Effects of High-Speed Railway Transfers in the Early COVID-19 Cross-Province Transmission in Mainland China. *International Journal of Environmental Research and Public Health* 18 (12):6394.

Chan, M., and M. A. Johansson. 2012. The Incubation Periods of Dengue Viruses. *PLOS ONE* 7 (11):e50972.

Chen, S., Z. Zhang, J. Yang, J. Wang, X. Zhai, T. Bärnighausen, and C. Wang. 2020. Fangcang shelter hospitals: A novel concept for responding to public health emergencies. *The Lancet* 395 (10232):1305-1314.

Cliff, A. D., and P. Haggett. 2006. A swash–backwash model of the single epidemic wave. *Journal of Geographical Systems* 8 (3):227-252.

Cliff, A. D., P. Haggett, J. K. Ord, and G. R. Versey. 1981. *Spatial Diffusion: An Historical Geography of Epidemics in an Island Community*: CUP Archive.

Cohen, J., and K. Kupferschmidt. 2020. Countries test tactics in ‘war’against COVID-19. *Science* 367 (6484):1287-1288.

Coltart, C. E. M., B. Lindsey, I. Ghinai, A. M. Johnson, and D. L. Heymann. 2017. The Ebola outbreak, 2013–2016: Old lessons for new epidemics.

doi:10.6342/NTU202101357

Philosophical Transactions of the Royal Society B: Biological Sciences 372 (1721):20160297.

Cooper, B. S., R. J. Pitman, W. J. Edmunds, and N. J. Gay. 2006. Delaying the international spread of pandemic influenza. *PLOS Medicine* 3 (6):e212.

Dai, D., and F. Wang. 2011. Geographic disparities in accessibility to food stores in southwest Mississippi. *Environment and Planning B: Planning and Design* 38 (4):659-677.

Danguy des Déserts, M., Q. Mathais, A. Luft, J. Escarment, and P. Pasquier. 2020. Conception and deployment of a 30-bed field military intensive care hospital in Eastern France during the 2020 COVID-19 pandemic. *Anaesthesia, Critical Care and Pain Medicine* 39 (3):361-362.

Day, T., A. Park, N. Madras, A. Gumel, and J. Wu. 2006. When is quarantine a useful control strategy for emerging infectious diseases? *American journal of Epidemiology* 163 (5):479-485.

De Ridder, D., J. Sandoval, N. Vuilleumier, S. Stringhini, H. Spechbach, S. Joost, L. Kaiser, and I. Guessous. 2020. Geospatial digital monitoring of COVID-19 cases at high spatiotemporal resolution. *The Lancet Digital Health* 2 (8):e393-e394.

Delamater, P. L. 2013. Spatial accessibility in suboptimally configured health care systems: A modified two-step floating catchment area (M2SFCA) metric. *Health & Place* 24:30-43.

Desjardins, M. R., A. Hohl, and E. M. Delmelle. 2020. Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. *Applied Geography* 118:102202.

Dony, C. C., E. M. Delmelle, and E. C. Delmelle. 2015. Re-conceptualizing accessibility to parks in multi-modal cities: A Variable-width Floating

doi:10.6342/NTU202101357

Catchment Area (VFCA) method. *Landscape and Urban Planning* 143:90-99.

El Bcheraoui, C., A. H. Mokdad, L. Dwyer-Lindgren, A. Bertozzi-Villa, R. W. Stubbs, C. Morozoff, S. Shirude, M. Naghavi, and C. J. L. Murray. 2018. Trends and patterns of differences in infectious disease mortality among US counties, 1980-2014. *JAMA* 319 (12):1248-1260.

Epstein, J. M., D. M. Goedecke, F. Yu, R. J. Morris, D. K. Wagener, and G. V. Bobashev. 2007. Controlling pandemic flu: The value of international air travel restrictions. *PlOS ONE* 2 (5):e401.

Ester, M., H.-P. Kriegel, J. Sander, M. Wimmer, and X. Xu. 1998. Incremental Clustering for Mining in a Data Warehousing Environment. Paper read at Very Large Data Base, at New York, USA.

Ester, M., H.-P. Kriegel, J. Sander, and X. Xu. 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Paper read at Knowledge Discovery and Data Mining.

Evans, J. P. 2011. Resilience, ecology and adaptation in the experimental city. *Transactions of the Institute of British Geographers* 36 (2):223-237.

Fagan, W. F., M. A. Lewis, M. Auger-Méthé, T. Avgar, S. Benhamou, G. Breed, L. LaDage, U. E. Schlägel, W. w. Tang, Y. P. Papastamatiou, J. Forester, and T. Mueller. 2013. Spatial memory and animal movement. *Ecology Letters* 16 (10):1316-1329.

Feldman, M. P., and R. Florida. 1994. The Geographic Sources of Innovation: Technological Infrastructure and Product Innovation in the United States. *Annals of the Association of American Geographers* 84 (2):210-229.

Fernandez, M. A. L., M. Schomaker, P. R. Mason, J. F. Fesselet, Y. Baudot, A. Boulle, and P. Maes. 2012. Elevation and cholera: an epidemiological spatial analysis of the cholera epidemic in Harare, Zimbabwe, 2008-2009. *BMC*

doi:10.6342/NTU202101357

Public Health 12:442.

Fong, M. W., H. Gao, J. Y. Wong, J. Xiao, E. Shiu, Y. C., S. Ryu, and B. J. Cowling. 2020. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—Social distancing measures. *Emerging Infectious Disease* 26 (5):976-984.

Fèvre, E. M., B. M. d. C. Bronsvoort, K. A. Hamilton, and S. Cleaveland. 2006. Animal movements and the spread of infectious diseases. *Trends in Microbiology* 14 (3):125-131.

Fransen, K., T. Neutens, P. De Maeyer, and G. Deruyter. 2015. A commuter-based two-step floating catchment area method for measuring spatial accessibility of daycare centers. *Health & Place* 32:65-73.

Fraser, C., S. Riley, R. M. Anderson, and N. M. Ferguson. 2004. Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences* 101 (16):6146-6151.

Gallotti, R., and M. Barthelemy. 2015. The multilayer temporal network of public transport in Great Britain. *Scientific Data* 2:140056.

Gatrell, A. C., T. C. Bailey, P. J. Diggle, and B. S. Rowlingson. 1996. Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology. *Transactions of the Institute of British Geographers* 21 (1):256-274.

Gatto, M., E. Bertuzzo, L. Mari, S. Miccoli, L. Carraro, R. Casagrandi, and A. Rinaldo. 2020. Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proceedings of the National Academy of Sciences* 117 (19):10484-10491.

Gensini, G. F., M. H. Yacoub, and A. A. Conti. 2004. The concept of quarantine in history: From plague to SARS. *Journal of Infection* 49 (4):257-261.

González, M. C., C. A. Hidalgo, and A.-L. Barabási. 2008. Understanding individual

doi:10.6342/NTU202101357

human mobility patterns. *Nature* 453 (7196):779-782.

Grenfell, B. T., O. N. Bjørnstad, and J. Kappey. 2001. Travelling waves and spatial hierarchies in measles epidemics. *Nature* 414 (6865):716-723.

Guagliardo, M. F. 2004. Spatial accessibility of primary care: concepts, methods and challenges. *International Journal of Health Geographics* 3:3.

Guida, C., and G. Carpentieri. 2021. Quality of life in the urban environment and primary health services for the elderly during the Covid-19 pandemic: An application to the city of Milan (Italy). *Cities* 110:103038.

Haggett, P., A. D. Cliff, and A. Frey. 1977. *Locational analysis in human geography*: Arnold London.

Halás, M., P. Klapka, and M. Erlebach. 2019. Unveiling spatial uncertainty: A method to evaluate the fuzzy nature of functional regions. *Regional Studies* 53 (7):1029-1041.

Hashtarkhani, S., B. Kiani, R. Bergquist, N. Bagheri, R. VafaeiNejad, and M. Tara. 2020. An age-integrated approach to improve measurement of potential spatial accessibility to emergency medical services for urban areas. *The International Journal of Health Planning and Management* 35 (3):788-798.

He, X., E. H. Y. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, X. Mo, Y. Chen, B. Liao, W. Chen, F. Hu, Q. Zhang, M. Zhong, Y. Wu, L. Zhao, F. Zhang, B. J. Cowling, F. Li, and G. M. Leung. 2020. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine* 26 (5):672-675.

Horner, M. W., M. D. Duncan, B. S. Wood, Y. Valdez-Torres, and C. Stansbury. 2015. Do aging populations have differential accessibility to activities? Analyzing the spatial structure of social, professional, and business opportunities. *Travel Behaviour and Society* 2 (3):182-191.

doi:10.6342/NTU202101357

Hu, L., C. Zhao, M. Wang, S. Su, M. Weng, and W. Wang. 2020. Dynamic healthy food accessibility in a rapidly urbanizing metropolitan area: Socioeconomic inequality and relative contribution of local factors. *Cities* 105:102819.

Indolfi, C., and C. Spaccarotella. 2020. The outbreak of COVID-19 in Italy: Fighting the pandemic. *Journal of the American College of Cardiology: Case Reports* 2 (9):1414-1418.

Jamtsho, S., R. Corner, and A. Dewan. 2015. Spatio-temporal analysis of spatial accessibility to primary health care in Bhutan. *ISPRS International Journal of Geo-Information* 4 (3):1584-1604.

Ji, Y., Z. Ma, M. P. Peppelenbosch, and Q. Pan. 2020. Potential association between COVID-19 mortality and health-care resource availability. *The Lancet Global Health* 8 (4):e480.

Jin, M., L. Liu, D. Tong, Y. Gong, and Y. Liu. 2019. Evaluating the spatial accessibility and distribution balance of multi-level medical service facilities. *International Journal of Environmental Research and Public Health* 16 (7):1150.

Joseph, A. E., and P. R. Bantock. 1982. Measuring potential physical accessibility to general practitioners in rural areas: a method and case study. *Social Science & Medicine* 16 (1):85-90.

Jun, S.-P., H. S. Yoo, and J.-S. Lee. 2021. The impact of the pandemic declaration on public awareness and behavior: Focusing on COVID-19 google searches. *Technological Forecasting and Social Change* 166:120592.

Kalnis, P., N. Mamoulis, and S. Bakiras. 2005. On Discovering Moving Clusters in Spatio-temporal Data. In *International symposium on spatial and temporal databases* 364-381. Springer, Berlin, Heidelberg.

Kan, C.-C., P.-F. Lee, T.-H. Wen, D.-Y. Chao, M.-H. Wu, N. H. Lin, S. Y.-J. Huang,

doi:10.6342/NTU202101357

C.-S. Shang, I.-C. Fan, P.-Y. Shu, J.-H. Huang, C.-C. King, and P. Lu. 2008. Two Clustering Diffusion Patterns Identified from the 2001–2003 Dengue Epidemic, Kaohsiung, Taiwan. *The American Journal of Tropical Medicine and Hygiene* 79 (3):344-352.

Keeling, M. J., and K. T. D. Eames. 2005. Networks and epidemic models. *Journal of The Royal Society Interface* 2 (4):295-307.

Khan, A. A. 1992. An integrated approach to measuring potential spatial access to health care services. *Socio-Economic Planning Sciences* 26 (4):275-287.

Kim, H., D. Kim, C. Paul, and C. K. Lee. 2020. The spatial allocation of hospitals with negative pressure isolation rooms in Korea: Are we prepared for new outbreaks? *International Journal of Health Policy and Management* 9 (11):475-483.

Klapka, P., and M. Halás. 2016. Conceptualising patterns of spatial flows: Five decades of advances in the definition and use of functional regions. *Moravian Geographical Reports* 24 (2):2-11.

Kuo, C.-L., and H. Fukui. 2007. Geographical structures and the cholera epidemic in modern Japan: Fukushima prefecture in 1882 and 1895. *International Journal of Health Geographics* 6:25.

Kupferschmidt, K., and J. Cohen. 2020. Can China's COVID-19 strategy work elsewhere? *Science* 367 (6482):1061-1062.

Kwan, M. P. 1998. Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis* 30 (3):191-216.

Lai, P. C., C. M. Wong, A. J. Hedley, S. V. Lo, P. Y. Leung, J. Kong, and G. M. Leung. 2004. Understanding the spatial clustering of severe acute respiratory syndrome (SARS) in Hong Kong. *Environmental Health Perspectives* 112

doi:10.6342/NTU202101357

(15):1550-1556.

Lee, G.-J., S.-I. Pak, K.-N. Lee, and S. Hong. 2019. Movement-based biosecurity zones for control of highly infectious animal diseases: Application of community detection analysis to a livestock vehicle movement network. *Sustainability* 11 (6):1642.

Leung, K., M. Jit, E. H. Lau, and J. T. Wu. 2017. Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong. *Scientific Reports* 7:7974.

Lian, M., R. D. Warner, J. L. Alexander, and K. R. Dixon. 2007. Using geographic information systems and spatial and space-time scan statistics for a population-based risk analysis of the 2002 equine West Nile epidemic in six contiguous regions of Texas. *International Journal of Health Geographics* 6:42.

Lima, A., M. De Domenico, V. Pejovic, and M. Musolesi. 2015. Disease containment strategies based on mobility and information dissemination. *Scientific Reports* 5:10650.

Liu, P., D. Zhou, and N. Wu. 2007. VDBSCAN: Varied Density Based Spatial Clustering of Applications with Noise. In *2007 International Conference on Service Systems and Service Management*.

Luo, J. 2014. Integrating the Huff model and floating catchment area methods to analyze spatial access to healthcare services. *Transactions in GIS* 18 (3):436-448.

Luo, W., and Y. Qi. 2009. An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health & Place* 15 (4):1100-1107.

Luo, W., and F. Wang. 2003. Measures of spatial accessibility to health care in a GIS

doi:10.6342/NTU202101357

environment: synthesis and a case study in the Chicago region. *Environment and Planning B: Planning and Design* 30 (6):865-884.

Luo, W., and T. Whippo. 2012. Variable catchment sizes for the two-step floating catchment area (2SFCA) method. *Health & Place* 18 (4):789-795.

Ma, L., N. Luo, T. Wan, C. Hu, and M. Peng. 2018a. An improved healthcare accessibility measure considering the temporal dimension and population demand of different ages. *International Journal of Environmental Research and Public Health* 15 (11):2421.

Mao, L. 2014. Modeling triple-diffusions of infectious diseases, information, and preventive behaviors through a metropolitan social network—An agent-based simulation. *Applied Geography* 50:31-39.

McGrail, M. R. 2012. Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements. *International Journal of Health Geographics* 11 (1):50.

McGrail, M. R., and J. S. Humphreys. 2014. Measuring spatial accessibility to primary health care services: Utilising dynamic catchment sizes. *Applied Geography* 54:182-188.

Meade, M. S., and M. Emch. 2010. *Medical Geography*: Guilford Press.

Paez, A., C. D. Higgins, and S. F. Vivona. 2019. Demand and level of service inflation in Floating Catchment Area (FCA) methods. *PLOS ONE* 14 (6): e0218773.

Parodi, S. M., and V. X. Liu. 2020. From containment to mitigation of COVID-19 in the U.S. *The Journal of the American Medical Association* 323 (15):1441-1442.

Peak, C. M., R. Kahn, Y. H. Grad, L. M. Childs, R. Li, M. Lipsitch, and C. O. Buckee. 2020. Individual quarantine versus active monitoring of contacts for

doi:10.6342/NTU202101357

the mitigation of COVID-19: A modelling study. *The Lancet Infectious Diseases* 20 (9):1025-1033.

Peak, C. M., A. Wesolowski, E. zu Erbach-Schoenberg, A. J. Tatem, E. Wetter, X. Lu, D. Power, E. Weidman-Grunewald, S. Ramos, S. Moritz, C. O. Buckee, and L. Bengtsoon. 2018. Population mobility reductions associated with travel restrictions during the Ebola epidemic in Sierra Leone: Use of mobile phone data. *International Journal of Epidemiology* 47 (5):1562-1570.

Pereira, R. H. M., C. K. V. Braga, L. M. Servo, B. Serra, P. Amaral, N. Gouveia, and A. Paez. 2021. Geographic access to COVID-19 healthcare in Brazil using a balanced float catchment area approach. *Social Science & Medicine* 273:113773.

Philbrick, A. K. 1957. Principles of areal functional organization in regional human geography. *Economic Geography* 33 (4):299-336.

Ram, A., A. Sharma, A. S. Jalal, R. Singh, and A. Agrawal. 2009. An Enhanced Density Based Spatial Clustering of Applications with Noise. In *2009 IEEE International Advance Computing Conference (IACC 2009)*.

Ranney, M. L., V. Griffeth, and A. K. Jha. 2020. Critical supply shortages—The need for ventilators and personal protective equipment during the Covid-19 pandemic. *New England Journal of Medicine* 382:e41.

Rosvall, M., and C. T. Bergstrom. 2008. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences* 105 (4):1118-1123.

Rubin, G. J., and S. Wessely. 2020. The psychological effects of quarantining a city. *The British Medical Journal* 368:m313.

Sabel, C. E., A. C. Gatrell, M. Löytönen, P. Maasilta, and M. Jokelainen. 2000. Modelling exposure opportunities: estimating relative risk for motor neurone

doi:10.6342/NTU202101357

disease in Finland. *Social Science & Medicine* 50 (7-8):1121-1137.

Sabel, C. E., D. Pringle, and A. Schoerstrom. 2009. Infectious Disease Diffusion. In *A Companion to Health and Medical Geography*, eds. T. Brown, S. McLafferty and G. Moon, 111-132: Wiley-Blackwell.

Schneider, C. M., V. Belik, T. Couronné, Z. Smoreda, and M. C. González. 2013. Unravelling daily human mobility motifs. *Journal of The Royal Society Interface* 10 (84): 20130246.

Shao, P. 2020. Impact of city and residential unit lockdowns on prevention and control of COVID-19. *medRxiv*.

Sharfstein, J. M., S. J. Becker, and M. M. Mello. 2020. Diagnostic testing for the novel coronavirus. *JAMA* 323 (15):1437-1438.

Sun, L., K. W. Axhausen, D.-H. Lee, and X. Huang. 2013. Understanding metropolitan patterns of daily encounters. *Proceedings of the National Academy of Sciences* 110 (34):13774-13779.

Tanser, F., B. Gijsbertsen, and K. Herbst. 2006. Modelling and understanding primary health care accessibility and utilization in rural South Africa: An exploration using a geographical information system. *Social Science & Medicine* 63 (3):691-705.

Tao, R., J. Downs, T. M. Beckie, Y. Chen, and W. McNelley. 2020. Examining spatial accessibility to COVID-19 testing sites in Florida. *Annals of GIS* 26 (4):319-327.

Tsai, Y.-S., C.-Y. Huang, T.-H. Wen, C.-T. Sun, and M.-Y. Yen. 2011. Integrating epidemic dynamics with daily commuting networks: Building a multilayer framework to assess influenza A (H1N1) intervention policies. *Simulation* 87 (5):385-405.

Velthuis, A. G. J., and M. C. M. Mourits. 2007. Effectiveness of movement-

doi:10.6342/NTU202101357

prevention regulations to reduce the spread of foot-and-mouth disease in The Netherlands. *Preventive Veterinary Medicine* 82 (3-4):262-281.

Verhetsel, A., J. Beckers, and M. De Meyere. 2018. Assessing daily urban systems: A heterogeneous commuting network approach. *Networks and Spatial Economics* 18:633-656.

Wallis, N., C. Gust, E. Porter, N. Gilchrist, and A. Amaral. 2020. Implementation of field hospital pharmacy services during the COVID-19 pandemic. *American Journal of Health-System Pharmacy* 77 (19):1547-1551.

Wan, N., B. Zou, and T. Sternberg. 2012. A three-step floating catchment area method for analyzing spatial access to health services. *International Journal of Geographical Information Science* 26 (6):1073-1089.

Wang, F. 2012. Measurement, optimization, and impact of health care accessibility: A methodological review. *Annals of the association of American Geographers* 102 (5):1104-1112.

————. 2020. From 2SFCA to i2SFCA: integration, derivation and validation. *International Journal of Geographical Information Science* 35 (3):628-638.

Wang, M., A. Wang, and A. Li. 2006. Mining Spatial-temporal Clusters from Geo-databases. In *Advanced Data Mining and Applications*, Xi'an, China.

Watve, M. G., and M. M. Jog. 1997. Epidemic diseases and host clustering: an optimum cluster size ensures maximum survival. *Journal of Theoretical Biology* 184 (2):165-169.

Weissman, G. E., A. Crane-Droesch, C. Chivers, T. Luong, A. Hanish, M. Z. Levy, J. Lubken, M. Becker, M. E. Draugelis, G. L. Anesi, P. J. Brennan, J. D. Christie, C. W. Hanson, M. E. Mikkelsen, and S. D. Halpern. 2020. Locally informed simulation to predict hospital capacity needs during the COVID-19

doi:10.6342/NTU202101357

pandemic. *Annals of Internal Medicine* 173 (1):21-28.

Wen, T.-H., and Y.-S. Tsai. 2015. Analyzing the Patterns of Space-Time Distances for Tracking the Diffusion of an Epidemic. In *Space-Time Integration in Geography and GIScience*, eds. M.-P. Kwan, D. Richardson, D. Wang and C. Zhou, 269-282: Springer.

Wilder-Smith, A., C. J. Chiew, and V. J. Lee. 2020. Can we contain the COVID-19 outbreak with the same measures as for SARS? *The Lancet Infectious Diseases* 20 (5):e102-e107.

Wu, J. T., K. Leung, and G. M. d. Leung. 2020. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *The Lancet* 395 (10225):689-697.

Xia, T., X. Song, H. Zhang, X. Song, H. Kanasugi, and R. Shibasaki. 2019. Measuring spatio-temporal accessibility to emergency medical services through big GPS data. *Health & Place* 56:53-62.

Xing, L., Y. Liu, X. Liu, X. Wei, and Y. Mao. 2018. Spatio-temporal disparity between demand and supply of park green space service in urban area of Wuhan from 2000 to 2014. *Habitat International* 71:49-59.

Yang, J., and L. Mao. 2018. Understanding temporal change of spatial accessibility to healthcare: An analytic framework for local factor impacts. *Health & Place* 51:118-124.

Yun, S. B., S. Kim, S. Ju, J. Noh, C. Kim, M. S. Wong, and J. Heo. 2020. Analysis of accessibility to emergency rooms by dynamic population from mobile phone data: Geography of social inequity in South Korea. *PLOS ONE* 15 (4):e0231079.

doi:10.6342/NTU202101357

附錄 A

作者：Fei-Ying Kuo, Tzai-Hung Wen, and Clive E. Sabel

文章標題：Characterizing Diffusion Dynamics of Disease Clustering: A Modified Space–Time DBSCAN (MST-DBSCAN) Algorithm

發表年分：2018

期刊名稱：Annals of the American Association of Geographers

卷號：108

期號：4

頁碼：1168-1186

數位物件識別號碼（DOI）：10.1080/24694452.2017.1407630

Annals of the American Association of Geographers

ISSN: 2469-4452 (Print) 2469-4460 (Online) Journal homepage: https://www.tandfonline.com/loi/raag21

Characterizing Diffusion Dynamics of Disease Clustering: A Modified Space–Time DBSCAN (MST-DBSCAN) Algorithm

Fei-Ying Kuo, Tzai-Hung Wen & Clive E. Sabel

To cite this article: Fei-Ying Kuo, Tzai-Hung Wen & Clive E. Sabel (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, DOI: 10.1080/24694452.2017.1407630

To link to this article: https://doi.org/10.1080/24694452.2017.1407630

Published online: 26 Jan 2018.

doi:10.6342/NTU202101357

Characterizing Diffusion Dynamics of Disease Clustering: A Modified Space–Time DBSCAN (MST-DBSCAN) Algorithm

Fei-Ying Kuo,* Tzai-Hung Wen,* and Clive E. Sabel†

*Department of Geography, National Taiwan University
†Department of Environmental Science, Aarhus University

Epidemic diffusion is a space–time process, and showing time-series disease maps is a common way to demonstrate an epidemic progression in time and space. Previous studies used time-series maps to demonstrate the animation of diffusion process. Epidemic diffusion patterns were determined subjectively by visual inspection, however. There currently are still methodological concerns in developing effective analytical approaches for profiling diffusion dynamics of disease clustering and epidemic propagation. The objective of this study is to develop a geocomputational algorithm, the modified space–time density-based spatial clustering of application with noise (MST-DBSCAN), for detecting, identifying, and visualizing disease cluster evolution, which takes the effect of the incubation period into account. We also map the MST-DBSCAN algorithm output to visualize the diffusion process. Dengue fever case data from 2014 were used as an illustrative case study. Our results show that compared to kernel-smoothed mapping, the MST-DBSCAN algorithm can better identify the evolution type of any cluster at any epoch. Furthermore, using only one two-dimensional map (and graphs), our approach can demonstrate the same diffusion process that time-series maps or three-dimensional space–time kernel plotting displays but in an easy-to-read manner. We conclude that our MST-DBSCAN algorithm can profile the spatial pattern of epidemic diffusion in detail by identifying disease cluster evolution. *Key Words: cluster evolution, DBSCAN, epidemic diffusion, geographical visualization, incubation period.*

流行病传播是一个时空过程，而显示时间序列的疾病地图，则是证实时空中的流行病进程之普遍方法。过往的研究运用时间序列地图来展现传播过程的状态。流行病传播模式，却是由视觉检验进行主观决定。建立有效的分析方法来描绘疾病集群与传染病繁殖的传播动态，至今仍有方法论上的疑虑。本研究的目标在于建立地理可计算化的演算法——一个改良的根据时空密度、并应用噪音的空间集群 (MST-DBSCAN)，用来侦测、指认并视觉化疾病集群的演化，并将孵化阶段纳入考量。我们同时绘製 MST-DBSCAN 演算法的结果，将传播过程进行可视化。自 2014 年以降的登革热案例数据，用来作为阐述的案例研究。我们的研究结果显示，相较于核平滑化的製图而言，MST-DBSCAN 演算法能够更佳地指认任何集群在任何时期的演化类型。此外，我们的方法仅运用一种二维地图 (和图表)，便能呈现时间序列图或三维时空核製图所展现的相同传播过程，但却是以简易阅读的方式。我们于结论中主张，我们的 MST-DBSCAN 演算法，通过指认疾病集群的演化，能够详细地描绘流行病传播的空间模式。 *关键词：集群演化，DBSCAN，流行病传播，地理可视化，孵化阶段。*

La difusión epidémica es un proceso del espacio–tiempo, y el exhibir mapas de series temporales de la enfermedad es una manera corriente de demostrar una progresión epidémica en el tiempo y en el espacio. Los estudios de tiempos pasados usaban mapas de series del tiempo para demostrar la animación del proceso de difusión. Sin embargo, los patrones de difusión epidémica se determinaban subjetivamente por inspección visual. Todavía en la actualidad hay preocupaciones metodológicas por desarrollar enfoques analíticos efectivos para reseñar la dinámica de difusión de la enfermedad en concentración espacial y la propagación epidémica. El objetivo de este estudio es desarrollar un algoritmo geocomputacional, la aplicación modificada espacio-temporal de agrupamiento espacial basada en densidad con ruido (MST-DBSCAN), para detectar, identificar y visualizar la evolución del agrupamiento de la enfermedad, que toma en cuenta el efecto del período de incubación. Nosotros también cartografiamos el producto del algoritmo MST-DBSCAN para visualizar el proceso de difusión. Los datos de 2014 del caso de la fiebre dengue se usaron como un estudio de caso ilustrativo. Nuestros resultados muestran que, comparado con el mapeo kernel suavizado, el algoritmo MST-DBSCAN puede identificar mejor el tipo de evolución de cualquier agrupamiento en cualquier época. Aún más, usando solo un mapa bi-dimensional (y gráficos), nuestro enfoque puede demostrar el mismo proceso de difusión que despliegan los mapas de series del tiempo o los despliegues trazados del kernel espacio-tiempo tridimensional, aunque de una manera

Annals of the American Association of Geographers, 108(4) 2018, pp. 1168–1186 © 2018 by American Association of Geographers
Initial submission, May 2017; revised submission, October 2017; final acceptance, October 2017
Published by Taylor & Francis, LLC.
doi:10.6342/NTU202101357

más fácil de leer. Concluimos que nuestro algoritmo MST-DBSCAN puede perfilar el patrón espacial de la difusión epidémica en detalle identificando la evolución del agrupamiento de la enfermedad. *Palabras clave: evolución del agrupamiento, DBSCAN, difusión epidémica, visualización geográfica, período de incubación.*

Epidemic diffusion refers to the phenomenon where an infectious disease spreads outward from its origin (Meade and Emch 2010). Through this process, an epidemic can evolve in space and time, with the emergence of a new case helping to reveal the diffusion process (Cliff and Haggett 2006). To prevent epidemic diffusion and unpack its etiological processes, understanding its spatial pattern is critical (Sabel, Pringle, and Schoerstrom 2009), and geographical visualization is a tool to help this dynamic phenomenon (Carroll et al. 2014). Cliff et al. (1981) developed a model for cluster evolution demonstrating the three broad diffusion patterns. Kalnis, Mamoulis, and Bakiras (2005) elaborated on this by proposing the term *moving cluster*, which focuses on the movement of a trajectory cluster over time. Its concept was similar to cluster evolution. These studies introduced the idea that cluster evolution could be a useful concept to demonstrate diffusion patterns. Further, Barreto et al. (2008), Kan et al. (2008), and Jeefoo, Tripathi, and Souris (2011) visualized geographically hypothetical concepts in real epidemic diffusion events. Thus, we propose that disease cluster evolution and geographical visualization are two important elements for realizing spatial patterns of epidemic diffusion.

Previous studies have discussed some possible types of disease cluster evolution such as describing trajectories of moving clusters and growth of an emerging cluster (Kalnis, Mamoulis, and Bakiras 2005; Lian et al. 2007; Barreto et al. 2008; Kan et al. 2008). Through the evolution process, a disease cluster might undergo distinct changes over time. Sometimes, it evolves singly; sometimes, it interacts with other processes. Needless to say, individual clusters could have distinct changes. There is still no study that systematically takes all of these characteristics into account together, however, and some situations are still ignored; for instance, the situation where a cluster center moves but the area becomes smaller or one cluster merges with a part of another cluster rather than its whole part. These shortcomings reflect the lack of a systematic classification of disease cluster evolution, which hinders us from completely depicting diffusion patterns. We propose here to provide a solution.

Methodologically, kernel density estimation (KDE), an exploratory spatial analysis method converting a point pattern data into a continuous risk surface that presents hot spots with high-density areas (Bailey and Gatrell 1995; Gatrell et al. 1996; Sabel et al. 2000), combined with time-series mapping, has some success in visualizing disease cluster evolution. KDE with time-series mapping, however, cannot "directly" demonstrate disease cluster evolution. In fact, KDE creates a risk surface, so time-series maps can only display the difference in risk area between epochs. When the change is very slight or the hot spot distribution is complex, it is very hard to observe characteristics and discover change by the naked eye. Different people might draw different conclusions when interpreting the same time-series maps.

Regarding spatial–temporal progression of infectious diseases, the incubation period for infectious diseases, a pivotal factor in epidemic diffusion (Gubler 1998; Chan and Johansson 2012), which is the period from being exposed to a pathogen to onset of illness and becoming infectious (Chan and Johansson 2012), is hard to incorporate in KDE with time-series mapping. In other words, two nearby cases cannot be in a transmission relationship unless there is a temporal relationship between the cases. Without an incubation period, a hot spot that KDE detects is only a concentration of event points, so the new cases that a hot spot contains are very likely not to be infected by the old cases in the same hot spot. Consequently, it is inappropriate to simply consider the difference between hot spots over time as a diffusion process.

This study thus aims to better geographically identify and visualize disease diffusion patterns by first describing a comprehensive disease cluster evolution classification and then implementing a suitable algorithm. We propose a new algorithm, the modified space–time density-based spatial clustering of application with noise (MST-DBSCAN), to combine clustering with the property of the incubation period. DBSCAN is a density-based algorithm that not only can detect irregularly shaped clusters but also does not need to predetermine the number of clusters (Ester et al. 1996; Birant and Kut 2007; Parimala, Lopez, and Senthilkumar 2011). Its feasibility for identifying clusters in geographical point pattern data has also been shown (Wang, Wang, and

doi:10.6342/NTU202101357

Li 2006; Birant and Kut 2007). Furthermore, we simplified time-series maps, making the diffusion process more easily understood. We present the effectiveness and feasibility of our methodology by applying it to the 2014 dengue fever epidemic, a mosquito-borne disease, in Kaohsiung City, Taiwan.

Related Work

Diffusion patterns have been categorized into three types: hierarchy, contagion, and relocation (Meade and Emch 2010), where the first two could also be considered as expansion diffusion (Cliff et al. 1981). Each type has its own epidemiological meaning and implies a distinct spatial transmission mechanism. Hierarchy describes transmission through an ordered sequence of settlements. Contagion depends on direct contact for infection. Finally, relocation concerns disease that shifts from a current location to another distant site by means of the movement of the host such as human or animal. In terms of geographical visualization of disease cluster evolution, Cliff et al. (1981) theoretically showed expansion diffusion as a process where a cluster has a fixed center but enlarges its area gradually. Relocation diffusion was illustrated as a process where a fixed-area cluster keeps moving. Finally, they described the case where these two processes might happen together, so a mixed process occurs where a cluster moves with its area expanding. Their work provided a fundamental understanding about disease cluster evolution, but it has remained largely theoretical.

The process of disease cluster evolution consists of the changes in single and interaction patterns. The aforementioned Cliff et al. (1981) study focused on the change in a cluster's center and area. Barreto et al. (2008) commented that the focus, or center, is an important characteristic about an epidemic; Kan et al. (2008) considered that area can reflect the severity of an epidemic. Thus, these two geographical characteristics can be used to profile the evolution of a single cluster in detail. On the other hand, Lian et al. (2007) mentioned that epidemics in different locations could merge and form a more serious epidemic; Kalnis, Mamoulis, and Bakiras (2005) considered that a cluster could split into two or several parts. Their work implied the existence of interaction between clusters. These studies revealed some types of single and interaction pattern of disease cluster evolution, but a systematic classification is still absent.

Thoroughly understanding spatial patterns of epidemic diffusion is vital to unpack etiological processes: Geographical visualization is a tool to help this dynamic phenomenon. Previous studies showed two categories of visualization (Bailey and Gatrell 1995; Cheng et al. 2013). First, exploratory visualization aims at displaying data in a visual form so that analysts can obtain some initial insights before analysis; the other one is summarized visualization, which presents the analytical results that analysts want to communicate with audiences. To geographically visualize disease cluster evolution, previous studies have used a two-stage exploratory approach. First of all, they detected hot spot distributions at each epoch by using techniques such as KDE. Then, they compared the risk surfaces between epochs to understand the differences between hot spots with a view to understand diffusion processes. After visualization, geographical context and environmental factors have been used to explain the cause of diffusion and some public health prevention advice was provided (Barreto et al. 2008; Kan et al. 2008; Jeefoo, Tripathi, and Souris 2011). This approach helped visualize dynamic procedures like diffusion and was regarded as time-series mapping because it takes a sequence of figures (Brunsdon, Corcoran, and Higgs 2007).

To better visualize temporal progression, Brunsdon (2001) developed a technique, *comap*, and combined it with KDE. Its concept is similar to time-series mapping, but the data sets between different epochs partially overlap. When the analysis procedure moves to the next epoch, some new data are included in the data set and some old data are excluded; that is, the data set is dynamic. This property enables readers to easily observe the difference between different epochs because two data sets share the overlapping data (Brunsdon, Corcoran, and Higgs 2007).

Kalnis, Mamoulis, and Bakiras (2005) have modified the DBSCAN algorithm to detect moving clusters. Because they dealt with trajectory data, there was no need to take the incubation period into account. In addition, they did not identify all possible patterns of evolution and visualize them. Their work, however, still revealed the potential of DBSCAN to properly visualize disease cluster evolution. Moreover, Ester et al. (1998) integrated the concept of a dynamic data set with DBSCAN, but their method is not suitable for disease data. In fact, they did not modify the algorithm but just implemented clustering once for each epoch. The dynamics came from the property of their data because it recorded the duration of each point. Each point might survive for a period of time, and they can know which point still exists at each epoch. Nevertheless, most disease data are a set of event

doi:10.6342/NTU202101357

points recording only the location and the onset time rather than how long they continue. Each point exists in only one epoch. Therefore, their method cannot capture the dynamic nature of this kind of disease data.

Methods

The aforementioned property of the incubation period highlights the need to consider the transmission relationship between cases when detecting clusters. Therefore, in this study, we defined a new concept of transmission cluster rather than conventional spatial–time clustering. The cases, which form a transmission cluster, are considered by not only spatial proximity in time and space with each other but also the transmission relationship, which represents that the neighbors of one case must appear spatially near but later than the case so that these neighbors could be infected by this case. Our proposed methodological framework included three stages (Figure 1). In the first stage, ten diffusion types of a transmission cluster evolution were defined to better describe the possible evolution patterns. Both single and interaction patterns were considered. The second stage introduced how MST-DBSCAN was modified from DBSCAN, including describing the formation of a dynamic data set, handling the property of the incubation period, and detecting the transmission cluster for identifying the evolutions. Finally, the third stage was to simplify time-series maps. We used the Louvain method, a grouping detection algorithm in network analysis (Blondel et al. 2008), to detect groups in the network whose nodes are administrative districts and links are diffusion similarity between districts. The districts in the same group have similar diffusion procedures, so we can use a thematic map to display the diffusion dynamics rather than a series of figures.

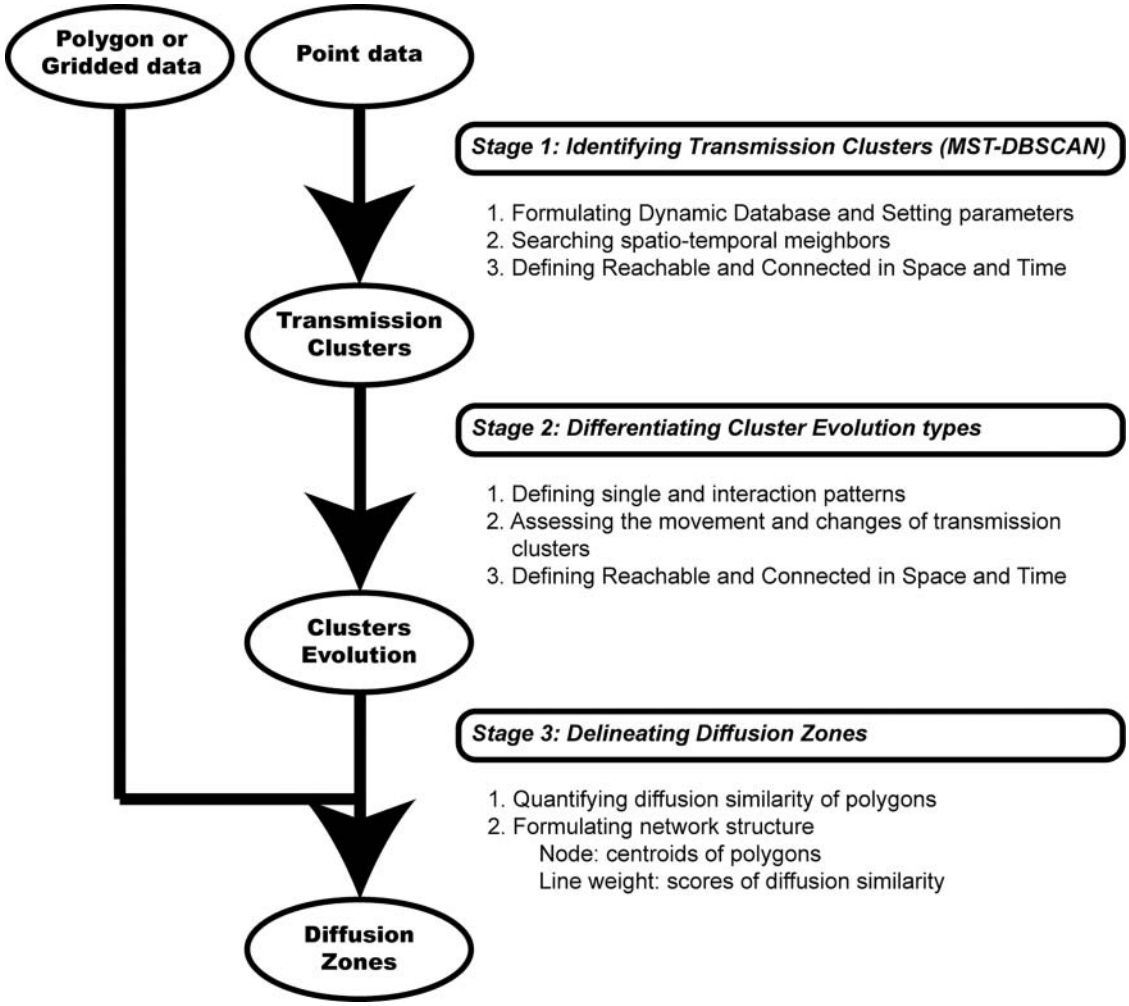


Figure 1. Methods framework.

doi:10.6342/NTU202101357

Classification of Disease Cluster Evolution

We propose a new classification of disease cluster evolution (Figure 2) separated into single and multiple disease clusters. The changes in the center and the area of a transmission cluster were used to monitor the single pattern. Each of these changes implies a significant property of diffusion, namely, that distinct strategies for public health are needed. For example, if the epidemic in a transmission cluster becomes more prevalent than before, the area might get larger. The health authority could focus its attention there. If a transmission cluster moves to another place, its center could move. The health authority could track its trajectory and try to stop its progression. By combining the changes in these two characteristics, six distinct types of single cluster evolution patterns could be discovered.

- Steady (Figure 2A): The transmission cluster does not move to another place and its influencing area remains the same. This type implies that although a cluster does not diffuse, it still keeps its current status and could diffuse in the future.
- Move (Figure 2B): The transmission cluster moves to another place but its influencing area remains the same. This type implies that although the sources of infection might be cleared at the original place, they have shifted to neighboring areas.
- Growth (Figure 2C): The transmission cluster does not move but its influencing area becomes larger. This type means that the sources of infection might not be thoroughly cleared at the original place, so they keep growing and continually infect healthy people around this area.
- Directional growth (Figure 2D): The transmission cluster moves and its influencing area increases. This type implies that the sources of infection not only remain at the original place but also shift to neighboring areas.
- Reduction (Figure 2E): The transmission cluster does not move but its influencing area becomes smaller. This type reflects that the sources of infection could be gradually wiped out.
- Directional reduction (Figure 2F): The transmission cluster moves and its influencing area decreases. This type means that although the sources of infection are being cleared, they still diffuse to neighboring areas.

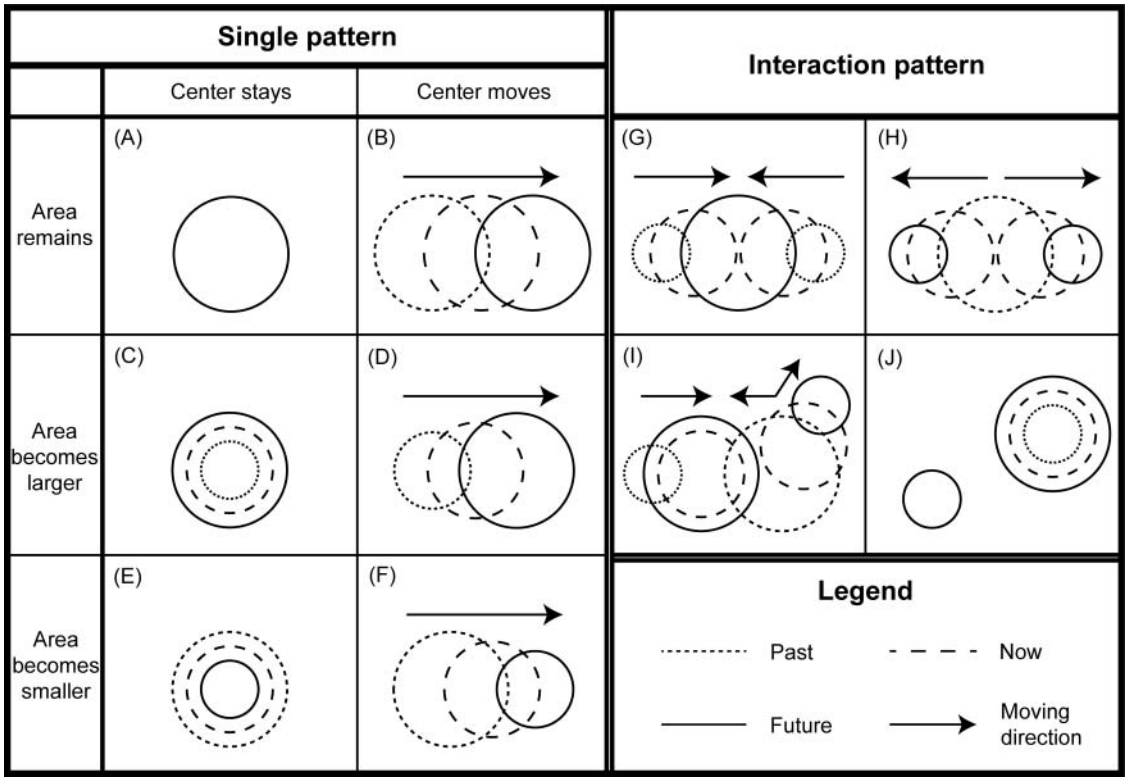


Figure 2. Ten cluster evolution types.

doi:10.6342/NTU202101357

Characterizing interaction cluster evolution patterns focuses on the changes resulting from the interaction between two or more transmission clusters. The first type is merge (Figure 2G). When two or more transmission clusters become closer or they expand toward each other, they could finally merge and form a bigger transmission cluster. This type implies that a series of outbreaks of an epidemic could come from the merging of several small transmission clusters. The second type is split (Figure 2H). A transmission cluster might split into two or more separate clusters that could undergo different developments in the future. The third type is split-merge (Figure 2I). This type aims at capturing the phenomena that a transmission cluster absorbs a part of another cluster rather than its whole part. The last type is emerge (Figure 2J). This means a new transmission cluster appears although several transmission clusters already exist. This type implies that the sources of infection are carried to a new place by the movement of the host, such as a human or an animal, so a transmission cluster can appear at a new site rather than being the expansion of an existing transmission cluster.

A Modified Space–Time DBSCAN

First of all, the concept of a dynamic data set requires that the data set should be clustered at each epoch but should be partially overlapping. To formulate a dynamic data set that is suitable for point pattern data, we designed a new parameter, *EpsT1*, to pick out the case points that should be included at each epoch (Equation 1).

$$D_t = \{p \mid t - EpsT1 \leq T(p) \leq t\}, \tag{1}$$

where D_t is the data set at time t, and $T(p)$ is the appearance time of case p. In Equation 1, $T(p) \leq t$ allows that the cases could be infected earlier than or at time t to be included in the data set; $(t - EpsT1) \leq T(p)$ keeps the cases that are still infectious and excludes the cases that appear too early to affect the current situation.

Second, the incubation period highlights the transmission relationship for a case point to find its neighbor. This modification necessitates redefining the "neighbor" in DBSCAN, which only focuses on the spatial adjacency between points. We extended the definition of neighbor and named it spatiotemporal neighbor. The parameter *Eps* originating from DBSCAN was used to set the threshold value of the spatial adjacency. We also renamed it as *EpsS* to distinguish it from the temporal parameters. On the time dimension, the incubation period is a rough range rather than a definite interval because it was hard to confirm (Gubler 1998; Chan and Johansson 2012). The length of the period often differs from person to person. Sometimes it is short, but sometimes it is much longer. Therefore, we considered that a time window is a much better way for a case to search its neighbors. *EpsT1*, which was used to create the dynamic data set, was also used here to set the threshold value of the possible longest transmission. Because a case cannot transmit a disease to those whose onset time is after the longest transmission, there might not be a problem in reusing *EpsT1*. Furthermore, a new parameter, *EpsT2*, was designed to set the threshold value of the possible shortest transmission. Equations 2 and 3 and Figure 3 demonstrate the quantitative definition of spatiotemporal neighbor.

$$\Delta T\ (i,j) = T(j) - T(i) \tag{2}$$

$$STNB_{i,t} = \{\ j \in D_t \mid \mathrm{dist}(i,j) \leq EpsS\ \cap\ EpsT2 \leq \Delta T(i,j) \leq EpsT1\ \cap\ i \neq j\ \}, \tag{3}$$

where $\Delta T(i,j)$ denotes the difference between the appearance time of i and j, dist(i,j) denotes the spatial distance from i to j, and $STNB_{i,t}$ denotes the spatiotemporal neighbors of i at time t. Notice that, without this modification, a neighbor only considered spatial adjacency, so it was bidirectional. Nonetheless, when the concept of transmission was added, the adjacency was

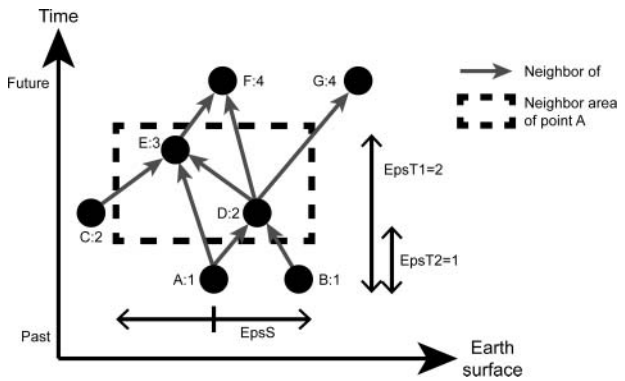


Figure 3. Spatiotemporal neighbors. The number behind each letter is the appearance time of the point. The dashed rectangle is the neighbor area of point A. That is, the points inside this area, D and E, are point A's neighbors. On the other hand, although point B is spatially adjacent to A, it is not infected by A due to the same appearance time.

doi:10.6342/NTU202101357

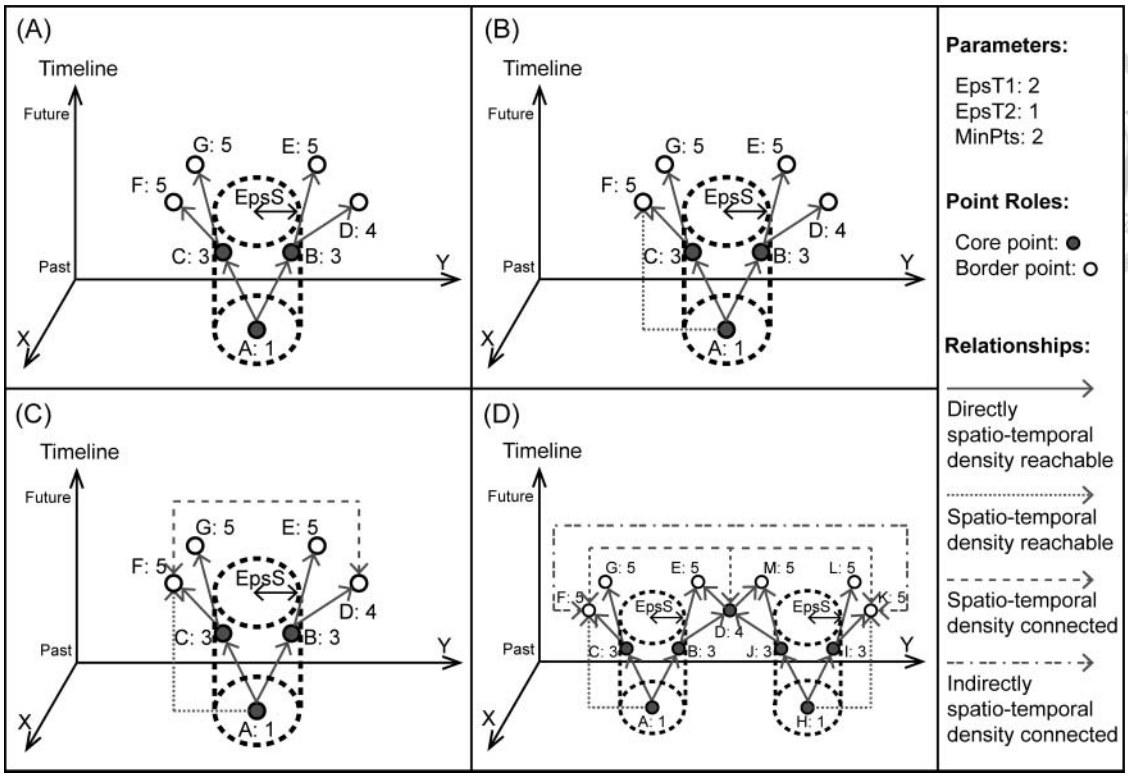


Figure 4. Spatiotemporal relationship. The number behind each letter is the appearance time of the point.

transformed to be unidirectional in this study; that is, two cases cannot infect each other.

In DBSCAN, the number of neighbors reflects a point's density. A high number means that a case has many neighbors within a given area, so the density is high. According to this relationship, DBSCAN classifies all points into three different roles, including core, border, and noise points (Ester et al. 1996). Our study followed the same criteria for classifying points. A parameter, minimum points (*MinPts*), was used to set the threshold value of how many neighbors a core point must have as a minimum. Core point means that it has enough spatiotemporal neighbors, so it can transmit disease to other case points and become the major structure of a transmission cluster with high disease incidence. A border point represents the border of a transmission cluster because its density is not high enough. It is like a border of a cluster where the epidemic diffusion is slowing down. A noise point acts as an outlier located in an area with low disease incidence.

The third modification of DBSCAN concentrates on the relationship between case points and the formation of a cluster. In fact, this modification was forced by the second modification because the relationship is tightly related to the way to find neighbors. The relationship further affects how to detect a cluster, so the redefinition was necessary. Moreover, DBSCAN has only three relationships. We added the fourth relationship to capture the interaction pattern between different clusters. "Spatiotemporal relationship" and "transmission cluster" were our new definitions, defined as follows:

- Directly spatiotemporally reachable (Figure 4A): Given two points p and q, if p is a neighbor of q and q is a core point, p is directly spatiotemporally reachable from q.
- Spatiotemporally reachable (Figure 4B): Given a point p and a core point q, p is spatiotemporally reachable from q if there is a chain-like link k_1, k_2, …, k_n, where $k_1 = q$ and $k_n = p$, such that k_{i+1} is directly spatiotemporally reachable from k_i, for $1 \leq i < n$, $k_i \in D_t$.
- Spatiotemporally connected (Figure 4C): Given two points p and q, if there is a point k such that both p and q are spatiotemporally reachable from k, p and q are spatiotemporally connected to each other.
- Indirectly spatiotemporally connected (Figure 4D): Given two points p and q, they are indirectly spatiotemporally connected to each other if there is a core point k such that p and q are spatiotemporally connected to k simultaneously.

doi:10.6342/NTU202101357

- Transmission cluster: A cluster C must satisfy the following two terms:
  - $\forall p, q \in D_t$, $p \in C$ if $q \in C$ and p is spatiotemporally reachable from q.
  - $\forall p, q \in C$, p is indirectly spatiotemporally connected to q.

After establishing the MST-DBSCAN and detecting the transmission clusters, the next task was to identify the diffusion type of each transmission cluster at every epoch. For single cluster evolution patterns, whether the area or the center of a transmission cluster changes was the criterion to identify the diffusion type. To check the changes, calculating the area and the center at each epoch was necessary. Because DBSCAN is essentially a data clustering method, the cluster to which each case point belongs at each epoch can be recorded. In other words, the case points that each cluster contains at each epoch can be known. Therefore, we defined the center of a transmission cluster as the spatial mean center of all cases in this cluster. Moreover, to calculate the area, we made buffers with a radius equal to the *EpsS* for all cases in the same cluster and then merged these buffers into a much bigger polygon. The area of this polygon represented the area of the transmission cluster. After that, we used Equation 4 to calculate the movement distance of the mean center. If the distance is larger than half of *EpsS*, the center is judged as moving; otherwise, it stays. Similarly, Equation 5 was used to calculate the proportion of the area change. If the proportion is larger than 0.1, the area becomes bigger; if the proportion is less than –0.1, the area becomes smaller; otherwise, it remains the same.

$$\Delta\, MC_{i,t} = \text{dist}\left(MC_{i,t},\ MC_{i,t-1}\right) \tag{4}$$

$$\Delta\, A_{i,t} = \frac{A_{i,t} - A_{i,t-1}}{A_{i,t-1}}, \tag{5}$$

where $MC_{i,t}$ and $A_{i,t}$ are the spatial mean center and the area of the transmission cluster i at time t, respectively. With Equation 4 and Equation 5, we could know whether the center or the area changes. By referring to Figure 2, the diffusion type of each transmission cluster at each epoch could be identified.

For interaction pattern, as mentioned earlier, MST-DBSCAN can store the component of each transmission cluster in each epoch. Therefore, whether the composition of a cluster changes between epochs can be detected and then the four types of interaction patterns can be distinguished.

Simplified Time-Series Maps

To simplify time-series maps, we first projected the evolution of transmission clusters onto the administrative districts. Technically, this approach allowed us to record what kind of cluster evolution each administrative district underwent at each epoch. Then, we compared any two districts and calculated how many times that they share the same types (Equation 6), which was defined as diffusion similarity in this study.

$$S_{i,j} = \frac{E_{i,j}}{E_a}, \tag{6}$$

where $S_{i,j}$ denotes the similarity of diffusion pattern between administrative districts i and j with values ranging between 0 and 1. The diffusion patterns between administrative districts i and j are regarded as identical if the value of $S_{i,j}$ is equal to one. E_a denotes the total number epochs of the epidemic. $E_{i,j}$ denotes the count of the epochs that i and j have the same type.

Based on the measurement of diffusion similarity, we used the Louvain method to group districts. The Louvain method, a grouping algorithm, operates on a network that contains a set of nodes and a set of links. Here, each administrative district can be regarded as a node and the score of the diffusion similarity between any two districts ($S_{i,j}$) is a link weight. Then, the Louvain method grouped districts into several groups by finding a set that has high similarities within each group and low similarities between groups. We defined each community as a diffusion zone because the districts in the same community had similar diffusion dynamics. With this property, we could use a thematic map to display the distribution of all diffusion zones. To further realize the detailed diffusion procedure, we also sketched a temporal graph of ten cluster evolution types for each zone.

The procedures of our proposed analytical framework are described in terms of pseudocode in Appendix A, including three major procedures of (1) MST-DBSCAN for detecting transmission cluster, (2) identifying evolution types, and (3) simplifying time-series maps. The time complexity of our core algorithm, MST-DBSCAN, is also assessed and described in detail in Appendix B. Finally, we compared performance and effectiveness of MST-DBSCAN with the original DBSCAN using our case study data.

doi:10.6342/NTU202101357

Application: Cluster Evolution and Diffusion Zone of the 2014 Dengue Fever in Kaohsiung, Taiwan

To demonstrate the feasibility of our proposed methodology, we illustrated our techniques with data on a 2014 dengue fever epidemic in Kaohsiung City. The city is the second largest municipality in Taiwan, with a population of 2.8 million in 2014. Twelve districts, part of Kaohsiung City displayed in Figure 5, made up our study area because this area is usually the major area of the dengue epidemic in southern Taiwan every year. Previous studies also chose the same area to study the dengue fever epidemic in Taiwan (Kan et al. 2008; Wen and Tsai 2015). Our data were collected through the national surveillance system of the Taiwan Centers for Disease Control (Taiwan CDC), which reported 13,606 confirmed dengue cases within these twelve districts from May 2014 to February 2015. The detailed epidemic curve is showed in Figure 6.

For the dengue epidemic in Taiwan, Taiwan CDC indicated that if a case appears within 150 m of an existing case, there might be a cluster. Based on this criterion, if the *EpsS* is set as 150 m, the *MinPts* value must be one. In other words, a case needs only one neighboring case for it to be a core point. Therefore, we set the *EpsS* as 300 m and the *MinPts* as three to preserve similar density. On the time dimension, because dengue fever is a vector-borne disease, extrinsic incubation periods (EIPs) and intrinsic incubation periods (IIPs) comprise its incubation periods (Chan and Johansson 2012). EIP is the time that a pathogen spends duplicating in a mosquito's body, and IIP is the time that it spends inside a person. When a person is infected, the time that he or she can transmit the virus to mosquitoes starts from the illness onset day but ends at an unfixed day (Gubler 1998). This period is the infectious period. We considered that it is much better to take these three periods into account when setting the value of *EpsT1* and *EpsT2* (Figure 7). Gubler (1998) reported that the end of the infectious period ranges from two days to ten days, so we took six days, the average of two and ten, to be the end of the infectious period in this study. Gubler also stated that the IIP ranges from four days to seven days. Chan and Johansson (2012) pointed out that the range of EIP is different under different temperature regimes. According to the historical climate data for Kaohsiung City, the monthly average temperature in August, September, October, and November in 2014, the dengue epidemic season, ranged from 25.2°C to 29.5°C. Referring to their study, the range of the EIP is two days to fifteen days. Guzman et al. (2010), however, commented that the EIP is about ten days and is much shorter at high ambient temperatures. Consequently, we combined the arguments of these two studies and used two days to ten days as the EIP. Based on these

Figure 5. Study area in Kaohsiung City.

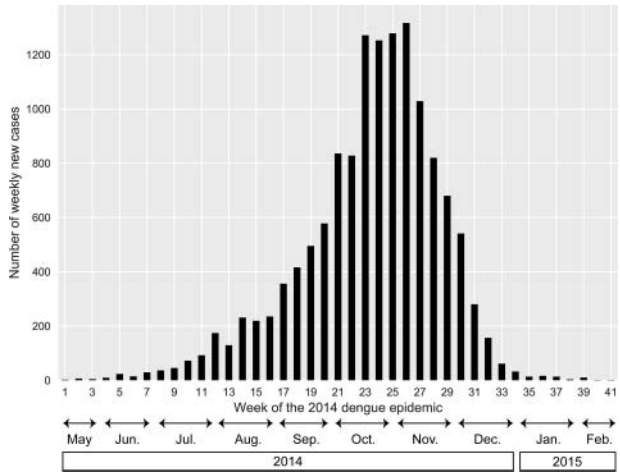


Figure 6. The temporal trend of dengue fever cases.

doi:10.6342/NTU202101357

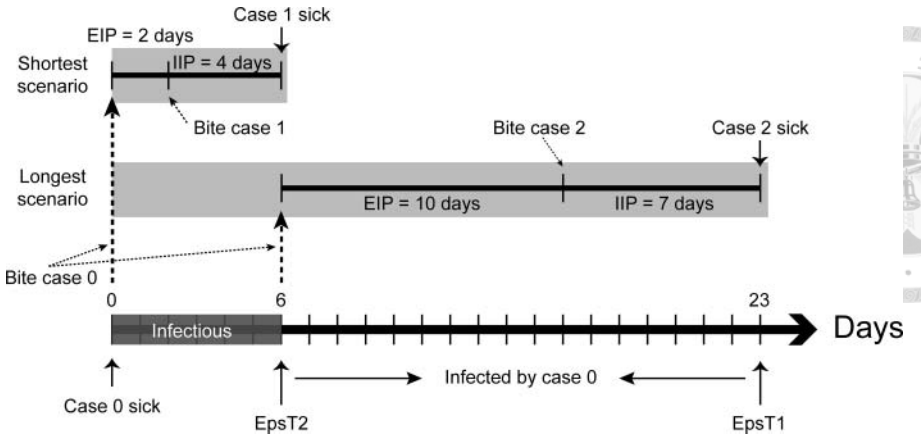


Figure 7. The time parameter settings were designed to capture the transmission based on the ranges of infectious periods, extrinsic incubation periods (EIPs), and intrinsic incubation periods (IIPs). *EpsT2* is the time that case 0 is sick and adds the minimum days for EIP and IIP. *EpsT1* ranges beyond *EpsT2* to the maximum days for EIP and IIP.

three periods, if a case shows symptoms, the earliest that it can be infectious is six days later (infectious period (0) + EIP (2) + IIP (4) = 6). Similarly, the last case must have an onset time twenty-three days later (infectious period (6) + EIP (10) + IIP (7) = 23). As a result, the time window can range from six to twenty-three days. Because our analysis unit of time was a week, we set *EpsT1* as two weeks and *EpsT2* as one week.

Figure 8 shows the result of MST-DBSCAN and the result of KDE for our study area. The parameters for the KDE analysis include the bandwidth, which was set to 300 m to match that of the MST-DBSCAN analysis. According to Figure 8B (KDE), the hot spots of the dengue epidemic were located in the central area in weeks 17 to 20. It also diffused to the north but gradually. The result of the MST-DBSCAN analysis (Figure 8A), however, not only illustrates the diffusion trend but also reveals the diffusion type of each transmission cluster. For example, in Figure 8A, a transmission cluster persistently merges with other clusters and becomes the largest cluster in week 20. Moreover, from week 18 to week 19, several clusters split and some new clusters appear. These clusters could be the cause triggering the epidemic to diffuse to the north. This comparison illustrates that, compared to KDE, MST-DBSCAN can not only reveal the process of cluster evolution but can also identify their evolution type(s) in any specific time period.

MST-DBSCAN quantitatively demonstrates the cluster evolution patterns of the epidemic diffusion, but it still relies on the time-series maps for visualization. Figure 9A displays the result of applying the Louvain method on the transmission clusters aggregated by administrative districts. In Figure 9A, villages are used as the spatial unit because it is the third-level administrative unit, which is the smallest aggregation unit for socioeconomic statistical data in Taiwan. The villages that have the same color have been assigned to share the same diffusion patterns. That is, these villages have similar diffusion procedures. Figure 9A shows that zones 2, 4, and 5 cover the most central area, which is the major area of the dengue epidemic; zones 1 and 3 surround the previous three zones around the border of our study area where the epidemic is not serious. Figure 9B further shows the major diffusion types of each diffusion zone over time. The two major diffusion types in zone 1 are steady and growth, which have the same characteristic, center remaining. Therefore, zone 1 can be identified as the epidemic source. The frequency of emergence in zone 3 is higher than that in the other zones, so it is named intensive occurrence. In addition, merge, split, and split-merge dominate the diffusion pattern in zones 2, 4, and 5. Although these three areas share the same major types, the start times of merge are a little distinct. Zone 2 has the earliest appearance of merge, so it is named early-stage expansion. Zone 4 follows zone 2 and has the second early appearance of merge. It can be categorized as middle-stage expansion. Finally, zone 5 is late-stage expansion because its merge appearance is the latest.

Based on the property of these five zones, we can depict more distinct space–time diffusion processes. The epidemic first appears in Figure 9 at zone 1 where the diffusion pattern tends to stay in a fixed area, so it does not affect many areas. Then, the epidemic

doi:10.6342/NTU202101357

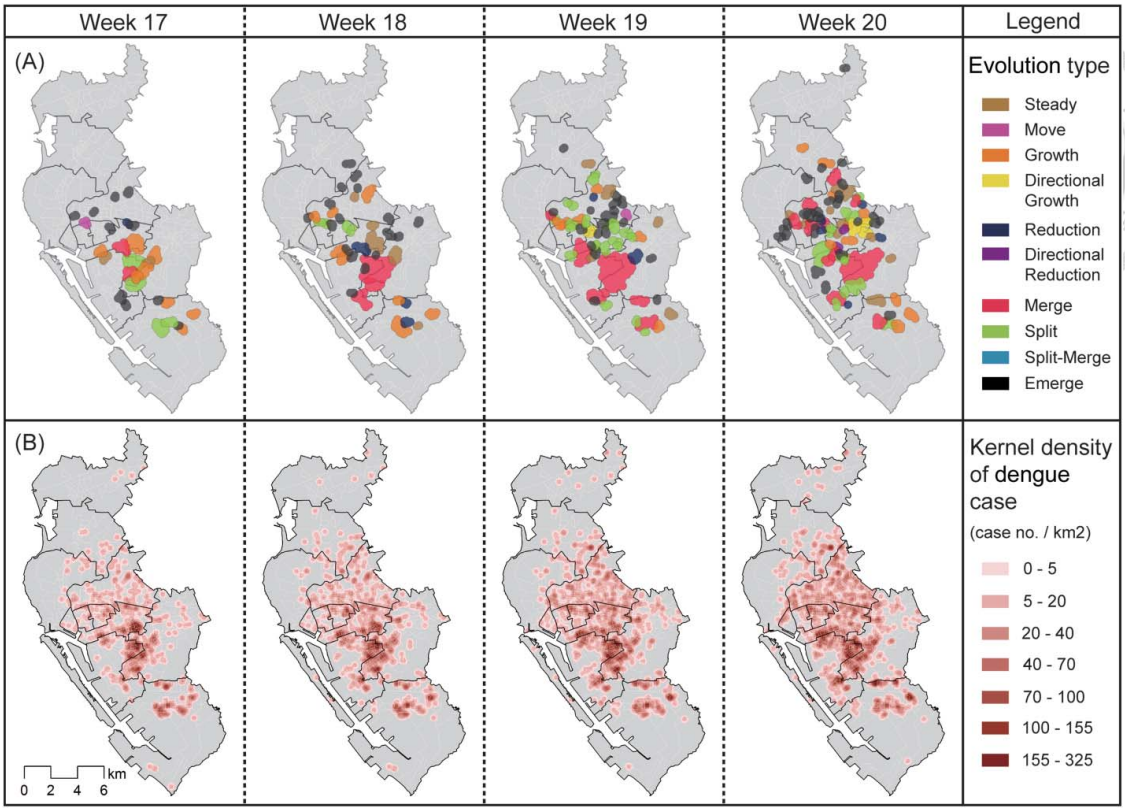


Figure 8. Comparison between (A) modified space–time density-based spatial clustering of application with noise and (B) kernel density estimation. (Color figure available online.)

appears at zone 2, the downtown of the city, where the interaction pattern dominates the diffusion and the first large-scale outbreak thus occurs. Later, the epidemic diffuses to zone 3, the border of the city. It creates many new and small clusters and enlarges the affected area sharply. Next, the epidemic moves to zone 4. The interaction pattern dominates the diffusion again and the second large-scale outbreak occurs. Finally, the diffusion in zone 5 is also dominated by the interaction pattern, which leads to the third large-scale outbreak. This process reveals a clear picture that the early epidemic resulted from a single pattern and the occurrence of an interaction pattern could cause large-scale outbreaks.

Figure 8A demonstrates that there are several split patterns in zone 4 (Figure 9A) in weeks 19 and 20; Figure 9B also illustrates that temporal progression of the split pattern in this zone starts at about week 19. Furthermore, the largest cluster in Figure 8A persistently merges with other clusters and is located in zone 2, the type of early-stage expansion. Figure 9B also shows that the merge pattern in this zone occurred about week 19. Therefore, we simplified the process of time-series maps and disease cluster evolution can be visualized in a much more easily readable way but still preserves its correctness.

To further prove the effectiveness of our simplified approach, we compared it with space–time kernel density estimation (ST-KDE). The parameters for the ST-KDE analysis include both space and time bandwidth. To match that of the MST-DBSCAN analysis, we set the space bandwidth as 300 m and the time bandwidth as two weeks. Figure 10 shows the ST-KDE of the dengue epidemic in a part of the study area. Hot spot I is the earliest one among these four hot spots and its location is zone 1. Figure 9B also shows that the diffusion in this area is the earliest. On the other hand, hot spots II and III are located in zone 4 and hot spot IV is located in zone 5. The emergence of hot spot IV is later than that of the previous two hot spots. A similar phenomenon can also be observed in Figure 9B. This comparison shows that our approach uses one two-dimensional (2D) map and graphs to interpret the result, which is more intuitive than the ST-KDE. Moreover, our approach can further demonstrate a dominated diffusion type in each hot spot of ST-KDE.

Finally, we compared the performance between MST-DBSCAN and the original DBSCAN. Both algorithms

doi:10.6342/NTU202101357

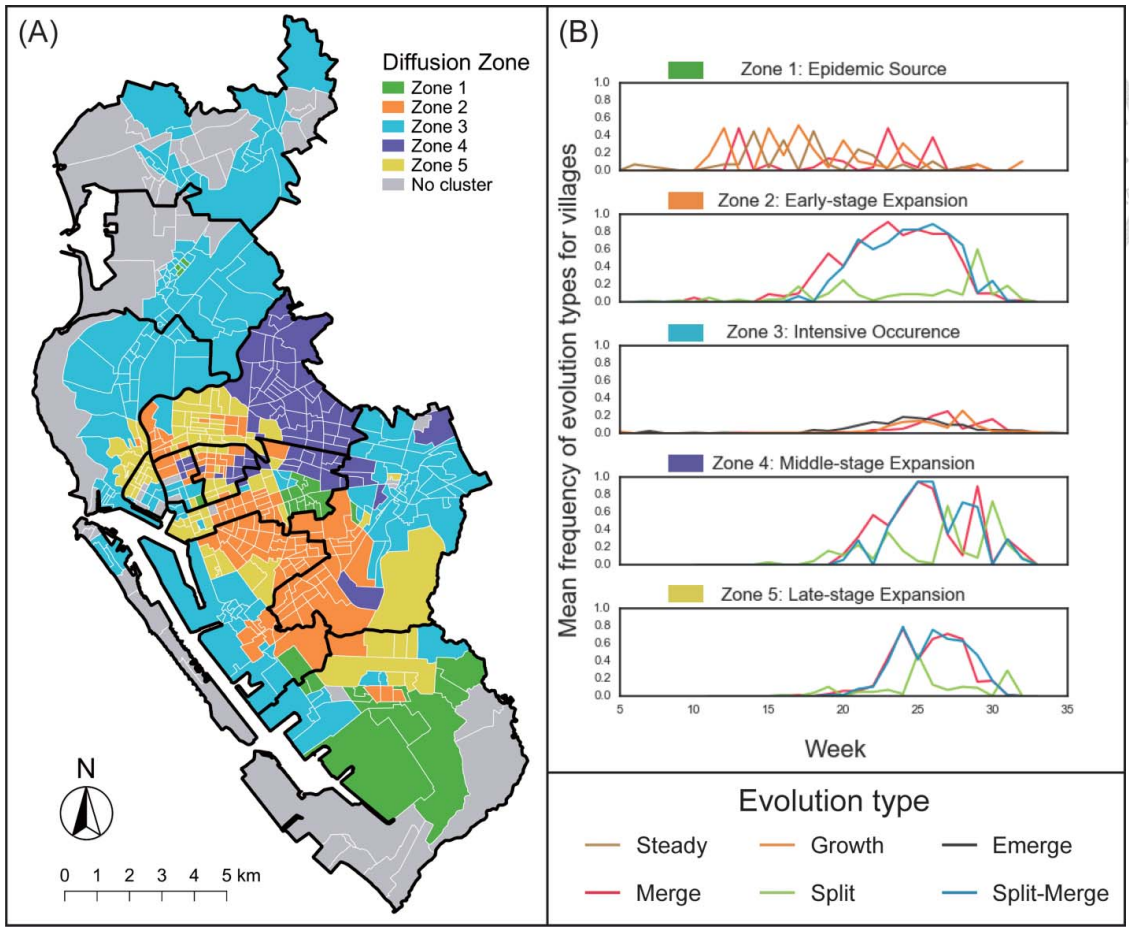


Figure 9. (A) Diffusion zone and (B) major evolution type. (Color figure available online.)

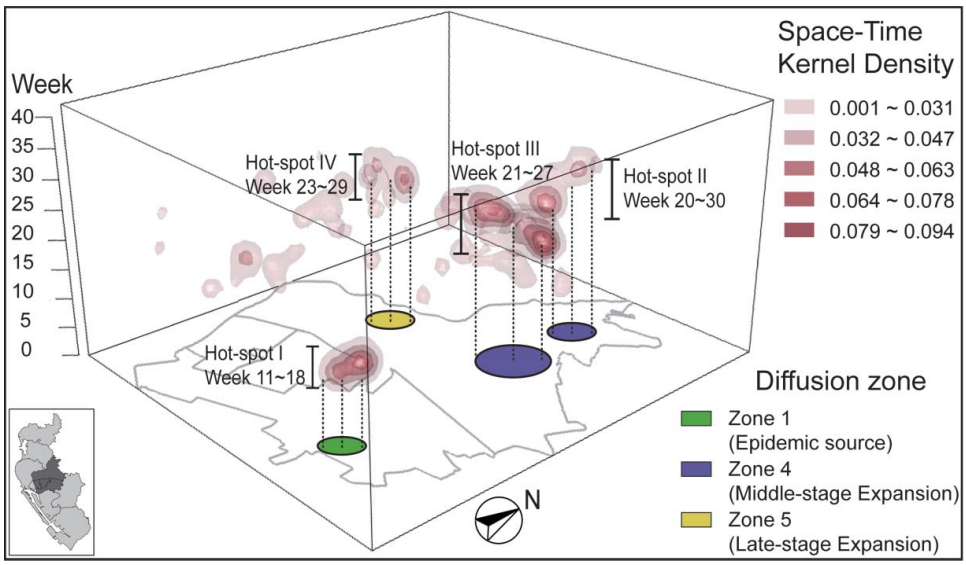


Figure 10. Comparing the results of modified space–time density-based spatial clustering of application with noise with space–time kernel density estimation. Hot spot I is the earliest one among these four hot spots, and its location is zone 1. Hot spots II and III are located in zone 4, and hot spot IV is located in zone 5. The emergence of hot spot IV occurs later than that of the previous two hot spots. (Color figure available online.)

doi:10.6342/NTU202101357

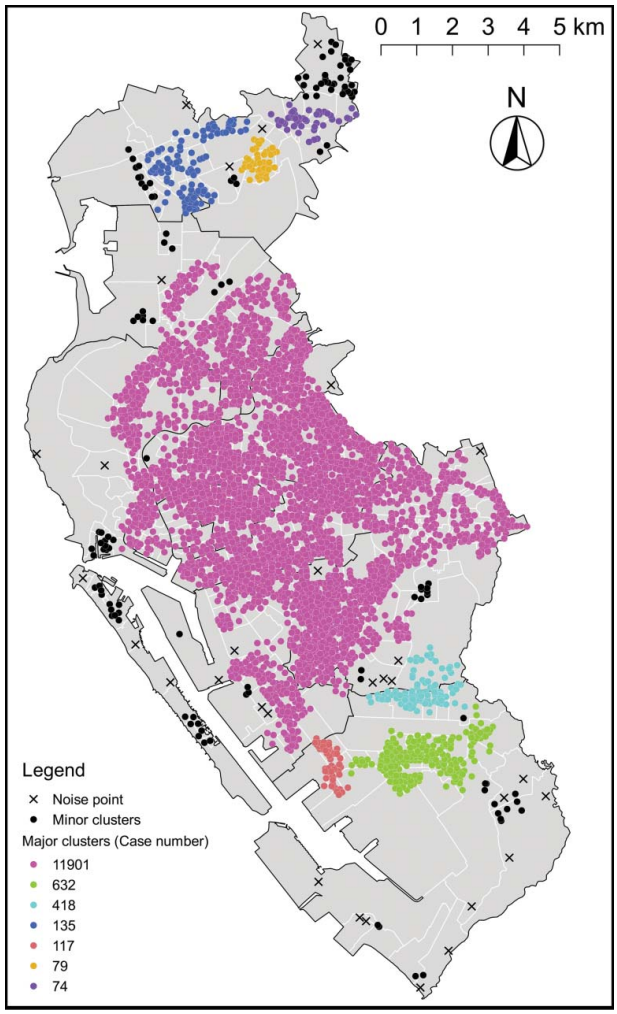


Figure 11. The result of applying density-based spatial clustering of application with noise to the dengue fever case. (Color figure available online.)

were implemented to analyze our case data in Python 3.6. All experiments were conducted on a 2.80-GHz Intel Core i7-7700HQ processor with 16 GB memory, running Microsoft Windows 10. In terms of execution time, with our data, MST-DBSCAN accelerated by KD-Tree took 6.00 seconds, and DBSCAN needed 0.23 seconds. Because MST-DBSCAN considers spatial–temporal neighbors, the time complexity of the MST-DBSCA algorithm is $O(N^2)$, and the original DBSCAN needs only O(NlogN). Without incorporating the temporal progression, the result of DBSCAN (Figure 11) shows a single large cluster at the central region of Kaohsiung. Using the MST-DBSCAN algorithm proposed in this study, the single cluster identified by DBSCAN was decomposed into more small zones (Figure 9A), which differentiates the spatial–temporal characteristics of disease cluster evolution. In this study, the grouping result from MST-DBSCAN could be also further be used for identifying cluster evolution types and to simplify time-series maps.

Discussion

In this article, inspired by Cliff et al. (1981), we have presented a method for classifying cluster evolution patterns, including single and interaction patterns. To directly identify and visualize them, we developed an algorithm, MST-DBSCAN, which integrates the concept of a dynamic data set with the incubation period in the disease transmission process. Moreover, we transformed the cluster, evolution into a network form that is composed of the administrative districts and their diffusion similarity. The Louvain method helps us depict diffusion zones and major diffusion types in each of them by detecting the communities of the districts. Requiring only one 2D map (and graphs; Figure 9), this is a new spatiotemporal visualization approach aiming to solve the problems that time-series maps and three-dimensional space–time kernel plots suffer from, namely, that the time series are hard to visualize.

Although Barreto et al. (2008), Kan et al. (2008), and Jeefoo, Tripathi, and Souris (2011) used KDE to understand disease diffusion processes, the innovation in this article is to enhance each cluster by identifying its evolution type at each epoch. In addition, the hot spots that KDE detects do not have a clear boundary, so it is hard to identify their area and the case points they contain. In other words, previous studies could not track the development of each cluster so they could not further quantitatively analyze the hot spots to profile the diffusion process in detail. This property allows us to aggregate the cluster evolution types into administrative districts and then measure the diffusion similarity between them, so our simplified approach is operable.

In our case study, the early epidemic follows the single pattern whereas the disease breaks out when the interaction patterns occur. Moving to the practical reality on the ground, the local government considers that several villages in central zone 2 are index or source villages because the first outbreak usually happens there every year. Our results also indicated the first outbreak in 2014 (the largest cluster in Figure 8A) appeared in the same place. Kan et al. (2008) found that this area was the early outbreak site of dengue fever in 2001 through 2003; Wen and Tsai (2015) also discovered

doi:10.6342/NTU202101357

the same phenomenon in 2009. Both of these studies attributed the early epidemic of dengue fever to "contagious diffusion," which spreads slowly and usually results from environmental characteristics such as poor garbage disposal or standing water. On the other hand, they ascribed the late outbreaks to "relocation diffusion," which rapidly intensifies the diffusion by demographic factors such as population mobility. Therefore, our single pattern might act like contagious diffusion and the interaction pattern might have the same effect as relocation diffusion.

This comparison shows that the interaction pattern might be more likely to cause the outbreaks than the single pattern. We considered it is because of the emergence pattern, a type of the interaction pattern. When a new cluster appears at a location, the people there have usually not been infected. That is, they do not have the antibody. The epidemic could spread quickly in a single pattern until most people have the antibody. Although the speed of the spread would thus slow down, the area of the cluster could become much larger than before. Bigger areas provide more opportunity for a cluster to merge with other clusters. During the merge pattern, the people who are located on the juncture of clusters could be infected by the people from any cluster. This complicates the diffusion process because it is hard to differentiate which person is infected by whom. If an infectious person within this big cluster migrates to another distant site, the emergence pattern can happen again. The cycle could also operate again, which intensifies the epidemic rapidly. Consequently, the critical point to prevent the diffusion seems to control the appearance of the emergence pattern. Because dengue fever is a vector-borne disease, the movement of both people and mosquitos should be monitored.

To incorporate the transmission properties of infectious diseases, the MST-DBSCAN algorithm needs four input parameters: *EpsS*, *EpsT1*, *EpsT2*, and *MinPts*. *EpsS* represents the geographical distance for one infected case transmitting disease to a susceptible person. *EpsT1* and *EpsT2* are the maximum and minimum temporal distances, which are calculated from the properties of the incubation and infectious periods. Finally, *MinPts* reflects the minimum number of cases forming a cluster. The issue on specifying these parameters appropriately, especially for *MinPts* in density-based algorithms, was addressed in recent studies (Liu, Zhou, and Wu 2007; Ram et al. 2009; Kriegel et al. 2011). In our study, each parameter of MST-DBSCAN is used for reflecting disease transmission properties and minimum cluster size of a specific disease. In other words, the parameters of MST-DBSCAN can be set up appropriately based on the incubation and infectious periods of a specific disease and the influential range of contact exposure. For example, measles is an airborne infectious disease spreading easily through the coughs and sneezes of those infected. The World Health Organization (2003) reported that the incubation period of measles ranges from ten days to twelve days, whereas the infectious period is from four days before to four days after the onset time. That is, the earliest transmission could happen on the sixth day ($10 - 4 = 6$) after the onset time and the latest transmission could be on the sixteenth day ($12 + 4 = 16$). Thus, *EpsT1* of MST-DBSCAN can be set as two weeks and *EpsT2* could be set to one week if the time unit for analysis is the week. Furthermore, Zipprich et al. (2014) considered that a measles outbreak results from three or more cases linked in time or space, so *MinPts* can be set as two. Finally, *EpsS* represents the possible geographical range of transmission from one infected case to another, which is difficult to measure quantitatively. Spatial pattern analysis methods are often used to quantify spatial influential zones of transmission. In recent studies on dengue fever transmission, Kan et al. (2008) used the K-order nearest neighbor (KNN) analysis to determine the most appropriate geographical bandwidth, which is regarded as the distance of human–mosquito transmission. Moreover, determining spatial influential range of dengue vector mosquitoes, Azil, Bruce, and Williams (2014) used incremental spatial autocorrelation analysis (i.e., global Moran's *I* at numerous distances) to identify the distance at which the spatial autocorrelation of mosquito population is the strongest. It can represent the home range of dengue vector mosquitoes.

This study took into account the effect caused by the incubation period. This was also a concern of previous studies that used mathematical differential equations (e.g., Susceptible-Exposed-Infected-Removed [SEIR] model) or agent-based simulation approaches to profile an epidemic diffusion (Li and Muldowney 1995; Lekone and Finkenstädt 2006; Degli Atti et al. 2008). The former, however, concentrated on the temporal progression of epidemics; the latter often used artificial data for experiments on hypothetical scenarios. The influence of the incubation period is not incorporated into spatial–time clustering methods, so the progression of clustering diffusion cannot be investigated quantitatively. The algorithm we propose fills this gap. In addition, although our main focus was on disease diffusion, the essence of our methods was the cluster evolution

doi:10.6342/NTU202101357

and visualization from any point pattern data. Other diffusion issues that have a time lag property similar to our incubation period, such as crime (Anselin et al. 2000) and innovation (Feldman and Florida 1994), could also be analyzed by our methods.

Although this study has classified, identified, and visualized cluster evolution types, some limitations still exist. First, MST-DBSCAN cannot differentiate the intensity of each identified cluster, whereas KDE can demonstrate the difference between hot spots. In MST-DBSCAN, although the clusters were identified in different areas, the number of cases in them was not used to reflect clusters' status. Some large clusters might be sparse because they contain just a few cases, whereas some small clusters might be dense due to many cases concentrating in a small area. In fact, this limitation is inherited from DBSCAN, which is also related to the aforementioned varying-density problem. Some studies attempted to address this limitation by proposing new methods such as VDBSCAN (Liu, Zhou, and Wu 2007) or an enhanced DBSCAN algorithm (Ram et al. 2009). Our study focuses on the geographic extent of cluster evolution rather than geographic variation within a cluster, such as varying-density phenomena. It still warrants further investigation considering spatial variation within a cluster in MST-DBSCAN, however.

Another limitation is the time complexity of MST-DBSCAN, which is $O(N^2)$, although we have adopted KD-Tree for searching spatial–temporal neighbors. In fact, KD-Tree only can find spatial adjacent neighborhood pairs, which requires our algorithm to further pick out the temporal adjacent pairs among those spatial adjacent pairs. Hence, the number of spatial adjacent pairs, which is proportional to N^2, dominates the time complexity in searching neighbors. The algorithm of Birant and Kut (2007), which also considered temporal neighbor search in DBSCAN, modified R-Tree so that spatial search and temporal search can be implemented simultaneously. The time complexity of their algorithm thus only costs $O(N*\log N)$. It could be an important issue to be further addressed for big-data spatial applications.

Although the focus of this article is an exploratory approach, a statistical delimitation of cluster boundaries is nevertheless very useful. In Sabel et al. (2000), the authors used KDE to produce surfaces of disease over space and time but delimited the cluster boundaries using Monte Carlo simulations. This remains a limitation of the MST-DBSCAN algorithm.

Conclusion

Disease cluster evolution is vital to understanding spatial patterns of epidemic diffusion, but a gap in the literature remains over a clear diffusion cluster classification. This study not only classified the evolution types but also developed an algorithm, MST-DBSCAN, to identify and visualize each cluster at each epoch. With a dynamic data set and the ability to consider the incubation period, the results of MST-DBSCAN more realistically depict the diffusion process. Moreover, because the evolution types that each cluster undergoes can be stored, further quantitative analysis is also available. We thus used this property to simplify the way to visualize a dynamic progression such as diffusion. With only one map and graphs, our simplified approach visualizes an epidemic diffusion in an easy-to-read manner. In summary, our MST-DBSCAN algorithm can profile the spatial–temporal pattern of epidemic diffusion in detail by identifying disease cluster evolution.

This study proposed an MST-DBSCAN algorithm to systematically differentiate the types of disease clustering evolution for identifying and visualizing each cluster at each epoch. Our algorithm, which incorporates dynamic data sets and effect of the incubation period, shows better ability to realistically and efficiently depict the diffusion process than conventional methods, including time-series maps and space–time kernel maps.

Funding

This research was supported by grants from the Ministry of Science and Technology (MOST) in Taiwan (MOST 106-2627-M-002-010; MOST 105-2410-H-002-150-MY3). The authors also acknowledge the financial support provided by the Taiwan Centers for Disease Control and Infectious Diseases Research and Education Center, Ministry of Health and Welfare, and National Taiwan University. The funders had no role in study design, data collection, analysis, or preparation of the article.

References

Anselin, L., J. Cohen, D. Cook, W. Gorr, and G. Tita. 2000. Spatial analyses of crime. *Criminal Justice* 4 (2):213–62.

Azil, A. H., D. Bruce, and C. R. Williams. 2014. Determining the spatial autocorrelation of dengue vector populations: Influences of mosquito sampling method,

doi:10.6342/NTU202101357

covariables, and vector control. *Journal of Vector Ecology* 39 (1):153–63.
Bailey, T. C., and A. C. Gatrell. 1995. *Interactive spatial data analysis*. Essex, UK: Longman Scientific & Technical.
Barreto, F. R., M. G. Teixeira, N. Costa Maria da Conceição, M. S. Carvalho, and M. L. Barreto. 2008. Spread pattern of the first dengue epidemic in the city of Salvador, Brazil. *BioMed Central Public Health* 8(1):51.
Birant, D., and A. Kut. 2007. ST-DBSCAN: An algorithm for clustering spatial–temporal data. *Data & Knowledge Engineering* 60 (1):208–21.
Blondel, V. D., J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and experiment* 2008 (10):P10008.
Brunsdon, C. 2001. The comap: Exploring spatial pattern via conditional distributions. *Computers, Environment and Urban Systems* 25 (1):53–68.
Brunsdon, C., J. Corcoran, and G. Higgs. 2007. Visualising space and time in crime patterns: A comparison of methods. *Computers, Environment and Urban Systems* 31 (1):52–75.
Carroll, L. N., A. P. Au, L. T. Detwiler, T. C. Fu, I. S. Painter, and N. F. Abernethy. 2014. Visualization and analytics tools for infectious disease epidemiology: A systematic review. *Journal of Biomedical Informatics* 51: 287–98.
Chan, M., and M. A. Johansson. 2012. The incubation periods of dengue viruses. *PLoS One* 7 (11): e50972.
Cheng, T., G. Tanaksaranond, C. Brunsdon, and J. Haworth. 2013. Exploratory visualisation of congestion evolutions on urban transport networks. *Transportation Research Part C: Emerging Technologies* 36: 296–306.
Cliff, A. D., and P. Haggett. 2006. A swash–backwash model of the single epidemic wave. *Journal of Geographical Systems* 8 (3):227–52.
Cliff, A. D., P. Haggett, J. K. Ord, and G. R. Versey. 1981. *Spatial diffusion: An historical geography of epidemics in an island community*. Cambridge, UK: CUP Archive.
Degli Atti, M. L. C., S. Merler, C. Rizzo, M. Ajelli, M. Massari, P. Manfredi, C. Furlanello, G. S. Tomba, and M. Iannelli. 2008. Mitigation measures for pandemic influenza in Italy: An individual based model considering different scenarios. *PLoS One* 3 (3):e1790.
Ester, M., H.-P. Kriegel, J. Sander, M. Wimmer, and X. Xu. 1998. Incremental clustering for mining in a data warehousing environment. Paper presented at Very Large Data Base, New York.
Ester, M., H.-P. Kriegel, J. Sander, and X. Xu. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. Paper presented at Knowledge Discovery and Data Mining, Portland, OR.
Feldman, M. P., and R. Florida. 1994. The geographic sources of innovation: Technological infrastructure and product innovation in the United States. *Annals of the Association of American Geographers* 84 (2): 210–29.
Gatrell, A. C., T. C. Bailey, P. J. Diggle, and B. S. Rowlingson. 1996. Spatial point pattern analysis and its application in geographical epidemiology. *Transactions of the Institute of British Geographers* 21 (1):256–74.
Gubler, D. J. 1998. Dengue and dengue hemorrhagic fever. *Clinical Microbiology Reviews* 11 (3):480–96.
Guzman, M. G., S. B. Halstead, H. Artsob, P. Buchy, J. Farrar, D. J. Gubler, E. Hunsperger, et al. 2010. Dengue: A continuing global threat. *Nature Reviews Microbiology* 8:S7–S16.
Jeefoo, P., N. K. Tripathi, and M. Souris. 2011. Spatio-temporal diffusion pattern and hotspot detection of dengue in Chachoengsao Province, Thailand. *International Journal of Environmental Research and Public Health* 8 (1):51–74.
Kalnis, P., N. Mamoulis, and S. Bakiras. 2005. On discovering moving clusters in spatio-temporal data. Paper presented at International Symposium on Spatial and Temporal Databases, Angra dos Reis, Brazil.
Kan, C.-C., P.-F. Lee, T.-H. Wen, D.-Y. Chao, M.-H. Wu, N. H. Lin, S. Y.-J. Huang, et al. 2008. Two clustering diffusion patterns identified from the 2001–2003 dengue epidemic, Kaohsiung, Taiwan. *The American Journal of Tropical Medicine and Hygiene* 79 (3):344–52.
Kriegel, H. P., P. Kröger, J. Sander, and A. Zimek. 2011. Density-based clustering. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1 (3):231–40.
Lekone, P. E., and B. F. Finkenstädt. 2006. Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. *Biometrics* 62 (4): 1170–77.
Li, M. Y., and J. S. Muldowney. 1995. Global stability for the SEIR model in epidemiology. *Mathematical Biosciences* 125 (2):155–64.
Lian, M., R. D. Warner, J. L. Alexander, and K. R. Dixon. 2007. Using geographic information systems and spatial and space–time scan statistics for a population-based risk analysis of the 2002 equine West Nile epidemic in six contiguous regions of Texas. *International Journal of Health Geographics* 6:42.
Liu, P., D. Zhou, and N. Wu. 2007. VDBSCAN: Varied density based spatial clustering of applications with noise. Paper presented at 2007 International Conference on Service Systems and Service Management, Chengdu, China.
Meade, M. S., and M. Emch. 2010. *Medical geography*. New York: Guilford.
Parimala, M., D. Lopez, and N. Senthilkumar. 2011. A survey on density based clustering algorithms for mining large spatial databases. *International Journal of Advanced Science and Technology* 31:59–66.
Ram, A., A. Sharma, A. S. Jalal, R. Singh, and A. Agrawal. 2009. An enhanced density based spatial clustering of applications with noise. Paper presented at the 2009 IEEE International Advanced Computing Conference (IACC 2009), Patiala, India.
Sabel, C. E., A. C. Gatrell, M. Löytönen, P. Maasilta, and M. Jokelainen. 2000. Modelling exposure opportunities: Estimating relative risk for motor neurone

doi:10.6342/NTU202101357

disease in Finland. *Social Science & Medicine* 50 (7–8):1121–37.
Sabel, C. E., D. Pringle, and A. Schoerstrom. 2009. Infectious disease diffusion. In *A companion to health and medical geography*, ed. T. Brown, S. McLafferty, and G. Moon, 111–32. New York: Wiley-Blackwell.
Wang, M., A. Wang, and A. Li. 2006. Mining spatial-temporal clusters from geo-databases. Paper presented at Advanced Data Mining and Applications, Xi'an, China.
Wen, T.-H., and Y.-S. Tsai. 2015. Analyzing the patterns of space–time distances for tracking the diffusion of an epidemic. In *Space–time integration in geography and GIScience*, ed. M.-P. Kwan, D. Richardson, D. Wang, and C. Zhou, 269–82. New York: Springer.
World Health Organization. 2003. *Surveillance guidelines for measles and congenital rubella infection in the European Region of WHO*. Copenhagen, Denmark: WHO Regional Office for Europe.
Zipprich, J., J. K. Hacker, E. L. Murray, D. Xia, K. Harriman, and C. Glaser. 2014. *Notes from the field: Measles—California, January 1–April 18, 2014*. Atlanta, GA: Centers for Disease Control and Prevention.

Appendix A: Pseudo-code

Part 1: MST-DBSCAN

Inputs

1. $D = \{p_1, p_2, \ldots, p_n\}$ Set of points
2. *EpsS*: Maximum spatial distance value
3. *EpsT1*: Maximum time distance value
4. *EpsT2*: Minimum time distance value
5. MinPts: Minimum number of points within *EpsS* and *EpsT1* and without *EpsT2*

Outputs

1. $TC = \{TC_1, TC_2, \ldots TC_k\}$, where $TC_t = \{C_1, C_2, \ldots, C_k\}$ is the set of transmission clusters at time step *t*.

Step 1. Build an edge-weighted digraph (G) whose vertices are *N* and edges are empty. Meanwhile, create an empty dynamic data set (*dd*).

Step 2. Create a KD-Tree for all points (*N*) in *D*.

Step 3. In the KD-Tree, find all neighborhood pairs whose spatial distance between the two points is equal to or smaller than *EpsS*.

Step 4. Among these spatial adjacent neighborhood pairs, pick out the pairs whose time lag is equal to or smaller than the following time: *EpsT1* – *EpsT2*. Additionally, mark the earlier point in each pair as the home point and the later point as the neighbor point. Then, insert these pairs into G as the in-flow edges of their neighbor points.

Step 5. For each time step *t*, repeat Steps 5a through 5f.

Step 5a. Remove the old points from *dd* if they appear before the following time: *t* – *EpsT1*.

Step 5b. From G, remove the out-flow edges whose home point is one of the old points.

Step 5c. Add the new points that appear at time *t* into *dd*.

Step 5d. For each new point, obtain its in-flow edges and the home point of each of these edges. Then, insert these edges into G as the out-flow edges of their home points, respectively.

Step 5e. For each point (*p*) in *dd*, get the out-flow edges of *p* in G. Then, if the number of these edges is equal to or larger than *MinPts*, mark *p* as core.

Step 5f. For each point (*p*) in *dd*, if *p* is core and has not been visited in the current time step *t*, use a depth-first search to find all points that can be spatio-temporally reached by *p* along the outflow direction. Then, assign these points and *p* to the same cluster and mark them as visited. The depth-first search will find one cluster each time and then start the next search from an unvisited core point until there is no core point in *dd*. Moreover, a core point being visited a second time or even more implies that it can be reached by two or more clusters, so the points belonging to these clusters will be assigned to the same cluster.

Part 2: Identify Cluster Evolution Types

Inputs

1. $TC = \{TC_1, TC_2, \ldots, TC_k\}$, where $TC_t = \{C_1, C_2, \ldots, C_k\}$ is the set of transmission clusters at time step *t*.

Outputs

1. $TEC = \{TEC_1, TEC_2, \ldots, TEC_k\}$, where $TEC_t = \{EC_1, EC_2, \ldots, EC_k\}$ is the set of evolution types of each cluster at time step *t*.

doi:10.6342/NTU202101357

Step 5. Calculate the mean geographical center and the area of each cluster at each time step t.

Step 6. For each time step t, repeat Steps 6a through 6b.

Step 6a. Compare each cluster in TC_t with each cluster in $TC_{t\text{-}1}$ to know whether clusters in different times share the same points.

Step 6b. Start to identify the evolution types:

1. If the points of a cluster in TC_t are not from each cluster in $TC_{t\text{-}1}$, mark the cluster emerge.
2. If these points are from a part of a cluster in $TC_{t\text{-}1}$, mark this cluster split.
3. If these points completely contain two or more clusters in $TC_{t\text{-}1}$, mark this cluster merge.
4. If these points not only completely contain two or more clusters in $TC_{t\text{-}1}$ but also contain points from a part of other clusters, mark this cluster split-merge.
5. If these points are from a part of two or more clusters, respectively, mark this cluster split-merge.
6. If these points completely inherit a cluster in $TC_{t\text{-}1}$, decide the single type of this cluster by comparing the difference of the area and the center between two clusters.

Part 3: Simplified Time-Series Maps

Inputs

1. A set of polygons of administrative districts.
2. $TEC = \{TEC_1, TEC_2, \ldots, TEC_k\}$, where $TEC_t = \{EC_1, EC_2, \ldots, EC_k\}$ is the set of evolution types of each clusters at time step t.

Outputs

1. The grouping result of the polygons

Step 7. For each polygon at each time step t, record the transmission clusters that intersect the polygon and record their evolution types.

Step 8. For any two polygons, calculate their similarity, which is the proportion of the week that they share the same evolution types.

Step 9. Build a network with nodes that are polygons and links that are the similarities between polygons.

Step 10. Apply the Louvain method on the network to group polygons.

Appendix B: Evaluation of the Time Complexity of MST-DBSCAN

Time complexity analyses begin with building an edge-weighted digraph that only has vertices but no edge (Step 1) according to $O(N)$. Meanwhile, building a KD-Tree (Step 2) entails $O(N\log N)$. Then, it takes $O(N\log N)$ to find all spatial adjacent neighborhood pairs (Step 3). Furthermore, the cost of picking out spatio-temporal neighborhood pairs among the preceding spatial adjacent pairs (Step 4) is $O(N^2)$ because we must check the time lag of the spatial adjacent pairs one by one, and the number of these pairs is proportional to $N(N-1)/2$. Step 5 is a repeated process. At each time step, all points in the dynamic data set must be checked once to remove old points (Step 5a). That is, a point is checked $EpsT1$ times during the whole process because $EpsT1$ is the time that a point can exist in the dynamic data set. Thus, Step 5a's time complexity is $O(N \times EpsT1)$. Moreover, because the old points at each time step have been picked out by Step 5a, removing old pairs, the out-flow edges of these old points, from G (Step 5b) only costs $O(N)$. Step 5c is to add newly appearing points into the dynamic data set. If storing all points by a dictionary data structure whose keys are all time steps, adding new points at each time step t only entails a constant time. Therefore, the total complexity of Step 5c is $O(T)$, where T is the duration of the data. In Step 5d, the total number of out-flow edges is equal to that of in-flow edges, so the time complexity is $O(N^2)$. The time complexity of Step 5e is same as that of Step 5a, which is $O(N \times EpsT1)$ because all points in the dynamic data set also have to be checked once at each time to find core points. Finally, the depth-first search will go through all edges in G at each time step. Because each edge can maximally exist in G for the time, $EpsT1 - EpsT2$ and the number of edges is proportional to $N(N-1)/2$, and the time complexity in Step 5f is $O(N^2 \times (EpsT1 - EpsT2))$. In summary, because the most

doi:10.6342/NTU202101357

important part of time complexity is the term that has the highest order, we concluded that the time complexity of MST-DBSCAN is $O\left(N^2 \times (EpsT1 - EpsT2)\right)$.

FEI-YING KUO is a PhD student in the Department of Geography, National Taiwan University, Taipei 10617, Taiwan. E-mail: fykuofelix@gmail.com. His research interests include diffusion dynamics of infectious diseases and human mobility in urban spatial structures.

TZAI-HUNG WEN is Professor in the Department of Geography, National Taiwan University, Taipei 10617, Taiwan. E-mail: wenthung@ntu.edu.tw. His research interests include modeling geospatial processes in the human environment and developing geocomputational methods for analyzing multidimensional spatial–temporal data from mobile geosensors, social media, and open data platforms.

CLIVE E. SABEL is Professor in the Department of Environmental Science, Aarhus University (Roskilde), Aarhus 8000, Denmark. E-mail: cs@envs.au.dk. He is a spatial scientist working with point-pattern data (often residential location) to reveal epidemiological relationships to environmental exposures or patterns in road traffic accidents and building whole life-course exposures to social and environmental sources to understand well-being in urban areas.

doi:10.6342/NTU202101357

附錄 B

作者：Fei-Ying Kuo, and Tzai-Hung Wen

文章標題：Regionalization for infection control: An algorithm for delineating containment zones considering the regularity of human mobility

發表年分：2021

期刊名稱：Applied Geography

卷號：126

文章編號：102375

數位物件識別號碼（DOI）：10.1016/j.apgeog.2020.102375

Applied Geography 126 (2021) 102375

Contents lists available at ScienceDirect

Applied Geography

journal homepage: http://www.elsevier.com/locate/apgeog

Regionalization for infection control: An algorithm for delineating containment zones considering the regularity of human mobility

Fei-Ying Kuo, Tzai-Hung Wen *

Department of Geography, National Taiwan University, Taiwan

ARTICLE INFO

Keywords:
Epidemic diffusion
Containment zone
Mobility regularity
Land use
Network community

ABSTRACT

Restricting human movement to decrease contact probability and frequency helps mitigate large-scale epidemics. Movement-based zoning can be implemented to delineate the boundaries for movement restrictions. Previous studies used network community detection methods, which capture cohesive within-region movements, to delineate containment zones. However, most people usually travel and spend most of their time in several fixed locations, which implies that an infected person could transmit the pathogens to only a specific group of people with whom s/he usually has a contact in frequently-visited locations. Existing network community detection methods cannot reflect the regularity of the flow of people; thus, this study aims to use land-use patterns to reflect trip purposes to measure the regularity of human mobility. We propose a novel network community detection method, the Human Mobility Regularity-based Zoning (HuMoRZ) algorithm, to delineate containment zones incorporating mobility regularity. The Taipei metropolitan area in Taiwan is used to demonstrate the feasibility of the proposed algorithm. The spatial diffusion of an emerging respiratory disease, novel influenza A/H1N1, is simulated for comparing three different quarantine zoning systems: (1) a minimum zoning unit, (2) optimal zoning without considering mobility regularity, and (3) optimal zoning considering mobility regularity. Two epidemiological performance indicators are used to compare simulation results: namely, the accumulated infected number (AN) on the 30th day, reflecting the severity of an epidemic, and the critical time (CT), the moment at which half of the population becomes infected, measuring the diffusion speed of an epidemic. To measure the variety of different facility types within a containment zone, we further use Shannon's entropy scores, representing a self-contained zone, and the boxplot of all zones' entropy scores, reflecting geospatial homogeneity of life functions across zones. Our results suggest that containment zones that incorporate mobility regularity could significantly delay the epidemic peak and critical time and decrease the severity of an epidemic. The zoning patterns proposed in our algorithm could also allow for more life functions in a zone and more evenly distributed life resources across zones than those of zones generated by other methods. These findings could provide insightful implications for fighting the COVID-19 pandemic.

1. Introduction

Epidemic diffusion is usually aggravated by human or animal movements because their movements increase the contact probability and the frequency between infected entities and healthy entities (Fèvre, Bronsvoort, Hamilton, & Cleaveland, 2006; Mao, 2014). Thus, restricting their movements to decrease such probability and frequency helps control and mitigate epidemics (Bajardi et al., 2011; Velthuis & Mourits, 2007). In practice, movement restriction is a common measure used to earn time to develop effective vaccines or drugs when an emerging infectious disease appears (Gensini, Yacoub, & Conti, 2004). The effectiveness of movement restrictions against global epidemics, including influenza A/H1N1 (Cooper, Pitman, Edmunds, & Gay, 2006; Epstein et al., 2007), SARS (T. Day, Park, Madras, Gumel, & Wu, 2006; Fraser, Riley, Anderson, & Ferguson, 2004), and Ebola virus (Peak et al., 2018), has also been proven in previous studies. Movement restrictions are usually separated into two different modes: isolation and quarantine (Barbera et al., 2001; T.; Day et al., 2006; Fraser et al., 2004). Both of these two modes control the movements of infected entities, but quarantine also restricts high-risk asymptomatic entities, who may have contacted infected entities, to their living neighborhoods. However, when the number of confirmed cases grows exponentially, it is difficult

* Corresponding author.
E-mail address: wenthung@ntu.edu.tw (T.-H. Wen).

https://doi.org/10.1016/j.apgeog.2020.102375
Received 30 April 2020; Received in revised form 3 November 2020; Accepted 24 November 2020

doi:10.6342/NTU202101357

to comprehensively trace all contacts for quarantine and further epidemiological investigations (Fong et al., 2020; Peak et al., 2020). In this situation, quarantine tends to focus on delineating a certain geographical area to cover all possible contacts. This mass quarantine approach, also named containment, is the strictest approach to preventing epidemic diffusion (Wilder-Smith, Chiew, & Lee, 2020).

Large-scale containment was implemented to control the diffusion of the COVID-19 epidemic in 2020. The Chinese government progressively implemented city lockdowns in Wuhan and nearby cities (Wu, Leung, & Leung, 2020). However, citywide containment could raise two concerns. First, widespread containment may frighten residents and cover many individuals with no significant risk of diseases (Barbera et al., 2001; Rubin & Wessely, 2020); thus, it could cause social panic and increase the infection probability of low-risk individuals within the city. High social panic may prompt people to resist quarantine policies or to flee (Coltart, Lindsey, Ghinai, Johnson, & Heymann, 2017), which generates loopholes in epidemic prevention. Second, containment requires enormous resources, including medical and security services, to sustain the lives of people in containment areas (Barbera et al., 2001). Although China has successfully implemented citywide containment in Wuhan, this approach may not be feasible in many other countries because few countries can completely imitate China's strategies due to different political systems (Kupferschmidt & Cohen, 2020).

On the other hand, megacities in China, such as Beijing and Shanghai, implemented another, less strict containment policy during the COVID-19 pandemic: "closed-off management" (Shao, 2020). A microdistrict (e.g., residential unit) is regarded as a control unit for restricting people's movement when a confirmed case appears in this microdistrict. However, it is difficult to effectively contain an epidemic to a small-scale containment zone, such as microdistrict closed-off management, because exposed contacts may occur in different microdistricts (Barbera et al., 2001). Furthermore, some infectious diseases, such as COVID-19 or MERS, can be transmitted asymptomatically (Omrani et al., 2013; Rothe et al., 2020). Thus, before being quarantined, an infected person may transmit infectious pathogens to many other individuals. These exposed individuals, who are not in containment zones, may become the next spreaders. In summary, neither the citywide level nor the microdistrict level may be the appropriate spatial scale for containment. An appropriate containment zone should include the places where an infected person usually stays and the places where exposed individuals frequently visit; it should also efficiently sustain the lives of people in quarantine.

Recent literature developed a movement-based zoning approach to delineating containment zones to confront foot-and-mouth disease (FMD) (G.-J. Lee, Pak, Lee, & Hong, 2019). A network structure was implemented to formulate livestock movement, where the nodes were livestock facilities, and the links were the amount of livestock transportation among different facilities. Then, a network community detection method was used to group livestock facilities into several communities. A community is a group of certain nodes that strongly interact with each other but weakly interact with the nodes in other communities (Malliaros & Vazirgiannis, 2013). If an outbreak of FMD occurs in a livestock facility, a network community could capture the infected facility and other high-risk livestock facilities. Thus, a network community can be used for delineating containment zones for FMD outbreaks.

Similar to livestock transportation, human movements from any one place to another could also be formulated as a network. This network reflects the aggregated movement behavior of a group of people, and a zone delineated by existing network community detection methods captures intensive human movements (Liu, Gong, Gong, & Liu, 2015; Zhong, Arisona, Huang, Batty, & Schmitt, 2014). A delineated community could include all possible contacts once an infected person identified, and this community may be a zone mixing high-risk and low-risk contacts. Thus, the size may be too large to be an appropriate containment zone for restricting human movement. Theoretically, in an urban area, a person may regularly encounter identical people, including his or her friends and some familiar strangers (Sun, Axhausen, Lee, & Huang, 2013). The reason is that most people usually travel short distances and return to a few frequently visited locations, such as their homes or workplaces, with a high probability (González, Hidalgo, & Barabási, 2008; Song, Koren, Wang, & Barabási, 2010). In other words, most people usually travel and spend most of their time in several fixed locations (Song, Qu, Blumm, & Barabási, 2010), thereby representing a spatially constrained movement pattern (Belik, Geisel, & Brockmann, 2011). Therefore, the regularity of flows of people implies that an infected person, before being quarantined, could transmit pathogens to only a specific group of people with whom s/he usually has contact in frequently visited locations.

Since existing network community detection methods cannot reflect the regularity of the flow of people, this study aims to incorporate the regularity of urban human mobility within a city into network community detection to delineate better containment zones for disease control. We propose a novel network community detection method, the Human Mobility Regularity-based Zoning (HuMoRZ) algorithm, for considering urban-scale daily routines. It is based on the map equation algorithm, a commonly used community detection method for directed and weighted graphs (Rosvall & Bergstrom, 2008). We use the population flow of the Taipei metropolitan area, one of the major East Asian cities, as a case study to demonstrate the effectiveness and feasibility of the proposed algorithm. An epidemic diffusion model is used to simulate the spatial dynamics of disease transmission with different movement restriction scenarios to compare the performance of zoning algorithms. As delineating containment zones for movement restrictions are usually considered to be implemented to block the spread of high-infectivity diseases, our simulation focuses on the intervention scenarios of high-infectivity disease transmission.

2. Methods

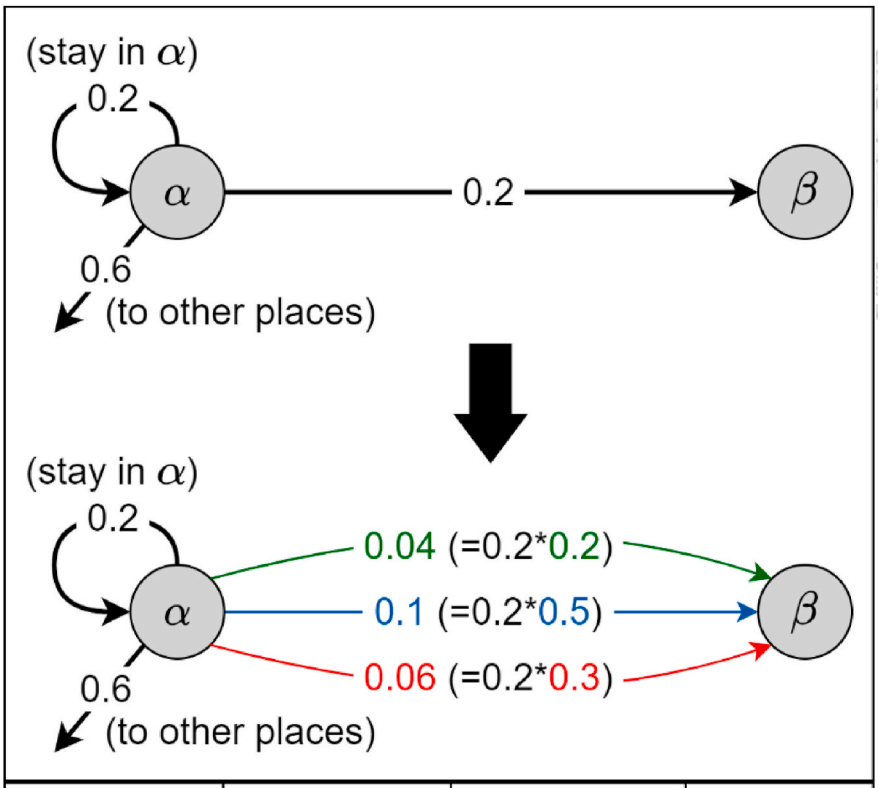
2.1. Regularity-based partitioning algorithm

We propose a mobility regularity-based zoning algorithm for detecting network communities. The concept is based on the random walker's long-term movement behaviors among the nodes of a human movement network to evaluate any given pattern of communities. In addition to strong interactions within a community and weak interactions between communities, our objective also pursues high levels of regular within-community movements. We adopt the optimization approach of the map equation algorithm to optimize the number of communities to which each node belongs. The procedure involves finding new network community patterns by following searching strategies and checking whether new patterns can improve the objective value. The final result is the one whose objective value is the smallest and cannot be further improved by new patterns. For detailed descriptions of the optimization procedure, refer to Rosvall, Axelsson, and Bergstrom (2009).

The proposed zoning algorithm incorporates mobility regularity into the objective function to find the optimal patterns of network communities. Regularity describes individuals' daily routine movement behaviors such as commuting or grocery shopping (Schneider, Belik, Couronné, Smoreda, & González, 2013). However, a population flow network cannot reflect individual-level trajectories, movement behaviors, and trip purposes. Recent studies have shown that land-use functions can be used to reveal trip purposes (M. Lee & Holme, 2015; Maat, Van Wee, & Stead, 2005); thus, we adopt the movement patterns within land-use functions as a proxy to measure mobility regularity. The details of measuring mobility regularity are as follows.

In Fig. 1, we illustrate the concept of estimating the probability of each specific purpose when the random walker moves from one place to another. The number 0.2 is the conditional probability that the random walker moves from α to β given that it is currently in α. This probability

doi:10.6342/NTU202101357



Land use type	Residential	Commercial	Industrial
Area proportion in β	0.2	0.5	0.3
Arrow mark	→	→	→

Fig. 1. An example demonstrating how we use different land-use types to reflect trip purposes in a human movement network. Suppose that the random walker moves to β with probability 0.2 when it leaves α; we separate this probability into three parts based on the area proportions of three different land-use types in β. Therefore, 0.04, 0.1, and 0.06, respectively, represent the probabilities of residential, commercial, and industrial purposes for the movement of the random walker from α to β.

is then separated into three parts based on the proportion in the area of three different land-use types in β. Each part represents a conditional probability of a specific trip purpose. Eq. (1) and Eq. (2) denote the general form of this estimation.

$$p^k_{\alpha\beta} = p_\alpha \omega_{\alpha\beta} z^k_\beta \tag{1}$$

$$\sum_{k=1}^{K} z^k_\beta = 1, \tag{2}$$

where α is an origin and β is a destination. p_α is the probability that the random walker visits node α. $\omega_{\alpha\beta}$ is the conditional probability that the random walker goes to β, given that it is currently in α. z^k_β is the area proportion of the kth land-use type in β. Here, $p_\alpha \omega_{\alpha\beta}$ is the total probability moving from α to β, and we separate this probability into K parts based on the area proportion of K land-use types so that each part can represent the probability of a specific trip purpose.

In this study, we focus on within-community movements to measure mobility regularity that results from daily routines. In other words, people repeatedly go to one place for the same purpose. For example, most people have only one fixed location for work, and people usually go to a few familiar restaurants to have a meal. Therefore, mobility regularity implies that travel for a specific purpose originating from one place should not be evenly distributed among other destinations but rather concentrate on just some of them. Our measurement is based on this implication.

Fig. 2 compares two different zoning scenarios. In scenario (A), when the random walker leaves node α, the probability of each travel purpose concentrates on one place. In contrast, each probability is evenly distributed on two destinations in scenario (B). To measure this concentration, we adopted Shannon's entropy score. A low value reflects a high level of concentration (Eq. (3) and Eq. (4)); furthermore, Eq. (5) denotes the measurement of mobility regularity.

$$\mathrm{H}\left(\vec{z}^k_\alpha\right) = -\sum_\beta \frac{p^k_{\alpha\beta}}{p^k_\alpha} \log_2 \left(\frac{p^k_{\alpha\beta}}{p^k_\alpha}\right), \ \alpha, \beta \in \text{the same community} \tag{3}$$

doi:10.6342/NTU202101357

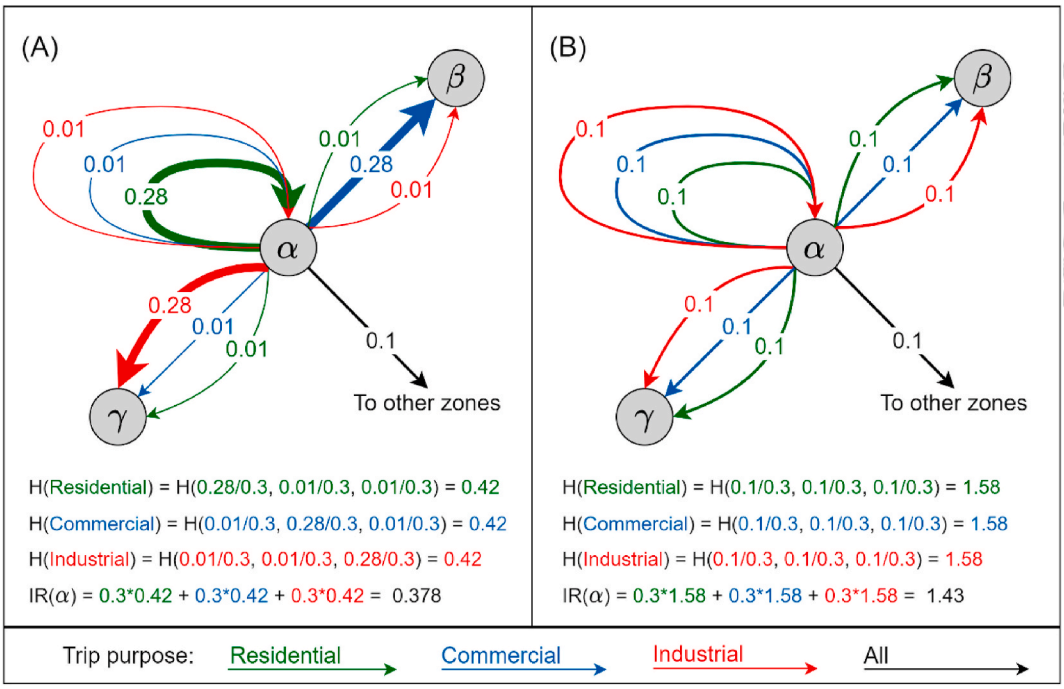


Fig. 2. An example demonstrating how we use entropy to measure mobility regularity for each node. In scenario (A), α, β, and γ compose a zone. When the random walker leaves α and goes to β or γ (i.e., it still stays in this zone) for a specific trip purpose, entropy helps to determine whether the random walker tends to go to one place or go to β and γ on average. The random walker tends to go to β for commercial purposes and go to γ for industrial purposes. This pattern displays typical movement behaviors. On the other hand, in scenario (B), α, δ, and ε compose another zone. For every trip purpose, no clear tendency indicates the preferred destinations when the random walker leaves α, so a regular pattern of movements does not exist. Our inverse-regularity index (IR(α)) in scenario (A) is lower than the index value in scenario (B), so this index can help indicate whether a within-zone movement pattern is regular.

$$p_\alpha^k = \sum_\beta p_{\alpha\beta}^k,\ \alpha,\ \beta \in \text{the same community} \tag{4}$$

where p_α^k represents the intracommunity probability that the random walker leaves α but stays within the same community because of the kth trip purpose. Note that we utilize only the flow originating from to its groupmates, the nodes belonging to the same community as α, to measure the entropy. Additionally, β can equal α if self-loops exist.

Each node contains at most K entropy values; thus, we use a weighted entropy score representing the compound effect of these K values. For each trip purpose, we weight its entropy value ($H(\mathscr{Z}_\alpha^k)$) by its intracommunity probability (p_α^k) and then summarize these weighted entropy values (Eq. (5)).

$$\mathrm{IR}(\alpha) = \sum_{k=1}^{K} p_\alpha^k \mathrm{H}(\mathscr{Z}_\alpha^k) \tag{5}$$

When the IR(α) value is larger; the movement behaviors of the random walker originating from α for each trip purpose show more destinations within a community. In contrast, a smaller value of IR(α) reflects that the movement behavior toward one specific destination is more concentrated, which represents mobility regularity. Therefore, IR(α) can be regarded as the *inverse regularity level*.

Finally, combining the original objective function of the map equation (L(M)) with the summation of the inverse regularity level of all nodes yields the objective function of our new algorithm (Eq. (6)). We minimize the value of the objective function to find the optimal pattern of network communities.

$$L_z(\mathrm{M}) = \mathrm{L}(\mathrm{M}) + \sum_\alpha IR(\alpha) \tag{6}$$

To compare the zoning performance of HuMoRZ algorithm, we compare our results with the random zoning method, which randomly assigns locations into different zones with the same zones number of HuMoRZ's result. The random zoning is repeatedly simulated 99 runs to capture the possible combinations of random assignment.

2.2. Comparing the performance of zoning methods

To check whether our HuMoRZ algorithm can delineate an effective zoning system for infectious disease containment, we compare three different zoning methods: (1) a minimum zoning unit (a traffic analysis zone (TAZ), noted as "TAZ-scale"), (2) optimal zoning without considering mobility regularity (the result of the map equation, noted as "without regularity"), and (3) optimal zoning considering mobility regularity (the result of our proposed HuMoRZ algorithm, noted as "with regularity"). Two dimensions are evaluated for the comparison. The first is the effect of different zoning methods on slowing down epidemic progression; the other is the level of self-containment (diversity of life functions) that different zoning methods allow zones to possess.

To evaluate the effect on blocking contagion spread in an epidemic, we simulate the spatial diffusion of an emerging respiratory disease, novel influenza A/H1N1. The simulation is implemented under different epidemic scenarios by a four-state susceptible-latent-infective-removed (SLIR) model, modified from Wen, Hsu, Sun, Jiang, and Juang (2018), as shown in Eq. (7) to Eq. (11).

$$\frac{dS_i(t)}{dt} = -\lambda_i(t)S_i(t) - \sum_{j\neq i}\lambda_j(t)S_i(t)\frac{W(j,i)}{N_j} \tag{7}$$

$$\frac{dL_i(t)}{dt} = \lambda_i(t)S_i(t) + \sum_{j\neq i}\lambda_j(t)S_i(t)\frac{W(j,i)}{N_j} - \theta L_i(t) \tag{8}$$

$$\frac{dI_i(t)}{dt} = \theta L_i(t) - \gamma I_i(t) \tag{9}$$

$$\frac{dR_i(t)}{dt} = \gamma I_i(t) \tag{10}$$

doi:10.6342/NTU202101357

$$\lambda_i(t) = \beta \times \frac{I_i(t)}{N_i} \tag{11}$$

where $S_i(t)$, $L_i(t)$, $I_i(t)$, $R_i(t)$ are the number of susceptible, latent, infective, and removed individuals at time t and place i. N_i is the population number of place i, and $W(j,i)$ is the amount of human movement from place j to place i. β is the average transmission rate defining the transmissibility of a disease; θ is the latent rate determining how fast a latent case can infect others; γ is the removed (recovery) rate representing the speed of transformation from infective to removed. We adopt the same three parameters values used in Wen et al. (2018): $\beta = 0.585/day$, $\theta = 0.32$/day, and $\gamma = 0.09$/day. We assume that the subpopulation number (N_i) in each place i does not change over time, which reflects urban daily routine movement such as commuting or going shopping. Most people moving to other places will return to their residential places within one day; thus, the subpopulation number may not significantly vary between different days. Recent studies also adopted this similar assumption (Gatto et al., 2020; Tang et al., 2020). Under this assumption, our Eqs. (7) and (11) capture the dynamics of possible pathogen transmissions when infected people move to other places in the daytime.

To examine all possible scenarios with different initial outbreak locations, we change the initial outbreak location by turns and repeat the epidemic simulation (a total of 498 runs, reflecting the number of locations in our study). In each simulation, the initial condition is composed of the initial outbreak location and five latent locations. The latent locations represent the locations with the undetected infected persons and these locations strongly interact with the origin of the outbreak. In our simulation, the quarantine intervention is implemented on only one specific containment zone where the initial outbreak occurs. Three different quarantine levels (80%, 90% and 99%) are tested to reflect different strictness of movement restrictions. The 99% level means that only 1% of the original travel amount is allowed to enter and leave the containment zone and is regarded as the strictest level in this study. Two epidemiological performance indicators are used to assess each simulated result. The first one is the accumulated infected number (AN) on the 30th day, reflecting the severity of an epidemic. Recent studies also adopted this similar indicator (Malavika et al., 2020). The other indicator is the critical time (CT), at which half of the population becomes infected. It reflects the diffusion speed of an epidemic.

Finally, we collect all categories of point-of-interest (POI) locations, such as government agencies, cultural and educational facilities, medical and funeral facilities, public and memorial places, life function facilities such as convenience stores and supermarkets, and transportation facilities. Every category is further divided into several different types, and a total of 83 subtypes are incorporated into our study. Shannon's entropy score is used to capture various types of POIs within a containment zone (Eq. (12)).

$$\mathrm{H}(P_i) = -\sum_k \frac{n_i^k}{N_i} \log_2 \left(\frac{n_i^k}{N_i}\right) \tag{12}$$

where n_i^k is the number of the kth POI type in zone i, and N_i is the total POI number in zone i.

A containment zone with a high entropy value indicates that the POIs within this zone could provide various life functions, thereby representing a high level of the self-contained zone. Thus, an appropriate zoning pattern should ensure most zones with high entropy values so that residents could maintain their everyday lives as much as possible, even being quarantined.

3. Case study: Taipei containment zones

The Taipei metropolitan area in Taiwan is used to demonstrate the feasibility of the proposed HuMoRZ algorithm. The Taipei metropolitan area is the political and economic center of Taiwan. This area covers approximately 2457.13 square kilometers of land and has a population of approximately 7.03 million, which is 30% of the total population of Taiwan. Public transportation systems include one international airport, three HSR (high-speed rail) stations, thirteen train stations, 125 mass rapid transit (MRT) stations, approximately 600 bus lines, and over 1000 public bicycle sharing (PBS) stations (Ministry of Transportation and Communications in Taiwan, 2020).

The land use data are from the National Land Surveying and Mapping Center (NLSC) of Taiwan for 2012. The spatial extent of this dataset covers the major part of the Taipei metropolitan area, specifically, approximately 1362.57 square kilometers of land and a population of approximately 6.34 million. This dataset is a three-level hierarchical classification system whose primary classification contains nine categories: agriculture, forest, transportation, water conservation, building, public facility, recreation, strip mining, etc. We adopt the finest-scale classifications because they reflect the most detailed human activity categories and trip purposes. A total of 79 land-use types are incorporated in this study.

The human mobility data are from the fourth version of the "Taipei Rapid Transit Systems Demand Model" (TRTS-IV), an official survey conducted by the Department of Rapid Transit Systems (DORTS) of the Taipei City Government in 2009. This dataset divides the entire Taipei metropolitan area into 508 traffic analysis zones (TAZs). The average area of a TAZ is 2.74 km^2, with an average population of 12,732 people. There are a total of 498 TAZs covered in our study, including downtown and suburban areas (Fig. 3). The transport authority investigates the daily number of people traveling from any one TAZ to another. It also provides the estimated amounts of travel in 2012 (Fig. 4); thus, these data correspond with the land use dataset on both spatial and temporal dimensions.

4. Results

Based on the human traveling data and land use data, our HuMoRZ algorithm separates 498 TAZs into 26 containment zones, averagely covering 19 TAZs with an average population density of approximately 8500 people per square kilometer (Fig. 4). Fig. 5 further displays the histogram of TAZ numbers and population density among 26 containment zones. It shows that only two zones are composed of over thirty TAZs, and eight zones possess a high population density of over 10,000 people per square kilometer. Moreover, the comparison between our zoning result and the random zoning method in Fig. 6 demonstrates that HuMoRZ can also reveal pivotal characteristics of network communities: strong within-community and weak between-community interactions.

Fig. 7A shows the spatial distribution of different boundary types of containment zones delineated by HuMoRZ. It shows that the main roads (avenues) partition different zones in downtown areas, while the suburban areas are partitioned by natural landforms, including mountain ridges and rivers. Moreover, Fig. 7B shows that the significant boundary types are Ridge, River, or Road. Fig. 8 further demonstrates the spatial distribution and the length proportion of different boundary types of TAZs. By comparing the proportions of boundary types between Figs. 7 and 8, the difference of boundary types indicates that the zoning result of HuMoRZ could highlight more barrier effects, including natural boundaries (River, and Ridge) and artificial boundaries (Road), to restrict human movement than TAZs.

Fig. 9 compares the zoning patterns generated by our HuMoRZ algorithm and the existing method, the map equation algorithm. The difference in zoning patterns arises from whether mobility regularity is considered in an algorithm. The map equation only considers flow intensity and cycling performance. Thus, the result shows that all downtown areas become one huge zone covering a large number of TAZs; the other TAZs located in peripheral urban areas constitute several small zones, which reflect isolated lifestyles in remote areas. The zoning pattern resulting from the map equation in Fig. 9B is its finest scale; in other words, the map equation cannot further divide the huge zone into

doi:10.6342/NTU202101357

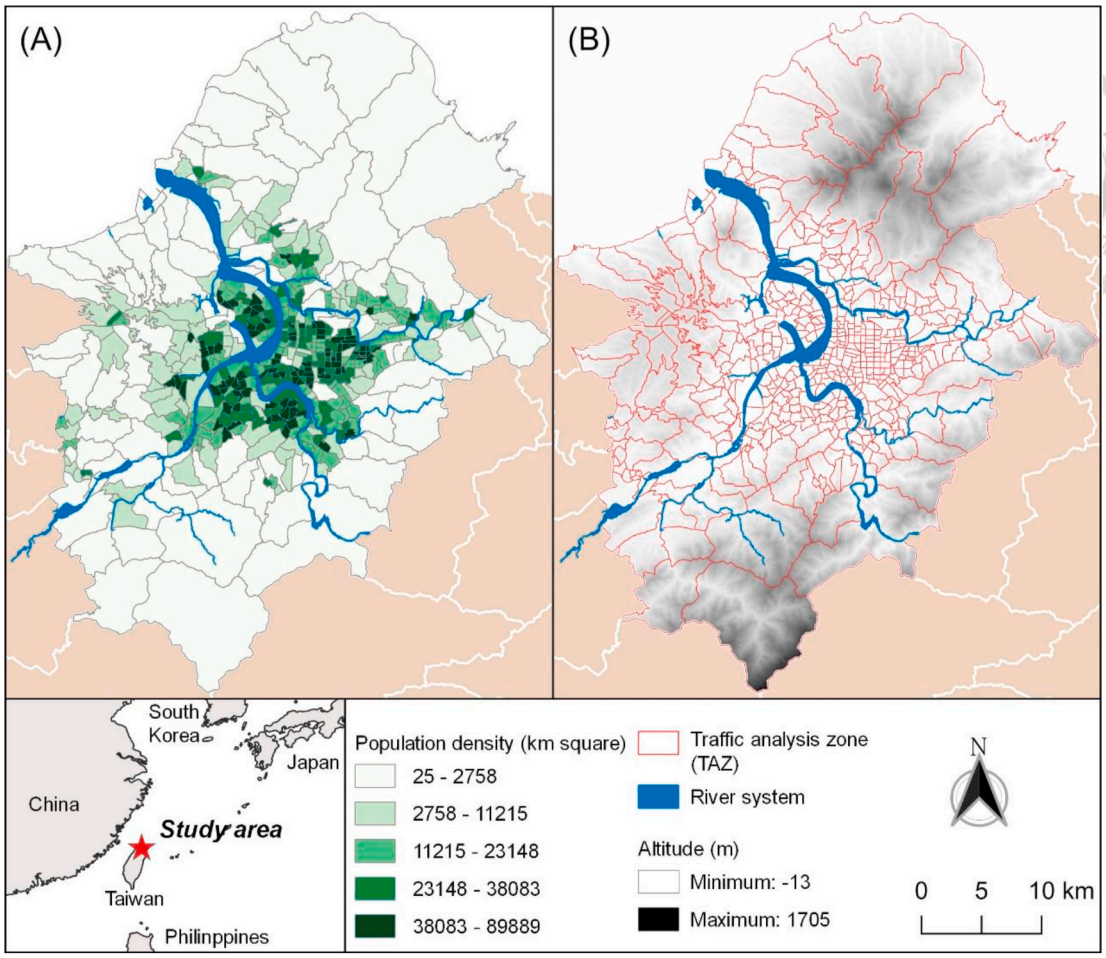


Fig. 3. (A) Population density and (B) landform of the study area.

different smaller zones. In contrast, our algorithm produces similar zoning patterns in the peripheral areas but partitions central downtown areas into several compact zones, each of which is connected by higher within-zone flows (Fig. 4).

We compare the progression dynamics of a respiratory disease based on three containment zoning modes and distinct quarantine levels in Fig. 10. The zoning pattern from the map equation, which does not consider mobility regularity, is analogous to a citywide lockdown because it groups the whole downtown area into one huge zone. In contrast, every TAZ as a containment zone is similar to microdistrict closed-off management. The scale of the zones generated by our HuMoRZ algorithm falls between that of these two zoning systems. The results show that our zoning mode can more effectively slow down the spread of an epidemic than the other two zoning modes. The effect of delaying the epidemic peak is more significant as the quarantine level becomes stricter. Under the strictest quarantine level (Fig. 10C), our zoning mode can delay the peak by approximately one week. Fig. 11 shows the comparison between three zoning modes based on epidemic severity (accumulated infected number on the 30th day, denoted as AN) and spreading speed (the time it takes to for half of the population to be infected, noted as CT). Through comparing the scenarios with different initial outbreak locations and quarantine levels, Fig. 11 indicates that the zoning result of HuMoRZ shows the lowest AN and the highest CT among the three zoning modes.

Fig. 12 displays distributions of POI entropy values of three different zoning modes, respectively. It shows that the regularity-based containment allows most zones to possess higher life functions (higher entropy value) than the microdistrict containment (TAZ-scale). Moreover, compared to the citywide containment (zoning without regularity) or the microdistrict containment, the regularity-based containment possesses the lowest spatial inequality of the life function diversity because of the shorter interquartile range (IQR) of the entropy values than other methods. In other words, neither microdistrict containment nor citywide containment can provide appropriate life functions within every zone and raises the issues of spatial inequality of life resources.

5. Discussion

In this study, we used land-use patterns to reflect trip purposes for measuring human mobility's regularity and proposed the HuMoRZ algorithm, which incorporates mobility regularity into network community detection, to delineate containment zones for infection control. Our results demonstrate more compact zones delineated by the HuMoRZ algorithm than those generated by existing algorithms and identify different types of barriers to restricting human movement. In other words, artificial infrastructures and natural landforms essentially shape the geographic extent of human mobility. Moreover, the results also show that our zoning system can significantly reduce epidemic severity in the early stage and has a positive effect on delaying the epidemic peak. Our movement restriction scenario helps figure out that quarantining only the initial outbreak location is not enough to contain an epidemic effectively. Our results also demonstrate that large-scale containment, such as quarantining all the city, is also not the optimal measure to contain an epidemic. It is challenging to implement effective individual quarantine for continuously increasing case incidence in the whole city. Therefore, our results show that containment zones

doi:10.6342/NTU202101357

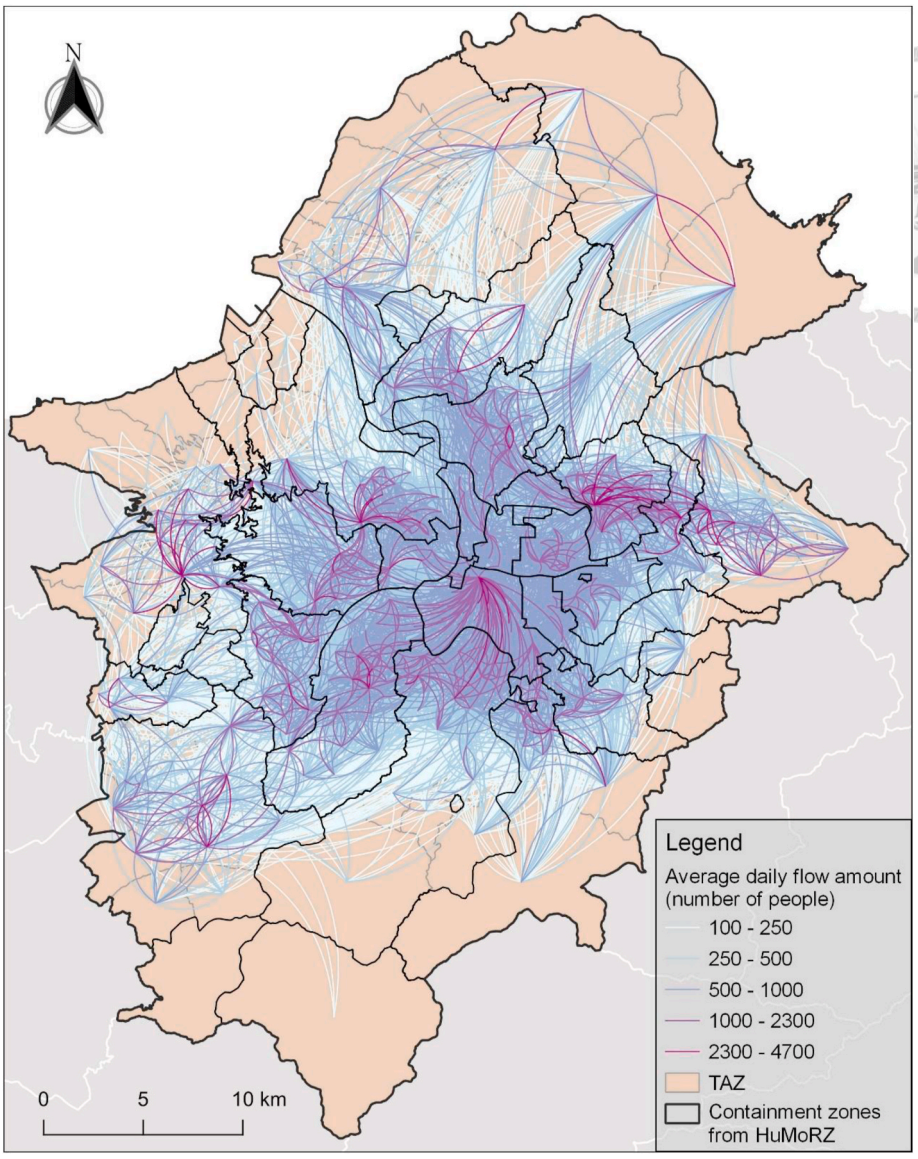


Fig. 4. The zoning result of HuMoRZ algorithm. The flows whose amount is less than one hundred are removed in this figure for better visualization. These flows are still used in analysis and zoning.

incorporating mobility regularity have the optimal effect on delaying peak time and decreasing the severity of an epidemic. Furthermore, the extent of our containment zones can also reflect the high diversity of POI types, allowing for various life functions and geographic homogeneity of life resources among zones. In sum, HuMoRZ's zoning pattern shows better outcomes than other existing containment approaches in terms of allowing a diversity of life functions and delaying epidemic peaks.

The sizes of the containment zones delineated by the HuMoRZ algorithm may also be more appropriate than those generated by existing methods. Barbera et al. (2001) argued that both large-scale and small-scale containment are not effective approaches. Large-scale containment, such as city lockdown represented by our without-mobility-regularity simulation scenario, makes it difficult to support the whole population in a city with medical and human resources. While this approach was successful in confronting the COVID-19 epidemic in Hubei Province, China, scholars argued that this approach might not be feasible in other countries because of different political institutions (Kupferschmidt & Cohen, 2020). On the other hand, small-scale containment, such as microdistrict closed-off management, may miss some asymptomatically infected people and allow disease transmissions to keep occurring outside containment zones; such a phenomenon has occurred in the COVID-19 epidemic in the U.S. (Parodi & Liu, 2020). The factor generating our compact zoning pattern is the inclusion of mobility regularity, which implies that people usually display spatially constrained movement patterns despite their potentially high mobility (Belik et al., 2011). Thus, these constrained movement behaviors spatially connect several local places and form compact zones (Farmer & Fotheringham, 2011).

In the COVID-19 epidemic, Italy was the first outbreak country in Europe, yet the government launched containment approaches too late (Indolfi & Spaccarotella, 2020). The early inaction may have been the result of economic considerations. The decision to implement a city lockdown poses a tradeoff between public health and the nation's economy (Auzan, 2020). Our HuMoRZ algorithm may provide a solution to this tradeoff problem. The epidemic simulation results (Figs. 10 and 11) demonstrate that the zoning pattern of HuMoRZ algorithm could significantly reduce the case number at the beginning of an epidemic and delay the time of the epidemic peak. This delay earns health authorities more time to prepare and deploy medical resources for implementing control measures against an epidemic (Desjardins, Hohl,

doi:10.6342/NTU202101357

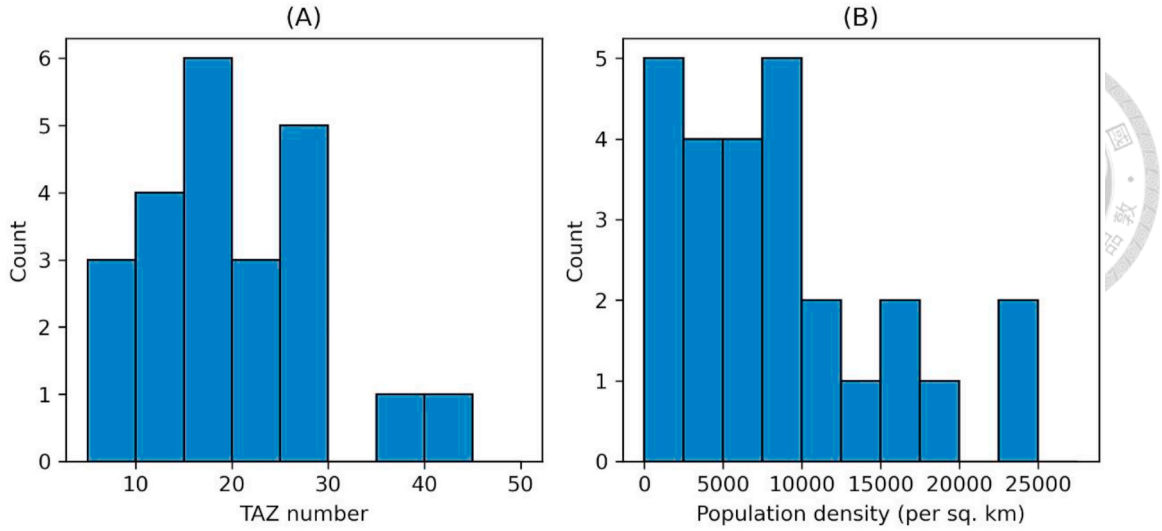


Fig. 5. Histograms of (A) TAZ number and (B) population density of HuMoRZ's zoning pattern.

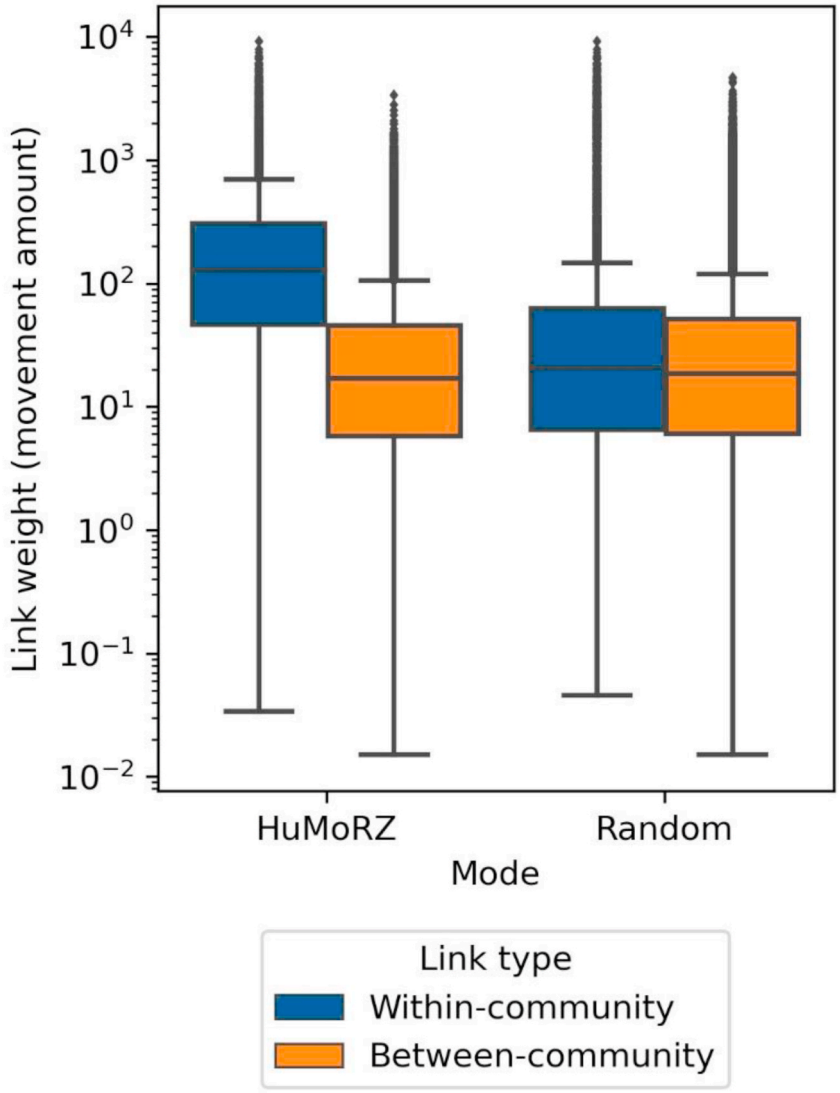


Fig. 6. Comparison of link weights between HuMoRZ and random zoning based on distributions of within- and between-community weights.

doi:10.6342/NTU202101357

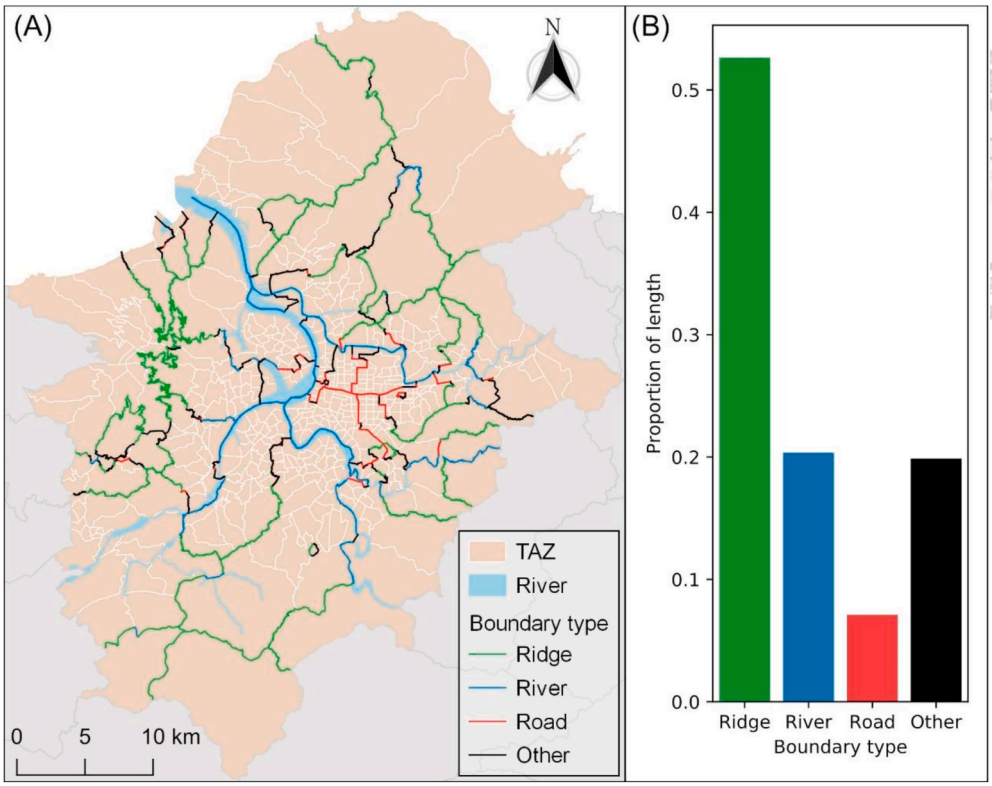


Fig. 7. (A) Spatial distribution and (B) Length proportion of different boundary types of the containment zones delineated by HuMoRZ.

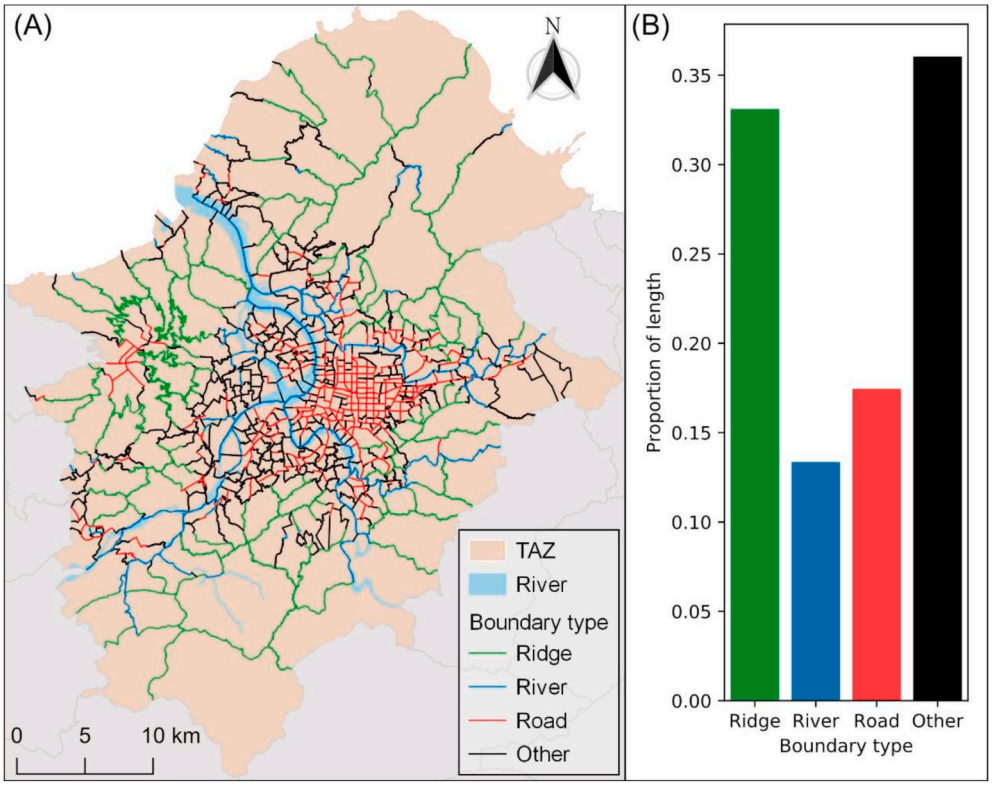


Fig. 8. (A) Spatial distribution and (B) Length proportion of different boundary types of the TAZs.

& Delmelle, 2020). Therefore, by blocking an appropriate zone at the initial epidemic stage, the Italian government could have gained more time to prepare prevention resources such as mask-wearing or environmental disinfection. Then, the epidemic could have been contained within the confinement zone, and the national economy may not have been dramatically affected.

doi:10.6342/NTU202101357

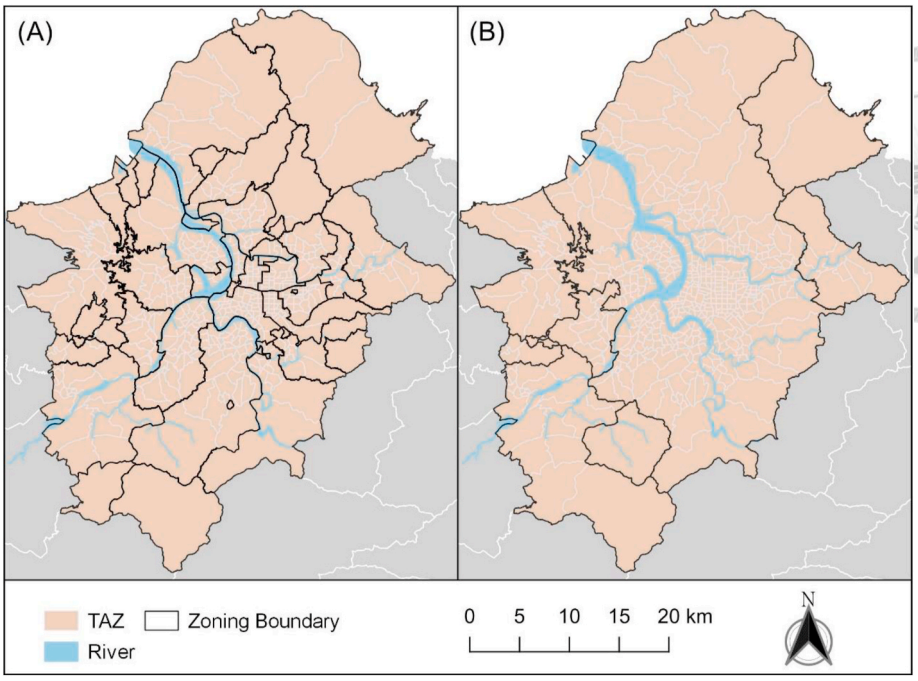


Fig. 9. Different zoning systems: (A) with-regularity (delineated by HuMoRZ) and (B) without-regularity (delineated by map equation).

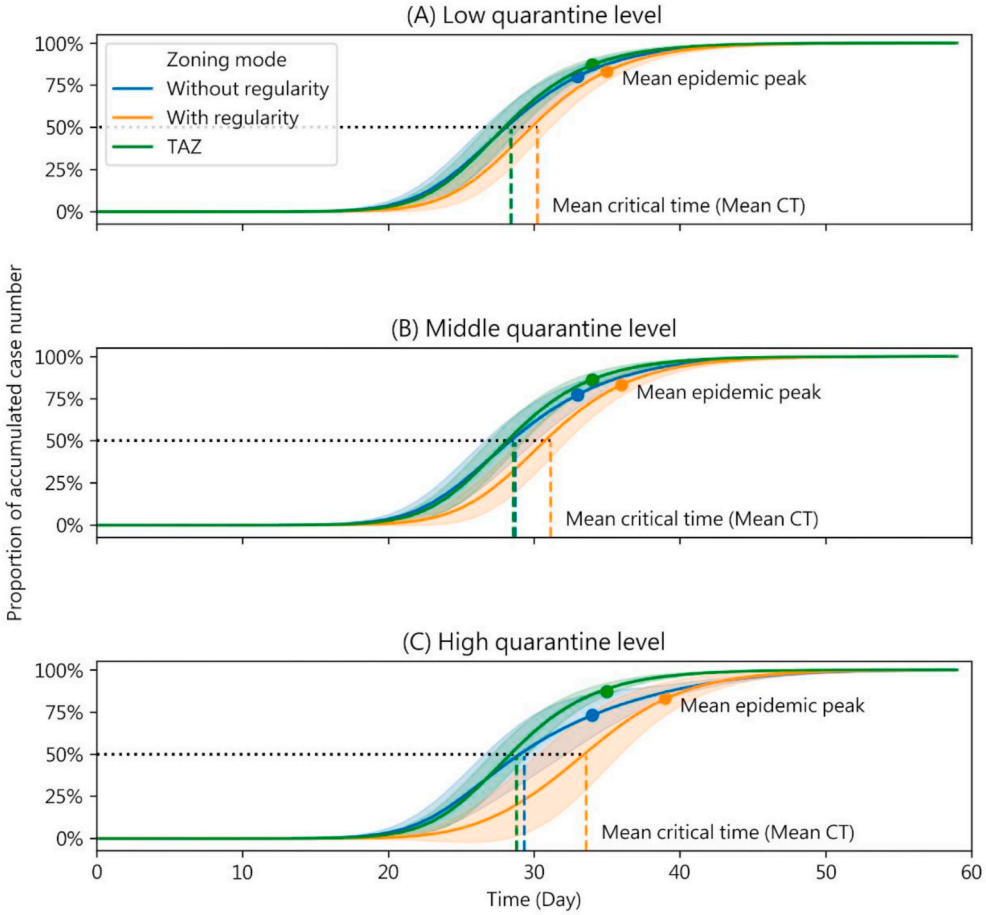


Fig. 10. Compare different zoning modes with SLIR simulation under three different quarantine levels: (A) 80%, (B) 90%, and (C)99%. Since we changed the initial outbreak places by turns. There are 498 TAZs (places); each curve represents the mean effect of 498 simulation results. Dash lines indicate every critical time (CT) showing when half of the population is infected; moreover, dots indicate every epidemic peak, which is when the prevalence rate reaches the highest level.

doi:10.6342/NTU202101357

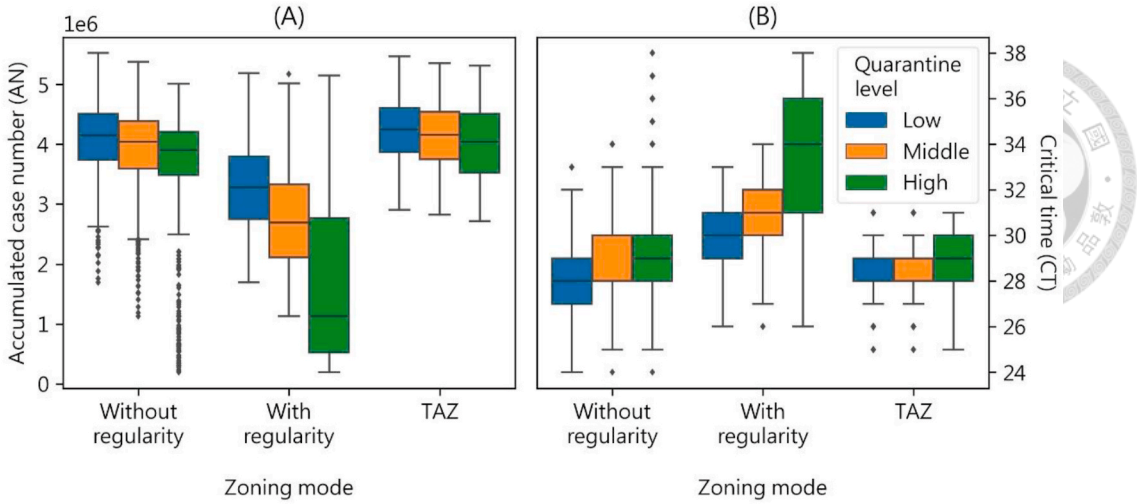


Fig. 11. Comparisons of epidemic performance indicators under three quarantine levels: Low 80%, Middle 90%, and High 99% among different zoning modes. (A) AN is the accumulated case number on the 30th day, reflecting the severity of an epidemic at the early stage. (B) CT is the critical time when half of the population once got infected, reflecting the spread speed of a disease. Each boxplot represents the distribution of 498 simulation results.

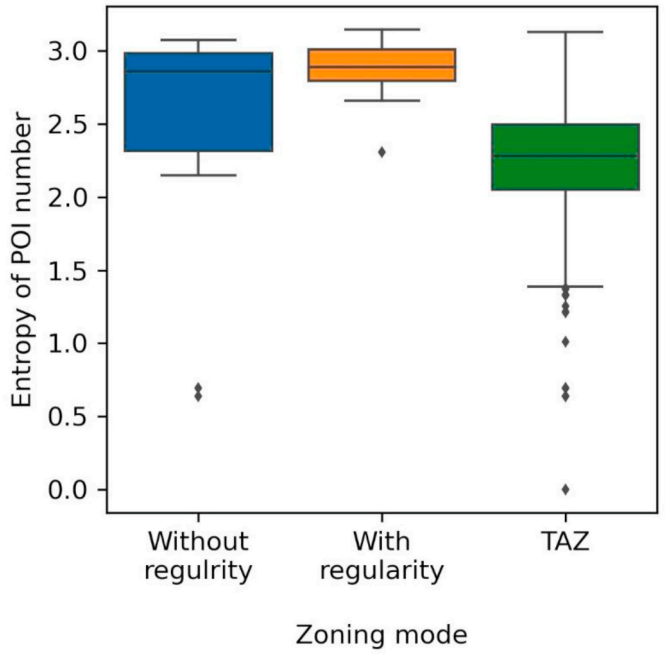


Fig. 12. Comparing the entropy value for each zone among different zoning modes.

Some infectious diseases, such as MERS or COVID-19, can transmit pathogens to susceptible hosts before symptom occurrence (Omrani et al., 2013; Rothe et al., 2020). In other words, a diagnosed case can infect others before the onset of illness or even any apparent symptoms; such invisible spreaders have caused social panic in Italy (M. Day, 2020). To identify these invisible spreaders, high numbers of testing per capita have been adopted during the COVID-19 pandemic in many countries, such as South Korea and Germany (Cohen & Kupferschmidt, 2020). Nonetheless, this approach involves cost-effectiveness issues, and not all countries or local governments can implement comprehensive screening and testing (Cohen & Kupferschmidt, 2020). Therefore, we argue that zoning based on regular contact patterns and delineation of the local spatial extent of contact among individuals could help narrow down possible close contacts and high-risk exposure to prioritize the deployment of screening and testing resources.

Our zoning pattern, which shows a self-contained life circle and is geographically homogeneous, echoes the concept of functional regions in the human geography literature. A functional region is described as a life circle consisting of many places with different land-use types where the residents can acquire almost all daily necessities (Philbrick, 1957). Based on this self-contained characteristic of their region, people seldom need to travel across boundaries to other regions, and a cohesive movement pattern exists within each functional region (Klapka & Halás, 2016). Such human movement patterns and land use diversity allow a functional region to act as an independent containment zone because people can survive well without interacting with others in another region. Moreover, while the level of self-containment increases as a functional region becomes larger, a large functional region will lose strong inner cohesion of movements (Halás, Klapka, & Erlebach, 2019). Therefore, a functional region cannot be either too large or too small, in line with the aforementioned suggestion of Barbera et al. (2001) regarding the size of a containment zone. As our HuMoRZ algorithm can divide a region into compact self-containment zones, it could also be appropriate for delineating functional regions.

A functional region could resemble a containment zone; however, existing methods focus only on cohesive within-region movements and weak between-region movements (Halás et al., 2019) and neglect diverse land-use types sustain daily lives. The land-use patterns could reflect different trip purposes of daily routines and are further used to measure mobility regularity. Thus, the HuMoRZ algorithm takes a step further to incorporate land use patterns into the partitioning of cohesive groups of a human mobility network.

Our study has some limitations. First, our zoning pattern is confined to TAZs, which are spatial units used in this study; these TAZs do not have equal size, shape, and land area. Although previous studies also delineate zones within an urban area based on small-scale administrative districts (Goddard, 1970), uniform grids could be a better resolution for calculating the movement amount between any two grids. However, it needs to aggregate detailed individual-level travel routes from GPS tracking data, which is difficult to access. Second, this study adopts agglomerated travel amounts to represent residents' level of interactions between two different locations. It cannot reflect individual-level travel frequency and regularity. Moreover, our travel amount only considers the two-dimensional surface area of land use and ignores the vertical dimensions, such as floor area in buildings and skyscrapers (Frank, Bradley, Kavage, Chapman, & Lawton, 2008). It could underestimate the travel amount related to in-door activities such as traveling to a department store for shopping or traveling to an office building for

doi:10.6342/NTU202101357

work. Therefore, floor area data should be included to measure mobility regularity better. Third, strangers may be in close proximity to each other for a long time on a bus or in a subway carriage, but such contacts do not exist if individuals use their private vehicles (Cooley et al., 2011). In other words, people have different contact patterns and epidemic risks based on their transport types (Andrews, Morrow, & Wood, 2013). It warrants further investigation considering various transport types. Last but not least, although most people move regularly, different demographic groups may have different movement behaviors. For example, the movements of women are usually shorter than those of men (Verhetsel, Beckers, & De Meyere, 2018), and low-income people seldom engage in long-distance travel due to the unaffordability of travel fees or costs (Casado-Díaz, 2000). Previous studies have shown that different social-demographic groups have functional regions of different sizes (Farmer & Fotheringham, 2011). Since a containment zone could be analogous to a functional region, the population structure should be another pivotal factor to be further considered in delineating containment zones.

6. Conclusion

Restricting human movement to decrease contact probability and frequency helps mitigate large-scale epidemics. We used land-use patterns to reflect trip purposes to measure the regularity of human mobility and proposed a novel network community detection method, the Human Mobility Regularity-based Zoning (HuMoRZ) algorithm incorporating mobility regularity to delineate containment zones. Our results suggest that containment zones that incorporate mobility regularity could significantly delay the epidemic peak and critical time and decrease the severity of an epidemic. The zoning patterns proposed in our algorithm could also allow for more life functions in a zone and more evenly distributed life resources across zones than those of zones generated by other methods. We conclude that zoning with mobility regularity could narrow down possible close contacts and high-risk exposure to prioritize the deployment of screening and testing resources. It could buy health authorities more time to prepare and deploy medical resources and implement control measures against an epidemic. These findings could provide insightful implications for fighting the COVID-19 pandemic.

CRediT authorship contribution statement

Fei-Ying Kuo: Formal analysis, Writing - original draft, conceived of the main conceptual ideas. developed the theory, analyzed the results, and wrote the manuscript, performed the experiments in discussion. **Tzai-Hung Wen:** Formal analysis, Writing - original draft, developed the theory, analyzed the results, and wrote the manuscript. performed the experiments in discussions.

Acknowledgments

This research was supported by the grant from the Ministry of Science and Technology in Taiwan [MOST 108-2638-H-002-002-MY2]. The authors also acknowledge the financial support provided by the Infectious Diseases Research and Education Center, the Ministry of Health and Welfare (MOHW) and National Taiwan University (NTU). The funders had no role in the study design, data collection and analysis, or manuscript preparation.

References

Andrews, J. R., Morrow, C., & Wood, R. (2013). Modeling the role of public transportation in sustaining tuberculosis transmission in South Africa. *American Journal of Epidemiology, 177*(6), 556–561. https://doi.org/10.1093/aje/kws331

Auzan, A. A. (2020). The economy under the pandemic and afterwards. *Population and Economics, 4*(2), 4–12. https://doi.org/10.3897/popecon.4.e53403

Bajardi, P., Poletto, C., Ramasco, J. J., Tizzoni, M., Colizza, V., & Vespignani, A. (2011). Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. *PloS One, 6*(1), Article e16591. https://doi.org/10.1371/journal.pone.0016591

Barbera, J., Macintyre, A., Gostin, L., Inglesby, T., O'Toole, T., DeAtley, C., et al. (2001). Large-scale quarantine following biological terrorism in the United States: Scientific examination, logistic and legal limits, and possible consequences. *The Journal of the American Medical Association, 286*(21), 2711–2717. https://doi.org/10.1001/jama.286.21.2711

Belik, V., Geisel, T., & Brockmann, D. (2011). Natural human mobility patterns and spatial spread of infectious diseases. *Physical Review X, 1*(1), Article 011001. https://doi.org/10.1103/PhysRevX.1.011001

Casado-Díaz, J. M. (2000). Local labour market areas in Spain: A case study. *Regional Studies, 34*(9), 843–856. https://doi.org/10.1080/00343400020002976

Cohen, J., & Kupferschmidt, K. (2020). Countries test tactics in 'war' against COVID-19. *Science, 367*(6484), 1287–1288. https://doi.org/10.1126/science.367.6484.1287

Coltart, C. E. M., Lindsey, B., Ghinai, I., Johnson, A. M., & Heymann, D. L. (2017). The Ebola outbreak, 2013–2016: Old lessons for new epidemics. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*, Article 20160297. https://doi.org/10.1098/rstb.2016.0297

Cooley, P., Brown, S., Cajka, J., Chasteen, B., Ganapathi, L., Grefenstette, J., et al. (2011). The role of subway travel in an influenza epidemic: A New York city simulation. *Journal of Urban Health, 88*(5), 982–995. https://doi.org/10.1007/s11524-011-9603-4

Cooper, B. S., Pitman, R. J., Edmunds, W. J., & Gay, N. J. (2006). Delaying the international spread of pandemic influenza. *PLoS Medicine, 3*(6), e212. https://doi.org/10.1371/journal.pmed.0030212

Day, M. (2020). COVID-19: Surge in cases in Italy and South Korea makes pandemic look more likely. *British Medical Journal, 368*, m751. https://doi.org/10.1136/bmj.m751

Day, T., Park, A., Madras, N., Gumel, A., & Wu, J. (2006). When is quarantine a useful control strategy for emerging infectious diseases? *American Journal of Epidemiology, 163*(5), 479–485. https://doi.org/10.1093/aje/kwj056

Desjardins, M. R., Hohl, A., & Delmelle, E. M. (2020). Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. *Applied Geography, 118*, 102202. https://doi.org/10.1016/j.apgeog.2020.102202

Epstein, J. M., Goedecke, D. M., Yu, F., Morris, R. J., Wagener, D. K., & Bobashev, G. V. (2007). Controlling pandemic flu: The value of international air travel restrictions. *PloS One, 2*(5), e401. https://doi.org/10.1371/journal.pone.0000401

Farmer, C. J. Q., & Fotheringham, A. S. (2011). Network-based functional regions. *Environment & Planning A: Economy and Space, 43*(11), 2723–2741. https://doi.org/10.1068/a44136

Fèvre, E. M., Bronsvoort, B. M.d. C., Hamilton, K. A., & Cleaveland, S. (2006). Animal movements and the spread of infectious diseases. *Trends in Microbiology, 14*(3), 125–131. https://doi.org/10.1016/j.tim.2006.01.004

Fong, M. W., Gao, H., Wong, J. Y., Xiao, J., Shiu, E., Y, C., et al. (2020). Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—social distancing measures. *Emerging Infectious Diseases, 26*(5), 976–984. https://doi.org/10.3201/eid2605.190995

Frank, L., Bradley, M., Kavage, S., Chapman, J., & Lawton, T. K. (2008). Urban form, travel time, and cost relationships with tour complexity and mode choice. *Transportation, 35*, 37–54. https://doi.org/10.1007/s11116-007-9136-6

Fraser, C., Riley, S., Anderson, R. M., & Ferguson, N. M. (2004). Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences, 101*(16), 6146–6151. https://doi.org/10.1073/pnas.0307506101

Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R., et al. (2020). Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proceedings of the National Academy of Sciences, 117*(19), 10484–10491.

Gensini, G. F., Yacoub, M. H., & Conti, A. A. (2004). The concept of quarantine in history: From plague to SARS. *Journal of Infection, 49*(4), 257–261. https://doi.org/10.1016/j.jinf.2004.03.002

Goddard, J. B. (1970). Functional regions within the city centre: A study by factor analysis of taxi flows in central london. *Transactions of the Institute of British Geographers, 49*, 161–182. http://www.jstor.org/stable/621647.

González, M. C., Hidalgo, C. A., & Barabási, A.-L. (2008). Understanding individual human mobility patterns. *Nature, 453*(7196), 779–782. https://doi.org/10.1038/nature06958

Halás, M., Klapka, P., & Erlebach, M. (2019). Unveiling spatial uncertainty: A method to evaluate the fuzzy nature of functional regions. *Regional Studies, 53*(7), 1029–1041. https://doi.org/10.1080/00343404.2018.1537483

Indolfi, C., & Spaccarotella, C. (2020). The outbreak of COVID-19 in Italy: Fighting the pandemic. *Journal of the American College of Cardiology: Case Reports, 2*(9), 1414–1418. https://doi.org/10.1016/j.jaccas.2020.03.012

Klapka, P., & Halás, M. (2016). Conceptualising patterns of spatial flows: Five decades of advances in the definition and use of functional regions. *Moravian Geographical Reports, 24*(2), 2–11. https://doi.org/10.1515/mgr-2016-0006

Kupferschmidt, K., & Cohen, J. (2020). Can China's COVID-19 strategy work elsewhere? *Science, 367*(6482), 1061–1062. https://doi.org/10.1126/science.367.6482.1061

Lee, M., & Holme, P. (2015). Relating land use and human intra-city mobility. *PloS One, 10*(10), Article e0140152. https://doi.org/10.1371/journal.pone.0140152

Lee, G.-J., Pak, S.-I., Lee, K.-N., & Hong, S. (2019). Movement-based biosecurity zones for control of highly infectious animal diseases: Application of community detection analysis to a livestock vehicle movement network. *Sustainability, 11*(6), 1642. https://doi.org/10.3390/su11061642

doi:10.6342/NTU202101357

Liu, X., Gong, L., Gong, Y., & Liu, Y. (2015). Revealing travel patterns and city structure with taxi trip data. *Journal of Transport Geography, 43*, 78–90. https://doi.org/10.1016/j.jtrangeo.2015.01.016

Maat, K., Van Wee, B., & Stead, D. (2005). Land use and travel behaviour: Expected effects from the perspective of utility theory and activity-based theories. *Environment and Planning B: Uraban Analytics and City Science, 32*(1), 33–46. https://doi.org/10.1068/b31106

Malavika, B., Marimuthu, S., Joy, M., Nadaraj, A., Asirvatham, E. S., & Jeyaseelan, L. (2020). Forecasting COVID-19 epidemic in India and high incidence states using SIR and logistic growth models. *Clinical Epidemiology and Global Health*. https://doi.org/10.1016/j.cegh.2020.06.006

Malliaros, F. D., & Vazirgiannis, M. (2013). Clustering and community detection in directed networks: A survey. *Physics Reports, 533*(4), 95–142. https://doi.org/10.1016/j.physrep.2013.08.002

Mao, L. (2014). Modeling triple-diffusions of infectious diseases, information, and preventive behaviors through a metropolitan social network—an agent-based simulation. *Applied Geography, 50*, 31–39. https://doi.org/10.1016/j.apgeog.2014.02.005

Ministry of Transportation and Communications in Taiwan. (2020). Taiwan transportation network geographic information warehousing system. In https://gist.motc.gov.tw/gist_web/.

Omrani, A. S., Matin, M. A., Haddad, Q., Al-Nakhli, D., Memish, Z. A., & Albarrak, A. M. (2013). A family cluster of Middle East Respiratory Syndrome Coronavirus infections related to a likely unrecognized asymptomatic or mild case. *International Journal of Infectious Diseases, 17*(9), e668–e672. https://doi.org/10.1016/j.ijid.2013.07.001

Parodi, S. M., & Liu, V. X. (2020). From containment to mitigation of COVID-19 in the U. S. *The Journal of the American Medical Association, 323*(15), 1441–1442. https://doi.org/10.1001/jama.2020.3882

Peak, C. M., Kahn, R., Grad, Y. H., Childs, L. M., Li, R., Lipsitch, M., et al. (2020). Individual quarantine versus active monitoring of contacts for the mitigation of COVID-19: A modelling study. *The Lancet Infectious Diseases, 20*(9), 1025–1033. https://doi.org/10.1016/S1473-3099(20)30361-3

Peak, C. M., Wesolowski, A., zu Erbach-Schoenberg, E., Tatem, A. J., Wetter, E., Lu, X., et al. (2018). Population mobility reductions associated with travel restrictions during the Ebola epidemic in Sierra Leone: Use of mobile phone data. *International Journal of Epidemiology, 47*(5), 1562–1570. https://doi.org/10.1093/ije/dyy095

Philbrick, A. K. (1957). Principles of areal functional organization in regional human geography. *Economic Geography, 33*(4), 299–336.

Rosvall, M., Axelsson, D., & Bergstrom, C. T. (2009). The map equation. *The European Physical Journal - Special Topics, 178*, 13–23. https://doi.org/10.1140/epjst/e2010-01179-1

Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences, 105*(4), 1118–1123. https://doi.org/10.1073/pnas.0706851105

Rothe, C., Schunk, M., Sothmann, P., Bretzel, G., Froeschl, G., Wallrauch, C., et al. (2020). Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *New England Journal of Medicine, 382*(10), 970–971. https://doi.org/10.1056/NEJMc2001468

Rubin, G. J., & Wessely, S. (2020). The psychological effects of quarantining a city. *British Medical Journal, 368*, m313. https://doi.org/10.1136/bmj.m313

Schneider, C. M., Belik, V., Couronné, T., Smoreda, Z., & González, M. C. (2013). Unravelling daily human mobility motifs. *Journal of The Royal Society Interface, 10*(84), Article 20130246. https://doi.org/10.1098/rsif.2013.0246

Shao, P. (2020). *Impact of city and residential unit lockdowns on prevention and control of COVID-19. medRxiv*. https://doi.org/10.1101/2020.03.13.20035253

Song, C., Koren, T., Wang, P., & Barabási, A.-L. (2010a). Modelling the scaling properties of human mobility. *Nature Physics, 6*, 818–823. https://doi.org/10.1038/NPHYS1760

Song, C., Qu, Z., Blumm, N., & Barabási, A.-L. (2010b). Limits of predictability in human mobility. *Science, 327*(5968), 1018–1021. https://doi.org/10.1126/science.1177170

Sun, L., Axhausen, K. W., Lee, D.-H., & Huang, X. (2013). Understanding metropolitan patterns of daily encounters. *Proceedings of the National Academy of Sciences, 110*(34), 13774–13779. https://doi.org/10.1073/pnas.1306440110

Tang, B., Wang, X., Li, Q., Bragazzi, N. L., Tang, S., Xiao, Y., et al. (2020). Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. *Journal of Clinical Medicine, 9*(2), 462. https://doi.org/10.3390/jcm9020462

Velthuis, A. G. J., & Mourits, M. C. M. (2007). Effectiveness of movement-prevention regulations to reduce the spread of foot-and-mouth disease in The Netherlands. *Preventive Veterinary Medicine, 82*(3–4), 262–281. https://doi.org/10.1016/j.prevetmed.2007.05.023

Verhetsel, A., Beckers, J., & De Meyere, M. (2018). Assessing daily urban systems: A heterogeneous commuting network approach. *Networks and Spatial Economics, 18*, 633–656. https://doi.org/10.1007/s11067-018-9425-y

Wen, T.-H., Hsu, C.-S., Sun, C.-H., Jiang, J.-A., & Juang, J.-Y. (2018). A location-based client-server framework for assessing personal exposure to the transmission risks of contagious diseases. In S.-L. Shaw, & D. Sui (Eds.), *Human dynamics research in smart and connected communities* (pp. 133–148). Cham, Switzerland: Springer Inc.

Wilder-Smith, A., Chiew, C. J., & Lee, V. J. (2020). Can we contain the COVID-19 outbreak with the same measures as for SARS? *The Lancet Infectious Diseases, 20*(5), e102–e107. https://doi.org/10.1016/S1473-3099(20)30129-8

Wu, J. T., Leung, K., & Leung, G. M.d. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in wuhan, China: A modelling study. *The Lancet, 395*(10225), 689–697. https://doi.org/10.1016/S0140-6736(20)30260-9

Zhong, C., Arisona, S. M., Huang, X., Batty, M., & Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science, 28*(11), 2178–2199. https://doi.org/10.1080/13658816.2014.914521

doi:10.6342/NTU202101357

附錄 C

作者：Fei-Ying Kuo, and Tzai-Hung Wen

文章標題：A Mathematical Model for Evaluating the Medical Resource Availability of COVID-19 in Time and Space

發表年分：2021

書籍名稱：Mapping COVID-19 in Space and Time, Human Dynamics in Smart Cities

主編：Shih-Lung Shaw, and Daniel Sui

出版社：Springer Nature

頁碼：（排版印刷中）

數位物件識別號碼（DOI）：10.1007/978-3-030-72808-3_15

Chapter 15
A Mathematical Model for Evaluating the Medical Resource Availability of COVID-19 in Time and Space

Fei-Ying Kuo and Tzai-Hung Wen

15.1 Introduction

Global pandemic COVID-19 rapidly spreads to many countries, causing enormous economic loss and many deaths in 2020 (Remuzzi & Remuzzi, 2020). Low accessibility and availability of medical resources (including screening/testing, prevention, and treatment facilities) for the at-risk populations is an urgent, important issue in preventing and controlling the pandemic (Ji, Ma, Peppelenbosch, & Pan, 2020). Thus, evaluating the accessibility and availability of medical resources is an essential topic for pandemic preparedness and responses. Healthcare accessibility and availability are determined by healthcare facilities' demand and service capacity (Wang, 2012). Floating catchment analysis (FCA) is a widely used model framework for measuring geographical accessibility (McGrail & Humphreys, 2014), and McGrail (2012) has comprehensively reviewed its derivative versions. In previous studies, the spatial distribution of demand is usually assumed as static and unchanged over time. However, epidemic-induced demand for medical resources, including the exposed and infected population, significantly varies over space and time (Weissman et al., 2020). Matching the availability of medical resources should consider the demand and supply at different epidemic stages simultaneously. Recent studies considered time-varying dynamic demand resulted from daily routines and human mobility for measuring accessibility (Ma, Luo, Wan, Hu, & Peng, 2018; Xia et al., 2019). Although these studies incorporated daily spatial behaviors of individuals reflecting the different spatial distributions of demand at distinct periods, these considerations cannot comprehensively capture the dynamic nature of epidemic-induced demand in time and space. In other words, the dynamics of exposed and infected populations reflect the time-varying medical demand, and spatial clusters of an epidemic may

F.-Y. Kuo · T.-H. Wen (✉)
Department of Geography, National Taiwan University
e-mail: wenthung@ntu.edu.tw

S. Shaw and D. Sui (eds.), *Mapping COVID-19 in Space and Time*, Human Dynamics in Smart Cities, https://doi.org/10.1007/978-3-030-72808-3_15
doi:10.6342/NTU202101357

also shift over time. Therefore, to avoid inappropriate resource allocation, it is necessary to incorporate the spatial–temporal dynamics of epidemic-induced demand into assessing medical resource availability.

The spatial mechanic dynamics of an epidemic is usually simulated by the family of susceptible-exposed-infectious-recovered (SEIR) meta-population models (Wang & Wu, 2018). An SEIR model separates the population into different stages of infection, and each individual will shift from one stage to the next due to epidemic propagation (Tsai, Huang, Wen, Sun, & Yen, 2011). The meta-population, the extended concept, highlights that the propagation speed and transmission risk of an epidemic would be different at distinct places due to local environmental heterogeneity and human movements among locations. These local risk factors may trigger an outbreak to diffuse across the surface (Lima, De Domenico, Pejovic, & Musolesi, 2015). Therefore, the SEIR meta-population model framework can be effectively used for representing the spatial–temporal dynamics of an epidemic. Nonetheless, the capacity and availability of medical resources for the exposed and infected population may influence consecutive epidemic propagation and resource accessibilities. High accessibility of medical resources in some areas may protect local at-risk populations and reduce disease transmission. Then, these resources in epidemic-controlled areas may provide neighboring regions to mitigate the prolongation of an epidemic spreading. On the other hand, low accessibility may aggravate an epidemic and enlarge the need and consumption of medical resources, which lowers resource availability.

As the dynamics of resource accessibility and the epidemic dynamic process interact, they cannot be measured independently. Therefore, we proposed an integrated mathematical model, named Epi-RA, to combine two components, including epidemic dynamic process and resource accessibility, to measure the medical resource availability of the COVID-19 epidemic. The epidemic dynamic process is developed based on the SEPIA model proposed by Gatto et al. (2020). The model considers the isolation measure, and the pre-symptomatic stage of infection, the essential characteristics of the COVID-19 epidemic. Moreover, resource accessibility is implemented from the two-step floating catchment area (2SFCA) model, which is widely used to measure spatial accessibility (Guagliardo, 2004; Wang, 2012). The community screening/testing capacity of COVID-19 is regarded as the medical resource availability in this study. The Taipei metropolitan area, one of the major cities in East Asia, is used as a study region to demonstrate the proposed model's feasibility and effectiveness.

15.2 Methods and Data

15.2.1 Integrated Model Framework

We proposed an integrated model, named Epi-RA, which combines epidemic dynamic process (Epi) and resource accessibility (RA), to evaluate the changing

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard

doi:10.6342/NTU202101357

ration levels of medical resources during different epidemic periods. Figure 15.1 illustrates the framework of our integrated model. It demonstrates how epidemic dynamic processes and resource accessibility interacted with each other. Our model divides the epidemic dynamic process into eight disease progression stages, including susceptible (S), exposed (E), pre-symptomatic (P), severe symptomatic (I), mild/no symptomatic (A), isolated (Q), recovered (R), and dead (D). As individuals at the infectious stages of P, I, and A can transmit the pathogen to others, those could reflect the epidemic-induced demand which needs to acquire the COVID-19 rapid tests. The spatial distribution of time-varying demand can be determined from the epidemic dynamic modeling process in time and space. Subsequently, resource accessibility modeling is used to estimate the resource availability of the epidemic-induced demand at a specific period. The matching results of resource availability influence the succeeding transmission propagation, which can be revealed from the following epidemic dynamic process.

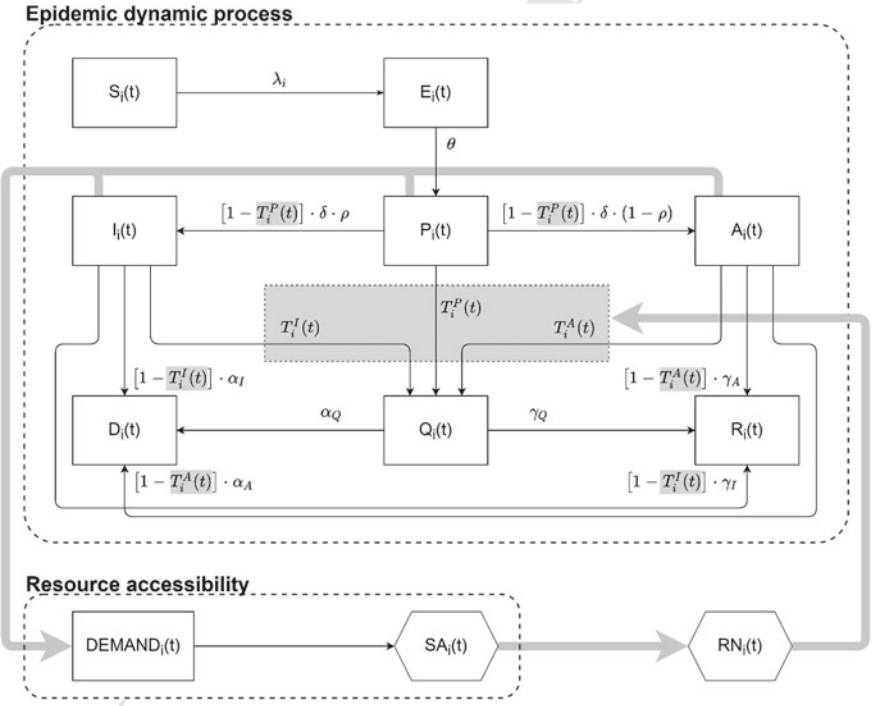


Fig. 15.1 The framework of the integrated model. The subscript *i* represents a specific location, and (t) represents the situation at time step *t*. The descriptions of each parameter are shown in Table 15.1

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard
doi:10.6342/NTU202101357

15.2.1.1 Resource Accessibility Sub-Model

The resource accessibility sub-model is implemented to evaluate the COVID-19 testing resource availability for epidemic-induced demand, including the individuals in pre-symptomatic, severe symptomatic, and mild/no symptomatic stages. The demand varies over space and time through epidemic dynamic process modeling (as shown in Eq. 15.1).

$$DEMAND_i(t) = v_P P_i(t) + v_I I_i(t) + v_A A_i(t), \tag{15.1}$$

where $P_i(t)$, $I_i(t)$, and $A_i(t)$ are the number of people in pre-symptomatic, severe symptomatic, and mild/no symptomatic stages at location i at time t; v_P, v_I, and v_A denote the proportion of people who are required to be tested from these three groups, respectively.

Based on these dynamic demands, we used the generalized 2SFCA (Wang, 2012) to measure the time-varying spatial accessibility (SA) for every location:

$$SA_i(t) = \sum_{j=1}^{n} \frac{S_j f(d_{ij})}{\sum_{k=1}^{m} DEMAND_k(t) f(d_{kj})} \tag{15.2}$$

where S_j denotes the amount of resource that facility j can provide; d_{ij} denotes the geographical distance between location i and facility j; $f(d_{ij})$ is a function transforming d_{ij} into a weight to represent the interaction strength between these two places. We adopted the negative-power transformation ($f(d_{ij}) = d_{ij}^{-2}$), which is one of the commonly used transformations in the literature (Kwan, 1998). According to Eq. 15.2, the accessibility varies across both the study region and the epidemic period.

15.2.1.2 Epidemic Dynamic Process Sub-Model

The epidemic dynamic process sub-model is to simulate the spatial–temporal dynamics of disease transmission for estimating epidemic-induced demand. Equation 15.3 shows details of our epidemic process component, which is developed based on the SPEIA model proposed by Gatto et al. (2020):

$$\frac{dS_i}{dt} = -\lambda_i(t) \cdot S_i \tag{15.3}$$

$$\frac{dE_i}{dt} = \lambda_i(t) \cdot S_i - \theta E_i$$

$$\frac{dP_i}{dt} = \theta \cdot E_i - \left\{T_i^P(t) + \left[1 - T_i^P(t)\right] \cdot \delta\right\} P_i$$

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard

doi:10.6342/NTU202101357

$$\frac{dI_i}{dt} = \left[1 - T_i^P(t)\right] \cdot \rho \cdot \delta \cdot P_i - \left\{T_i^I(t) + \left[1 - T_i^I(t)\right](\alpha_I + \gamma_I)\right\} I_i$$

$$\frac{dA_i}{dt} = \left[1 - T_i^P(t)\right] \cdot (1 - \rho) \cdot \delta \cdot P_i - \left\{T_i^A(t) + \left[1 - T_i^A(t)\right](\alpha_A + \gamma_A)\right\} A_i$$

$$\frac{dQ_i}{dt} = T_i^P(t) \cdot P_i + T_i^I(t) \cdot I_i + T_i^A(t) \cdot A_i - \left(\alpha_Q + \gamma_Q\right) \cdot Q_i$$

$$\frac{dR_i}{dt} = \left[1 - T_i^I(t)\right] \cdot \gamma_I I_i + \left[1 - T_i^A(t)\right] \cdot \gamma_A A_i + \gamma_Q Q_i$$

$$\frac{dD_i}{dt} = \left[1 - T_i^I(t)\right] \cdot \alpha_I I_i + \left[1 - T_i^A(t)\right] \cdot \alpha_A A_i + \alpha_Q Q_i$$

In the sub-model, susceptible people (S) become exposed to infection upon contact with infectious people (in pre-symptomatic (P), severe symptomatic (I), or mild/no symptomatic (A) stages). The force of infection at location i at time t is denoted as Eq. 15.4:

$$\lambda_i(t) = \sum\nolimits_j \frac{\beta_P \cdot P_j(t) + \beta_I \cdot I_j(t) + \beta_A \cdot A_j(t)}{S_j(t) + E_j(t) + P_j(t) + I_j(t) + A_j(t) + R_j(t)} \cdot \frac{w(j,i)}{N_j} \quad (15.4)$$

where β_P, β_I, and β_A are the transmission rates corresponding to three different infectious compartments, respectively; $w(j,i)$ denotes the daily average number of people traveling from location j to location i; N_j denotes the population number in location j. The fraction, $w(j,i)/N_j$, reflects how the epidemic in location j affects the epidemic in location i through residents' movements. In Eq. 15.4, j is allowed to equal to i, which reflects local transmissions occurring in location i.

In our study, exposed people (E) are assumed to be latently infected. In other words, they are not infectious until they become (at rate θ) pre-symptomatic (P) after the incubation period. For those non-isolated pre-symptomatic people, some of them then progress (at rate δ) to develop severe symptoms (I) with probability ρ, and some of them may show mild or still no apparent symptoms (A) with probability $1 - \rho$. The people in these two stages, despite they are isolated, will finally recover (at the rate γ_I or γ_A) or die (at the rate α_I or α_A). Similarly, isolated people (Q) also end the process of infection with recovery or death at the rate $\gamma_Q$orα_Q, respectively.

Subject to the accessibility of testing reagents, some infected people who need isolation could become non-isolated infectious people and influence the following epidemic propagation. The geographical accessibility sub-model dynamically decides these three floating proportions, including individuals in P, I, and A stages. The estimated spatial accessibility (Eq. 15.2) represents the average amount of resource that one individual in demand at location i at time t can acquire. This accessibility score may be more than one if a location is full of resource supply or no large demand. Thus, we transform this accessibility score into the ration level (RN) through Eq. 15.5:

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard

doi:10.6342/NTU202101357

$$RN_i(t) = \min(1, SA_i(t)) \tag{15.5}$$

This equation ensures that $RN_i(t)$ ranges from 0 to 1, with one means all demand for reagent resources can be satisfied. Then, the proportions that infected people shift from three infectious compartments to the isolation compartment are formulated as Eq. 15.6:

$$T_i^P(\mathrm{t}) = v_P \cdot RN_i(\mathrm{t}) \cdot \mathrm{u} \tag{15.6}$$

$$T_i^I(t) = v_I \cdot RN_i(\mathrm{t}) \cdot \mathrm{u}$$

$$T_i^A(t) = v_A \cdot RN_i(\mathrm{t}) \cdot \mathrm{u}$$

where v_P, v_I, and v_A, as stated above, represent the proportion of people who are required to be tested from P, I, and A groups; u is the accuracy of the COVID-19 reagent. The values of $T_i^P(\mathrm{t})$, $T_i^I(t)$, and $T_i^A(t)$ not only differ among distinct places but also dynamically change over time. Thus, Eq. 15.6 integrates spatiotemporal dynamics of accessibility into the epidemic dynamic process. Table 15.1

Table 15.1 Descriptions and values of model parameters

Parameter	Description	Value	Reference
β_P	Transmission rate from P to S	4.8014	*
β_I	Transmission rate from I to S	0.1632	*
β_A	Transmission rate from A to S	0.1585	*
$1/\theta$	Incubation period	3.32 days	*
$1/\delta$	Latent period	0.75 days	*
ρ	Probability to be severe symptomatic (I)	0.25	*
γ_I	Recovery rate of I	0.0698	*
γ_A	Recovery rate of A	0.1396	*
γ_Q	Recovery rate of Q	0.1047	*
α_I	Death rate of I	0.0413	*
α_A	Death rate of A	0.0207	-
α_Q	Death rate of Q	0.0207	-
v_P	Demand proportion of P	0.1	-
v_I	Demand proportion of I	0.5	-
v_A	Demand proportion of A	0.3	-
u	Accuracy of reagent	0.9	-

Note parameter values with * mark refer to Gatto et al. (2020). The other parameter values are assumed by the authors

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard

doi:10.6342/NTU202101357

presents detailed descriptions of parameter settings, and the settings of the COVID-19 epidemic process refer to the study that proposed SEPIA model (Gatto et al. 2020).

15.2.2 Materials

15.2.2.1 Study Area

To demonstrate our Epi-GA model's feasibility, we used the Taipei metropolitan area, one of the major cities in East Asia, as the case study. Our study area covers approximately 1362.57 square kilometers of land and has a population of approximately 6.34 million, 27.5% of Taiwan's total population. In 2015, the Department of Rapid Transit Systems (DORTS) of the Taipei City Government divided this area into 543 traffic analysis zones (TAZs). The average size of a TAZ is about 2.52 square kilometers, and the average population is 12,500. These 543 TAZs are used as spatial units of demand locations in this study. Figure 15.2a displays spatial distributions of population densities. The downtown area with high population density locates at the center of our study area, and most peripheral areas have low population density.

15.2.2.2 Data

The supply facilities are represented by 22 community COVID-19 testing stations established by Taiwan CDC (Fig. 15.2a), and these locations are district hospitals or regional hospitals. In Taiwan, people who once contacted COVID-19 confirmed cases or show significant symptoms, such as fever or dyspnea, can only be tested in these official testing stations. Before the testing process is finished, people have to be temporarily quarantined. If the testing result is negative, the person can leave the station; otherwise, s/he has to be isolated immediately until recovering. The testing process usually takes a few hours, which is shorter than the time unit in our model (day); thus, we ignored this waiting time. We assumed that each testing station provides 500 reagents for COVID-19 testing each day.

The human movement data is from an official survey conducted by DORTS in 2015, the fourth version of the "Taipei Rapid Transit Systems Demand Model" (TRTS-IV). Based on the 543 TAZs, the daily average number of residents moving from any one TAZ to another were surveyed (Fig. 15.2b). According to the official report (Chou, Ko, Lee, Chen, & Lee, 2018), the error of movement amount on most flow links are smaller than 10%, and only a few links have larger errors yet remaining within 20%.

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard

doi:10.6342/NTU202101357

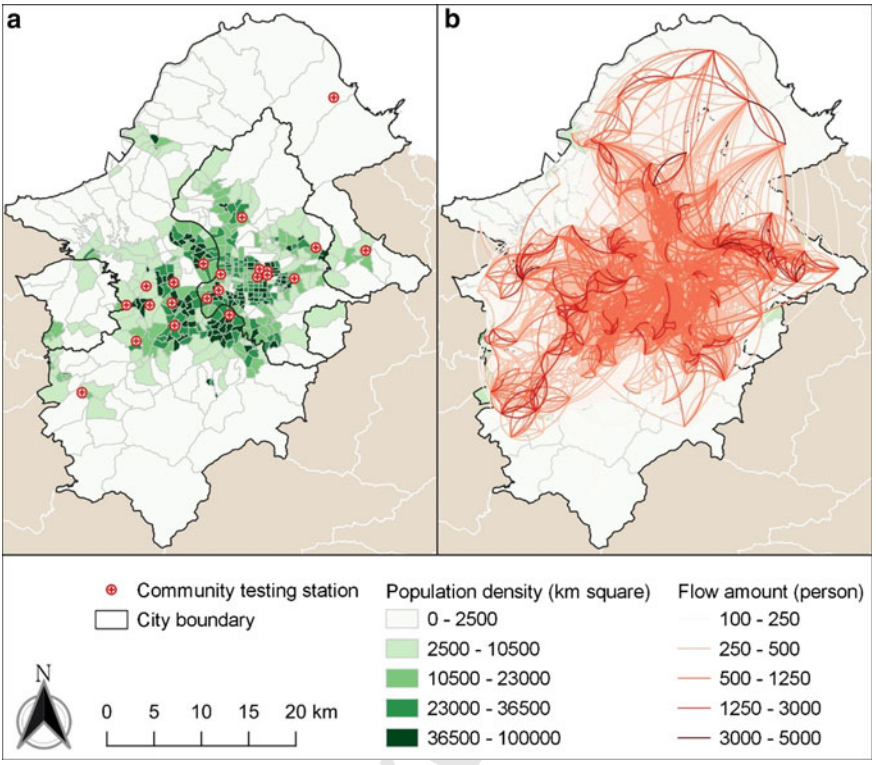


Fig. 15.2 Spatial distributions of **a** population density and locations of community COVID-19 testing stations, and **b** Population daily movements

15.3 Results

We presented the temporal progression of epidemic-induced demand for COVID-19 testing reagents, including individuals in the pre-symptomatic (P), severe symptomatic (I), or mild/no symptomatic (A) stages, estimated from the Epi-RA model (Fig. 15.3a). The temporal pattern shows the exponential growth of the demand is initiated after the 20th day of the first infected person and a significant decrease in the demand after the 26th day. Figure 15.3b, c, and d illustrates the spatial distributions of the demand on the 20th (exponential growth), 26th (epidemic peak), and 32th (significant decrease) day. These figures illustrate how the demand dynamically changes over space. In the initial stage of exponential growth, the demand concentrated on Taipei downtown areas (Fig. 15.3b). As the epidemic peak is reached on the 26th day, not only in downtown areas but also in peripheral urban areas have a high demand for testing reagents (Fig. 15.3c). In the last stage of the epidemic, the demand in downtown areas declines, but the demand in peripheral urban areas remains high.

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard

doi:10.6342/NTU202101357

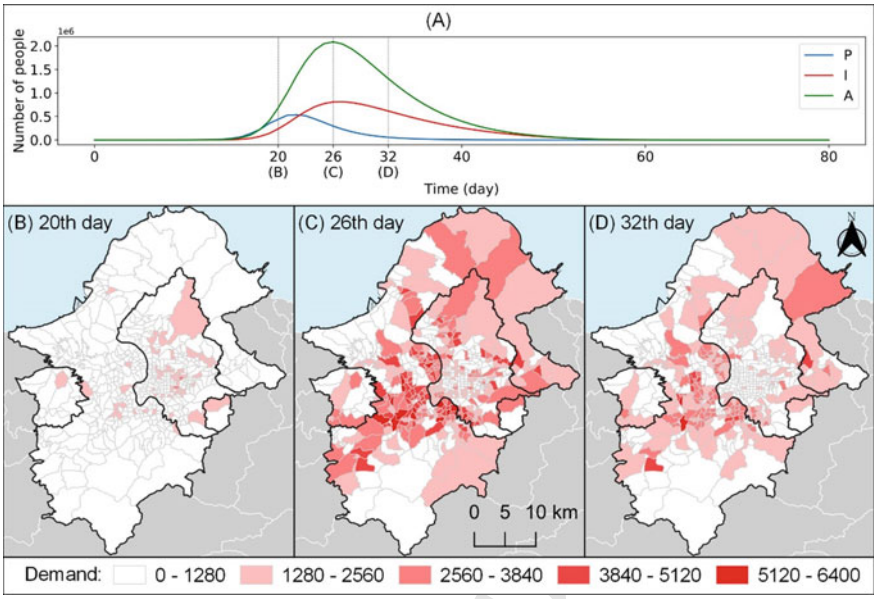


Fig. 15.3 Temporal progression and spatial distributions of epidemic-induced demand for COVID-19 testing reagents, including individuals in the pre-symptomatic (P), severe symptomatic (I), or mild/no symptomatic (A) stages

Figure 15.4 displays spatial distributions of the ration level of testing reagents in three periods of an epidemic. As the epidemic exponentially grows in the initial stage, the ration level in peripheral urban areas drops immediately, but testing reagents in the city centers is sufficient (Fig. 15.4a). As the epidemic reaches the peak, the demand

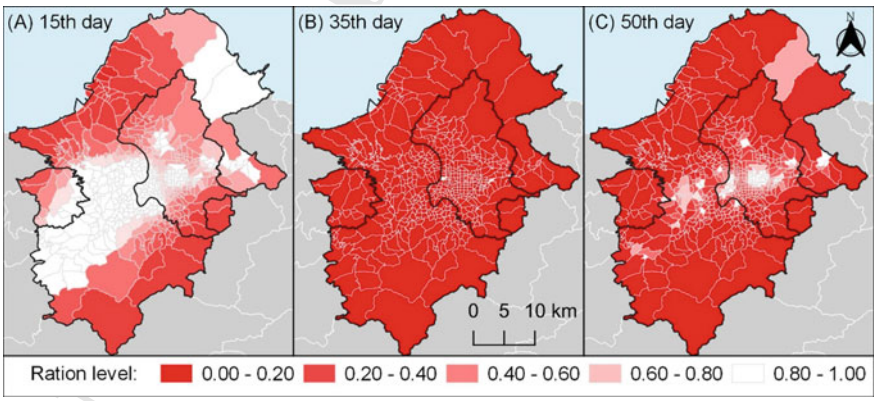


Fig. 15.4 Spatial distributions of the ration level of testing reagents in three periods of an epidemic: **a** the 15th day, **b** 35th day, and **c** 50th day

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard

doi:10.6342/NTU202101357

is more significant than the scheduled volume of testing reagents in all the areas. The volume of reagents is challenging to be provided to all the infectious people at the epidemic stage. The deficit in testing reagents could underestimate cumulative cases and intensify the severity of the following epidemic. In the last stage of the epidemic, the downtown areas start the recovery and return to a high ration level. Most urban peripheral areas are the latest ones recovering from resource deficiency (Fig. 15.4c). Our results capture each TAZ's ration level over time, and the changing rate is also different geographically.

We used two indicators to measure the temporal variations of ration levels of testing reagents in each TAZ (Fig. 15.4): insufficient onset time and the deficit duration. The insufficient onset time is defined as when the TAZ's ration level begins to be lower than 100%. The deficit duration represents the duration of the ration level in a TAZ to recover to 100%. These two indicators are highly negatively correlated (Pearson's correlation coefficient $r = -0.82$), namely that a TAZ with a longer duration of resource deficit usually encounters insufficient resources earlier. Figure 15.5 depicts the spatial distributions of insufficient onset time and the deficit duration across the study region. Our results indicate that regions with the earlier insufficient onset and longer duration of the deficit are mostly located in the outskirt of the

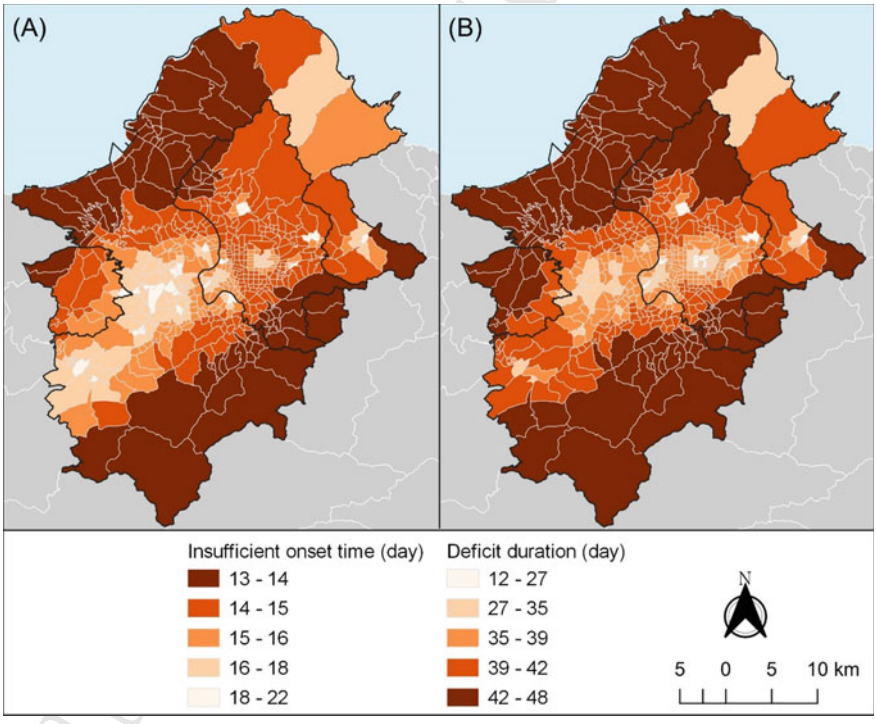


Fig. 15.5 **a** The spatial distributions of insufficient onset time and **b** the deficit duration across the study region

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard
doi:10.6342/NTU202101357

metropolitan. However, some downtown areas (north-west corner in Fig. 15.5) could also have the potential of resource deficits (earlier onset and longer duration), which could be more vulnerable than peripheral urban areas.

15.4 Discussions and Conclusion

We developed an integrated mathematical model, named Epi-RA, which combines epidemic dynamic process and resource accessibility to assess the availability of COVID-19 testing reagents in space and time. Our model framework captures the dynamic nature of an epidemic. It demonstrates that the resource demand for testing reagents and each TAZ's ration level varies over space and time. The specific regions with the earlier insufficient onset and longer duration of the deficit are identified. These areas are located in the outskirt of the metropolitan and some of the downtown areas. The results could support the health authority to implement the resource allocation schemes for confronting possible pandemic.

Methodologically, our Epi-RA model captured the spatial–temporal dynamics of epidemic-induced demand and proposed the integrated accessibility modeling to deal with time-varying demand. Previous studies on measuring geographic accessibility often used population census as the potential demand for resources, which is regarded as static spatial distribution and assumed unchanged over time (Yun et al., 2020). Population census represents population distribution in the nighttime situation and cannot reflect appropriately spatial behaviors of people seeking resources in the daytime (Fransen, Neutens, De Maeyer, & Deruyter, 2015). Recent studies highlighted the importance of considering people's routine mobility to capture more realistic resource-seeking behaviors in measuring geographical accessibility (Ma et al., 2018; Xia et al., 2019). However, their time-varying demand didn't take the frequency of people's contacts into account to not capture the dynamics of disease transmission. Our results from the Epi-RA model represent more reasonable spatial distributions of the time-varying ration level of COVID-19 testing reagents. It could provide critical information for the health authority to decide how effectively allocating community testing reagents in time and space.

Our results identified some specific locations are more vulnerable to infections due to the deficit of community testing reagents (Fig. 15.5). The low density of testing facilities and small community testing reagents in metropolitan areas' outskirts makes it reasonable that exposed individuals in these areas are likely to encounter the shortage of testing reagents. On the other hand, some downtown areas with a high density of testing facilities also can encounter testing resource shortages. Our results from the epidemic dynamic process reveal the areas where people more easily get infected could result from many contacts happening in public places (Leung, Jit, Lau, & Wu, 2017). In other words, the exponential growth of an epidemic causes a surge of infected populations and the need for pathogen testing in a short time. Therefore, some downtown areas may require more testing reagents than we ever

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard
doi:10.6342/NTU202101357

expected. The finding is also consistent with the study on the accessibility of Seoul city's emergency rooms in South Korea (Yun et al., 2020).

The deficit in medical resources could increase the severity of disease transmission. Matching the availability of medical resources should consider the demand and supply over space and time simultaneously. During the COVID-19 epidemic, the health authority must allocate medical resources appropriately. Our integrated modeling framework differentiates well-supplied and under-resourced areas. Geographically, infectious disease is quickly diffused to locations which well connected through spatial interactions. Implementing appropriate resource allocation from well-supplied to under-resourced areas could prevent vulnerable areas from high-damage epidemic centers. Similar suggestions to confront COVID-19 are also discussed in recent studies (Ranney, Griffeth, & Jha, 2020). The under-resourced areas identified in this study could be enhanced by more capacity of medical resources against a large-scale epidemic in advance. For example, locating the appropriate locations for building field hospitals (or makeshift hospitals) are widely used in many countries to house more infected persons (Arango, 2020; Chen et al., 2020; des Déserts, Mathais, Luft, Escarment, & Pasquier, 2020; Wallis, Gust, Porter, Gilchrist, & Amaral, 2020). In summary, these findings from the Epi-RA model provide more comprehensive insights into mobilizing urgent medical resources against large-scale epidemics in time and space.

There are some limitations in this study. First, the age-structured population is not considered in our model framework. The elderly people who are vulnerable groups possess lower accessibility than young people because of low motilities (El Bcheraoui et al., 2018; Horner, Duncan, Wood, Valdez-Torres, & Stansbury, 2015). Thus, distinguishing the demand for medical resources among different age groups could improve the measurement of spatiotemporal accessibility. Previous studies have incorporated the age-structure into either the FCA model (Hashtarkhani et al., 2020) or SEIR meta-population model (Tsai et al., 2011). It could be helpful to extend our model framework. Second, socioeconomic status and transportation mode may influence the individual mobility behaviors to access the medical resource (Kim, Kim, Paul, & Lee, 2020) and the risk of exposure to infection (Cooley et al., 2011; Scargrough, Holt, Hill, & Kafle, 2019). Incorporating high-resolution geo-demographics and contact structures into our integrated model warrants further investigation. Last but not least, the amount of supply is assumed to be fixed in this study. The health authority may dynamically raise the amount according to the situation of epidemic progression (Tanne et al., 2020), thus our results may underestimate the availability of medical resources.

Acknowledgements This research was supported by grants from the Ministry of Science and Technology in Taiwan (MOST 108-2638-H-002-002-MY2) and the National Health Research Institutes (MR-109-CO-14). The authors also acknowledge the financial support provided by the Infectious Diseases Research and Education Center, the Ministry of Health and Welfare (MOHW), and National Taiwan University (NTU). The funders had no role in the study design, data collection and analysis, or manuscript preparation.

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard

doi:10.6342/NTU202101357

References

Arango, C. (2020). Lessons learned from the coronavirus health crisis in Madrid, Spain: How COVID-19 has changed our lives in the last 2 weeks. *Biological Psychiatry, 88*(7), e33–e34.

Chen, S., Zhang, Z., Yang, J., Wang, J., Zhai, X., Bärnighausen, T., & Wang, C. (2020). Fangcang shelter hospitals: A novel concept for responding to public health emergencies. *The Lancet, 395*(10232), 1305–1314.

Chou, W.-R., Ko, C.-H., Lee, S.-W., Chen, Y.-Y., & Lee, T.-Y. (2018). A study of Taipei rapid transit system model version 4S (TRTS-4S). *Journal of Rapid Transit Systems and Technology, 53,* 1–45.

Cooley, P., Brown, S., Cajka, J., Chasteen, B., Ganapathi, L., Grefenstette, J., et al. (2011). The role of subway travel in an influenza epidemic: A New York City simulation. *Journal of Urban Health, 88*(5), 982–995.

El Bcheraoui, C., Mokdad, A. H., Dwyer-Lindgren, L., Bertozzi-Villa, A., Stubbs, R. W., Morozoff, C., et al. (2018). Trends and patterns of differences in infectious disease mortality among US counties, 1980–2014. *JAMA, 319*(12), 1248–1260.

Fransen, K., Neutens, T., De Maeyer, P., & Deruyter, G. (2015). A commuter-based two-step floating catchment area method for measuring spatial accessibility of daycare centers. *Health and Place, 32,* 65–73.

Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R., & Rinaldo, A. (2020). Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proceedings of the National Academy of Sciences, 117*(19), 10484–10491.

Guagliardo, M. F. (2004). Spatial accessibility of primary care: Concepts, methods and challenges. *International Journal of Health Geographics, 3,* 3.

Hashtarkhani, S., Kiani, B., Bergquist, R., Bagheri, N., VafaeiNejad, R., & Tara, M. (2020). An age-integrated approach to improve measurement of potential spatial accessibility to emergency medical services for urban areas. *The International Journal of Health Planning and Management, 35*(3), 788–798.

Horner, M. W., Duncan, M. D., Wood, B. S., Valdez-Torres, Y., & Stansbury, C. (2015). Do aging populations have differential accessibility to activities? Analyzing the spatial structure of social, professional, and business opportunities. *Travel Behaviour and Society, 2*(3), 182–191.

Ji, Y., Ma, Z., Peppelenbosch, M. P., & Pan, Q. (2020). Potential association between COVID-19 mortality and health-care resource availability. *The Lancet Global Health, 8*(4), e480.

Kim, H., Kim, D., Paul, C., & Lee, C. K. (2020). The spatial allocation of hospitals with negative pressure isolation rooms in Korea: Are we prepared for new outbreaks? *International Journal of Health Policy and Management, 9*(11), 475–483.

Kwan, M. P. (1998). Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis, 30*(3), 191–216.

Leung, K., Jit, M., Lau, E. H., & Wu, J. T. (2017). Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong. *Scientific Reports, 7,* 7974.

Lima, A., De Domenico, M., Pejovic, V., & Musolesi, M. (2015). Disease containment strategies based on mobility and information dissemination. *Scientific Reports, 5,* 10650.

Ma, L., Luo, N., Wan, T., Hu, C., & Peng, M. (2018). An improved healthcare accessibility measure considering the temporal dimension and population demand of different ages. *International Journal of Environmental Research and Public Health, 15*(11), 2421.

McGrail, M. R. (2012). Spatial accessibility of primary health care utilising the two step floating catchment area method: An assessment of recent improvements. *International Journal of Health Geographics, 11*(1), 50.

McGrail, M. R., & Humphreys, J. S. (2014). Measuring spatial accessibility to primary health care services: Utilising dynamic catchment sizes. *Applied Geography, 54,* 182–188.

Ranney, M. L., Griffeth, V., & Jha, A. K. (2020). Critical supply shortages—The need for ventilators and personal protective equipment during the Covid-19 pandemic. *New England Journal of Medicine, 382,* e41.

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard

doi:10.6342/NTU202101357

Remuzzi, A., & Remuzzi, G. (2020). COVID-19 and Italy: What next? *The Lancet, 395*(10231), 11–17.

Scargrough, A. W., Holt, M. M., Hill, J., & Kafle, R. C. (2019). Is there a relationship between income and infectious disease: Evidence from Cameron County. *International Journal of Community Well-Being, 2,* 3–13.

Tanne, J. H., Hayasaki, E., Zastrow, M., Pulla, P., Smith, P., & Rada, A. G. (2020). COVID-19: How doctors and healthcare systems are tackling coronavirus worldwide. *BMJ, 368,* m1090.

Tsai, Y.-S., Huang, C.-Y., Wen, T.-H., Sun, C.-T., & Yen, M.-Y. (2011). Integrating epidemic dynamics with daily commuting networks: Building a multilayer framework to assess influenza A (H1N1) intervention policies. *SIMULATION, 87*(5), 385–405.

Wallis, N., Gust, C., Porter, E., Gilchrist, N., & Amaral, A. (2020). Implementation of field hospital pharmacy services during the COVID-19 pandemic. *American Journal of Health-System Pharmacy, 77*(19), 1547–1551.

Wang, F. (2012). Measurement, optimization, and impact of health care accessibility: A methodological review. *Annals of the Association of American Geographers, 102*(5), 1104–1112.

Wang, L., & Wu, J. T. (2018). Characterizing the dynamics underlying global spread of epidemics. *Nature Communications, 9,* 218.

Weissman, G. E., Crane-Droesch, A., Chivers, C., Luong, T., Hanish, A., Levy, M. Z., et al. (2020). Locally informed simulation to predict hospital capacity needs during the COVID-19 pandemic. *Annals of Internal Medicine, 173*(1), 21–28.

Xia, T., Song, X., Zhang, H., Song, X., Kanasugi, H., & Shibasaki, R. (2019). Measuring spatio-temporal accessibility to emergency medical services through big GPS data. *Health and Place, 56,* 53–62.

Yun, S. B., Kim, S., Ju, S., Noh, J., Kim, C., Wong, M. S., & Heo, J. (2020). Analysis of accessibility to emergency rooms by dynamic population from mobile phone data: Geography of social inequity in South Korea. *PLoS ONE, 15*(4), e0231079.

511208_1_En_15_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:25/3/2021 Pages: 14 Layout: T1-Standard

doi:10.6342/NTU202101357