University of Cyprus

DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING

FROM ONLINE TRAFFIC MAPS TO HUMAN MOBILITY PATTERNS: ANALYSING AND VISUALISING SPATIO-TEMPORAL DYNAMICS IN URBAN AREAS

DOCTOR OF PHILOSOPHY DISSERTATION

VANA GKANIA

2020

University of Cyprus

DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING

FROM ONLINE TRAFFIC MAPS TO HUMAN MOBILITY PATTERNS: ANALYSING AND VISUALISING SPATIO-TEMPORAL DYNAMICS IN URBAN AREAS

VANA GKANIA

A Dissertation Submitted to the University of Cyprus in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

November 2020

VALIDATION PAGE

Doctoral Candidate: Vana Gkania

Doctoral Thesis Title: From Online Traffic Maps to Human Mobility Patterns: Analysing and Visualising Spatio-Temporal Dynamics in Urban Areas

The present Doctoral Dissertation was submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy at the ***Department of Civil and Environmental Engineering*** *and was approved on the …….. [date of approval] by the members of the* ***Examination Committee****.*

Examination Committee:

Research Supervisor:___ Asst. Prof. Loukas Dimitriou_____

(Name, position and signature)

Committee Member: _____Prof. Symeon Christodoulou_____

(Name, position and signature)

Committee Member: _____Assoc. Prof. Dimos Charmpis_____

(Name, position and signature)

Committee Member: _____Asst. Prof. Andreas Savvides_____

(Name, position and signature)

Committee Member:_____ Asst. Prof. Andreas Gregoriades_____

(Name, position and signature)

DECLARATION OF DOCTORAL CANDIDATE

The present doctoral dissertation was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy of the University of Cyprus. It is a product of original work of my own, unless otherwise mentioned through references, notes, or any other statements.

……Vana Gkania…….[Full Name of Doctoral Candidate]

…….[Signature]

ABSTRACT [in Greek language]

Η κατανόηση του τρόπου με τον οποίο οι άνθρωποι μετακινούνται μέσα στον αστικό χώρο μαζί με τις αλληλεπιδράσεις που αναπτύσσονται σε αυτό απασχόλησε έντονα πλήθος ερευνητών σε διάφορους επιστημονικούς τομείς όπως, η συγκοινωνιολογία, ο πολεοδομικός σχεδιασμός και η επιδημιολογία. Εντούτοις, η μοντελοποίηση και προσομοίωση της αστικής κινητικότητας παραμένει ένα δύσκολο εγχείρημα καθώς απαιτεί πληθώρα δεδομένων που σχετίζονται με τις μετακινήσεις των ανθρώπων μέσα στον αστικό ιστό. Αρχικά, η συλλογή αυτών των δεδομένων γινόταν κυρίως μέσα από χρονοβόρες και δαπανηρές έρευνες. Τις τελευταίες δεκαετίες όμως, οι "αναδυόμενες τεχνολογίες" και η πρόσβαση σε ανοιχτά δεδομένα αύξησαν το φάσμα των δυνατοτήτων για την μοντελοποίηση και την χαρτογράφηση της. Σε αυτή τη νέα πηγή πληροφοριών συγκαταλέγονται δεδομένα από κινητά τηλεφώνα, έξυπνες κάρτες και από τα μέσα κοινωνικής δικτύωσης, τα οποία προσφέρουν άπειρες ψηφιακές εγγραφές μετακινήσεων, επιτρέποντας στους ερευνητές να κατανοήσουν καλύτερα τους νόμους που διέπουν την κινητικότητα των ανθρώπων στις αστικές περιοχές. Ωστόσο, οι νέες αυτές πηγές δεδομένων κινητικότητας παρουσιάζουν ορισμένα μειονεκτήματα καθώς μπορεί να μην ανταποκρίνονται σε αντιπροσωπευτικά δείγματα του πληθυσμού ή η πρόσβαση σε αυτά είναι να είναι περιορισμένη.

Στην παρούσα διατριβή χρησιμοποιείται για πρώτη φορά μια εναλλακτική πηγή δεδομένων για την εκτίμηση και την ερμηνεία της αστικής κινητικότητας και συγκεκριμένα οι διαδικτυακοί χάρτες κυκλοφορίας, οι οποίοι παρέχουν πληροφορίες για την οδική κυκλοφορία σε πραγματικό χρόνο. Στα πλεονεκτήματα των διαδικτυακών χαρτών κυκλοφορίας συγκαταλέγονται η διάχυτη παρουσία τους καθώς και η διαδραστικότητα και η ελεύθερη πρόσβαση που προσφέρουν σε σύγκριση με άλλες πηγές. Συνεπώς, είναι σημαντικό να διερευνηθεί η αξία τους για περαιτέρω εφαρμογές πέρα από τις πληροφορίες κυκλοφορίας σε πραγματικό χρόνο και τους χρόνους διαδρομής.

Βασικός στόχος είναι η διερεύνηση της αξίας της πληροφορίας που απεικονίζεται στους διαδικτυακούς χάρτες κυκλοφορίας πέρα από την πρώτη οπτική ερμηνεία και το όριο που θέτει η ανθρώπινη όραση. Σημειώνεται ότι, η δημιουργία των διαδικτυακών χαρτών κυκλοφορίας απαιτεί τη συλλογή τεράστιου όγκου λεπτομερών δεδομένων κινητικότητας σε πραγματικό χρόνο, τα οποία συμπυκνώνονται για σκοπούς αναγνωσιμότητας. Με την εξαγωγή αυτής της συμπυκνωμένης πληροφορίας, χρησιμοποιώντας τεχνικές επεξεργασίας εικόνας δίνεται η δυνατότητα για την διακριτοποίηση του αστικού χώρου σε μικρότερα στοιχεία απεικόνισης (pixels), στα οποία καταγράφεται η κατάσταση της κυκλοφορίας με

βάση κάποιο συγκεκριμένο χρωματικό κώδικα, σε μια κατάλληλη δομή δεδομένων πολύτιμη για μετα-ανάλυση και ερμηνεία των μοτίβων αστικής κινητικότητας.

Το πρώτο μέρος της διατριβής εστιάζει στον έλεγχο της εγκυρότητας των δεδομένων που αντλούνται από την επεξεργασία των διαδικτυακών χαρτών κυκλοφορίας, μέσα από την εκτίμηση θεμελιωδών σχέσεων κυκλοφορίας, όπως τα Μακροσκοπικά Θεμελιώδη Διαγράμματα της Κυκλοφορίας ή τον εντοπισμό άλλων δυναμικών φαινόμενων, όπως η κυκλοφοριακή υστέρηση σε επίπεδο δικτύου.

Σε δεύτερο στάδιο, διερευνάται η χρονική και χωρική διάσταση της αστικής κινητικότητας με βάση τα δεδομένα που λήφθηκαν από την επεξεργασία των διαδικτυακών χαρτών κυκλοφορίας. Το προτεινόμενο μεθοδολογικό πλαίσιο βασίζεται σε αλγόριθμους μηχανικής όρασης, στην εκτίμηση διαγραμμάτων πυκνότητας και σε διάφορες τεχνικές ομαδοποίησης. Μέσα από την εφαρμογή στην περιοχή μελέτης, η οποία αποτελείται από δεκαοκτώ διαφορετικές πόλεις σε όλο τον κόσμο, αποκαλύπτονται κανονικότητες και επαναλαμβανόμενα μοτίβα καθώς και χρήσιμες πληροφορίες σχετικά με τη δυναμικά χαρακτηριστικά της αστικής κινητικότητας, συμβάλλοντας στην ευρύτερη κατανόηση του πολύπλοκου και πολυπαραγοντικού αυτού φαινομένου.

Εν κατακλείδι, η περαιτέρω ανάπτυξη της προτεινόμενης μεθόδου μπορεί να οδηγήσει σε ένα νέο χρήσιμο εργαλείο, χαμηλού κόστους για την αντιμετώπιση σύνθετων αστικών ζητημάτων, πολύτιμο για τους πολεοδόμους, τους υπεύθυνους χάραξης πολιτικής αλλά και για τα κέντρα διαχείρισης της κυκλοφορίας.

ABSTRACT

The understanding of how people move along with the relationship between urban forms has been a vivid research topic in several scientific fields such as transportation engineering, urban planning, and epidemiology. Hence, urban mobility modeling remains a challenging task as it requires a plethora of mobility data. In the early years, this data was derived from costly manual surveys but recently the emerging technologies and the data availability enhanced the spectrum of possibilities in mobility modeling and mapping. This new source of information such as mobile phone data, smart card data, floating car data, and web-based sources services that create digital records enabled researchers to better understand the governing laws in human mobility within urban areas. Although, these new sources of mobility data can also be significantly biased or have other drawbacks such as limited access.

In this thesis, an alternative source of data, namely aggregated traffic information, broadcasted by online traffic maps are utilized for the first time to interpret urban mobility dynamics. The merits of online traffic maps lie on the ubiquitous and low-cost characteristics of this type of opensource data compared to other sources and thus it is crucial to explore their value for further applications beyond real-time traffic information and travel times. The main objective is to investigate the value of traffic information that is depicted at online traffic maps beyond the bounds of information that a human eye can perceive. Online traffic maps facilitate the collection of the vast amount of extremely detailed mobility data in real time that is aggregated for the sake of readability. The extraction of this aggregated information using image processing techniques enables to perform discretization of the urban space in seamless pixels, capturing the traffic state in each pixel based on the colour code in a suitable data structure valuable for meta-analysis and patterns interpretation.

Initially, the validity of the simplified/coded information that dynamic traffic maps provide in terms of traffic operational characteristics is investigated. The results revealed that fundamental traffic relationships, such as the Macroscopic Fundamental Diagrams (MFDs), hold even in the case of the abstracted information broadcasted from online traffic maps while in a meta-analysis stage it was able to capture spatio-temporal phenomena of urban mobility, like concentration and homogeneity.

As a second challenge the temporal and spatial dimension of urban mobility is explored based on the information retrieved from online traffic maps. The proposed methodological framework to estimate and interpret the spatiotemporal mobility patterns is built on Computer Vision algorithms, Kernel Density estimation, and clustering techniques. The

application of the method on over fifteen cities around the world revealed regularities and useful insights regarding the urban mobility dynamics, enhancing our understanding in the particularly complex and multifaceted world of human mobility in urban areas.

Last, further development of the proposed method can lead to a new useful and low-cost tool to confront complex urban issues valuable for urban planners, policy makers and traffic management' authorities.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my supervisor Asst. Prof. Loukas Dimitriou for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

Next, I would like to thank my thesis committee: Prof. Symeon Christodoulou, Assoc.Prof. Dimos Charmpis, Asst. Prof. Loukas Dimitriou (supervisor) for their insightful comments and encouragement.

My sincere thanks to the personnel of the Public Works Department of the Ministry of Transport, Communications, and Works in Cyprus for giving me access to "Diavlos" data, used in this thesis.

Finally, I would like to thank my family, my friends, and my fellow labmates for supporting me spiritually throughout this research journey.

Table of Contents

Chapter 1 : Introduction 1

1.1 Motivation 1

1.2 Scope and Contribution 3

1.2 Research Objectives 4

1.3 Organization of the Thesis 4

Chapter 2 : Literature Review 6

2.1 Network-Level Traffic Relationships 6

2.1.1 Earlier works with empirical observations 6

2.1.2 Recent empirical studies on urban MFD 7

2.1.3 An alternative source of empirical data for the MFD estimation 8

2.2 Human Mobility Modeling and Mapping 9

2.2.1 Terms and motivation 9

2.2.2 Mobility mapping eras and milestones 10

2.2.3 Categorizing human mobility studies based on data source 14

2.3 Chapter Summary 34

PART I 35

Chapter 3 : Traffic Flow Fundamentals 35

3.1 Fundamental Characteristics of Traffic Flow 35

3.2 Macroscopic Stream Models 38

3.2.1 Calibration of the Greenshields's model 40

3.2.2 Other macroscopic stream models 43

3.3 Traffic Flow Dynamics 44

3.4 Macroscopic or Network Fundamental Diagrams (MFD-NFD) 46

3.5 Chapter Summary 47

Chapter 4 : Experimental Setup 49

4.1 Study Area Characteristics 49
4.2 Data Collection 53
4.3 Data Processing 55
4.3.1 Raster Images 55
4.3.2 Image Processing 56
4.4 Data Analysis 59
4.4.1 Urban coverage 59
4.4.2 Timeseries of traffic through pixels 62
4.5 Chapter Summary 72
Chapter 5 : Validation of class-type Information from Online Traffic Maps 73
5.1 Methodological Framework 73
5.1.1 Assumptions 74
5.1.2 Estimation of MFD based on traffic maps data 76
5.2 Calibration with Empirical Data 77
5.3 MFD based on the Greenshields model for the city center of Nicosia 81
5.4 Case Studies for the Different Urban Networks and Insights for Generalization 84
5.4.1 Application based on traffic information from Yandex maps 84
5.4.2 Network Operations and Performance – Phenomena of Traffic Hysteresis 86
5.4.3 Application based on traffic information from Google, Bing, and Here maps 91
5.5 MFD estimation based on the Greenberg Model 102
5.6 Chapter Summary 106
PART II 107
Chapter 6 : Human Mobility Patterns based on Information from Online Traffic Maps 107
6.1 Urban Morphology 107
6.1.1 London 107
6.1.2 Paris 108
6.1.3 Berlin 108

6.1.4 Moscow 109
6.1.5 Nicosia 109
6.1.6 New York 109
6.1.7 Los Angeles 110
6.1.8 Toronto 110
6.1.9 Sao Paulo 111
6.1.10 Buenos Aires 111
6.1.11 New Delhi 112
6.1.12 Riyadh 112
6.1.13 Singapore 112
6.1.14 Beijing 113
6.1.15 Tokyo 113
6.1.16 Johannesburg 114
6.1.17 Sydney 114
6.1.18 Istanbul 114
6.2 Methodological Approach and Outcomes 115
6.2.1 Quantification of variability in human mobility 115
6.2.2 Spatio-temporal visualization of human mobility 129
6.3 Dynamic Clustering and Propagation of Congestion 137
6.4 Chapter Summary 141
Chapter 7 : Conclusions and Future Directions 142
7.1 Conclusions 142
7.2 Future Directions 145
Publications 147
References 149
Appendix A 159
Appendix B 176

Appendix C 183
Appendix D 189
Appendix E 203

List of Figures

Figure 2.1:Milestones in human mobility mapping during the 19th century. 13
Figure 2.2:Contemporary visualizations of human mobility based on different source of data. 13
Figure 3.1:Illustration of relation between fundamental parameters of traffic flow 37
Figure 3.2: Basic Stream Flow Diagrams 39
Figure 3.3: Effect of measurement location on data obtained. 42
Figure 3.4: (a) Greenberg's logarithmic model; (b) Underwood's exponential model 44
Figure 3.5: The capacity drop in the flow-density diagram. 45
Figure 3.6: Hysteresis loop observed by Newell (Newell, 1962) from volume-density plot. 46
Figure 4.1: Spatial distribution of the selected cities. 49
Figure 4.2: Data collection 53
Figure 4.3: Decomposition of a raster image of size 394×394 pixels, to the pixels' intensities in a small area and the RGB intensities for the three planes of one pixel. 56
Figure 4.4: (a) Original image size (width x height) 1920 x 1080; (b) Cropped image size (width x height) 873 x 639; c) Final image size (width x height) 873 x 639 58
Figure 4.5:Traffic coverage for the city of Paris from Google maps (left) and Bing maps (right). 60
Figure 4.6: Traffic coverage for the city of Paris from Here maps (left) and Yandex maps (right). 60
Figure 4.7: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Paris; (b) London; (c) Istanbul from Google maps. 64
Figure 4.8: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) New Delhi; (b) Moscow; (c) New York from Google maps. 65
Figure 4.9: Timeseries of pixels and boxplot per traffic layer for one weekday for Los Angeles from (a) Los Angeles; (b) Sao Paulo; (c) Singapore from Google maps. 66
Figure 4.10: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Sydney; (b) Johannesburg; (c) Toronto from Google maps. 67
Figure 4.11: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Berlin; (b) Buenos Aires; (c) Riyadh from Google maps. 68
Figure 4.12: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Tokyo form Google maps and (b) Beijing from Baidu maps. 69

Figure 4.13: Timeseries of pixels and boxplot per traffic layer for Nicosia for (a) Tuesday; (b) Saturday; (c) whole week from Google maps. 70
Figure 4.14: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Paris; (b) Istanbul; (c) Moscow from Yandex maps. 71
Figure 5.1: Location of the loop detectors in Nicosia, Cyprus 78
Figure 5.2: Comparison between the proposed model and the empirical data from the loop detectors at each location 1004(a), 1005(b), 1006(c), and 1010(d) for one week (11-17 February 2019). 80
Figure 5.3: Comparison between the proposed models and the empirical data from the all the available loop detectors and for the whole network for one week (11-17 February 2019). 81
Figure 5.4: Nicosia's MFDs, for the selected week, (5-minute intervals): (a) average flow-average density diagram; (b) average speed-average density diagram; (c) average speed-average flow diagram. 82
Figure 5.5: The relationship between average network speed(a); flow(b) versus the total number of vehicles, and network's traffic state for selected values of average accumulation during February 11th, 2019 (5-minute intervals). 83
Figure 5.6: MFDs (average flow-average density diagram) for one typical weekday (5-minute intervals): (a) Paris; (b) Moscow; (c) Istanbul. 85
Figure 5.7: The relationship between average network flow and accumulation for (a), Kj= 150 veh/km and Uf= 80Km/h, and (b) Kj= 150 veh/km and Uf= 90Km/h (5-minute intervals) for the city of Moscow. 86
Figure 5.8: The relationship between average network flow (a), speed (b) versus the total number of vehicles, and network's traffic state for selected values of average accumulation during a weekday (5-minute intervals) for the city of Paris. 87
Figure 5.9: The relationship between average network flow (a), speed (b) versus the total number of vehicles, and network's traffic state for selected values of average accumulation during a weekday (5-minute intervals) for the city of Moscow. 88
Figure 5.10: The relationship between average network flow (a), speed (b) versus the total number of vehicles, and network's traffic state for selected values of average accumulation during a weekday (5-minute intervals) for the city of Istanbul. 89
Figure 5.11: Average traffic flow vs. the total number of vehicles diagram during a weekday (5-minute intervals) for Paris: (a) anti-clockwise hysteresis; (b) clockwise hysteresis and for Istanbul; (c) clockwise hysteresis; (d) anti-clockwise hysteresis. 90

Figure 5.12: Paris's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram. 92
Figure 5.13: Istanbul's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram. 93
Figure 5.14: Moscow's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram. 94
Figure 5.15: Average network flow versus accumulation(total number of vehicles) for the city of (a) Paris; (b) London; (c) Istanbul (5-minute intervals): (left column) from Google; (middle column) from Bing; and (right column) from Here maps. 97
Figure 5.16: Average network flow versus accumulation(total number of vehicles) for the city of (a) New Delhi; (b) Moscow; (c) New York (5-minute intervals): (left column) from Google; (middle column) from Bing; and (right column) from Here maps. 98
Figure 5.17: Average network flow versus accumulation(total number of vehicles) for the city of Los Angeles; (b) Sao Paulo; (c) Singapore (5-minute intervals): (left column) from Google; (middle column) from Bing; and (right column) from Here maps. 99
Figure 5.18: Average network flow versus accumulation(total number of vehicles) for the city of (a) Sydney; (b) Johannesburg; (c) Toronto (5-minute intervals): (left column) from Google; (middle column) from Bing; and (right column) from Here maps. 100
Figure 5.19: Average network flow versus accumulation(total number of vehicles) for the city of (a) Berlin; (b) Buenos Aires; (c) Riyadh (5-minute intervals): (left column) from Google; (middle column) from Bing; and (right column) from Here maps. 101
Figure 5.20: Average network flow versus accumulation(total number of vehicles) for the city of (a) Tokyo from Google; (b) Beijing from Bing and (c) Beijing from Baidu maps. 102
Figure 5.21: Paris's MFDs, for a typical weekday, (5-minute intervals) from Yandex maps based on (a) Greenshields model and (b) Greenberg model; (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram 103

Figure 5.22: Moscow's MFDs, for a typical weekday, (5-minute intervals) from Yandex maps based on (a) Greenshields model and (b) Greenberg model; (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram. 104
Figure 5.23: Istanbul's MFDs, for a typical weekday, (5-minute intervals) from Yandex maps based on (a) Greenshields model and (b) Greenberg model; (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram. 105
Figure 5.24: Average network flow versus accumulation(total number of vehicles) for the city of (a) Paris; (b) Moscow and (c) Istanbul (5-minute intervals) from Yandex maps based on Greenshields model. 105
Figure 5.25: Average network flow versus accumulation(total number of vehicles) for the city of (a) Paris; (b) Moscow and (c) Istanbul (5-minute intervals) from Yandex maps based on Greenberg model. 106
Figure 6.1:SSIM index for the city of (a) Paris, London, Istanbul; (b) Istanbul, New Delhi Sao Paulo; (c) New York, Los Angeles, Tokyo. 118
Figure 6.2: SSIM1 index for the city of (a) Singapore, Sydney, Johannesburg ; (b) Toronto, Berlin, Buenos Aires ; (c) Riyadh, Nicosia, Beijing. 119
Figure 6.3: SSIM1 index for the city of Paris, Moscow and Istanbul from 120
Figure 6.4: SSIM1 index for the city of Paris, Moscow and Istanbul from (a) Bing; (b) Here and (c)Yandex maps 121
Figure 6.5: Variability of SSIM1 for the selected cities a) Boxplot graph; b) Violin graph 122
Figure 6.6: SSIM2 index for the city of (a) Paris, London, Istanbul; (b) Istanbul, New Delhi Sao Paulo; (c) New York, Los Angeles, Tokyo. 124
Figure 6.7: SSIM2 index for the city of (a) Singapore, Sydney, Johannesburg ; (b) Toronto, Berlin, Buenos Aires ; (c) Riyadh, Nicosia, Beijing. 125
Figure 6.8: SSIM2 index for the city of Paris, Moscow and Istanbul from (a) Google and (b) Bing maps 126
Figure 6.9: SSIM2 index for the city of Paris, Moscow and Istanbul from (a) Here and (b) Yandex maps 127
Figure 6.10: Variability of SSIM2 for the selected cities a) Boxplot graph; b) Violin graph 128

Figure 6.11:Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) Paris; (b) London; (c) Moscow. 131
Figure 6.12:Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) New York; (b) Los Angeles; (c) Tokyo. 132
Figure 6.13: Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) Istanbul; (b) New Delhi; (c) Sao Paulo. 133
Figure 6.14: Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) Singapore; (b) Sydney; (c) Johannesburg 134
Figure 6.15: Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) Toronto; (b) Berlin; (c) Buenos Aires 135
Figure 6.16: Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) Riyadh; (b) Nicosia; (c) Beijing 136
Figure 6.17: Steps followed to create congestion patterns: (Step 1) Heavy traffic layers isolation; (Step 2) Image segmentation to meaningful atomic regions by clustering pixels; (Step 3) Setting the color of each superpixel region; (Step 4) Congestion patterns visualization. 138
Figure 6.18: Spatio-temporal congestion patterns for Moscow (left) and Istanbul (right) during morning peak hours (07:00-08:00 a.m.). 139
Figure 6.19: Spatio-temporal congestion patterns for New York (left) and Los Angeles (right) during morning peak hours (07:00-08:00 a.m.). 140

List of Tables

Table 2.1:Related Studies based on Census Data and Surveys. 24
Table 2.2: Related Studies based on Mobile Phone Data 26
Table 2.3: Related Studies based on GPS and Smart Card Data 28
Table 2.4: Related Studies based on Online Social Network Data 31
Table 3.1:Framework for Fundamental Characteristics of Traffic Flow 35
Table 4.1:Cities characteristics 51
Table 4.2: Cities characteristics (continued) 52
Table 4.3: Data summary collection 54
Table 4.4: Urban coverage per city and provider. 61
Table 5.1: Nomenclature 73
Table 5.2: Calibration Results 79

List of Notations

S	Scale of the thematic map
p_c	Pixels coloured with specific color c
U_c	Representative value of average speed according to color c from the speed map legend, in Km/h
c	Color index (i.e. g for green, y for yellow, o for orange and r for red)
j	Link index
i	Pixel index
m	Total number of links in one image
n	Total number of pixels in one image
l_j	Length of link j
t	time interval
U^t	Average network speed across all the links of a given network, in Km/h at time interval t
Q^t	Average network flow across all the links of the network, in veh/h at time interval t
K^t	Average network density across all the links of the network, in veh/km at time interval t
A^t	Total number of vehicles (accumulation) on the network at time interval t
U_f	Free-flow speed of a network
U_o	Optimum speed of a network
K_{jam}	Jam Density of a network

Chapter 1 : Introduction

1.1 Motivation

Maps as mediators between an inner mental world and an outer physical world are fundamental tools helping the human mind make sense of its universe at various scales. Thus, the use of map is both extremely ancient and extremely widespread. It could be said that, maps constitute a specialized graphic language, an instrument of communication that has influenced behavioral characteristics and the social life of humanity (Freundlich, 2011). The widespread use of maps lies on the immediacy about the message in a map that makes it more readily perceived than knowledge encoded in other ways. The most significant attribute of a map is that it can be taken in quickly by the eye, contributing to the potency of cartographic images. Thus maps have been associated with cultures that differ widely in social or technological development while an early psychological research has shown that children can derive meaning from maps (and indeed draw them) from young age (Piaget and Inhelder, 1967). The variety of its occurrences is vast as anything that can be spatially conceived can be mapped. There are political maps, physical maps, topographic maps, climate maps, economic or resource maps, road maps, thematic maps etc. A road map is one of the most widely used map types.

In the 20th century the use of aerial photography changed the types of data that could be used to create maps. Furthermore, the use of satellite imagery aided in showing large areas in detail. Over and above, GIS technology introduced a new era in cartography by allowing the creation of many different types of maps using various types of data. Likewise, the digital revolution of the last decades has seen a major shift from paper to electronic maps with the rise of mobile technology. MapQuest (MapQuest, 2020) was the first well-known online mapping service launched on the Internet on February 5, 1996 and changed the way people obtain street maps and directions forever. Since then, several other Internet companies, including Google, Yahoo and Bing have come online with mapping and routing services.

Online traffic maps provide the means for augmenting valuable traffic information including the spatial element, resulting to an important instrument to organize, observe and communicate valuable (but complex) information for making decisions, in a relevant and comprehensible manner. The recent technological developments in the traffic sensor and surveillance systems and especially in information and communication technologies,

facilitate the collection of vast amounts of extremely detailed mobility data in real-time, making possible the development of dynamic traffic and mobility maps, able to be broadcasted online to authorities and–up to a certain degree-to interested other users (professionals and individual travellers). This type of products and services has become extremely valuable in everyday life and thus reasonably massively popular.

The standard that is used for depicting traffic in roadmaps makes use of a colour coding scheme representing alternative traffic states in links, while in some cases the links' line width is also used for classifying road type or provide additional information (e.g. traffic volume). By this manner, it can compose a valuable message about the network conditions, communicated in the form of visual signal (image) and easily comprehended by both experts and ordinary travellers. Although traffic maps have become an important tool for monitoring first and then manage and control of traffic systems, a question raises about the value of the information that is broadcasted in terms of traffic operational characteristics and if the information reduction that is performed in the colour-coding process still possess valid traffic characteristics, especially in the scale of large and complex urban areas. This thesis addresses these questions through a thorough investigation of the traffic information that is broadcasted in online traffic maps.

At first, online traffic maps are processed as raster images, providing a way to discretize the urban space (and the road network therein) in seamless micro-segments (pixels), capture the traffic states in each pixel based on the colour-code of each of them and come up with a suitable data structure able to organize the spatio-temporal information of urban traffic networks for meta-analysis. The theoretical foundations used for the meta-analysis are based on the fundamental relationships of traffic flow, both at the microscopic as well as at its generalization on the macroscopic scale, especially with the well-defined Macroscopic Fundamental Diagram-(MFD). The estimation and the investigation of the MFD properties are performed on commercial freely available online traffic maps that offer global coverage across many cities.

As a second challenge the temporal and spatial dimension of urban mobility is explored based on the information retrieved from online traffic maps. The proposed methodological framework to estimate and interpret the spatiotemporal mobility patterns is built on Computer Vision algorithms, Kernel Density estimation, and clustering techniques. The application of the method on over fifteen cities around the world revealed regularities

and useful insights regarding the urban mobility dynamics, enhancing our understanding in the particularly complex and multifaceted world of human mobility in urban areas.

Together the proposed techniques enable both monitoring and control of urban road networks while a concrete connection between human mobility research and real-time traffic information is established.

1.2 Scope and Contribution

The scope of this thesis is two-fold. Firstly, it attempts to investigate the validity of the simplified/coded information that dynamic traffic maps provide in terms of traffic operational characteristics. Secondly, it intends to apply the retrieved information for the estimation of spatiotemporal mobility patterns. Towards the research scope, the following questions were formulated:

1. What type of information can be retrieved from online traffic maps that the human eye cannot perceive?
2. Is there a connection between the real urban space and the seamless micro-segments (pixels) of the traffic maps images?
3. Do fundamental traffic relationships such as the MFD or traffic hysteresis phenomena hold even in the case of abstracted information broadcasted in online traffic maps?
4. How this type of information can be further applied on urban mobility modeling for cities that have online traffic coverage?

The contribution of the thesis lies on the thorough investigation of the traffic information displayed on traffic maps both on the micro and the macro level along with the deep understanding of urban mobility dynamics at a high level using quantified indices and mesmerizing visualization tools and models. Specifically, in **Chapter 5** the capacity of online traffic maps is investigated in terms of traffic operational characteristics. Here, a connection between the average network flow, the average network speed, the average network density, and the number of pixels that belong to distinctive operating class of online traffic maps is established exploiting the fundamental linear speed-density relation proposed by Greenshields. Moreover, **Chapter 6** introduces alternative and innovative methods in understanding movement within the urban fabric at a macroscopic level, based on suitably developed quantified indices that may be applied to any city that has online traffic coverage.

Ultimately, this work introduces a new perspective on tackling urban mobility challenges, utilizing the ubiquitous aggregated data from online traffic maps.

1.2 Research Objectives

The research objectives of the thesis can be summarized into the following topics:

a. To gain a deep understanding of online traffic maps by investigating fundamental traffic flow characteristics and relationships that can be extracted by the aggregated type of traffic information used in them,
b. To form the mathematical connection between the discretized/digitized map images on seamless pixels and the aggregated traffic variables across entire urban areas,
c. To describe urban mobility through quantified indices and models that can be applied to every city that has online traffic coverage,
d. To establish a concrete connection between human mobility research and real-time traffic information.

Towards these objectives, data in the form of raster images were collected for one-week period for over fifteen cities around the world from different online providers/sources. Additional data from loop detectors were collected, for validation testing for the city of Nicosia, Cyprus.

1.3 Organization of the Thesis

The rest of the thesis is organized as follows:

Chapter 2 provides an extent review of related work into two main categories following the two-fold scope of the thesis. Firstly, a review among network-level traffic relationships and models is provided, that focuses on earlier works with empirical data and landmark studies on the concept of the macroscopic traffic flow analysis. In the current thesis, the feasibility of using online traffic maps information for capturing network traffic phenomena (e.g. MFD estimation) is investigated as a validation tool for further meta-analysis and urban mobility modeling. Thus, recent empirical studies on urban MFD and their application are further discussed. The rest of the literature review is devoted on the human mobility and mapping studies. The motive of this review is to give a short resume of the status of our knowledge and understanding in human mobility, and the way it is visualized. Mobility mapping areas and milestones are initially introduced while significant human mobility studies are presented based on the data source.

Chapter 3 introduces the fundamental characteristics of traffic flow and the relationships between these parameters. To figure out the exact relationship between the traffic parameters, a great deal of research has been done over the past several decades. Most important among them is the relation between speed and density. The results of this research yielded many mathematical models. Thus, emphasis is given on the most known models such as the single regime models of Greenshields, Greenberg, and Underwood, that dominated for many years, Afterwards, Macroscopic Fundamental Diagrams that are employed in Chapter 5 as a validation tool are further discussed.

Chapter 4 is devoted on the experimental setup. Starting with the presentation of the study area, the spatial distribution of the cities under investigation is provided along with the basic characteristics such as population, land, population density. After the selection of the study area, the description of the data collection follows. Then the data processing section presents the steps followed for the traffic information extraction while a preliminary analysis in the form of timeseries is provided at the end of the chapter.

In **Chapter 5** the validity of the simplified/coded information that dynamic traffic maps provide is investigated in terms of traffic operational characteristics. Particularly, the estimation and the investigation of the Macroscopic Fundamental Diagram-MFD properties are performed utilizing the captured pixels properties for the selected study area. The validity of the method is tested by comparing the estimated MFDs to ground-truth MFD obtained using empirical data from loop detectors. Further, other macroscopic models are demonstrated for comparative reasons.

Chapter 6 extends the work in chapter 4, starting with a closer look to urban morphology of the cities in the study area while estimation of the temporal and spatio-temporal patterns follow. The chapter ends with the dynamic clustering and propagation of congestion analysis, providing dynamic congestion patterns for the selected urban areas.

Chapter 7 ends the thesis by offering some concluding remarks and emphasizing the original contribution of the dissertation. Several issues that were not investigated in this work are included in this chapter as future work. The list of future directions presented in this chapter is not exhaustive, but it is meant to serve as a guide towards some interesting directions that warrant investigation.

Chapter 2 : Literature Review

The work described in this thesis ultimately has two intertwined goals: (1) to investigate the validity of traffic information that is broadcasted in dynamic online maps through fundamental network-traffic relationships, and (2) to apply this extracted traffic information in a meta-analysis stage on human mobility modeling and mapping for urban areas. In this Section a review of related work is provided both for network-level traffic relationships (subsection 2.1) and for human mobility modeling and mapping (subsection 2.2).

2.1 Network-Level Traffic Relationships

The starting point for observing and modeling road traffic at networks started at the case of specific locations/cross-sections and date back to the seminal work of Greenshields (Greenshields et al., 1934, Greenshields 1935), who carried out tests to measure traffic flow, traffic density and speed using photographic measurement methods for the first time. The conversion from single-entity level characteristics of traffic flow to comparable system-level characteristics was first introduced by Lighthill and Whitham (Lighthill and Whitham, 1955) and Richards (Richards, 1956) who independently proposed a simple continuum model to describe the characteristics of traffic flow, known as the LWR model. Their study attracted attention as essential features of traffic flow, such as wave formation and propagation, could be qualitatively well reproduced with the LWR model (Transportation Research Board, 2011).

2.1.1 Earlier works with empirical observations

Moving to the network-wide traffic modeling, in the 1960s several landmark studies appeared exploring network-level traffic flow relationships. (Smeed 1966&1968, Thomson, 1967, Wardrop, 1968, Godfrey, 1969). Later Zahavi (Zahavi, 1972) worked on the relationship between network-level parameters of traffic intensity, road density, and the weighted space mean speed using data from England and the United States.

In the late 1970s, Prigogine and Herman (Herman and Prigogine, 1979) proposed the two-fluid theory in order to characterize the flow of vehicles in the urban network. The two-fluid model assumes that vehicular traffic in an urban network can be distinguished into stopped vehicles and running vehicles. The model is based on two assumptions: first, the average speed in a road network is proportional to the fraction of vehicles that are moving and second the fractional stop time of a test vehicle circulating in a network is equal to the average fraction of the vehicles stopped during the same period (Gartner et al., 2001).

Empirical validation of the two-fluid theory appeared in the 1980s utilizing data obtained using chase-car techniques, supplemented by aerial photography (Herman and Ardekani, 1984, Ardekani et al., 1985, Ardekani and Herman, 1987). Mahmassani et al. (1984a) and Williams et al. (1985) during their study of the two fluid models related the three fundamental traffic variables speed-flow-concentration at a network level based on data from simulation and indicated that they were similar to those on individual road facilities. A comprehensive review of these macroscopic flow models is included in the Traffic Flow Theory monograph revised in 2001 (Gartner et al., 2001).

2.1.2 Recent empirical studies on urban MFD

More than a decade later, Daganzo (Daganzo, 2007) revisited the concept of the macroscopic traffic flow analysis as part of an urban traffic dynamics model for improving city mobility through gridlock control. Traditionally, a fundamental diagram reveals the traffic condition on a short road section, relating traffic variables such as flow, density and speed. Although, Geroliminis and Dagazo (Geroliminis and Daganzo, 2008) in a field experiment in Yokohama, Japan, expanded the concept of the fundamental diagram and proved that average flow and density are indeed related by a reproducible curve for a complete network, which has come to be known as the Network or Macroscopic Fundamental Diagram (NFD or MFD). The data on the network level are easily generated by averaging the link data while a sufficient number of observations under various traffic conditions is required to ensure all states in the MFD can be demonstrated.

So far, many researchers have examined several different applications of the MFD for improving traffic control. For instance studies including gating strategies (Keyvan-Ekbatani et al., 2012), pricing strategies (Geroliminis and Levinson, 2009, Amirgholy and Gao, 2017) or routing strategies (Knoop, Hoogendoorn and Van Lint, 2012; Yildirimoglu, Ramezani and Geroliminis, 2015) that are sensitive to the functional form of the MFD. The factors that influence the shape of an MFD also drew research attention. Geroliminis and Sun (Geroliminis and Sun, 2011) explored the effect of the spatial variability of vehicle density on the shape and the scatter of a well-defined MFD. In (Ji and Geroliminis, 2012) a partitioning mechanism based on the criteria of a well-defined MFD in the urban transportation networks was designed. The effects of adaptive driving on network capacity and traffic instability in a simulated network model was examined in (Saberi, Mahmassani and Zockaie, 2014). The impacts of locally adaptive traffic signals on network stability and the MFD were studied by (Gayah et al., 2014). Furthermore, a connection between network-

wide travel time reliability and the MFD concepts have been investigated by (Mahmassani et al., 2013, Gayah et al., 2014, Yildirimoglu et al., 2015). Geroliminis et al. (Geroliminis et al., 2014) extended the modeling and the application of the single-model MFD to a bi-modal (bus and cars) one, with the consideration of passenger flows and traffic performance of each mode. Thus, a three-dimensional Macroscopic Fundamental Diagram (3D-MFD) for mixed bi-modal urban networks was investigated based on simulation data while the first empirical estimate of the 3D-MFD can be found in (Loder et al., 2017). Last, another recent study explores the existence and the characteristics of the pedestrian MFD (Hoogendoorn et al., 2017).

2.1.3 An alternative source of empirical data for the MFD estimation

While it is convenient to use an MFD to describe the traffic status across a network and design traffic control strategies the data needed to plot the MFD are not always readily available. Up to now, there are typically two empirical data sources considered as viable for the estimation of the MFD: Loop Detector Data (LDD) and Floating Car Data (FCD) or fusing both data sources to estimate a more accurate MFD. Specifically, a combination of LDD flows and FCD speeds partly eliminates key drawbacks of the two data sources. A preliminary study (Ambühl and Menendez, 2016) has shown that applying a data fusion algorithm increases the accuracy of the MFD estimation.

Throughout the literature, empirical studies on urban MFD estimation were conducted for the city of Toulouse, France (Buisson and Ladier, 2009), Rome, Italy (Bazzani et al., 2011), Brisbane, Australia (Tsubota et al., 2014), Shenzhen, China (Ji et al., 2014), Sendai, Japan (Wang et al., 2015), Chania, Greece (Ampountolas and Kouvelas, 2015), Changsha, China (Beibei et al., 2016) and Zurich, Switzerland (Ambühl et al., 2016) utilizing either LDD or FCD.

In (Guenan, 2014) the limitations of each data source are pointed out while a combination of them is proposed following an approach by Leclercq et al. (Leclercq, Chiabaut and Trinquier, 2014) that leads to a well- defined MFD, reducing key drawbacks of each data source. Loop detectors are usually installed close to traffic signals and on links with greater congestion probability while for a reliable MFD, loop detectors must be positioned uniformly within the links across the network. This implies that density and congestion levels are more likely to be overestimated. FCD, on the other hand, is more reliable for average traffic states during daytime and on main roads with good coverage of probe vehicles. Although, probe vehicles may have specific behaviour that can also

introduce additional bias, e.g. taxis. To eliminate this bias, in a recent study (Knoop *et al.*, 2018), authors utilised floating car data of private vehicles from Google for the empirical estimation of MFD for the city of Amsterdam. Although, Google shares aggregate data, after applying some "noise", and only shares that information with few institutions in the world (university, institutes or transportation centre's) (García-Ramírez, 2020). On the other hand, Google codes the numerical information using four colours and gives it for free in the form of online traffic maps through its platforms (web and mobile app) and this colour-coded traffic (live and typical) is available in several cities worldwide. Thus, it is crucial to investigate, the feasibility of using this free online traffic information broadcasted through maps for capturing network traffic phenomena (e.g. MFD estimation) for several cities and from different platforms.

Compared to several studies that utilize crowdsourcing data, in this thesis, the acquisition of the vital traffic information is achieved by performing a back-engineering approach based on image processing and using fundamental microscopic traffic relationships to estimate macroscopic phenomena. The main difference is that the free and condensed information form online traffic maps is utilized instead of the raw/numerical data. In the light of other data sources limitation (access, cost, bias, coverage) to provide a network-level perspective, the capturing of macroscopic phenomena utilizing online traffic maps' information worldwide has great value especially in the era of Big Data, both for traffic monitoring and control.

In the next section, a thorough review in human mobility research follows, presenting mobility mapping eras and milestones and recent remarkable studies and findings.

2.2 Human Mobility Modeling and Mapping

2.2.1 Terms and motivation

The term "mobility" has multiple connotations, as it can refer to the movement of human, goods, and ideas depending on the approach of each academic community. Regarding the human mobility term, it can be said that it refers to the movement of human beings (individuals as well as groups) in space and time. In space dimension, human mobility can occur over a large variety of distances, within the city, among cities, within a country or among continents while in terms of time it can be short-term or long-term depending on the mode of transport. Regarding the level of human mobility, it can be either individual or aggregated to flows of people leading to a variety of models for reproducing and forecasting.

Human mobility can be further studied based on the mode of transfer (e.g. sea mobility, air mobility) or within the whole city (urban mobility).

Considering all these aspects of human mobility, the depiction of mobility patterns becomes a challenging task, as huge spatiotemporal information at different scales should be conveyed in a human friendly and understandable way. Arguably, the science of cartography was the first to explore mobility data and gave answers to human questions through maps in an efficient way. Despite the difficulties in producing these mobility maps due to the time-consuming data collection methods and the lack of computation means, pioneer scientist produced the first flow maps during the early 1800s. Since then, the emerging technologies and the data availability enhanced the spectrum of possibilities in mobility mapping.

In the years to come, where future mobility may be completely different compared to the traditional scheme of car ownership, new challenges will arise both in understanding, modeling and visualizing urban mobility dynamics. Thus, the motive of the current review is to give a short resume of the status of our knowledge and understanding in human mobility, and the way it is visualized.

2.2.2 Mobility mapping eras and milestones

In this sub-section, the mobility mapping eras unfold as a timeline that includes selected milestones, which paved the way for a better understanding of human movements. As a starting point, the twenty-year period 1835-1855 is chosen, where almost every know technique for representing population numbers, distribution, density and movements came into being. The earliest visualization of traffic flows was constructed in 1837 by Henry Harness for the British Army prior to the decision to build a railway in Ireland. The Harness' Passenger Conveyance Map (Figure 2.1) showed the relative number of passengers in different direction throughout Ireland by regular public conveyances. Interestingly, for a map prepared a long time ago, the flow lines were drawn between connecting points, without showing the actual route, as contemporary origin-destination lines, while the varying widths are properly proportional to the numbers that appear along them (Robinson, 1955). Another early pioneer of data-centric mapping is Charles Joseph Minard, best known for his acclaimed depiction of Napoleon's ill-fated invasion of Russia (Figure 2.1) that incorporates six types of data (the number of Napoleon's troops; distance; temperature; the latitude and longitude; direction of travel; and location relative to specific dates) (Claudel, Nagel and Ratti, 2016). Minard created several flow maps, from passenger traffic on European railroads

to the overseas movement of French wines and paved the way in thematic cartography in the sense that he had almost no pattern to follow. The width of flow lines on a Minard map was strictly proportional to the magnitude it represents while the scales of the geographical features on his maps were forced to fit the data being portrayed (Robinson, 1967). Scales and shapes for maps were also transformed for a variety of purposes. For instance, in the anamorphic map by Émile Cheysson, (Figure 2.1) deformations of spatial size were used to show a quantitative variable (e.g. the decrease in time to travel from Paris to various places in France), (Friendly, 1975). Harry Beck's original Tube map designed in 1933 also didn't give emphasis on geographical accuracy. Although, his clear and comprehensible chart became an essential guide and a template for transport maps all over the world (Beck, 2016).

The first attempt not only to visualize but more to explain and predict mobility patterns in terms of migration is dated at 1885, known as the Laws of Migration by Ernst Georg Ravenstein. Ravenstein considered not only the effect of distance but also the gender and the age of people as explanatory variables based on data from birth tables published in the British Censuses (Grigg, 1977). Ravenstein's 1885 paper (Ravenstein, 1885) also includes a map named "Currents of Migration", which illustrates migration patterns around the British Isles, shown in Figure 2.1. Although today, we are all familiar with such representation of mobility (radial or network flow map), it is still surprising how these pioneer geographers produced these maps considering the scarcity of data and the lack of computational means.

Moving towards the twentieth century one encounters breakthrough developments that shaped the map history. In that era, we can trace the origins of web mapping and advances including the global positioning system (GIS) and the use of satellite imagery in creating maps with greater accuracy and richness of detail. The advent of online mapping enabled users to retrieve and share mapping data while Web.2.0 interactive mapping allowed users' active contribution to the maps and the knowledge it was shared (Veenendaal, 2016). The user interaction was further enhanced by two waves of web map design revolution. The first regards the virtual globes development begging with the release of Goggle Earth by Google in 2005 while the second was the mobile mapping on smart phones that incorporates Global Positioning System (GPS) location tracking. (Tsou, 2011).

The continuous blend of improvements in technologies, growth of data and users' interaction enabled contemporary visualizations of flow maps in a faster and more accurate way. Examples of these representations are illustrated in Figure 2.2. Starting from the left,

the first image is an outgoing migration map from Colorado based on the US Census data for 1995-2000 records from all the county to county migrations. Here, the authors' goal is to produce flow maps to visualize networks and other kinds of flow data, by using a flow map layout that allows a user to see the differences in magnitude among the flows with a minimum of clutter (Phan *et al.*, 2005). The next image shows running routes around London based on data from the application "RunKeeper"(Yau, 2007). These mobile fitness apps and online fitness networks use the built-in GPS capabilities of smartphones to track distance and speed of activities of the smartphone owner such as running sessions, bike rides and walks (Stragier and Mechant, 2013). Although the data from such source does not represent all the runners in the city, useful insights can be gained by exploring their mapping. Mobile data were also used to show the distribution of users around Rome's Termini station (Ratti *et al.*, 2006), in the third image. In the fourth image, a mesmerizing dynamic map of global refuges flows for year 2014 is shown. The red dot is equal to seventeen refugees arriving in a country, while yellow dots represent refugees leaving their home country behind. The data source, behind this map is the United Nations High Commissioner for Refugees (UNHCR)(Weller, 2017). The last map in Figure 2.2, is inspired from interactive wind maps and shows movement in Manhattan based on tweets. Here, tweets sent by the same person within a 4-hour time-window were used as samples of speed and direction. The average flow of people within the area and total tweet density over the space were used to simulate the movement of people (Clark, 2012). Other contemporary visualizations of mobility can be found in (Andrienko *et al.*, 2017) that provides a survey of literature from the visual analytics domain with respect to the different types of transportation data, movement and its relationship to infrastructure and behaviour.

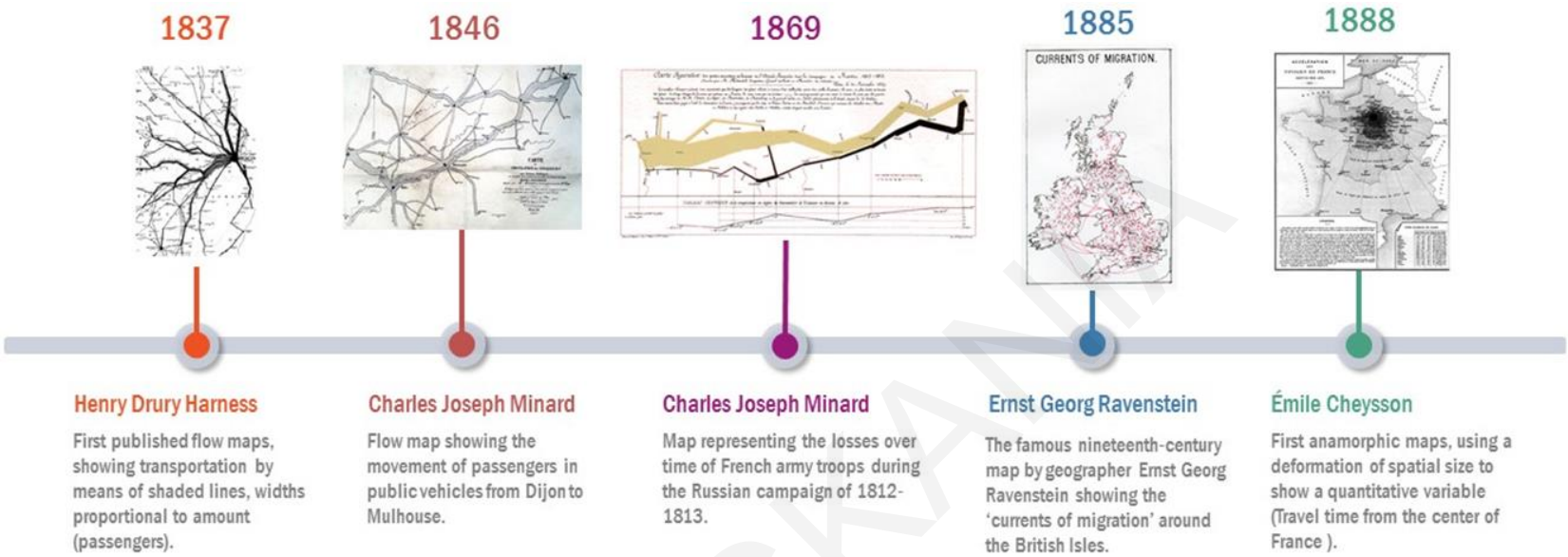


Figure 2.1:Milestones in human mobility mapping during the 19th century.

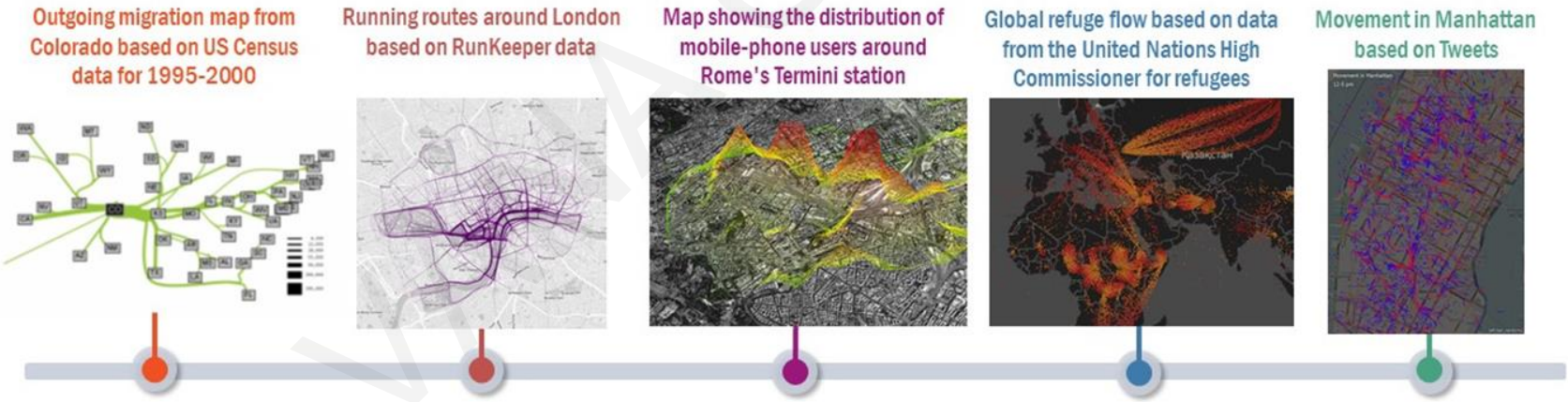


Figure 2.2:Contemporary visualizations of human mobility based on different source of data.

2.2.3 Categorizing human mobility studies based on data source

The historical availability of empirical data has been vital to both aggregate and individual mobility research. The main sources of data are outlined in this section and used as a key parameter to cluster the relevant research endeavors. Following a chronological order, the first cluster encapsulates mobility studies based on census data, the yesterday's big data. The second data source and perhaps the most used in human mobility investigation is the mobile phone Call Detail Records (CDRs). The GPS and smart card data follow while the section ends with studies based on online data from Online Social Network (OSN) or Location-Based Social Network (LBSN) services. Although, the empirical data for mobility studies are not limited to these four categories, the frequency of data used was the main criterion for selection. Another criterion that was applied is the existence of visualization of human mobility through maps or infographics.

In the following sub-sections each data source is described while four summarized tables with the selected studies are given at the end (Table 2.1,Table 2.2, Table 2.3 and Table 2.4). Each column of the table refers to certain characteristics of the study. The first column groups the studies according to the data source used. In many cases, multiple sources were used and thus one study may belong to more than one category. The second column provides the authors names and the year of the study. In the third column each study is characterized by the mobility level that can be either population or individual level according to the analysis. In the same column the movement analysis task is provided borrowed from the taxonomy techniques described in (Andrienko *et al.*, 2011), where according to the fundamental constituents of movement (space, time, object) three different foci are possible. The first type refers to studies that focus on object (movers, events, trajectories), characteristics of objects in terms of space and time and relation to location, times and other objects. The second type focus on space and the third foci focus on time respectively. The fourth column outlines the data and the study area while the fifth describes the visualization method used to illustrate human mobility. The last column highlights the main findings or applications of each study in terms of human mobility.

Census Data & Surveys

Census data derive from periodical national surveys in which householders are questioned regarding the socio-demographic and economic status of the household members. Commuting flows or internal migration flows within a country can be estimated through

questions in census related to the place of birth, the place of previous residence, and the location of workplace. Usually, national censuses are held in most countries typically every ten years, as they have been singled out as the most elaborate, complex and costly data collection activity. Although, population census is in fact one of the earliest systematic catalogue of data produced on a continuous basis and it is broadly used in mobility studies from the nineteenth century till now. One the other hand, travel surveys investigate the individual travel behaviour and thus travel trajectories data can be extracted such as origin and destination and departure and arrive time. Both type of sources was primarily used in migration and human mobility research.

As previously mentioned, E.G. Ravenstein in his spermatic work the Laws of Regression, used birth tables published in the British Censuses of 1871 and 1881 in order to explain the migration patterns within and between countries (Ravenstein, 1885). Considering the effect of distance from place of birth to place of enumeration, as well as the gender and the age of a migrant as primary factors, he posited the seven laws of migration that paved the way for subsequent research. A map titled "Currents of Migration" accompanied his paper but was not mentioned in the text. In this map, arrows were used to illustrate migration from country to country while some cities were named and shown as circles. The map conveyed the migration patterns in a simple and clear way. A similar visualization technique is used in (Bell and Ward, 1998) where flow maps were used to illustrate the temporary population movements in both coastal and inland locations. Here, the size of the arrows is proportional to the number of movers while movements are explored by comparing the data on place of usual residence with place of enumeration at the time of the Census. In a more recent study (Verhetsel and Vanelslander, 2010), the authors took full advantage of the last census in Belgium to get a geographically complete image of commuting in Flanders and Brussels. Data from nearly 1.2 million Flemish commuters that were available from the processing of the 2001 census were aggregated to the level of neighbourhood. This data entailed information regarding: the neighbourhood of the workplace and the place of departure, the distance and frequency of the movements, the means of transportation generally used for commuting, the times of departure and arrival for the outward journey and the return journey and last the number of cars that the household had at its disposal. The resulting clusters based on characteristics of commuters are depicted on a map that gives a geographically complete image of commuting in Flanders and Brussels.

In contemporary research, census and survey data are mainly used for evaluation and comparison. For instance, in (Guo and Zhu, 2014), authors used data from the US census

2000 to evaluate a universal model for mobility and migration patterns, the so-called radiation model. A mapping of mobility fluxes is included based on these data while the study focuses on movers in terms of movement analysis. In another study (Csáji *et al.*, 2013) home and office locations were identified by using communication data and the results compared with official census data. In that case the movement analysis focuses on space while a map showing commuters patterns between home and office positions is provided. Respectively, in (Steiger *et al.*, 2015) tweet patterns were correlated with official census data for the case study of London, showing an overall strong positive correlation in comparison with workplace population census data. Here a weighted graph was used to visualize the frequency between home and workplaces. Likewise taxi data was combined with U.S migration data set from the 2000 census in (Guo and Zhu, 2014) to demonstrate a new method for origin-destination flow density estimation and for flow map generalization. The study includes visually legible flow maps that faithfully represent the major flow patterns. Last, traffic census data were also used for validation in (Kashiyama, Pang and Sekimoto, 2017) that proposes a novel dataset creation approach (called Open PFLOW) that continuously reports the spatiotemporal positions of all individual's in urban areas based on open data. Noteworthy animated maps, representing time series population distributions are included in the study.

Overall, both census data and traffic surveys are a valuable source for further research especially in case of mobility and migration patterns. The merit of these source lies on the scale that is equal to population rather than a sample and thus a complete picture of mobility can be achieved. Although, the cost and time limitations led researchers and practitioners to seek new data sources, as the following.

Mobile Phones Data

The advent of mobile telecommunications and the huge amount of data derived both from network operator and smartphone sensors played a critical role in inferring human mobility, especially at the individual level. As mobile phones are (typically) personal devices and are mostly carried by a single person, one's trajectories can be matched to the company's cell tower location that routing the communication. Thus, an immerse dataset is produced that includes individual human movement in an urban area, valuable for microscopic research. Although, CDRs provide an unpresented spatio-temporal resolution compared to traditional census and surveys, it is still a sample of population (mobile phone holders) and usually restricted due to privacy concerns. An a overview of the Mobile Data Challenge can be found

in (Laurila *et al.*, 2013) where a thorough review is provided for the Lausanne Data Collection Campaign, an initiative to collect unique longitudinal smartphone dataset.

Early endeavours to employ such data can be found in projects of the MIT Senseable City Lab (MIT Senseable City Lab, 2020), such as the Mobile Landscape Graz (2005), Real Time Rome (2006), Real Time Copenhagen (2007), A Tale of Many Cities (2014), and Friendly Cities (2018). In the first project, telecommunication data were used to produce three real-time maps of Graz that display cell phone traffic intensity, traffic migration (handovers) and traces of registered users as they move through the city (Ratti *et al.*, 2007). Accordingly, the second project presented real-time urban monitoring for the city of Rome, based on the use of anonymous real-time data collected from cellular phone and locational data from the public transportation system in the city (Calabrese *et al.*, 2011). A similar approach to understand city's dynamic pulse and get insight in individual movement traces followed for the city of Copenhagen that combined mobile phone records and GPS data from volunteers (Outram, Ratti and Biderman, 2010). The "Tale of many cities" project presented a technique and a visualization tool to gain information in human activity across four global cities that was based on aggregated activity measures of mobile networks (Kondor *et al.*, 2015). The last project investigates spatiotemporal patterns of share urban space between friends while the proposed framework is applied to a CDR dataset collected in Singapore (Xu *et al.*, 2017). All the above projects are mainly visualization projects and have introduced a variety of prototypes in urban data visualization.

As mobile phone data can provide information for human trajectories, in many studies it was applied for research on individual level. For instance in (González, Hidalgo and Barabási, 2008), the authors studied trajectories of 100,000 anonymous mobile phone users and concluded that the human trajectories show a high degree of temporal and spatial regularity. In another studied human trajectories from mobile phone records used to relate human mobility with two factors the city shape and size (Kang *et al.*, 2012). To validate the representativeness of mobile phone users, authors in (Calabrese *et al.*, 2013) compared mobility measures generated from mobile phone traces with mobility measures computed using odometer readings. The results showed that mobile-phone-based mobility measures have similar spatial distribution patterns as the odometer-reading-based mobility measures and thus mobile phone traces represent a reasonable proxy for individual mobility. A relevant study investigated the error between real human trajectory and the one obtained through mobile phone data using different interpolation methods (linear, cubic, nearest interpolations) taking into consideration mobility parameters (Hoteit *et al.*, 2014). From the

same data authors extracted real hotspot positions and compare them with the estimated positions that one can get by applying the interpolation methods. Decomposing the state of Massachusetts into census blocks, they computed the real load of each block in the region and visualized the results on a heat map. Using both mobile phone and GPS data, authors in (Pappalardo *et al.*, 2015) discovered the existence of two distinct classes of individuals: returners and explorers. Also, a visualization of the complexity of the explored mobility patterns is provide in the study.

Mobility information gathered at the individual level can be aggregated to study the flows of individuals moving from one region to another at different spatio-temporal scales. A vast body in the literature focus on human mobility on a population level. As in (Bayir, Demirbas and Eagle, 2010) where authors designed a complete framework the Mobility Profiler for discovering mobility profiles based on cell phone log data. A map on end locations and top mobility paths for one user is given in the study. The space-time structure of human mobility is explored in (Sun *et al.*, 2011). Here, the original data were collected from cellular networks in a southern city of China, recording the population distribution by dividing the city into thousands of pixels. Also, the population distribution is illustrated on map divided into pixels. Results of the study indicated four underlying rules of urban dynamics: low intrinsic dimensionality, three categories of common patterns, dominance of periodic trends, and temporal stability. A methodology to extract Origin-Destinations (OD) trips by purpose and time of day from CDR data is presented in (Alexander *et al.*, 2015). Trips reported in local and national surveys are used for validation while home-work flow maps illustrate the extracted trips (CTPP, 2018). OD travel flows are also estimated by massive cellular signalling data in (Ni, Wang and Chen, 2018) and combined with other explanatory features of urban regions in a spatial econometric model. A map that illustrates travel flows between the 51 Traffic Analysis Zones (TAZs) of the study area is also provided.

Although this review does not cover all the human mobility studies that employed mobile phone data, a representative view is given regarding its utility in movement analysis and visualization projects. Another valuable source that enables human tracing with a high degree of accuracy is data from GPS enabled devices and records of smart cards transactions in public transport that follows in the next sub-section.

GPS and Smart Card Data

GPS technology has been widely used in transportation science as it provides highly accurate trajectories data. Initially, GPS devices has been utilized in travel surveys to improve the

accuracy of travel data (Shen and Stopher, 2014). The potential use of GPS increased rapidly with the light GPS devices, cellular phones with GPS-device, and GPS-enabled PDA (Personal Digital Assistant) (Wachowicz, 2010). Many research projects emerged that used GPS data to understand individual human movement. For instance in (Giannotti *et al.*, 2011) a project on mobility data mining is presented called M-Atlas. The authors conducted a large-scale experiment, based on the detailed trajectories of tens of thousands private cars with on-board GPS receivers to unveil the complexity of human mobility. Various visualization techniques were used to depict the mobility patterns of the study area.

GPS trajectories from taxis was also widely used in literature as they also encode information about movement. A taxi trip is associated with pickup and drop-off locations, travel times. Also, it contains other attributes such as the taxi id, the distance travelled, fare and tip amount, which further enable, for example, the study of the economics of fare structure and optimal fleet size. In human mobility studies, taxi GPS data was used for exploring both individual and population level of mobility. For instance in (Jiang, Yin and Zhao, 2009), researchers illustrated that the human mobility pattern, or the Lévy flight is mainly attributed to the underlying street network by using 72000 peoples' moving trajectories, obtained from 50 taxicabs during a six-month period. A heat map visualizing the density of the Origins and Destinations (O/D), as a possible mechanism behind the power law behaviour of trail length is also provided in the study.

For a better personalization recommendation in geographic information service, the authors in (Yuan *et al.*, 2014) proposed a method to mine interesting locations and the frequent travel sequences in a given geo-spatial region by taking the users' historic travel experiences into account as well as the correlation between locations. Two real GPS trajectories dataset were used in their study, namely CD consisted of a series of regular car routes generated from July to September 2012 in Chengdu city of China and the BJ dataset that contained 17,621 trajectories with a total distance of about 1.2 million km and a total duration of 48,000 hours.

Regarding mobility on the population-level, taxi trips from New York City were utilized in many research endeavours. As an example, (Ferreira *et al.*, 2013) presents a new system that supports visual exploration of big origin-destination and spatio-temporal data by using a large data set consisting of over 520 million taxi trips in NYC. Taxi data from 29 million district trip records was used for validation of their proposed methodology for the prediction of spatial-temporal activities as human mobility in (Guo and Karimi, 2017).

Another study (Dimitriou *et al.*, 2016) presents dynamic demand patterns in New York City along with the identifications of taxi operators characteristics that optimize fleet performance. The study provides also a mapping of pickup and drop-off locations per taxi vendor. A similar mapping of clusters of pickup and drop-off locations collected from more than 1100 drivers is illustrated in (Tang *et al.*, 2015) but for the city of Harbin, China. With a view of spatial interactions represented by taxi trips, another study (Liu *et al.*, 2015) revealed a two-level, hierarchical, polycentric structure of Shanghai based on GPS trajectories of more than 6600 taxis. A mapping of the estimated travel patterns was also included in the study.

In some cases taxi trips were combined with other source of data to explore human mobility, as in (Jiang *et al.*, 2017) that innovatively merged the displacement data of taxi GPS trajectories and smart card transaction from subway and bus from Beijing, China. Another study (Ma *et al.*, 2017) combined transit smart card data and survey of travel behaviour to identify transit commuters as well as to extract individual-level residence and workplace. The study also provided visualization of the commuting spatial patterns of transit commuters in a map-based platform.

Smart card data usually derived from public transport systems was widely used on human mobility studies. Research endeavours that exploit smart card data mainly focus on time or space in terms of movement analysis. As an example in (Sun and Axhausen, 2016) a probabilistic factorization framework on multi-way transit trip records was applied to reveal the spatial-temporal patterns of urban mobility. The transit trip records were derived from a large-scale public transport smart card transactions dataset (14 million transit journeys) collected in Singapore, which included both bus and metro modes. In terms of mapping, the study provides illustrations of spatial interaction (O/D) for different temporal patterns.

In (Zhong *et al.*, 2016) a comparative study for three megacities revealed that variability of temporal patterns increases with increased temporal resolution following a negative exponential function rather than a random distribution based on one-week smart card data. A relevant study based on Singapore Public Transaction data revealed that human daily movements with public transportation can be mostly described with twelve frequent movement rhythms. The study also offers an interactive visual interface that presented frequent human movement rhythms.

Although both GPS trajectories data and smart card data have been widely used in human mobility studies due to their high accuracy, they lack in coverage as they represent a biased sample of a city's population. Particularly, compared to mobile phone data which can provide information on the mobility of millions of users GPS and smart card datasets typically feature a smaller number of individual users. Thus, the last decade researchers turned to data from Online Social Network (OSN) or Location-Based Social Network (LBSN) services that attract hundreds of millions of users worldwide. Selected studies that have used such data along with their visualization methods follows on the next sub-section.

Online Social Network data

The advent of the Online Social Network (OSN) and their Location-Based Social Network (LBSN) services provided a unique opportunity to study the social and temporal characteristics of how people use these services and to model patterns of human mobility. Indeed, services such as Twitter, Facebook, Foursquare and Flickr collect geotagged data every time a user enables localization for the content being posted and thus fine granularity data about human movements is becoming available. Moreover, the worldwide adoption of these tools implies that the scale of the datasets is global, and this enable to focus on human mobility patterns in many cities across the world. On the other hand, the scientist had to face the challenge of these huge datasets in terms of mapping and visualization. Selected studies that have utilized this source of data follow along with their visualization methods used for human mobility mapping.

To begin with studies that utilize global datasets, in (Cheng *et al.*, 2011) authors used 22 million check-ins across 220,000 users from Foursquare, Twitter, Gowalla, Echofon, and Gravity to assess global human mobility patterns. The study also provides maps representing these check-ins across the globe, the United States and the New York City along with a venue cloud of check-ins. Another research endeavour that utilize global datasets can be found in (Hawelka *et al.*, 2014). In this case, almost one billion tweets were used to uncover global mobility patterns and compare mobility characteristics of different nations. The results showed that increased mobility (measured in terms of the probability of travel, diversity of destinations, and geographical spread of travels) is characteristic of West European and other developed countries. The study provides multiple infographics while a visualization tool called Circos is used to depict the Top 30 country-to-country estimated flows of visitors. In another large-scale experiment (Noulas *et al.*, 2012), authors studied the movements of Foursquare users across 5 million places in 34 metropolitan cities that span

four continents and eleven countries to perform an empirical validation of past theories on the driving factors of human movements. Thermal maps of the density of places within cities was used as a visualization method in the same study.

Twitter data dominates compared to data from other social media services. This social media tool enables users to post short messages up to 140 characters, called "Tweets" in Twitter and when permissions are given by the users, each of their tweets are attached with a corresponding geo-location. Furthermore, Twitter allows its users to post statuses from third-party "check-in" services (e.g. Foursquare). Thus, this combined information makes Twitter a promising proxy for understanding the social dynamics in geographic spaces. In a relevant study (Hasan, Zhan and Ukkusuri, 2013), urban human mobility and activity patterns are analysed using location-based data collected from Foursquare and Twitter. The authors first characterize aggregate activity patterns by finding the distributions of different activity categories over a city geography and then determine the purpose-specific activity distribution maps. In (Luo *et al.*, 2016) urban mobility for the city of Chicago was investigated at the individual level, utilizing 300 million records from Twitter. Similar as previous findings based on other sources of data (e.g., cell phone calling logs) the human mobility measures generally follow the power law distribution. The study further grouped Twitters users based on three demographic factors (race/ethnicity, gender and age) by analysing the names provided in their profiles. By comparing the human mobility across different demographic groups, they found that the human mobility measures of each group still generally follow the power law distribution but demonstrating obvious differences across demographic groups. The results were illustrated on maps showing the spatial distribution of activity centers in the Chicago area for different demographic groups. A similar study demonstrates (Jurdak *et al.*, 2015) that Twitter is a suitable proxy for studying human mobility and further identifies two types of Twitter users in terms of the predictability of tweeting locations, a group that is highly persistent and predicable and another group that is much more diverse and less predictable. The study utilizes a Twitter dataset with more than six million geotagged tweets posted in Australia and maps differences in tweet spatial distributions as the radius of gyration varies.

In (Manca *et al.*, 2017) a workflow to mine urban mobility patterns from social media is presented that uses geolocated tweets from the city of Barcelona as a case-study. The authors make use of maps to plot users' paths and to highlight the different mobility behaviour between tourists and locals. The study also focuses on the research challenges that

still exist on social media data in terms of low frequency of data sharing, data sampling, data collection from third parties, privacy issues, and Big data issues.

To confront the distinct characteristics of location-based social media data, that pose computational challenges to spatiotemporal analysis, authors in (Cao *et al.*, 2015) presented a framework that transforms the massive, dynamic and unstructured location-based social media data into flexible geospatial datasets. The system architectures and implementation details were presented based on a public Twitter feed collection of the United States. The authors also developed an online interactive visual analytical interface where a flow mapping service is provided based on the spatiotemporal data cube for visual exploration of movement dynamics at multiple spatiotemporal scales.

Another location-based data type, not a typical example of movement data in the research literature, refers to the georeferenced photos made publicly available on the photo-sharing web site Flickr. In (Andrienko *et al.*, 2011) authors proposed a conceptual framework that describes in a systematic and comprehensive way the possible types of information that can be extracted from movement data and on this basis defines the respective types of analytical tasks. To illustrate their framework, they use a dataset from the Flickr photo-sharing web site referring to the territory of Switzerland and the period from January 1, 2005 to September 30, 2009. In terms of their general framework, the Flickr users that publish georeferenced photos are the movers. The sequence of the photo taking events of one mover makes the trajectory of this mover. The locations are the spatial positions of the events, which are originally specified as points. As a next step they describe the types of tasks for which the Flickr photos data can be used (e.g. to explore the spatio-temporal distribution of a set of photos or the variety of spatial characteristics of trajectories). Remarkably visualization methods are used to illustrate their movement analysis such as a Growth Ring Map showing the spatio-temporal distribution of the Flickr photos and flow map showing the clusters of trajectories by route similarity.

Overall, it can be said that location-based social media data provide a new set of lenses and tremendous opportunities to examine urban mobility dynamics and complex social activity. Thus, location-based data has received increasing attention in the research community compared to the conventional spatiotemporal data, as this new modality of data is dynamic, massive and provides the whereabouts of daily activities.

Table 2.1:Related Studies based on Census Data and Surveys.

Data Source	Author(s) (Year)	Level of Mobility/ Movement analysis tasks	Data /Study Area	Visualization method description	Main finding(s) or applications
Census Data and Surveys	1. Ravenstein (1985)	Population-level/focus on objects (movers, events, trajectories)	Birth tables published in the British Censuses of 1871 and 1881/ British Isles	Flow Map	The laws or principles of migration considering the age and the gender of a migrant as primary factors
	2. Bell and Ward (1998)	Population-level/focus on time	Data from 1991 Australian Census/Australia	Flow Map	Temporary mobility is selective of certain groups and there are substantial concentrations of temporary movers in both coastal and inland locations
	3. Verhetsel A. and Vanelslander T. (2010)	Neighborhood-level/ focus on objects (movers, events, trajectories)	Individual census data from nearly all 1,2 million Flemish commuters /Brussels and Flanders, Belgium	Map of clusters based on characteristics of commuters	People working and/or living in areas near railway stations, public transport junctions, urban areas and areas with a high economic density travel less distance, especially by car, and make more use of public transport and slow modes
	4. Simini et al. (2012)	Population-level/focus on objects (movers, events, trajectories)	Census, mobile phones, tax documents/America, Europe	Mapping of mobility fluxes	A radiation model that predicts mobility patterns in good agreement with mobility and transport patterns

Table 2.1 (continued)

Census Data and Surveys	5. Csáji B. et al. (2013)	Population-level/focus on space	Data from the Portuguese National Institute of Statistics (INE) and a sample of 100,000 mobile phone customers /Portugal	Commute map for the sample of users	Identification of home and office locations and comparing the results with official census data
	6. Guo D. and Zhu X. (2014)	Population-level/ focus on objects (movers, events, trajectories)	U.S. migration data set from the 2000 Census and a taxi data set for millions of taxi riders / United States	Net Migration Flow Map and Stratified Migration Maps	An innovative approach for the computational analysis and flow mapping of large spatial mobility data
	7. Steiger E. et al. (2015)	Population-level/ focus on objects (movers, events, trajectories)	UK census data and data from 476,071 individual users of twitter /London	Frequencies of work and home topics as a weighted graph network	Investigated the correlation between tweet activity clusters and census population densities
	8. Kashiyama T. et al. (2017)	Population-level/ focus on objects (movers, events, trajectories)	Open data and census data/Tokyo metropolitan area and Chukyo area, Japan	Map of human flow in time series trips	Creation and evaluation of an open dataset for typical people mass movement in urban areas

Table 2.2: Related Studies based on Mobile Phone Data

Data Source	Author(s) (Year)	Level of Mobility/ Movement analysis tasks	Data /Study Area	Visualization method description	Main finding(s) or applications
Mobile Phone Data	1. Ratti C. et al. (2007)	Population-level/ focus on objects (movers, events, trajectories)	Cellphone data and user movement sample data/ Graz, Austria	Mapping of cellphone traffic intensity, traffic migration and traces of registered users as they move through the city	Three different real-time maps of Graz produced: cellphone traffic intensity, traffic migration (handovers) and traces of registered users as they move through the city
	2. González M. et al. (2008)	Individual-level/ focus on objects (movers, events, trajectories)	Trajectory of 100,000 anonymized mobile phone users	Figure of basic human mobility patterns	Human trajectories show a high degree of temporal and spatial regularity. Human movement is characterized by a time-independent characteristic travel distance and a significant probability to return to a few highly frequented locations
	3. Bayir M. et al. (2010)	Population-level/focus on time and space	Cell phone log data/Boston	Mapping of time spend on end-locations and top mobility paths for one user	Design of a complete framework, the Mobility Profiler, for discovering mobility profiles from raw cell tower connection data
	4. Calabrese F. et al. (2011)	Population-level/ focus on objects (movers, events, trajectories)	Data from mobile cellular networks and GPS data from buses and taxis/Rome	Various visualizations of the data on maps	A visualization tool for urban monitoring for the city of Rome

Table 2.2 (continued)

Mobile Phone Data	5.	Sun J. et al. (2011)	Population-level/ focus on space and time	Data from cellular networks/ Shenzhen, China	Map of the area divided in pixels and heat maps for mobility visualization	Resulting into four underlying rules of urban dynamics: low intrinsic dimensionality, three categories of common patterns, dominance of periodic trends, and temporal stability
	6.	Kang C. et al. (2012)	Individual-level/focus on space	Basic geographical dataset and the MPRs (mobile phone records) dataset /Eight cities in Northeast China	Plot of the radius of gyration (ROG) of a mobile user's trajectory	Results showed that the city size and shape are two important factors affecting intra-urban human mobility
	7.	Calabrese F. et al. (2013)	Individual-level/ focus on objects (movers, events, trajectories)	Mobile phone trace data and vehicle safety inspection data /Boston Metropolitan Area	Spatial distribution of daily mobility at block group level on a map	Results showed that mobile phone traces represent a reasonable proxy for individual mobility

Table 2.3: Related Studies based on GPS and Smart Card Data

Data Source	Author(s) (Year)	Level of Mobility/ Movement analysis tasks	Data /Study Area	Visualization method description	Main finding(s) or applications
GPS Data and Smart Card Data	1. Jiang B. et al. (2008)	Individual-level /focus on objects (movers, events, trajectories)	72000 people's moving trajectories, obtained from 50 taxicabs/ Gävle, Sandviken, Storvik, and Hofors, Sweden	Heat map visualizing the density of the origins and destinations (O/D)	The human mobility pattern, or the Lévy flight behavior, is mainly attributed to the underlying street network
	2. Giannotti F. et al. (2011)	Population-level/ focus on objects (movers, events, trajectories)	GPS data sets obtained from private vehicles and survey data (Milano Survey) collected in 2005-2006 / Milan and Pisa	Mapping of trajectories	A querying and mining language and system M-Atlas, that provides the mechanisms to master the complexity of transforming raw GPS tracks into mobility knowledge
	3. Ferreira N. et al. (2013)	Population-level/ focus on objects (movers, events, trajectories)	540 million taxi trips providedby Taxi and Limousine Commission of New York City/ New York	Visual query model for taxi trips	A system that supports visual exploration of big origin-destination and spatio-temporal data.
	4. Yuan H. et al. (2014)	Individual-level/ focus on objects (movers, events, trajectories)	Two real GPS trajectory datasets, namely CD and BJ datasets/ Chengdu and Beijing city of China	Mapping of experiment trajectories in the Google earth system	The proposed method can be used to detect a traveler's frequent paths as well as fixed territories for better personalized geographic recommendation

Table 2.3 (continued)

GPS Data and Smart Card Data	5. Tang J. et al. (2015)	Population-level/ focus on objects (movers, events, trajectories)	Taxi GPS data collected from more than 1100 drivers/Harbin, China	Mapping the clusters of Pickup and Drop-off locations	The distribution of taxi trips in occupied status include two patterns: ascending part and descending part while in non-occupied status, there only exists a monotonically pattern
	6. Liu X. et al. (2015)	Population-level/focus on space and time	GPS trajectories of more than 6600 taxis/Shanghai	Mapping of travel patterns	The study revealed a two-level, hierarchical, polycentric structure of Shanghai with a view of spatial interactions represented by taxi trips.
	7. Sun L. and Axhausen K. (2016)	Population-level/focus on space and time	14 million transit journeys extracted from smart card transactions / Singapore	Figure of Spatial interaction (origin × destination) for different temporal patterns	Insights for spatial-temporal urban dynamics through collective transit mobility
	8. Zhong C. et al. (2016)	Population-level/ focus on time	One-week of smart-card data /London, Singapore and Beijing	Mapping of regularity ranking for the three cities	In all cities, variability of temporal patterns increases with increased temporal resolution following a negative exponential function rather than a random distribution.
	9. Dimitriou L et al. (2016)	Population-level/focus on time and space	GPS dataset for 401856 taxi trips/ New York, USA	Mapping of Pickup and Drop-off locations per taxi vendor for one day	Dynamic demand patterns in Mega-cities as this is exposed in taxi operations and identification of taxi operations characteristics that optimize fleet performance

Table 2.3 (continued)

GPS Data and Smart Card Data	10. Zeng W. et al. (2017)	Population-level/focus on time	Singapore Public Transportation data and the MIT reality mining dataset / Singapore	An interactive visual interface that presents frequent human movement rhythms	Human daily movements with public transportation can be mostly described with 12 frequent movement rhythms
	11. Ma X. et al. (2017)	Population-level/focus on time and space	One-month transit smart card data and survey of travel behavior via social media / Beijing	Visualization of the commuting spatial patterns of transit commuters in a map-based platform	The proposed framework can identify transit commuters by mining spatiotemporal travel regularities over continuous long-term observation, as well as extract individual-level residence and workplace
	12. Guo Q. and Karimi H. (2017)	Population-level/ focus on objects (movers, events, trajectories)	Taxi data from 29 million distinct trip records/New York	Mapping of clustered neighborhoods with similar spatial latent features.	A new methodology for the prediction of spatial-temporal activities as human mobility
	13. Jiang S. et al. (2017)	Individual-level/ focus on objects (movers, events, trajectories)	Taxi GPS trajectories, smart card transaction data of subway and bus from Beijing, China	Figures fitting distributions on the empirical data	The study innovatively merges the displacement data from different modes together to explore human mobility

Table 2.4: Related Studies based on Online Social Network Data

Data Source	Author(s) (Year)	Level of Mobility/ Movement analysis tasks	Data /Study Area	Visualization method description	Main finding(s) or applications
Online Social Network Data	1. Cheng Z. et al. (2011)	Individual-level/ focus on objects (movers, events, trajectories)	22 million check-ins across 220,000 users from Foursquare, Twitter, Gowalla, Echofon, and Gravity/ across the globe	Mapping of Global Distribution of Check-ins and a Venue Cloud for Check-ins	Location sharing services (LSS) users follow simple reproducible patterns; Social status, in addition to geographic and economic factors, is coupled with mobility; and Content and sentiment-based analysis of posts can reveal heretofore unobserved context between people and locations
	2. Adriano G. et al. (2011)	Population-level/ focus on objects (movers, events, trajectories)	Geographically referenced photos from the Flickr photo-sharing Web site/ Switzerland	Growth Ring Map showing the spatio-temporal distribution of the Flickr photos, Flow map showing the clusters of trajectories by route similarity	A conceptual framework describing in a systematic and comprehensive way the possible types of information that can be extracted from movement data and the respective types of analytical tasks. An example of movement data is given based on images from the Flickr photo-sharing Web site

Table 2.4 (continued)

Online Social Network Data	3.	Noulas A. et al. (2012)	Population-level/ focus on objects (movers, events, trajectories)	35,289,629 movements of 925,030 users of Foursquare/ 34 cities	Thermal maps of the density of places within cities	A universal law for human mobility is identified, which isolates as a key component the rank-distance, factoring in the number of places between origin and destination, rather than pure physical distance
	4.	Hasan S. et al. (2013)	Aggregate and individual mobility/ focus on objects (movers, events, trajectories)	A large-scale check-in data from Twitter and Foursquare/ New York, Chicago and Los Angeles	Mapping of Check-in Density for Different Activity Categories	Spatio-temporal patterns of aggregate and individual mobility in a city using online social media data
	5.	Hawelka B. et al. (2014)	Population-level/focus on time and space	A billion tweets recorded in 2012/ Global	A visualization tool called Circos depicts the Top 30 country-to-country estimated flows of visitors	Increased mobility (measured in terms of the probability of travel, diversity of destinations, and geographical spread of travels) is characteristic of West European and other developed countries
	6.	Cao G. et al. (2015)	Population-level/ focus on objects (movers, events, trajectories)	Public data stream of Twitter feeds /North America	Multiple-source flow maps of the travel flows between major cities in North America	An online interactive flow mapping service based on the spatiotemporal data cube model to effectively represent the movement dynamics of groups of social media users

Table 2.4 (continued)

Online Social Network Data	7.	Jurdak R. et al. (2015)	Individual-level/ focus on objects (movers, events, trajectories	Twitter dataset with more than six million geotagged tweets/ Australia	Mapping of tweet spatial distributions for various radius of gyration	The study demonstrates that Twitter is a suitable proxy for studying human mobility and identified two types of Twitter users in terms of the predictability of tweeting locations, a group that is highly persistent and predicable and another group that is much more diverse and less predictable
	8.	Luo F. et al. (2016)	Individual-level/ focus on objects (movers, events, trajectories)	300 million records from Twitter/ Chicago	Mapping of the spatial distribution of the radius of gyration of Chicago Twitter users	Spatiotemporal characteristics of urban human mobility were explored using Twitter posts. Similar as previous findings the human mobility measures generally follow the power law distribution. Race/ethnicity has the largest and gender has the least impact on human mobility, in the three demographic factors of study
	9.	Manca M. et al. (2017)	Population-level/ focus on objects (movers, events, trajectories)	1, 225, 199 geolocated tweets/Barcelona	Maps of Barcelona displaying mobility patterns of tourists and locals	The authors explore how social media data can be used to infer knowledge about urban dynamics and mobility patterns by presenting a survey of the state of the art and a case-study based on the city of Barcelona

Overall, it can be said that flow maps dominate as a visualization technique in the previous studies, especially on the population level category. Although, beyond flow maps, a variety of visualization techniques were developed in the recent years. An overview of methods and tools for analysis of movement data can be found in (Andrienko and Andrienko, 2013). Here, authors present an illustrated structured survey of the state of the art in visual analytics concerning the analysis of movement data. Another survey describes the recent developments in visual analytics that are related to the study of movement and transportation systems (Andrienko *et al.*, 2017). Here, authors highlight the need for gaining better understanding of the new or changed problems, which is leading to new opportunities that arise due to the availability of large amounts of data that did not exist or were scarce in the past. This includes not only data that describe the movement of people, but also data referring to population mobility activities and lifestyle.

2.3 Chapter Summary

Through the literature overview, it was revealed that the availability of new source of empirical data shaped the future research directions both in the network-wide traffic studies and in the urban mobility field. While researchers have explored a wide variety of different mobility data, most of the studies focus on disaggregated data from traffic surveillance systems or mobile phones that are not available or free to all interested parties. To bridge this gap, this thesis explores for the first time the potentiality to employ aggregated information from traffic maps for understanding urban mobility dynamics. As, online traffic maps display condensed traffic information for the sake of readability, open research questions remain to be answered regarding the reduction of information that is performed in this display process. If valid traffic characteristics still possess, that means that the data retrieved from online maps can be further utilised in urban mobility modeling worldwide, especially in the scale of large and complex urban areas. The merits of online traffic maps lie on the ubiquitous and low-cost characteristics of this type of opensource data compared to other sources and thus it is crucial to explore their value for further applications beyond real-time traffic information and travel times.

PART I

Chapter 3 : Traffic Flow Fundamentals

As discussed in the scope of the thesis in Chapter 1, the validity of the simplified/coded information that dynamic traffic maps provide will be investigated in terms of traffic operational characteristics. Thus, in this chapter the fundamental characteristics of traffic flow and the relationships between these parameters will be presented while Macroscopic Fundamental Diagrams (MFD) that are employed in Chapter 5 as a validation tool will be further discussed.

3.1 Fundamental Characteristics of Traffic Flow

Traffic flow represents the traffic load on the transportation systems while the interaction between these "loadings" and the facility capacity determines the operational performance of the system. The fundamental characteristics of traffic flow are flow, speed, and density which can be observed and studied either at the microscopic or at the macroscopic level. The microscopic approach focuses on the movement of each individual vehicle and on the vehicle behaviour with respect to others. On the other hand, the macroscopic analysis refers to the study of the behaviour of groups of units. A framework for these characteristics is presented in the following Table 3.1 (May, 1990).

Table 3.1:Framework for Fundamental Characteristics of Traffic Flow

Traffic Characteristics	**Microscopic (Individual Units)**	**Macroscopic (Groups of Units)**
Flow	Time headways	Flow rates
Speed	Individual speeds	Average speeds
Density	Distance headways	Density rates

Flow rate is an important macroscopic flow characteristic and is defined as the number of vehicles passing a point in each period usually expressed as an hourly flow rate. Flow rate is related to the microscopic flow characteristic in the following manner:

$$q_{60} = \frac{3600}{\bar{h}} \quad (3.1)$$

Where q_{60} is the hourly flow rate

3600 is the number of seconds per hour

$\bar{h}$ is the average time headway (seconds per vehicle)

The measurement of traffic flow has many uses in the planning, design and operation of road facilities. In planning, flow measurements are used in the classification of streets, in origin-destination studies, and in revenue predictions while the design of new facilities and the re-design of the existing ones require traffic flow information. Traffic flow rates vary over time (monthly, daily, hourly, and within-hour) and over space (linear, network, directional, and lane use when several lanes are available). It could be said that, traffic flow along urban radial routes is analogous to water flow along a river (May, 1990).

Macroscopic speed characteristics refer to the vehicle groups passing a point or short segment during a specified of time or travelling over longer sections of roads. In the literature, the distinction has frequently been made between different ways of calculating the average speed of a set of vehicles. The first way of calculating speeds, namely taking the arithmetic mean of the observation is termed the time mean speed, because it is an average of observations taken over time:

$$\bar{u}_t = \frac{1}{N}\sum_{i=1}^{N} u_i \qquad (3.2)$$

The second term that is used in the literature is space mean speed, but there are a variety of definitions for it, not all of which are equivalent. There appear to be two main types of definition. One definition is the speed based on the average time taken to cross a given distance, or space, D:

$$\bar{u}_s = \frac{D}{\frac{1}{N}\sum_i t_i} \qquad (3.3)$$

where t_i is the time for vehicle i to cross distance D

$$t_i = \frac{D}{u_i} \qquad (3.4)$$

The second principal type of definition of space mean speed involves taking the average of the speeds of all the vehicles on a section of road at one instant of time. Under conditions of stop-and-go traffic, as along a signalized street or a badly congested freeway, it is important to distinguish between these two mean speeds. For freely flowing freeway traffic, however, there will not be any significant difference between the two ($\bar{u}_t$, $\bar{u}_s$) (Gartner N., Messer C., Rathi A., 2001).

Last, traffic density is a fundamental macroscopic characteristic of traffic flow. Traffic density is defined as the number of vehicles occupying a length of road-way. The length is usually specified as 1 mile or 1 Km and normally a single lane is considered. The easiest way to visualize traffic density is to consider an aerial photograph of a section of a road and to count the number of vehicles in a single lane having a length of 1 mile or 1 Km. Traffic densities vary from zero (the absence of vehicles) to values representing vehicles bumper to bumper which are completely stopped. This upper limit, called jam density, is normally on the order of 185 to 250 vehicles per lane-mile, depending on the length of vehicles and the distance gaps between vehicles. The relationship between traffic density and average distance headway can easily be obtained from the following equation:

$$k = \frac{5280}{\bar{d}} \tag{3.5}$$

Where k is density (vehicles per lane-mile)

$\bar{d}$ is the average distance headway (feet per vehicle)

Optimum density is defined as the density level that exists when the lane of traffic is flowing at capacity (May, 1990).

The relationship between these fundamental variables of traffic flow, (namely speed, flow, and density) is called the fundamental relation of traffic flow and can be derived by a simple concept. Let there be a road with length v km and assume all the vehicles are moving with v km/hr (Figure 3.1).

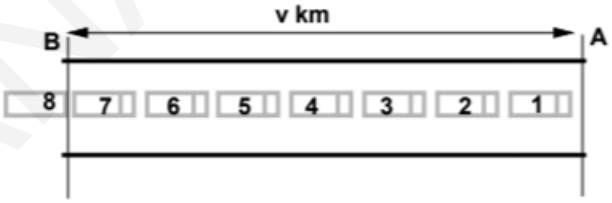


Figure 3.1:Illustration of relation between fundamental parameters of traffic flow

Let the number of vehicles counted by an observer at A for one hour be n1. The number of vehicles counted in one hour is flow(q). Therefore,

$$n_1 = q \tag{3.6}$$

Similarly, by definition, density is the number of vehicles in unit distance. Therefore, number of vehicles n2 in a road stretch of distance v1 will be density × distance. Therefore,

$$n_2 = q \times v \tag{3.7}$$

Since all the vehicles have speed v, the number of vehicles counted in 1 hour and the number of vehicles in the stretch of distance v will also be same. (i.e. n1 = n2). Therefore,

$$q = k \times u \qquad (3.8)$$

This is the fundamental equation of traffic flow. It is important to note that, speed in the above equation refers to the space mean speed (Mathew and Rao, 2006).

3.2 Macroscopic Stream Models

Macroscopic stream models represent how the behavior of one parameter of traffic flow changes with respect to another. To figure out the exact relationship between the traffic parameters, a great deal of research has been done over the past several decades. Most important among them is the relation between speed and density. The results of these researches yielded many mathematical models. The first and most simple relation between them is proposed by Greenshields, who assumed a linear speed-density relationship as shown in Figure 3.2. The equation for this relationship is shown below:

$$u = u_f - \left(\frac{u_f}{k_j}\right) k \qquad (3.9)$$

This equation is often referred as the Grieenshields' model and indicates that when density becomes zero, speed approaches free-flow speed.

The three flow characteristic presented in the same figure are defined as follow:

Flow (q) is the number of vehicles passing a specifying point or short section in each period in a single lane. Flow is expressed a s an hourly rate on a per lane basis (veh/hr/lane). One unique flow parameter is maximum flow or capacity (q_m). Speed (u) is defined as the average rate of motion and is expressed in miles per hour (mi/hr) or kilometer per hour (km/hr). From a theoretical perspective, space mean speed rather than time-mean-speed should be employed. The two unique speed parameters are free-flow speed (u_f) which exists when flow approach zero under free flow conditions while the second is optimum speed (u_o) which exists under maximum flow conditions. Density (k) is defined as the number of vehicles occupying a section of roadway in a single lane. Density is expressed on a per mile and a per lane basis, as previously mentioned. The two unique density parameters are jam density (k_j) that occurs when both flow and speed approach zero and optimum density k_o that occurs under maximum flow conditions.

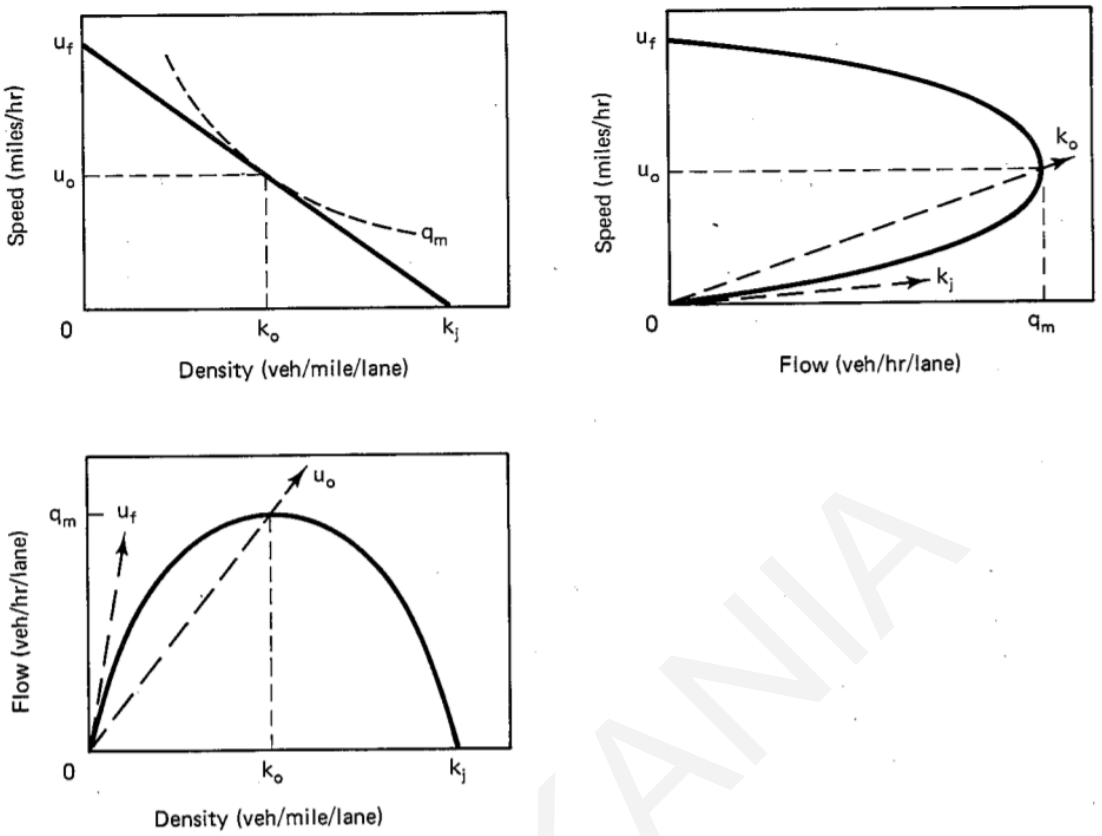


Figure 3.2: Basic Stream Flow Diagrams

Once the relation between speed and flow is established, the relation with flow can be derived. This relation between flow and density is parabolic in shape and is shown in Figure 3.2. Also, we know that

$$q = k \times u \qquad (3.10)$$

Now substituting equation (3.9) in equation (3.10) we get

$$\mathrm{q} = \mathrm{u_f}\,\mathrm{k} - \left(\frac{\mathrm{u_f}}{\mathrm{k_j}}\right)k^2 \qquad (3.11)$$

Similarly, we can find the relation between speed and flow. By substituting $k = q/u$ in equation (3.9) and solving we get:

$$\mathrm{q} = \mathrm{k_j}\,\mathrm{u} - \left(\frac{\mathrm{k_j}}{\mathrm{u_f}}\right)u^2 \qquad (3.12)$$

This relationship is again parabolic and is shown in Figure 3.2. Once the relationship between the fundamental variables of traffic flow is established, the boundary conditions can be derived. The boundary conditions that are of interest are jam density, free flow speed, and

maximum flow. To find density at maximum flow, differentiate equation (3.11) with respect to k and equate it to zero:

$$k = \frac{k_j}{2}$$

Denoting the density corresponding to maximum flow as k_0,

$$k_0 = \frac{k_j}{2} \quad (3.13)$$

Therefore, density corresponding to maximum flow is half the jam density. Once we got k_0, we can derive for maximum flow, q_{max}. Substituting equation (3.13) in equation (3.11):

$$q_{max} = \frac{u_f\, k_j}{4} \quad (3.14)$$

Thus, the maximum flow is one fourth of the product of free flow and jam density. Last, we get the speed at maximum flow, u_0, by substituting equation (3.13) in equation (3.9) and solving:

$$u_0 = \frac{u_f}{2} \quad (3.15)$$

Therefore, speed at maximum flow is half of the free speed. At that point, it is important to remember that the last three equations are all based on a linear speed-density relationship.

The three diagrams shown in Figure 3.2 are redundant, as it is obvious that if one relationship is known, the other two are uniquely defined. However, each relationship has a particular purpose and use. For instance, the speed-density relationship is used in more theoretical work as there is a single-valued speed for each single-valued density, which is not true in the other two relationships. Moving to the flow-density relationship, this one is used as the basis for freeway control systems. Under low-density conditions no control is needed because all the demand is being satisfied at a high level of service. As density is observed to increase, control is required to maintain densities below the optimum density value. On the other hand, the speed-flow relationship is used in design to identify the trade-off between level of service (speed) and the level of productivity (flow) (May, 1990).

3.2.1 Calibration of the Greenshields's model

In order to use the Greenshields's model for any traffic stream, the boundary values, especially free flow speed (u_f) and jam density (k_j) should be known. This can be obtained

by field survey and this is called calibration process. Although it is difficult to determine exact free flow speed and jam density directly from the field, approximate values can be obtained from several speed and density observations and then fitting a linear equation between them (Rao and Mathew, 2007). Let the linear equation be $y = a + bx$ such that y is density k and x denotes the speed u. Using regression method, coefficients a and b can be solved as:

$$a = \bar{y} - b\bar{x} \qquad (3.16)$$

$$b = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i}^{n}(x_i - \bar{x})^2} \qquad (3.17)$$

Where x_i and y_i are the samples, n is the number of samples, and $\bar{x}$ and $\bar{y}$ are the mean of x_i and y_i respectively.

A novel calibration approach for single-regime models was recently proposed in (Qu, Wang and Zhang, 2015). Seeing that existing single-regime models calibrated by the least square method (LSM) could not fit the empirical data consistently well (both in light-traffic/free-flow conditions and congested/jam conditions), the authors point out that the inaccuracy of single-regime models is not caused solely by their functional forms, but also by the sample selection bias. As a solution they proposed the application of a weighted least square method (WLSM) that addresses the sample selection bias problem.

At that point is crucial to mention the importance of field location to obtain meaningful speed-flow-density measurements. This issue can be explained more easily with an example. Considering the simple representation of the speed-flow curve as shown in Figure 3.3, for three distinct sections of roadway. The underlying curve is assumed to be the same at all three locations. Between locations A and B, a major entrance ramp adds considerable traffic to the road. If location B reaches capacity due to this entrance ramp volume, there will be a backup of traffic on the mainstream, resulting in stop-and-go traffic at location A. These vehicles can be in a queue, waiting their turn to be served by the bottleneck section immediately downstream of the entrance ramp. The data superimposed on graph A reflect the situation whereby traffic at A had not reached capacity before the added ramp volume caused the backup. There is a good range of uncongested data (on the top part of the curve), and congested data concentrated in one area of the lower part of the curve. The volumes for that portion reflect the capacity flow at B less the entering ramp flows.

At location B, the full range of uncongested flows is observed, right out to capacity, but the location never becomes congested, in the sense of experiencing stop-and-go traffic. It does, however, experience congestion in the sense that speeds are below those observed in the absence of the upstream congestion. Drivers arrive at the front end of the queue moving very slowly, and accelerate away from that point, increasing speed as they move through the bottleneck section. Consequently, the only data that will be observed at B are on the top portion of the curve, and at some speed in the queue discharge segment.

If the exit ramp between B and C removes a significant portion of the traffic that was observed at B, flows at C will not reach the levels they did at B. If there is no downstream situation like that between A and B, then C will not experience congested operations, and the data observable there will be as shown in Figure 3.3. None of these locations taken alone can provide the data to identify the full speed-flow curve. Location C can help to identify the uncongested portion, but cannot deal with capacity, or with congestion. Location B can provide information on the uncongested portion and on capacity (Hall, 1992).

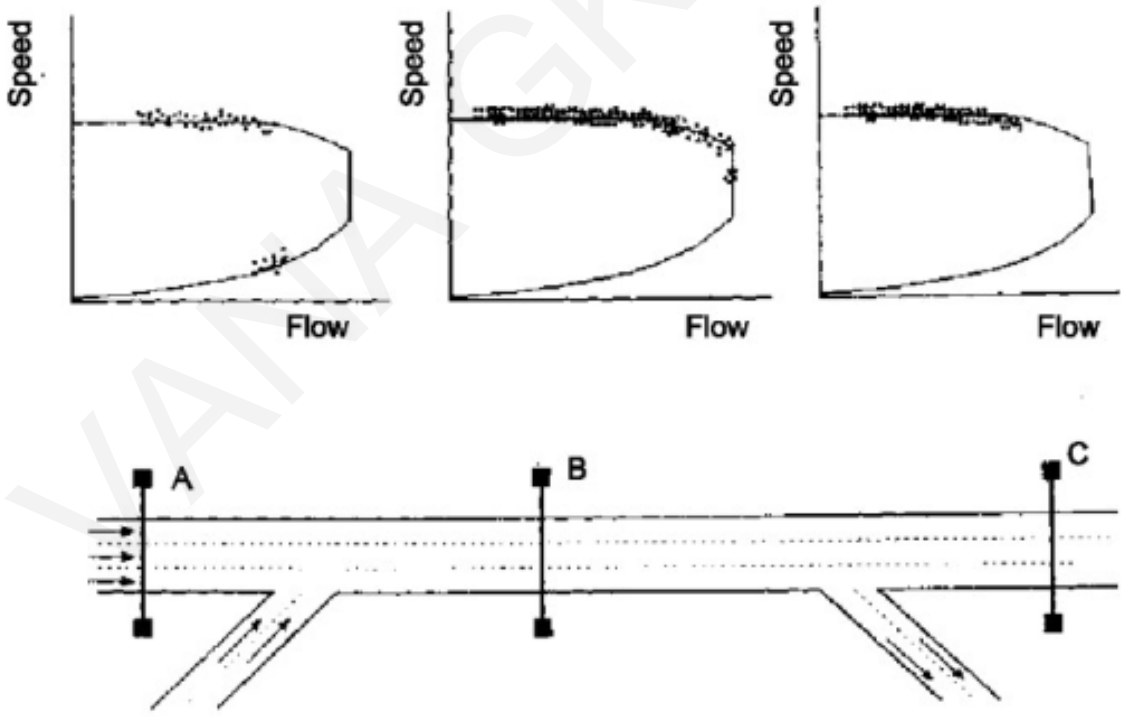


Figure 3.3: Effect of measurement location on data obtained.

From the previous example, it is possible that the apparent need for several different models, or for different parameters for the same model at different locations, or even for discontinuous models instead of continuous ones, result because of the nature (location) of

the data each was using. The second important point is that in validating traffic stream models, the data sets used may influence the results and the comparison between models.

3.2.2 Other macroscopic stream models

Over the years many traffic stream models have been proposed. These can be distinguished into two broad categories the single-regime models that assumed a single regime phenomenon over the complete range of flow conditions (free-flow and congestion) and the two regime models that separate free-flow regime and congested-flow regime.

The first single regime model was developed by Greenshields who concluded that speed was a linear function of density according to equation (3.9). The Greenberg model was the second single-regime model that was proposed. Greenberg concluded that a nonlinear model might be more appropriate and using a hydrodynamic analogy he combined the equations of motion and continuity for one-dimensional compressible flow and derived the following equation:

$$u = u_o \ln\left(\frac{k_j}{k}\right) \quad (3.18)$$

The Greenberg model (graphically expressed in Figure 3.4a) requires knowledge of the optimum speed and jam density parameters. Like the Greenshields model, jam density is difficult to observe in the field, and estimating optimum speed is even more difficult than estimating free-flow speed. A crude estimation is that the optimum speed is approximately one-half the design speed. Another drawback of this model is that free-flow speed is infinity. Thus later, Edie (Edie, 1961) recognizing this disadvantage, proposed a two-regime modeling approach with the Greenberg model being used for the congested regime.

Trying to overcome the limitation of Greenberg's model, Underwood put forward an exponential model as shown below:

$$u = u_f \, e^{-k/k_0} \quad (3.19)$$

This formulation (graphically expressed in Figure 3.4 b) requires knowledge of the free-flow speed, which is fairly easy to observe, and the optimum density, which is difficult to observe and varies depending on the road environment. Another disadvantage of this model is that speed never reaches zero and jam density is infinity. Again Edie, recognizing this drawback, proposed a two-regime modeling approach with the Underwood model being used for the free-flow regime (May, 1990).

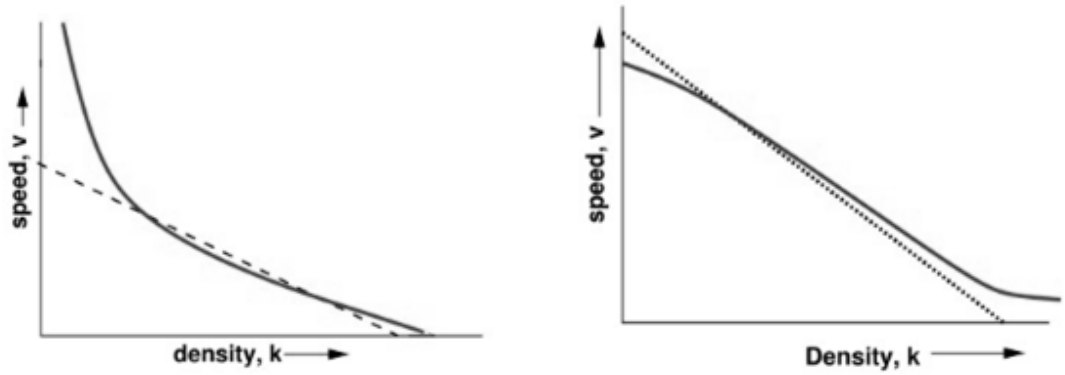


Figure 3.4: (a) Greenberg's logarithmic model; (b) Underwood's exponential model

The preceding single-regime speed-density models that dominated for many years will be also applied on data retrieved from online traffic maps while field observation from loop detectors will be used for their calibration, in order to link the microscopic traffic flow mechanics with the macroscopic phenomena, in Chapter 5.

3.3 Traffic Flow Dynamics

So far, the main macroscopic characteristics of traffic flow were investigated, mainly focusing on the static characteristics. However, there are different characteristics, which are dynamic in nature or have to do with the dynamic properties of traffic flow, such as the capacity drop and traffic hysteresis.

Starting with the capacity drop that describes the fact that, once congestion has formed, drivers are not maintaining a headway that is as close as it was before the speed breakdown. Therefore, the road capacity is lower. The capacity drop hypothesis was confirmed for the first time in 1991(Banks, 1991; Hall and Agyemang-Duah, 1991). Since then, a large amount of empirical observations of capacity drop can be found in the literature. The magnitude of capacity drop varies over a wide range depending on the local traffic conditions. The effect of the capacity drop is illustrated in Figure 3.5(Hoogendoorn and Knoop, 2012).

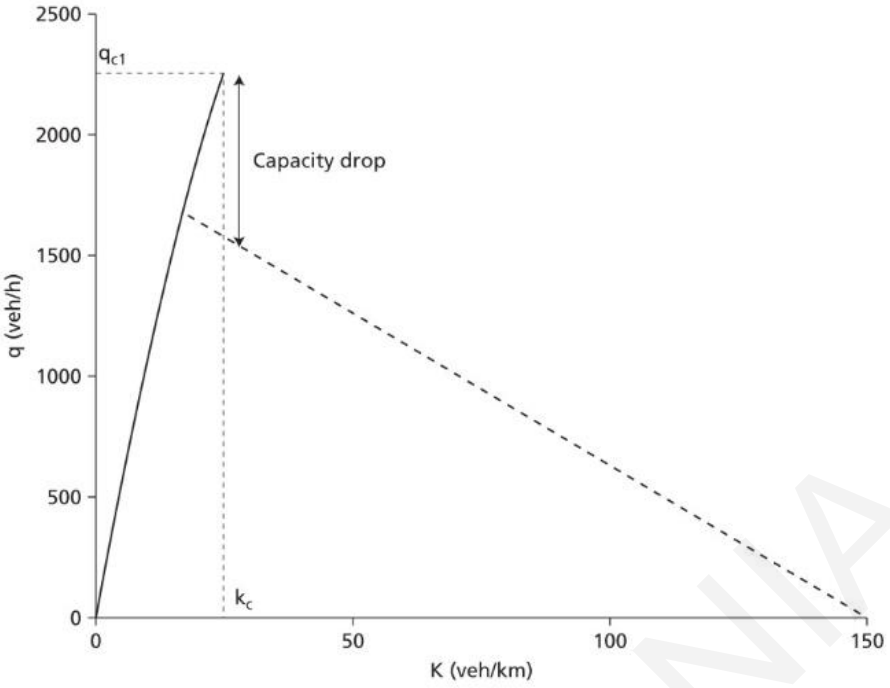


Figure 3.5: The capacity drop in the flow-density diagram.

From real-life observations of traffic flow, it can be observed that several data points collected are not on the fundamental diagram. While some of these points can be explained by stochastic fluctuations (e.g. vehicles have different sizes, drivers have different desired speeds and following distances), some can be structural, and stem from the dynamic properties of traffic flow. That is, they reflect so-called transient states, that is, changes from congestion to free flow (acceleration phase) or from free flow to congestion (deceleration phase) in traffic flow.

In other words, if we consider the average behavior of drivers (assuming stationary traffic conditions), observed mean speeds will generally not be equal to the 'equilibrium' speed. The term 'equilibrium' reflects the fact that the observed speeds in time will converge to the equilibrium speed, if the average conditions remain the same. That is, the average speed does not adapt instantaneously to the average or equilibrium speed. This introduces traffic hysteresis, which means that for the same distance headway drivers choose a different speed during acceleration from that chosen during deceleration (Hoogendoorn and Knoop, 2012).

Traffic hysteresis was first observed by Newell (Newell, 1962). He conjectured the existence of two different congested branches in the fundamental diagram as shown in Figure 3.6. When the acceleration branch stays above the deceleration branch in a flow-density

diagram, it is known as a positive hysteresis, and the opposite is termed as a negative hysteresis (Laval, 2011).

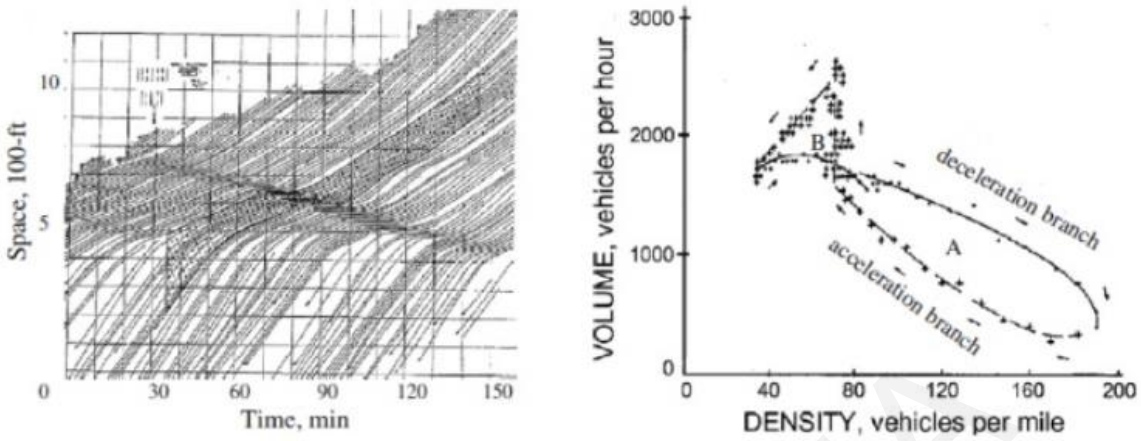


Figure 3.6: Hysteresis loop observed by Newell (Newell, 1962) from volume-density plot.

3.4 Macroscopic or Network Fundamental Diagrams (MFD-NFD)

In the preceding sections, some of the main traffic flow characteristics have been presented. Using the microscopic and macroscopic models discussed, flow operations on simple infrastructure elements can be explained and predicted. Predicting flow operations in a network is, obviously, more elaborate. A remarkable simple relation that can describe the overall dynamics of a traffic network, referred to as the macroscopic or network fundamental diagram (MFD or NFD), that was initially introduced in Chapter 2 (Section 2.1 Network-Level Traffic Relationships).

The MFD provides a relationship between network-wide averages of flow and density, both of which can be used to describe operating conditions within a network, in contrast to the Fundamental Diagrams (FD) (see Figure 3.2) that refer to specific road sections (links). Similarly, to a conventional link fundamental diagram, relating the local flow and density, three states are demonstrated on an MFD/NFD. When only a few vehicles use the network, the network is in the free flow condition and the outflow is low. With an increase in the number of vehicles, the outflow rises to the maximum. Like the critical density in a link fundamental diagram, the value of the corresponding accumulation when maximum outflow is reached is also an important parameter (Hoogendoorn and Knoop, 2012).

The network wide averages of speed, flow and density can be estimated according to the pioneer work of (Mahmassani, Williams and Herman, 1984) as following:

For a given period of observation, t average speed, $\overline{U}$, can be taken as the ratio of total vehicle-miles to total vehicle-hours over the network during period t (yielding an average speed in miles per hour or Km per hour). This average is taken both over time and over all vehicles in the network.

The average concentration, $\overline{K}$, for the same period, t is the time average of the number of vehicles per unit lane length. Letting $N(t)$ denote the number of vehicles at time t, and L the lane-miles of roadway, the average concentration $\overline{K}$, can be expressed as:

$$\overline{K} = \left({}^{1}/_{t}\right) \int_{t_0}^{t+t_0} \left[\frac{N(t)}{L}\right] dt \qquad (3.20)$$

where t_0 is the beginning of the observation period..

The definition of an average flow variable, $\bar{Q}$, is somewhat less obvious; the one pursued here considers average network flow to be the average number of vehicles per unit time that passes through an "average" point of the network. Letting q_i and l_i , respectively, denote the average flow $\bar{Q}$ over the observation period and the length of link $i, i = 1,2, \ldots . M$, where denotes the total number of links,

$$\overline{Q} = \left(\sum_{i=1}^{M} l_i \, q_i\right) / \left(\sum_{i=1}^{M} l_i\right) \qquad (3.21)$$

Authors examined also, the principal relationship between these three variables as:

$$\bar{Q} = \overline{K}\overline{U} \qquad (3.22)$$

which is fundamental for arterials but unverified at the network level. According to the simulation results, for properly defined averages of the three quantities, the fundamental relationship could be expected to hold.

A similar approach is followed in Chapter 5, to estimate network wide averages of flow and density, by extrapolating Greenshields model from links to network and utilizing the network average speed retrieved from online traffic maps.

3.5 Chapter Summary

Traffic flow theory and modeling are important in order to design comfortable and safe roads, to solve road congestion problems and to design traffic management measures. Traffic flow theory entails knowledge of the fundamental characteristics of traffic flows. In traffic flow theory a basic distinction is made between microscopic and macroscopic traffic flow variables. Microscopic traffic flow variables focus on individual drivers. Macroscopic traffic flow variables reflect the average state of the traffic flow. The fundamental diagram in traffic

flow theory describes a statistical relation between the macroscopic flow variables of flow, density and speed. The basic premise underlying the fundamental diagram is that under similar traffic conditions drivers will behave in a similar way. Traffic flow models can be used to simulate traffic, for instance to evaluate in advance the use of a new part of the infrastructure. Models can be categorized based on the representation of the traffic flow, in terms of flows (macroscopic), groups of drivers (macroscopic) or individual drivers (microscopic). Single regime speed-density models dominated for many years due to their mathematical elegance.

The overall dynamics of a traffic network can be described using a remarkably simple relation, referred to as the macroscopic or network fundamental diagram (MFD-NFD). This relation shows one of the most important properties of network traffic operations, namely that their performance decreases when the number of vehicles becomes greater. Through the literature review in Chapter 2, we saw that the empirical existence of the MFD (in small networks) was recently proved by Geroliminis and Daganzo (Geroliminis and Daganzo, 2008) , revealing the relation between the outflow and accumulation in the network.

In Chapter 5, the empirical estimation of MFD for different urban networks around the word, will be demonstrated for the first time, based on aggregated information from online traffic maps. Prior to that, in Chapter 4, the experimental setup for collecting this information from online traffic maps along with the traffic data acquisition procedure will be presented and analyzed.

Chapter 4 : Experimental Setup

This chapter is devoted to the experimental setup of the thesis. Initially the selected study area is introduced. Then, the data collection and processing are described. Last, in the data analysis section, a preliminary analysis of the extracted traffic information from maps is provided. In the following chapters, the extracted traffic information in the form of structured database is applied both for MFD estimation and mobility patterns identification.

4.1 Study Area Characteristics

To collect a representative dataset from online traffic maps, eighteen geographically and culturally different cities were chosen. Regarding the spatial and cultural distribution of the cities under investigation (Figure 4.1), five cities were chosen from Europe (Paris, London, Berlin, Moscow, Nicosia), three from North America (New York, Los Angeles, Toronto), two from Latin America (Sao Paulo, Buenos Aires), five from Asia (New Delhi, Singapore, Beijing, Riyadh, Tokyo), one from Africa (Johannesburg), one from Australia (Sydney) and last Istanbul that one part of the city lies in Europe and the other part lies in Asia.

Figure 4.1: Spatial distribution of the selected cities.

For consistency, the basic characteristics of the selected cities such as population, land, population density (Cox, 2019) and Gross Domestic Product (GDP) of the country (*IMF: World Economic Outlook (WEO), October 2019 - knoema.com*, 2019) are

summarized in Table 4.1. The population, land area, and population density data refer to the metropolitan area and thus, significantly differs from the main city's municipal boundaries, as the metropolitan area includes not only the principal built-up urban area but also the economically and infrastructurally connected suburbs and rural areas. According to the data from Table 4.1, Tokyo (including Yokohama) ranks first with estimated population over 38 million while, New York, New Delhi, Moscow and Sao Paulo follow with populations over 19 million. Los Angeles, Buenos Aires, Istanbul, Paris, Johannesburg' s population range between 10 and 16 million people while the rest of the cities' populations are under 6 million people. Last, follows the city of Nicosia in Cyprus which has a smaller number of residences compared to the rest of the cities. Nicosia was added in the dataset, at a later stage as both traffic data from online traffic maps and loop detectors were available for comparative reasons and calibration purposes.

In terms of population density (rounded to the nearest 100 per square kilometre) the cities' ranking varies. Singapore, Istanbul and New Delhi rank first with densities of over 10,000 persons per square kilometre, while Sao Paulo, Beijing, Buenos Aires, Paris, Johannesburg, Riyadh, Berlin range between 3,000 to 7,000 persons per square kilometre. The remaining cities follow with much more lower densities under of 3,000 people per square kilometre. The population density is strongly connected with the road network and the mobility patterns in a city, especially with the public transport, as none-public transport system can efficiently cope with low density.

Moving to the GDP per country, the United States is the world's largest economy, while Japan and European countries follow. Last, the Asian countries and Brazil present the lowest GDP, according to the same table.

Regardless of their apparent differences in terms of road network morphology, climate conditions, physical distances, population densities and economic output, one may expect similarities between these cities in terms of traffic and mobility profiles and thus 'universal' mobility patterns may be identified, facilitating comparative analysis across geographically and culturally diverse locations.

Table 4.1:Cities characteristics

Country	City	Population Estimation	Land (Km^2)	Population Density (per Km^2)	GDP (Country) (Billion USD)
Japan	Tokyo	38,505,000	8,223	4,700	5,154.5
India	New Delhi	28,125,000	2,240	12,600	2,935.6
United States	New York	21,045,000	11,875	1,700	21,439.5
Brazil	Sao Paulo	20,935,000	3,043	6,900	1,847.0
China	Beijing	19,430,000	4,144	4,700	14,140.2
Russia	Moscow	16,555,000	5,698	2,900	1,637.9
United States	Los Angeles	15,440,000	6,299	2,300	21,439.5
Argentina	Buenos Aires	15,130,000	3,212	4,700	445.5
Turkey	Istanbul	13,860,000	1,360	10,200	743.7
France	Paris	10,960,000	2,845	3,700	2,707.1
United Kingdom	London	10,470,000	1,378	5,600	2,743.59

Table 4.2: Cities characteristics (continued)

South Africa	Johannesburg	9,335,000	2,590	3,600	358.8
Canada	Toronto	6,630,000	2,300	2,800	1,730.9
Saudi Arabia	Riyadh	6,050,000	1,658	3,600	779.3
Singapore	Singapore	5,670,000	518	10,900	362.8
Australia	Sydney	4,515,000	2,179	2,000	1,376.3
Germany	Berlin	4,060,000	1,347	3,000	3,863.3
Cyprus	Nicosia	300,000	122	2,500	24.3

4.2 Data Collection

After the selection of the study area, traffic maps in the form of raster images for all the eighteen cities were collected every five minutes, using screen capture software, for the period of one week, the last week of May 2017. This period (end of spring semester) the computer laboratory of the Department of Civil and Enviromental Enginnering at Strovolos (Arsalidou Building) was available for data collection, as shown in Figure 4.2. Online traffic coverage for the city of Nicosia was not available at this period, so map images for Nicosia were collected at a later stage (February 2019). In total, around 50 desktop computers were required to capture map images from all the different providers (Google, Bing, Here, Baidu, Yandex) for the study area. As the available desktop computers in the laboratory were less than the required, an algorithm (Algorothm 1 Appendix A) was developed, that enables to open a specific map for each city in the browser and then capture and save the map image. To secure the update of the displayed online traffic maps, an extension of the browser that refresh pages after a set number of seconds was also enabled. The computers sprcificication used for the data collection are the following: Xeon E-2124G, 3.4GHz-4.50GHZ, 4 Cores, 32GB (2 X 16GB), 2.5'' SATA SSD, SSD SATA III 512GB, DVD SuperMulti SATA slim, Intel I219LM, 10/100/1,000 MBit/s, RJ45, NVIDIA Quadro P1000 4GB, Display Port / DVI-D adapter cable, Windows 10 Pro 64.

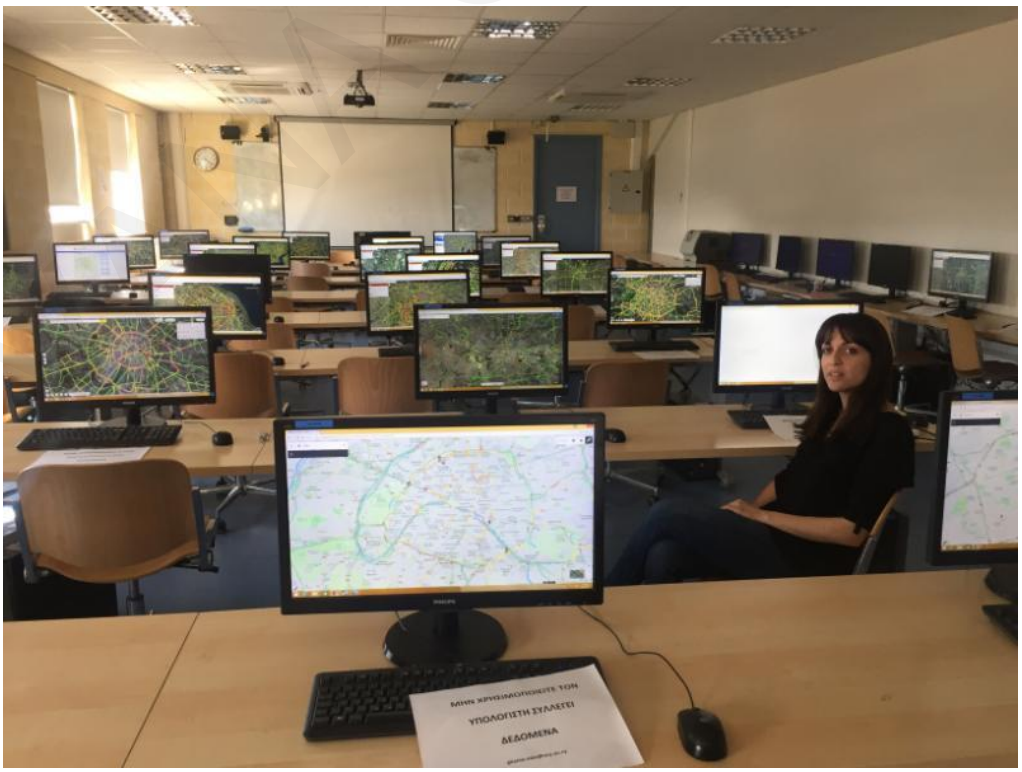


Figure 4.2: Data collection

Table 4.3: Data summary collection

No	City	Google maps			Bing Maps			Here Maps			Yandex Maps			Baidu Maps		
		(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
1	Paris	2016	1939	96.18	2016	1973	97.87	2016	1792	88.88	2016	1967	97.56	-	-	-
2	London	2016	1939	96.18	2016	1951	96.78	2016	1785	88.54	-	-	-	-	-	-
3	Istanbul	2016	1939	96.18	2016	1957	97.07	2016	1778	86.19	2016	1981	98.26	-	-	-
4	New Delhi	2016	1939	96.18	2016	1961	97.27	2016	1771	87.84	-	-	-	-	-	-
5	Moscow	2016	1932	95.83	2016	1965	97.47	2016	1764	87.50	2016	1988	98.61	-	-	-
6	Tokyo	2016	1967	97.56	-	-	-	-	-	-	-	-	-	-	-	-
7	New York	2016	1932	95.83	2016	1956	97.02	2016	1792	88.88	-	-	-	-	-	-
8	Los Angeles	2016	1939	96.18	2016	1950	96.73	2016	1841	91.31	-	-	-	-	-	-
9	Sao Paulo	2016	1925	95.48	2016	1955	96.97	2016	1778	88.19	-	-	-	-	-	-
10	Singapore	2016	1643	81.50	2016	1983	98.36	2016	1925	95.48	-	-	-	-	-	-
11	Sydney	2016	1932	95.83	2016	2002	99.30	2016	1932	95.83	-	-	-	-	-	-
12	Johannesburg	2016	1440	71.43	2016	1795	89.04	2016	1932	95.83	-	-	-	-	-	-
13	Toronto	2016	1385	68,70	2016	1974	97.91	2016	1932	95.83	-	-	-	-	-	-
14	Berlin	2016	1179	58.48	2016	1961	97.27	2016	1938	96,13	-	-	-	-	-	-
15	Buenos Aires	2016	869	43.10	2016	1906	94.54	2016	1932	95.83	-	-	-	-	-	-
16	Beijing	-	-	-	2016	1928	95.63	-	-	-	-	-	-	2016	2014	99.99
17	Riyadh	2016	882	43.35	2016	1979	98.16	2016	1933	95,88	-	-	-	-	-	-
18	Nicosia	2016	1964	97,42	-	-	-	-	-	-	-	-	-	-	-	-

(1) Pre-Cleaning Images
(2) Post Cleaning Images
(3) Data Retained %

In Table 4.3 a summary of the data collection is given. The first column refers to the selected cities of the study area. The rest columns refer to the five different providers. For comparative reasons providers that have online traffic coverage for most of the cities in the study area were selected. For three cities, named Paris, Moscow, Istanbul data from four different providers were available (Google, Bing, Here and Yandex). Google maps is the only one that provides traffic information for Tokyo and Nicosia. Traffic data were available for the city of Beijing from Baidu and Bing maps. The rest of the cities have online traffic coverage from three different providers (Google, Bing and Here maps). Under each online traffic provider, there are three separate columns labelled with numbers (1,2,3) that provide information regarding the pre-cleaning images collected, the total number of remained images after filtering and the percentage of the data retained, accordingly. In total, over 103,000 images were included in the original dataset while the filtering process, used to diminish web page disturbances, resulted in 92% valid images. For three cities from Google maps (Buenos Aires, Riyadh and Berlin) almost 50 percent of the dataset was invalid (black images) due to refresh problems during the collection procedure.

4.3 Data Processing

This section presents the steps followed for the traffic information extraction. Prior to that, some basic characteristic of the raster images/graphics are introduced.

4.3.1 Raster Images

Digital images can be stored either as a raster image or as a vector image. By default, a digital image usually refers to a raster image only. Raster image are composed of a set of digital values called picture elements or pixels. Pixels form the smallest building block of a digital image that can be processed. Each pixel stores information regarding the intensity of a given colour at any specific point in the image. The whole image is stored as a two-dimensional array, consisting of rows and columns of these pixels. Such a method of storing a colour picture in a digital computer is known as bitmap representation of an image. Bitmap images can be classified based on the colour values of these pixels as: binary, grayscale and colour. Binary images consist of only two colours (like only black and white). Grayscale images consist of shades of grey from pure white to pure black while colour images consist of multi-colour information (Bhattacharya, 2016). Colours in a raster image are represented by a colour model. Common colour models include RGB (red, green, blue) and CMYK (cyan, magenta, yellow, black). Using an RGB model, each pixel of the image has a colour

that is described by the amount of red, green and blue in it. Such an image is a composition of three matrices; representing the red, green and blue values for each pixel (Figure 4.3).

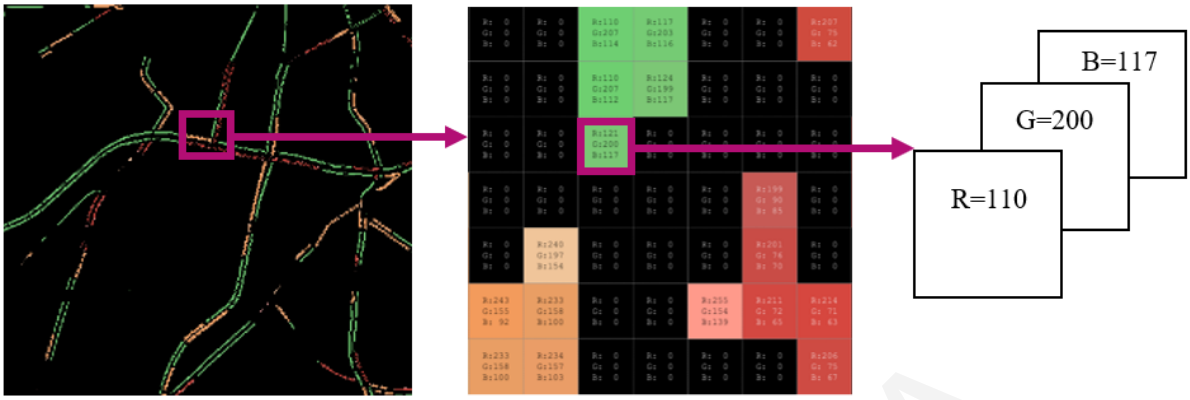


Figure 4.3: Decomposition of a raster image of size 394×394 pixels, to the pixels' intensities in a small area and the RGB intensities for the three planes of one pixel.

A raster image can be stored in a variety of formats, such as JPEG, PNG, TIFF, GIV etc. Some formats are created by a process that compresses the amount of data found in the file. This compression leaves only the data that describes the actual colour of each pixel and thus enable a faster transfer of the data. For the selected dataset the PNG (Portable Network Graphics) format was chosen as the PNG format comes close to TIFF in quality and is ideal for complex images. Unlike JPEG, PNG uses a lossless compression algorithm in order to preserve as much quality in the image. Last, PNG provides better compression because it works in a way that acknowledges the two-dimensional structure of the image, rather than treating it simply as a one-dimensional sequence of bytes (Lina J., 2009).

4.3.2 Image Processing

After the filtering of the dataset the number of valid images was equal to 97676 (total size approximately 537 GB). As the original image was captured from a screen capture software the redundant information was removed by cropping the images at the same size. An algorithm (Algorithm 2 Appendix A) was developed that crops the original image into a smaller one, as it can be seen in Figure 4.3(b).

Moving to the traffic information acquisition, the first step to extract traffic information from map images here relies on distinguishing and then capturing the traffic layers from the rest of the map. The approach followed for achieving this is based on raster image processing of online traffic maps and make use of some important image features:

i. The discretization of areas into seamless sections/digits (pixels, in raster images),

ii. Each pixel has selected dimensions (length and width) representing an area at a specific location, and

iii. Each pixel captures-in a seamless manner- the traffic characteristics of the area it represents.

There are many ways for extracting color layers from raster images (Hai and Bao, 2008; Chiang and Knoblock, 2009; Petrovska and Stevanovic, 2015). Here the most straightforward and reliable way was preferred, which is based on identifying the colors codes that are used for traffic depiction, in a color mode. In the current application the Red-Green-Blue/RGB color model is used, while four colors (green, orange, red, and dark red) used for coloring the traffic layers according to the legend of the online map. Then, the extraction of the traffic layers was performed, by selecting and extracting the pixels belonging to one of those colors. The following Algorithm summarizes the steps followed to process one map image.

```
Algorithm: Map Image Processing for Traffic Layer Extraction
```

```
Load image matrix img (r, c, 3)
Create B (r, c,3) (empty matrix) same size as image
for i=1: r
     for j=1:c
        if img(i,j,1)>R1min and img(i,j,1)>R1max or img(i,j,2)<G1min and
img(i,j,2)>G1max  or img(i,j,3)<B1min and img(i,j,3)>B1max
                img=B
            end
        elseif img(i,j,1)>R2min and img(i,j,1)>R2max or img(i,j,2)<G2min
and img(i,j,2)>G2max  or img(i,j,3)<B2min and img(i,j,3)>B2max
               img=B
           end
      elseif img(i,j,1)>R3min and img(i,j,1)>R3max or img(i,j,2)<G3min
and img(i,j,2)>G3max or img(i,j,3)<B3min and img(i,j,3)>B3max
               img=B
          end
     elseif img(i,j,1)>R4min and img(i,j,1)>R4max or img(i,j,2)<G4min
and img(i,j,2)>G4max  or img(i,j,3)<B4min and img(i,j,3)>B4max
            img=B
          end
  end
 end
```

Where r is the number of image rows, c is the number of image columns and 3 corresponds to the number of color planes for the RGB color model. In order to isolate each traffic layer that represents a certain class of traffic state, different range of values for the [R1min-R1max, B1min-B1max, G1min-G1max], [R2min-R2max, B2min-B2max, G2min-G2max], [R3min-R3max, B3min-B3max, G3min-G3max], and [R4min-R4max, B4min-B4max, G4min-G4max] were set. Consecutively, the isolated colored road network was achieved by adding the four separated images, one for each traffic layer as is shown in Figure 4.3 (c).

Figure 4.4: (a) Original image size (width x height) 1920 x 1080; (b) Cropped image size (width x height) 873 x 639; c) Final image size (width x height) 873 x 639

It should be noted that when collecting online images, depending on the PC configuration and the graphics card there could be differences (or even discrepancies) in selecting the colour layers. For selecting each colour, a range of RGB format is used (see

Algorithm 3 Appendix A), ensuring that all pixels belonging to the same colour layer are identified and selected. The total number of pixels per traffic layer and per time interval is also calculated during this step in a separate matrix.

Noted that traffic maps provide additional information (names, locations, labels, points of interest etc.) and thus the selection of the colour layers in pixels could have slight discontinuities. Since these discontinuities appear in the same places in the maps, the overall selection of the pixels can be regarded as consistent.

4.4 Data Analysis

As a result of the above data processing, a matrix containing the characteristic values of the depicted traffic variables is produced, where each non-empty entry (pixel) of the matrix can be viewed as a notional 'detector' of traffic states across the map's space, while the network properties are retained since each pixel contain the traffic information of a specific link's lane part/stretch. Subsequently, a typical map which is processed as an image, provides a discretized version of the network with many thousands of notional detectors created with a complete system coverage (at least in all locations that traffic information is broadcasted), offering a large dataset of abstracted (but dynamic) traffic information.

4.4.1 Urban coverage

Initially a connection between the number of pixels depicted in a raster image for the same zoom level, and the corresponding real linear distance on the road network, was achieved by using the following equation:

$$S = \text{Distance on the maps} / \text{Real Distance} = \text{Number of Pixels} / 1\text{km} \quad (4.1)$$

Where S represents the scale of the map while the number of pixels corresponds to the total number of pixels of the scale bar of the map. In raster images, the distance displayed on the scale bar of the maps is equal to the number of pixels between two points on the image, thus the scale of the image is actual the analogy between the number of pixels and the real distance in kilometres. By applying the above equation on the dataset, the urban coverage can be quantified by estimating the total road length in kilometres (Km) for each city. It is crucial to mention here that this coverage relates strictly to the traffic coverage of each online provider, which usually includes main and secondary roads in an urban area (see Figure 4.5, Figure 4.6). In the following Table the scale, the total road length and the relationship between pixels and real distance per city and provider is given. According to Table 4.4 and for the same city, Yandex maps provide greater urban coverage while Google maps and Bing

maps follow. Here maps provide significantly lower urban coverage, as it is also observed from Figure 4.5 and Figure 4.6.

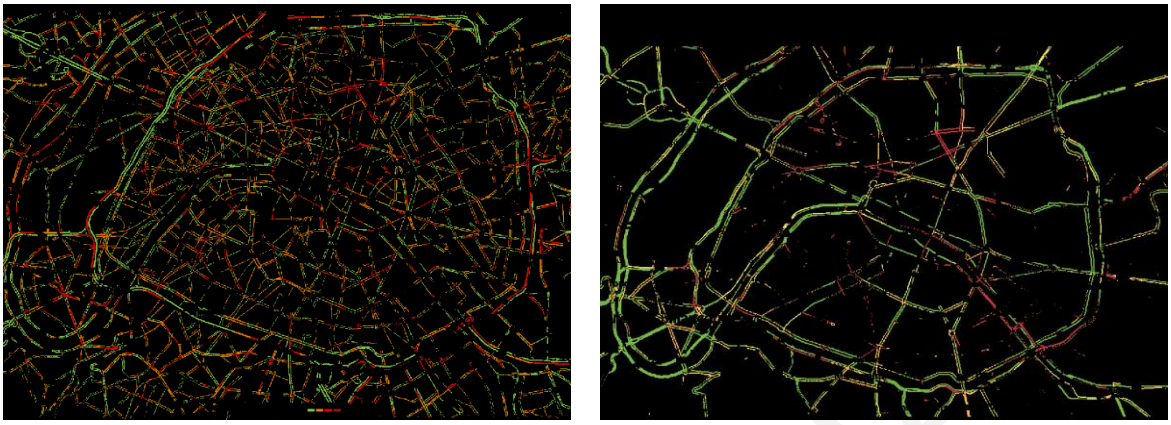


Figure 4.5:Traffic coverage for the city of Paris from Google maps (left) and Bing maps (right).

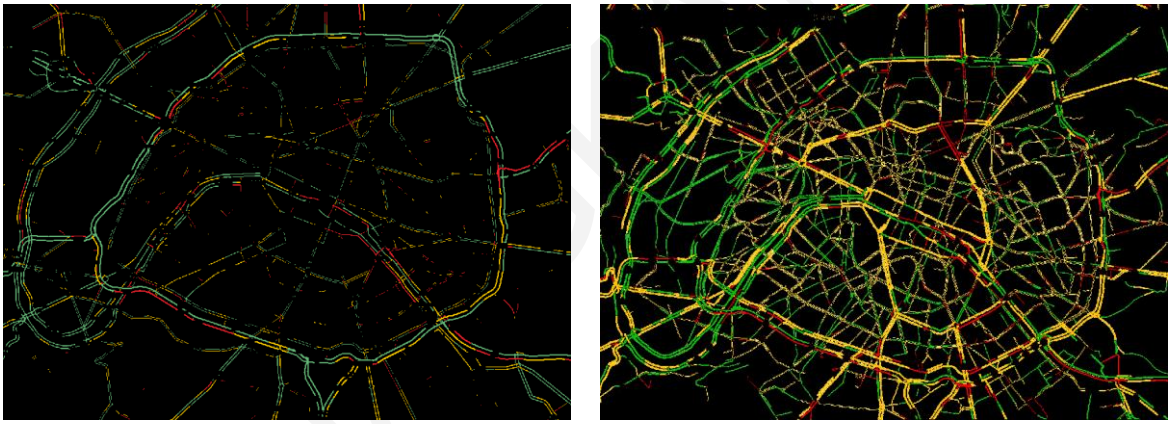


Figure 4.6: Traffic coverage for the city of Paris from Here maps (left) and Yandex maps (right). Paris

According to the same table, network traffic coverage from Google maps appeared higher for the city of Paris, Tokyo, Istanbul, Los Angeles, New York, Johannesburg, and Sao Paulo (over 500Km) compared to the rest of the cities. Notably, Nicosia differs as the business center of the city was chosen. Regarding Bing maps, greater network coverage (over 800Km) appeared for the city of Beijing, Istanbul, Johannesburg, New Delhi, Riyadh, Los Angeles, and New York. Last, Yandex maps, gave significantly higher values of estimated road length from pixels (over 1000Km) while Here maps the lowest (below 500 Km) for most of the cities. The estimated road length is indicative as the width of lines is quite different for each provider.

Table 4.4: Urban coverage per city and provider.

No	City	Google maps			Bing Maps			Here Maps			Yandex Maps			Baidu Maps		
		(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
1	Paris	0.011	707	1km=88p	0.012	588	1km=81p	0.012	322	1km=83p	0.04	1,193	1km= 83p	-	-	-
2	London	0.011	446	1km=91p	0.012	494	1km=85p	0.011	103	1km=88p	-	-	-	-	-	-
3	Istanbul	0.013	685	1km=76p	0.014	1353	1km=71p	0.014	257	1km=73p	0.04	1,819	1km=70p	-	-	-
4	New Delhi	0.015	472	1km=65p	0.016	917	1km=61p	0.016	382	1km=63p	-	-	-	-	-	-
5	Moscow	0.010	494	1km=101p	0.010	779	1km=94p	0.010	175	1km=97p	0.04	765	1km=95	-	-	-
6	Tokyo	0.014	962	1km=70p	-	-	-	-	-	-	-	-	-	-	-	-
7	New York	0.013	566	1km=75p	0.014	816	1km=71p	0.014	272	1km=73p	-	-	-	-	-	-
8	Los Angeles	0.014	665	1km=69p	0.015	828	1km=65p	0.015	242	1km=67p	-	-	-	-	-	-
9	Sao Paulo	0.010	505	1km=104p	0.017	772	1km=59p	0.010	180	1km=118p	-	-	-	-	-	-
10	Singapore	0.010	305	1km=113p	0.010	361	1km=54p	0.010	168	1km=108p	-	-	-	-	-	-
11	Sydney	0.014	496	1km=69p	0.015	667	1km=65p	0.015	304	1km=67p	-	-	-	-	-	-
12	Johannesburg	0.015	551	1km=64p	0.017	1180	1km=60p	0.016	404	1km=62p	-	-	-	-	-	-
13	Toronto	0.012	435	1km=79p	0.013	419	1km=74p	0.013	93	1km=76p	-	-	-	-	-	-
14	Berlin	0.010	372	1km=102p	0.011	689	1km=87p	0.011	277	1km=90p	-	-	-	-	-	-
15	Buenos Aires	0.014	462	1km=69p	0.015	645	1km=65p	0.015	339	1km=67p	-	-	-	-	-	-
16	Beijing	-	-	-	0.050	2209	1km=36p	-	-	-	-	-	-	0.047	1,821	1km=21 p
17	Riyadh	0.013	292	1km=75p	0.017	886	1km=59p	0.016	214	1km=61p	-	-	-	-	-	-
18	Nicosia	0.038	48	1km=265p	-	-	-	-	-	-	-	-	-	-	-	-

(1) Scale of the map (S)
(2) Estimated road length from pixels (km)
(3) Relationship between real distance and pixels (1 km = number of pixels)

4.4.2 Timeseries of traffic through pixels

As a second step, the correlation of the coloured traffic layers, that represent different traffic states during the diurnal cycle for all the selected cities, was investigated. In the following figures, the percentage of pixels for each traffic layer is plotted for one typical weekday using 5-minute intervals along with the boxplot of each traffic layer for comparative purposes. The right corner of the figures depicts the isolated traffic layers for each city. Regarding the number of traffic layers, four different traffic layers are used from Google, Bing, Yandex and Baidu maps to display urban traffic (light, moderate, heavy traffic and congestion) while Here maps use only three, omitting a separate layer for congestion. Slightly different colours are used from each provider. For instance, green colour is used for light traffic and red or dark red for congestion, while moderate traffic is displayed either with yellow (Bing, Here and Yandex maps) or orange colour (Google and Baidu maps).

For most of the cities, traffic coverage is provided from three different online traffic maps, Google, Bing and Here maps, according to Table 4.4. In Figures 4.7-4.11, the timeseries of pixels per traffic layer is depicted for Paris, London, Istanbul, New Delhi, Moscow, New York, Los Angeles, Sao Paulo, Singapore, Sydney, Johannesburg, Toronto, Berlin, Buenos Aires and Riyadh based on information retrieved from Google maps. Respectively, timeseries of pixels for the same cities from Bing maps (Appendix B) and Here maps (Appendix C) were also estimated for the same day. Traffic coverage for Paris, Istanbul, and Moscow is further provided from Yandex maps, (Figure 4.14), with greater urban coverage, as it can be seen from the image in the right corner of the same figure. On the other hand, the traffic state for the city of Tokyo is given only from Google maps while Baidu maps provide traffic information for the city of Beijing (Figure 4.12).

A clear correlation between the traffic layers of all the cities is evident; when the percentage of green pixels (Light traffic) drops rapidly during the morning and evening peak hours, a rise in the percentages of the remaining layers appears, as expected, at the same time. The same applies to all cities; however, interestingly in Tokyo and in Johannesburg the drop in the percentage of green pixels is significantly less compared to the rest cities (less than 50%). Regarding the dark red traffic layer (congestion), Moscow and Los Angeles record the greatest percentage during the day (over 10%). Sharp peaks of red and dark traffic layers can be identified in both cities in the United States while the cities of Paris and Istanbul follow. Comparing morning and evening peak hours, Paris presented higher percentage of red pixels (heavy traffic) during the morning peak hours while New York, Istanbul, Delhi,

Sao Paulo, Singapore and Berlin during the evening peak hours. A quite different pattern is observed for the city of London and Moscow where approximately the same percentage of red pixels remains from morning till evening peaks hours.

A quick view of the variation of each traffic layer for the same day is given by the accompanied boxplots in the same figures. Again, Moscow score highest since red and dark red traffic layers vary the most compared to other cities, while Los Angeles follows. Tokyo shows the least variation for all the traffic layers during the day, which suggests a more stable temporal mobility pattern. Comparing the median values of each traffic layer per city, the first three depicted in Figure 4.7 share the same value (less than 60%) for the green traffic layer in contrast to the city of Singapore and Johannesburg where median values are over 80% while the rest cities range between 60-80%. A higher median value for the green traffic layer implies that a greater part of the network within the city remains uncongested during the day. Regarding the minimum values for the same layer, Paris, Moscow, Los Angeles and Sao Paulo presented values less than 40%. Moderate traffic (orange traffic layer) rises over 30% (median value) in Paris and in New Delhi while the remaining cities follow with smaller values around 25%. Median values less than 20% are observed for three cities, named Singapore, Sydney and Johannesburg. Regarding the maximum values for the same layer, Sao Paulo and New Delhi rank first (over 40%). Last, maximum values for heavy traffic (red traffic layer) during the same weekday appeared for the city of Paris, London, Moscow, New York Los Angeles, Sao Paulo and Berlin (over 20%). Here, it should be noted that disturbances appearing in the time-series of New York, Los Angeles, and Sao Paulo are due to data filtering, as already stated.

For the city of Nicosia, a different scale of the urban area was chosen for the analysis. The selected area focuses on the Business District Area (BDA) and Figure 4.13 presents the daily traffic pattern both for a typical weekday (Tuesday) and Saturday and further for a whole week. Morning peak appears at 8:00 a.m. and evening peak at 18:00 p.m. with the percentage of red pixels to reach 15%. On Saturday the city centre remains uncongested as the percentage of red pixels is lower than 10% and the average value of light traffic is 75% according to the boxplot. Last, the weekly traffic pattern of the city clearly shows the defences between weekday and the weekend.

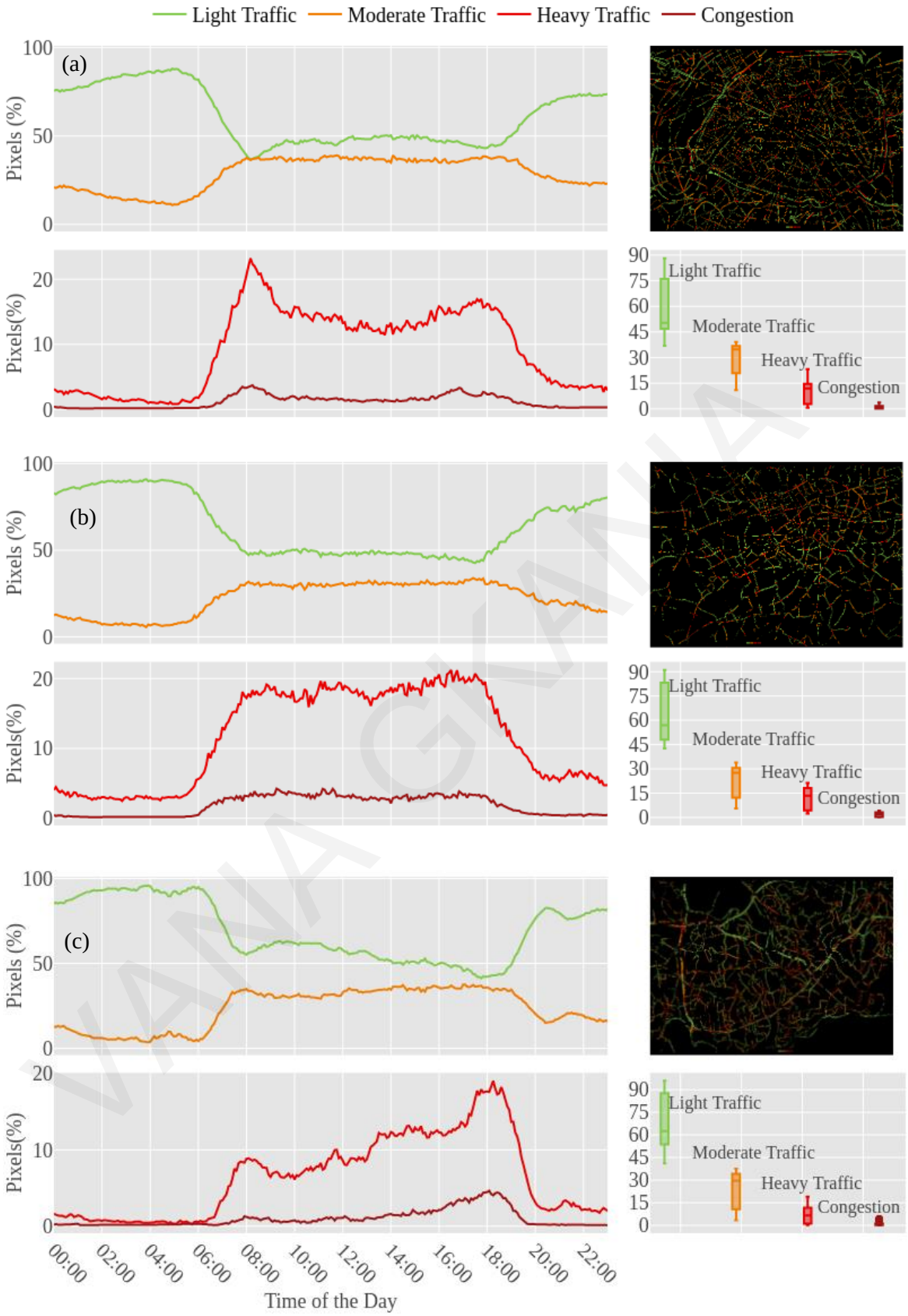


Figure 4.7: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Paris; (b) London; (c) Istanbul from Google maps.

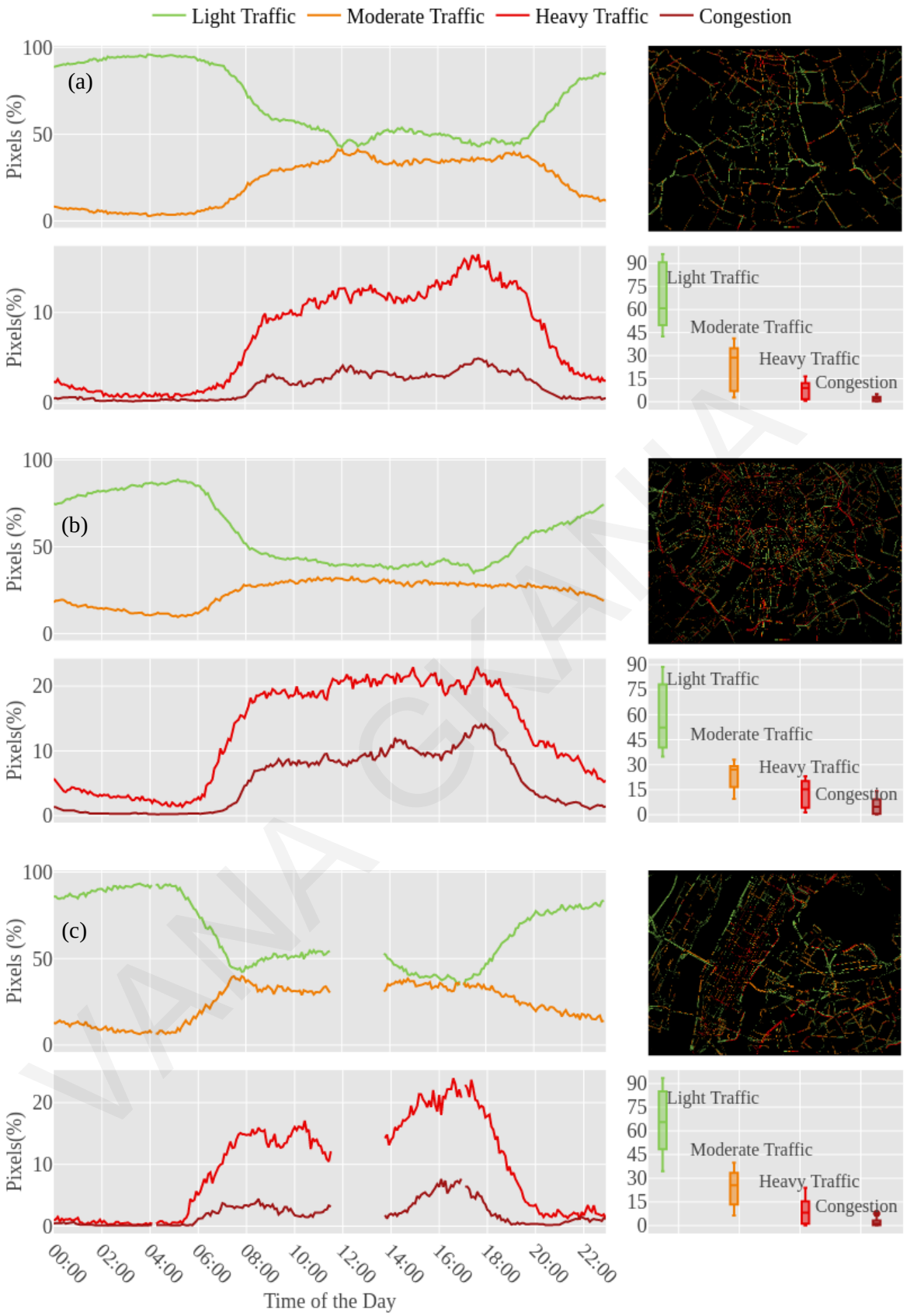


Figure 4.8: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) New Delhi; (b) Moscow; (c) New York from Google maps.

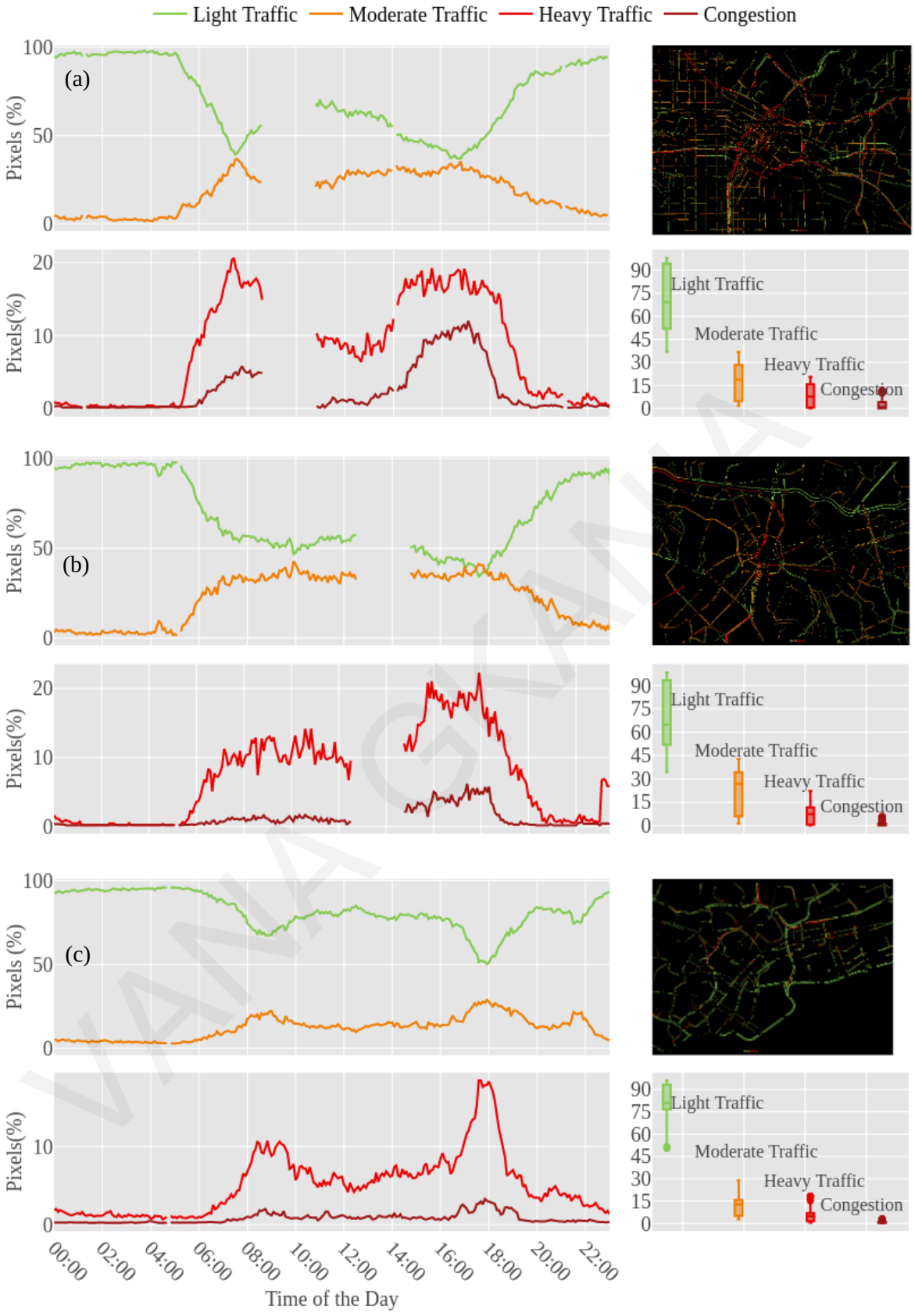


Figure 4.9: Timeseries of pixels and boxplot per traffic layer for one weekday for Los Angeles from (a) Los Angeles; (b) Sao Paulo; (c) Singapore from Google maps.

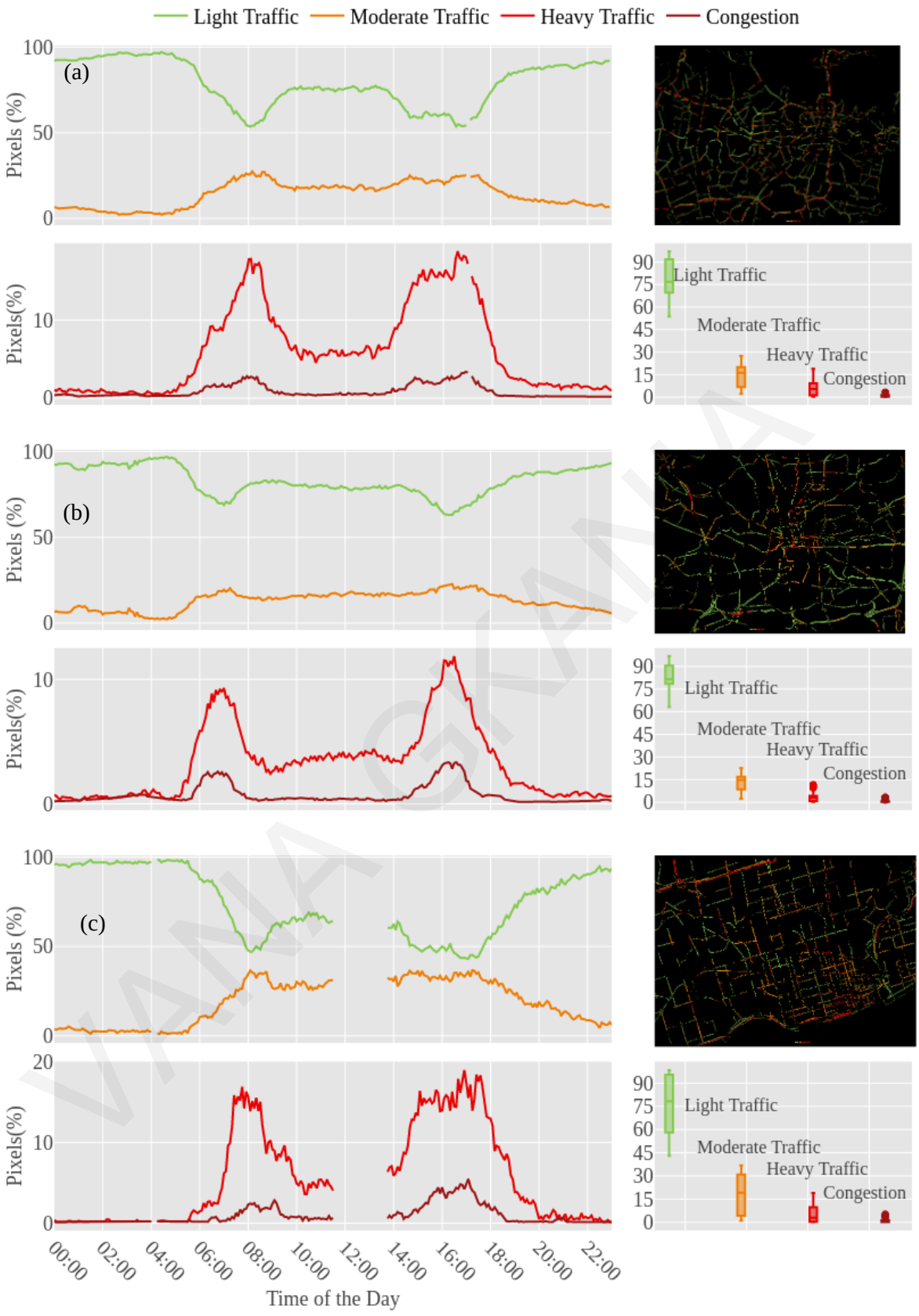


Figure 4.10: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Sydney; (b) Johannesburg; (c) Toronto from Google maps.

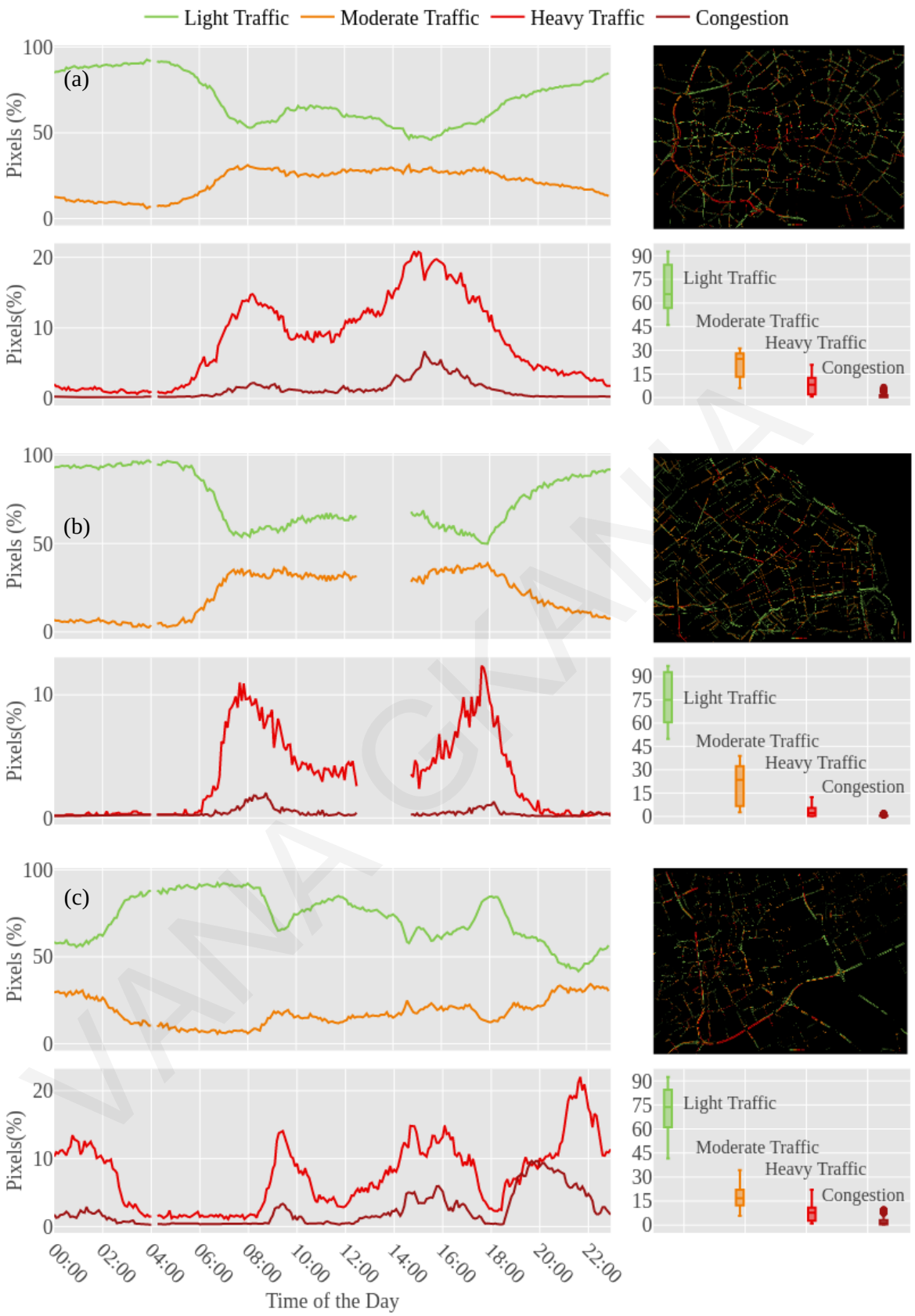


Figure 4.11: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Berlin; (b) Buenos Aires; (c) Riyadh from Google maps.

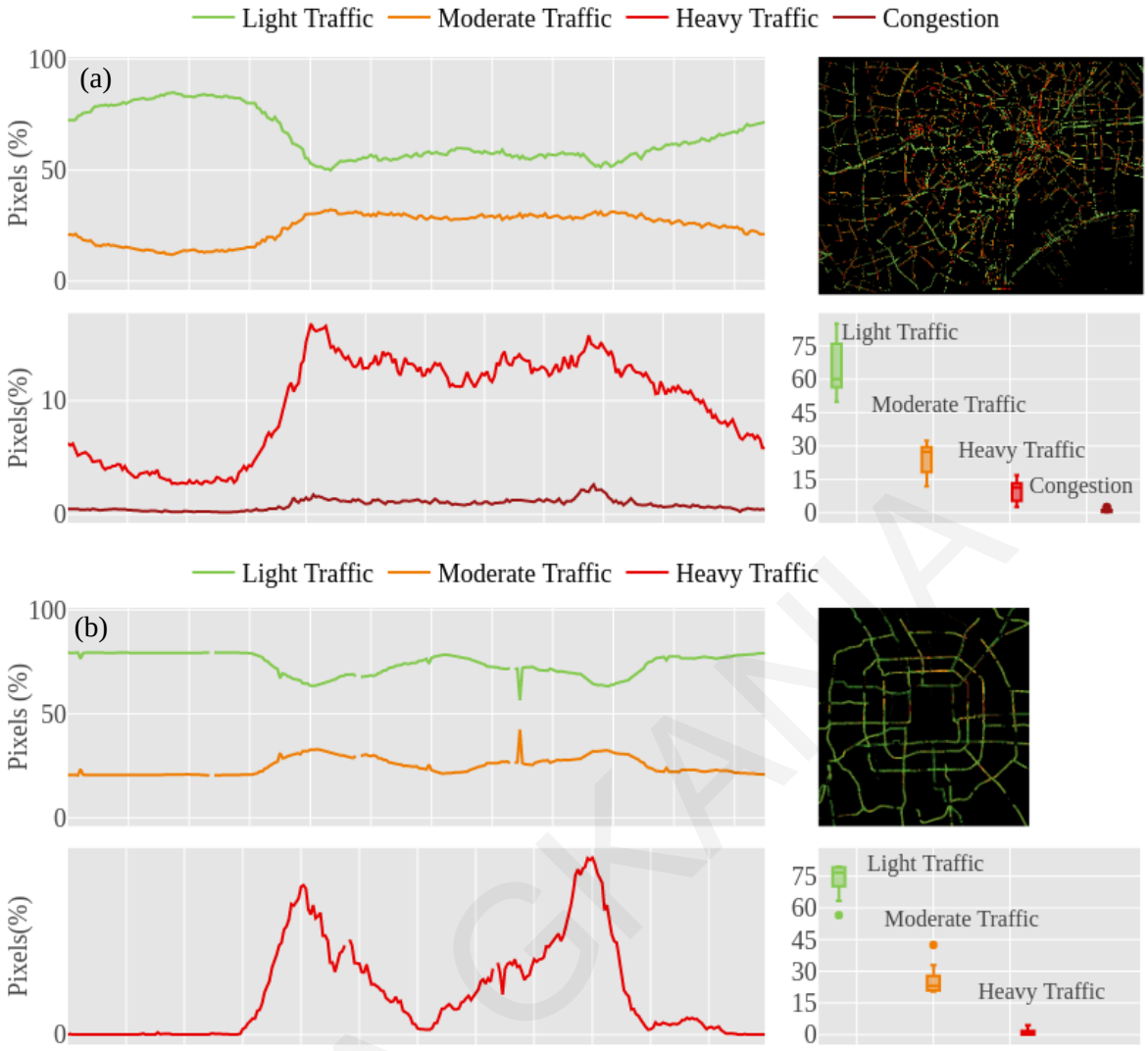


Figure 4.12: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Tokyo form Google maps and (b) Beijing from Baidu maps.

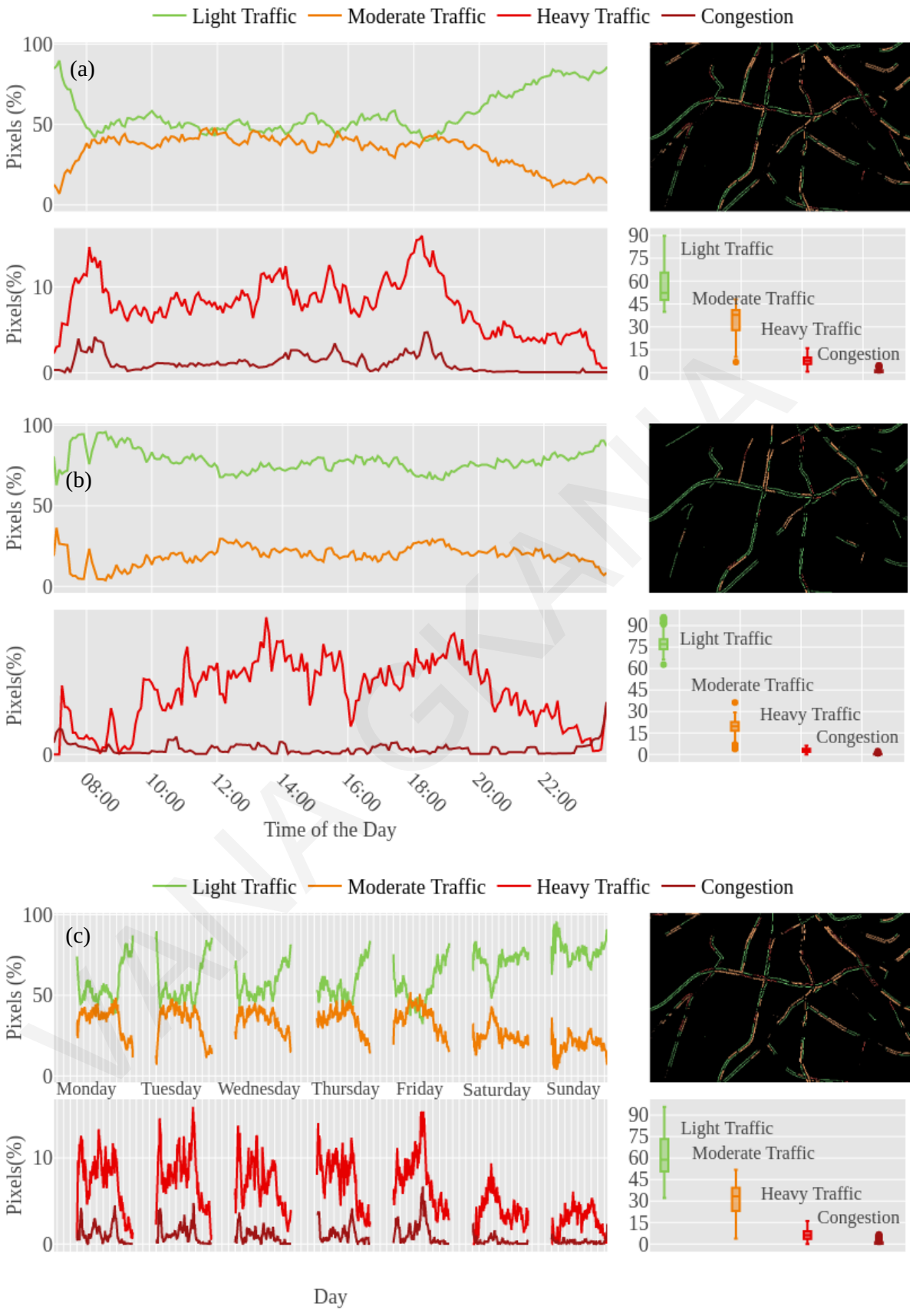


Figure 4.13: Timeseries of pixels and boxplot per traffic layer for Nicosia for (a) Tuesday; (b) Saturday; (c) whole week from Google maps.

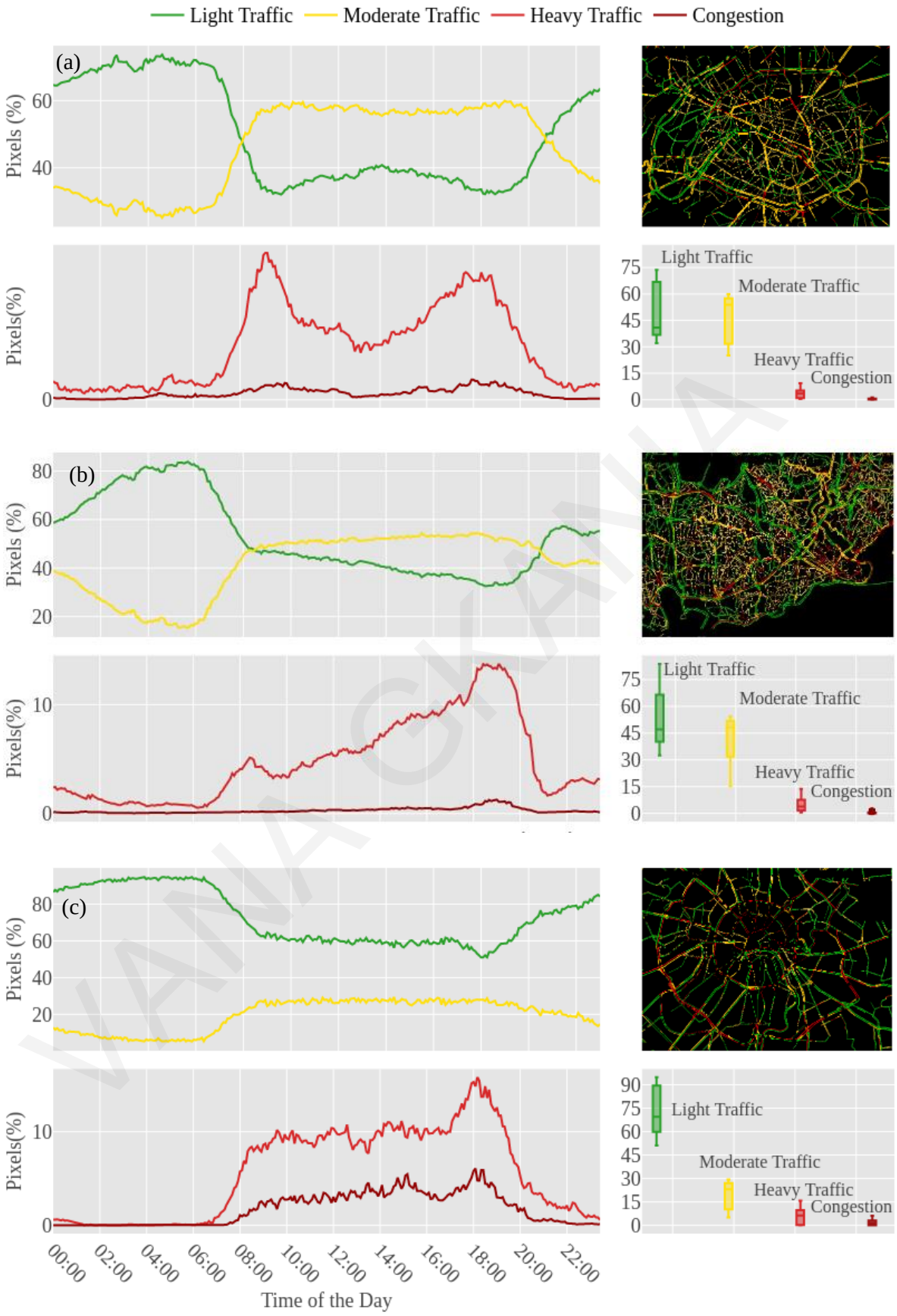


Figure 4.14: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Paris; (b) Istanbul; (c) Moscow from Yandex maps.

This first analysis provides good insights on the temporal dimension of mobility within the city besides the total percentage of space that remains uncongested across the cities. The alternation of the percentage of pixels for each traffic layer reveals the proportion of the road network that remains congested through the day. Final, it was proved that the estimated timeseries of traffic through raster images resembles known traffic timeseries estimated from traffic surveillance systems.

4.5 Chapter Summary

In this chapter the experimental setup for collecting aggregated traffic information from online traffic maps along with the traffic data acquisition procedure was presented and analyzed. Initially the study area and the basic characteristics of the selected cities were introduced. The data collection and processing followed, while through the preliminary analysis of the data, insights regarding the urban coverage and the pattern of traffic state through pixels were given. A thorough examination of the extracted data follows in the next chapter.

Chapter 5 : Validation of class-type Information from Online Traffic Maps

In the current chapter, the estimation and the investigation of the Macroscopic Fundamental Diagram-MFD properties are performed utilizing the captured pixels properties for the selected study area. The validity of the method is tested by comparing the estimated MFDs to ground-truth MFD obtained using empirical data from loop detectors. Further, other macroscopic models are demonstrated for comparative reasons.

5.1 Methodological Framework

The MFD provides a relationship between network-wide average values of traffic flow, speed, and density, as was previously mentioned in Chapter 3 and thus it is claimed that it represents the capacity of a road network in terms of vehicle density. The remainder of this section describes how the MFD can be estimated using class-type information from online traffic maps along with the assumptions of the proposed method. Table 5.1 summarizes the notation of all the symbols and variables used in this chapter

Table 5.1: Nomenclature

Symbol	Description	Unit
$\boldsymbol{S}$	Scale of the thematic map	-
$\boldsymbol{p_c}$	Pixels colored with specific color c	-
$\boldsymbol{U_c}$	Representative value of average speed according to color c from the speed map legend	Km/h
$\boldsymbol{c}$	Color index (i.e. g for green, y for yellow, o for orange and r for red)	-
$\boldsymbol{j}$	Link index	-
$\boldsymbol{i}$	Pixel index	-
$\boldsymbol{m}$	Total number of links in one image	-
$\boldsymbol{n}$	Total number of pixels in one image	-
$\boldsymbol{l_j}$	Length of link j	Km
$\boldsymbol{t}$	time interval	-
$\boldsymbol{U^t}$	Average network speed across all the links of a given network	Km/h at time interval t

$\boldsymbol{Q^t}$	Average network flow across all the links of the network	Veh/h at time interval t
$\boldsymbol{K^t}$	Average network density across all the links of the network, in	Veh/km at time interval t
$\boldsymbol{A^t}$	Total number of vehicles (accumulation) on the network at time interval t	Veh at time interval t
$\boldsymbol{U_f}$	Free-flow speed of a network	Km/h
$\boldsymbol{U_o}$	Optimum speed of a network	Km/h
$\boldsymbol{K_{jam}}$	Jam Density of a network	Veh/km

5.1.1 Assumptions

Prior the estimation of the fundamental variables (average network speed, flow and density) based on data from traffic maps, the following assumptions of the proposed method are introduced:

1) The length of each road network link (j) on the map is estimated by using the total number (n) of colored pixels p_{ci} (where c represented one of the four colour codes and pixel p_{ci} where $i \in j$) based on the following equation:

$$l_j = \sum_{i=1}^{n} \sum_{c=1}^{4} (p_{ci} \times S) \qquad (5.1)$$

2) Although, a pixel has both horizontal and vertical dimensions, for the proposed application the pixel dimensions are strictly related to the length of a link and not to its width.

3) Each pixel is also related to the spatial dimension of traffic and can be regarded as a virtual loop detector providing information about traffic conditions at each location.

At that point, it is important to mention also the methodology that online providers follow to depict current traffic conditions in a city. Starting with the Yandex maps, where data are collected through the Yandex.Maps app or the Yandex.Navigator app. After activating the app's "Send traffic information" option, the user starts sending every few seconds their geographic coordinates, direction and speed to the automated analytical system

of the Yandex.Traffic service. All information sent from each mobile device is non-personal. Yandex.Traffic's automated analyzer then integrates the speed and coordinate information from all participating vehicles driving along the same route into unified traffic patterns – tracks. In addition to contributions from private users, Yandex.Traffic also receives information from companies that have fleets of vehicles operating in the city's streets on a regular basis.

To make a track for a moving car, Yandex.Traffic uses a number of geographic coordinates, which are delivered by the driver's GPS device and sent to the service through the Yandex.Maps app. The GPS accuracy, however, has the error margin from one to ten meters in all directions, which may result in positioning a car on a sidewalk or rooftop of the nearest building. To solve this problem, GPS coordinates are mapped to the digital map of the city, which accurately displays roads and streets with all the markings, buildings, parks, and other urban facilities. This detailed mapping allows the system correct the course of a car based on the real physical layout even if the GPS coordinates say that the car is on the wrong side of the road or has cut through a building instead of following road markings and turning around the corner. Another important issue is to understand how useful the speed information received from the driver is, as it may or may not truthfully reflect the real situation on the road. If all other cars sending information to Yandex.Traffic proceed as normal on the same route, the system ignores the rogue track and this data isn't considered for the general traffic evaluation. This is exactly why the number of Yandex.Traffic users matters. The more drivers send information to Yandex.Traffic via Yandex.Maps, the more accurate the picture of the real-time traffic situation is. After combining the tested tracks, the algorithm analyzes them and sets "green", "yellow" and "red" ratings to the corresponding sections of roads. Next comes aggregation - the process of combining information. Every two minutes, Yandex.Maps aggregates, like a jigsaw puzzle, all information from all users of the Yandex.Maps app and maps the results on the Traffic Jams layer, both in the mobile application and on the desktop service (Yandex, 2014).

Similarly, Google utilize crowdsourcing to improve the accuracy of its traffic predictions. When Android phone users turn on their Google Maps app with GPS location enabled, the phone sends back bits of data, anonymously, to Google that let the company know how fast their cars are moving. Google Maps continuously combines the data coming in from all the cars on the road and sends it back by way of those colored lines on the traffic layers (Barth, 2009).

5.1.2 Estimation of MFD based on traffic maps data

To explore the capacity of online traffic maps for estimating network-wide phenomena (e.g. in the detail that MFDs are offering), fundamental traffic variables were first estimated utilizing measurements from the whole network. For achieving that, a connection between the average network flow, the average network speed, the average network density, and the number of pixels that belong to distinctive operating class of online traffic maps was established. The mathematical expressions used for achieving this follow.

Firstly, by considering that typical online maps offer information about links' average speed per coloured segment (complete links or link's stretches), we can estimate the network average speed U^t at time interval t as:

$$U^t = \frac{\sum_{j=1}^{m}(l_j \times U_{ci}^t)}{\sum_{j=1}^{m} l_j} = \frac{\sum_{i=1}^{n}(p_{ci}^t \times S \times U_{ci}^t)}{\sum_{i=1}^{n}(p_{ci}^t \times S)} \quad (5.2)$$

Where l_j is estimated using equation (5.1) as the scale of the map (S) is known, while the number of pixels $\mathrm{p_c}$ coloured with specific colour c is calculated based on the proposed image processing techniques (presented in Chapter 4). Regarding the representative value of average speed (U_{ci}) according to colour c, rational assumptions for each threshold can be used as green refers to free-flow, orange to moderate traffic, red to heavy traffic and dark red to congestion based on the map's legend. Usually, online traffic maps' providers do not give precise ranges for the color-coding scheme that has been adopted and as so ground truth data form loop detectors were used to assign representative values of average speed through empirical calibration, as described in the following section.

Then to estimate the average network density K^t that is defined as the average number of the vehicles per unit length at time interval t , a parametric approach was followed based on the fundamental linear speed-density relation proposed by Greenshields (Greenshields, 1935), that follows:

$$U^t = U_f - \left(\frac{U_f}{K_{jam}}\right) K^t \quad (5.3)$$

where U_f stands for the free-flow speed and K_{jam} is the jam density of the network.

The free-flow speed is relatively easy to estimate, as it generally lies between the speed limit and the design speed of the roadway. On the other hand, a reliable estimation of jam density, being representative enough of the vehicles using the network of each given time, is much more difficult. In our analysis the representative values both for free-flow

speed and density jam of the network were estimated through a calibration process utilizing empirical data from loop detectors for the selected road network.

As the variable U^t is also known from online traffic maps (see Eq. (5.2)) an estimation of average network density (K^t) can be achieved for known free-flow speed and jam density by substituting U^t into Eq. (5.3) and solving for K^t:

$$K^t = K_j - \left(\frac{K_{jam}}{U_f}\right) U^t \tag{5.4}$$

The equation of flow-density relationship can be derived by substituting $\frac{Q}{K}$ for U (based on the fundamental traffic relationship), into Eq. (5.3) and solving for Q, as:

$$Q^t = U_f K^t - \left(\frac{U_f}{K_{jam}}\right)(K^t)^2 \tag{5.5}$$

Additionally, the total number of vehicles (accumulation) on the network A^t at time interval t, can be estimated by multiplying the average network density with the total length of the links in the network using Eq. (5.6).

$$A^t = \sum_{j=1}^{m} l_j \times K^t = \sum_{i=1}^{n} (p_{ci}^t \times S) \times K^t \tag{5.6}$$

As it can be noticed by the above equations, a relationship between microscopic traffic information (captured in each pixel) with macroscopic (network-wide) metrics has been established. This interplay between microscopic and macroscopic variables will be extended, enhanced and used in the later sections. It is explicitly noted here that the fundamental formula Q=K×U used for integrating pixels/locations traffic variables can be regarded as simplified, reflecting the assumption that in the complete functional form where E(Q)=E(K)×E(U)+COV(K,U) the term COV(K,U)=0 (E standing for expected value). Though, the implementation made here reflects an average functional among the three fundamental traffic variables, adequate for the purposes of the current analysis.

5.2 Calibration with Empirical Data

To evaluate the proposed method, traffic data from online traffic maps for the city of Nicosia (see Figure 4.24) and empirical data from available loop detectors for the same network were used. Through this process the values of jam density and free flow of the network as well as the representative values of speed per coloured traffic layer were selected and applied on the Eqs. (5.2), (5.4), (5.5), and (5.6).

As a first step, traffic data for the city of Nicosia were collected for the same week (11-17 February 2019), from four available loop detectors, named 1004, 1005, 1006, 1010, located mainly in bidirectional cross-sections of major arteries in Nicosia. as Figure 5.1 shows. These loop detectors provide observations of average speed and traffic flow for 5-minute intervals. The observations of the loop detectors were used to calibrate the proposed models through optimization process. Particularly, the objective was to minimize the Sum Absolute Error (SAE) between the estimated values from Eqs. (5.2) and (5.5) and the empirical observations, at the location of the detectors. The constraints that were set for the six variables, U_g, U_o, U_r,U_{dr}, K_{jam},U_f follow:

- $0 \leq U_g \leq 60\ Km/h$, $0 \leq U_o \leq 50\ Km/h$, $0 \leq U_r \leq 40\frac{Km}{h}$, $0 \leq U_{dr} \leq 20\ Km/h$
- $100 \leq K_{jam} \leq 180\ veh/km$
- $40 \leq U_f \leq 90\ km/h$

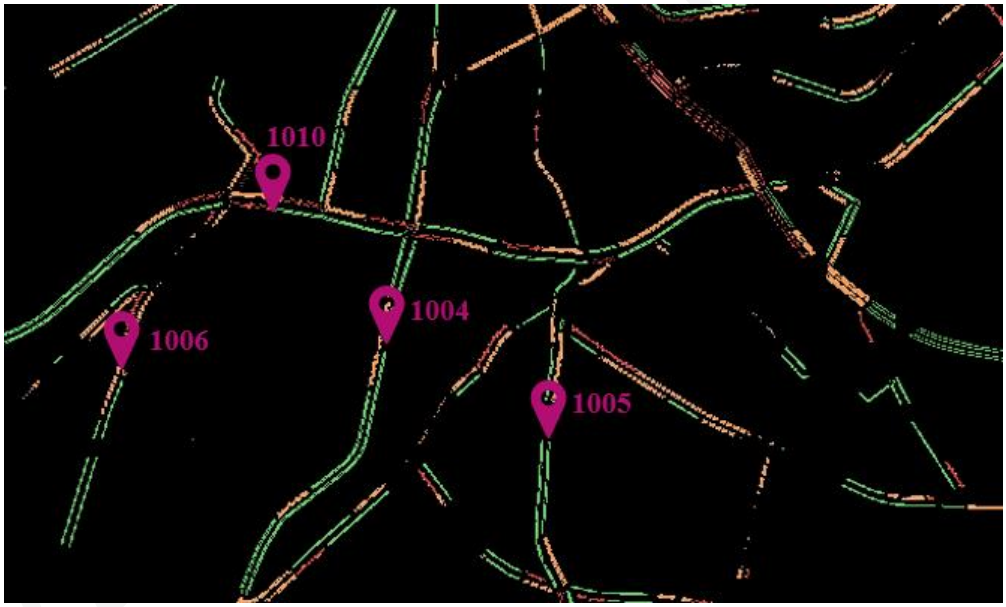


Figure 5.1: Location of the loop detectors in Nicosia, Cyprus

A summary of the calibration results is provided in Table 5.2. At each location of the four loop detectors (1010, 1005, 1004 bidirectional) five different models were tested. In the first model the objective function was to minimize the SAE both for average flow and speed at each location by searching the best fit for all the six variables. In the second model the same objective function was used while values for jam density and free flow remained constant, as in the third model but for different value of free-flow speed. In the fourth model the objective function was to minimize SAE only for average speed while in the fifth model only for flow, respectively.

Table 5.2: Calibration Results

Row	Model U_g(km/h)	U_o(km/h)	U_r(km/h)	U_{dr}(km/h)	K_{jam}(veh/km)	U_f(km/h)	SAE
	Location:1010 Westbound						
1	**45**	**35**	**31**	**20**	**105**	**56**	**54,357(1)**
2	43	37	37	19	150	50	55,192(2)
3	52	48	40	20	150	60	65,562(3)
4	51	42	40	20	100	63	58,756(4)
5	45	36	32	20	100	80	146,748(5)
	Location: 1010 Eastbound						
7	**45**	**41**	**40**	**20**	**100**	**57**	**48,142(1)**
8	43	41	40	20	150	50	48,897(2)
9	52	50	40	11	150	60	62,857(3)
10	45	41	40	20	100	57	48,156(4)
11	45	41	38	16	100	80	150,292(5)
	Location: 1006 Southbound						
12	**42**	**41**	**40**	**20**	**100**	**45**	**12,115(1)**
13	48	47	40	20	150	50	27,049(2)
14	58	50	40	3	150	60	58,387(3)
15	42	41	40	20	100	45	12,136(4)
16	45	43	40	20	100	80	116,518(5)
	Location: 1005 Northbound						
17	**45**	**43**	**40**	**20**	**100**	**51**	**25,012(1)**
18	46	45	40	20	150	50	26,127(2)
19	56	50	40	5	150	60	37,053(3)
20	42	40	40	20	100	48	25,888(4)
21	45	42	33	20	100	80	149,022(5)
	Location: 1005 Southbound						
22	**46**	**42**	**40**	**20**	**100**	**51**	**20,501(1)**
23	47	45	40	20	150	50	22,604(2)
24	57	50	40	20	150	60	53,130(3)
25	45	42	40	20	100	50	20,569(4)
26	48	42	38	20	100	80	142,911(5)
	Location: 1004 Northbound						
27	**43**	**42**	**40**	**20**	**100**	**47**	**18,612(1)**
28	47	47	40	19	150	50	28,616(2)
29	57	50	40	20	150	60	67,675(3)
30	42	41	40	20	100	46	19,060(4)
31	50	48	40	0	100	80	122,231(5)
	Location: 1004 Southbound						
32	**48**	**46**	**40**	**20**	**101**	**56**	**23,285(1)**
33	45	44	40	20	150	50	24,176(2)
34	55	50	40	20	150	60	31,755(3)
35	42	40	40	20	100	50	25,834(4)
	48	47	40	20	100	80	96,913(5)
	Location: 1010, 1006, 1005, 1004						
36	**44**	**42**	**40**	**20**	**100**	**51**	**272,267(1)**
37	**46**	**45**	**40**	**4**	**150**	**50**	**283,342(2)**
38	55	50	40	20	150	60	410,927(3)
39	40	38	40	20	100	47	282,949(4)
40	46	43	40	20	100	80	933,457(5)

Note

(1) Minimize SAE (flow and speed) by changing values in all variables

(2) Minimize SAE (flow and speed) by changing values in speed variables, K_{jam}=150veh/h and U_f=50vkm/h.

(3) Minimize SAE (flow and speed) by changing values in speed variables, K_{jam}=150veh/h and U_f=60vkm/h.

(4) Minimize SAE (flow) by changing values in all variables

(5) Minimize SAE (speed) by changing values in all variables

At the location 1004, both for southbound and northbound direction the best fit was given for the first model (minimum SAE in bold). Thus, the values from row 27 and 32 of Table 5.2 were used to estimate average speed and flow from Eqs. (5.2) and (5.5) illustrated in Figure 5.2 (a) (black and red triangles). To better visualize the validity of the proposed method, data from the loop detector 1004 (red color for southbound and black color for northbound) are plotted at the same figure, as well. The small number of observations from the model for the two direction is due to the small number of pixels in the image that corresponds to the location of the detector. As a result, the pixels had only one color at each time interval and according to Eq. (5.2) the value of average speed was determined for example only from the value of U_g.

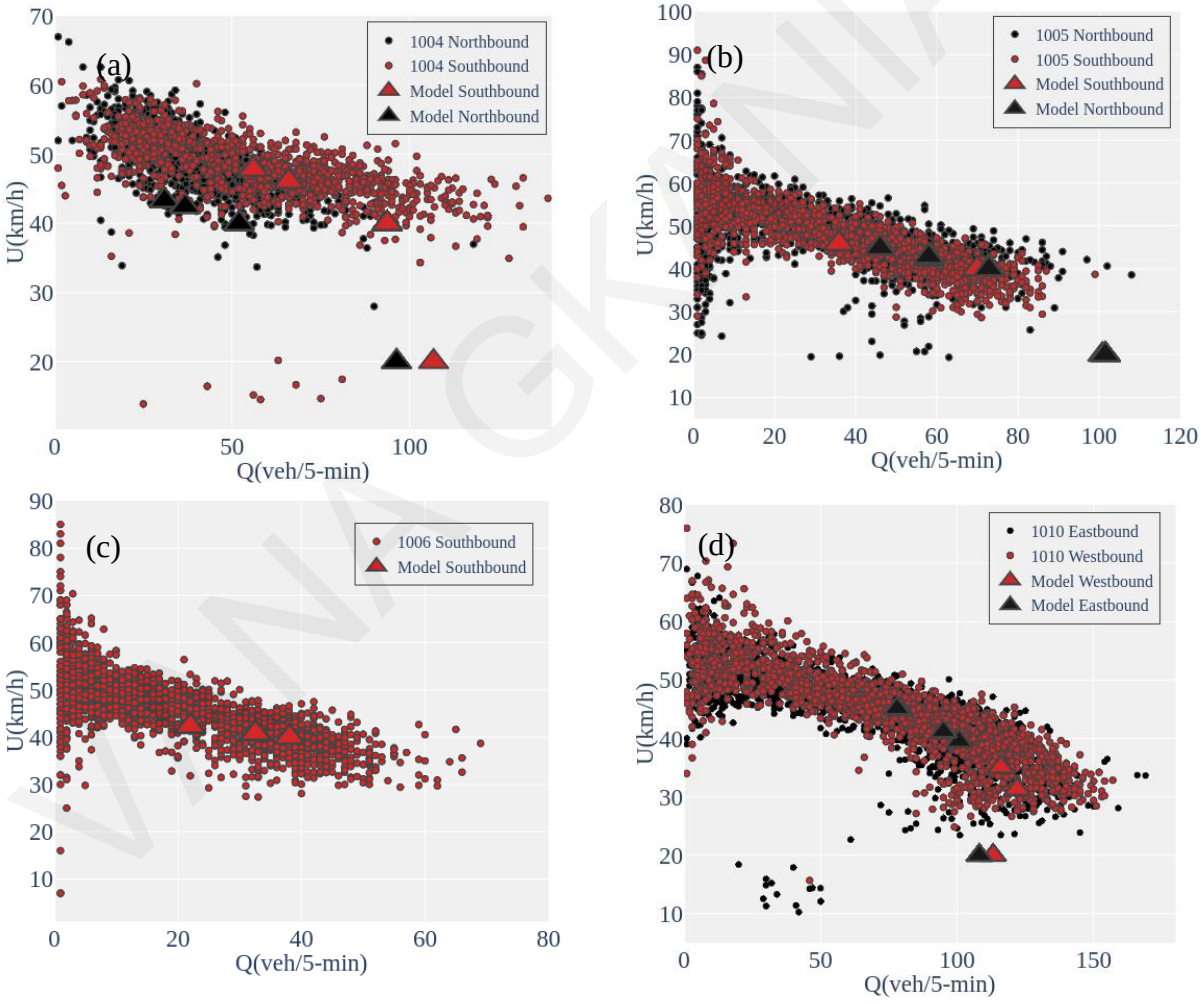


Figure 5.2: Comparison between the proposed model and the empirical data from the loop detectors at each location 1004(a), 1005(b), 1006(c), and 1010(d) for one week (11-17 February 2019).

Regarding the rest locations in the network, the first model also gave better fitting results. Thus, the values of rows 22, 17, 12, 7, 1 (see Table 5.2) were used to estimated

average speed and flow at the rest locations, depicted on Figure 5.2. (b), 5.2 (c), 5.2 (d,) respectively. For a consistent comparison between the estimated values and empirical data, average speed and flow from the loop detectors (red and black circles) are also depicted in the same figures.

So far, the optimization process was applied at one location of the network at each time. At the last part of Table 5.2 all the available locations of the detectors were utilized in the calibration process. In that case, the target was to minimize SAE both for average flow and speed and for all the four locations (in all the five models). Figure 5.3 illustrates estimated values (triangles) from all models (1,2,3,4,5) and the loop detectors' data(circles) for the same week, for comparative purposes. According to Table 5.2 the first model (row 36) presented a slight lower SAE compared to the second one (row 37), but the shape of the model two (blue color, Figure 5.3) is much closer to the theoretical one. It is of interest, how well the second model fits to the empirical data (almost in the middle). Thus, values from row 37 (Table 5.2) were chosen for the estimation of the MFDs for the whole network that follows in the next section.

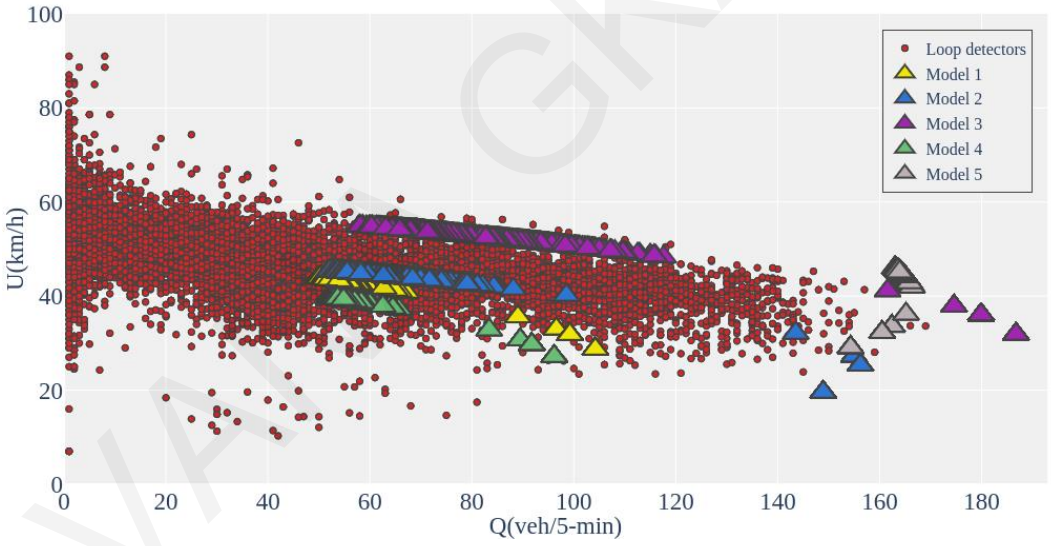


Figure 5.3: Comparison between the proposed models and the empirical data from the all the available loop detectors and for the whole network for one week (11-17 February 2019).

5.3 MFD based on the Greenshields model for the city center of Nicosia

MFD is an instrument that encompasses the most (if not all) the important network-wide information related to traffic operations. As so, the estimation of the MFD based on the earlier described image-processing technique and its macroscopic traffic flow interpretation are presented in this section.

For the city of Nicosia, the reference method is implemented, and results are shown in Figure 5.4. The plotted data set was calculated from traffic maps' images, utilizing solely pixels' information for the selected week in February 2019, using a 5-minutes time interval t. After the empirical calibration described in the previous section, the following values were selected and applied in Eqs. (5.2), (5.4), (5.5), (5.6) to estimate U^t, K^t, Q^t and A^t, respectively:

- K_j = 150 veh/km and U_f = 50Km/h
- Applied values for each color code are: U_g =46Km/h, U_y =45Km/h, U_r =40Km/h, U_{dr} =4Km/h
- S=0.038 according to Table 4.3

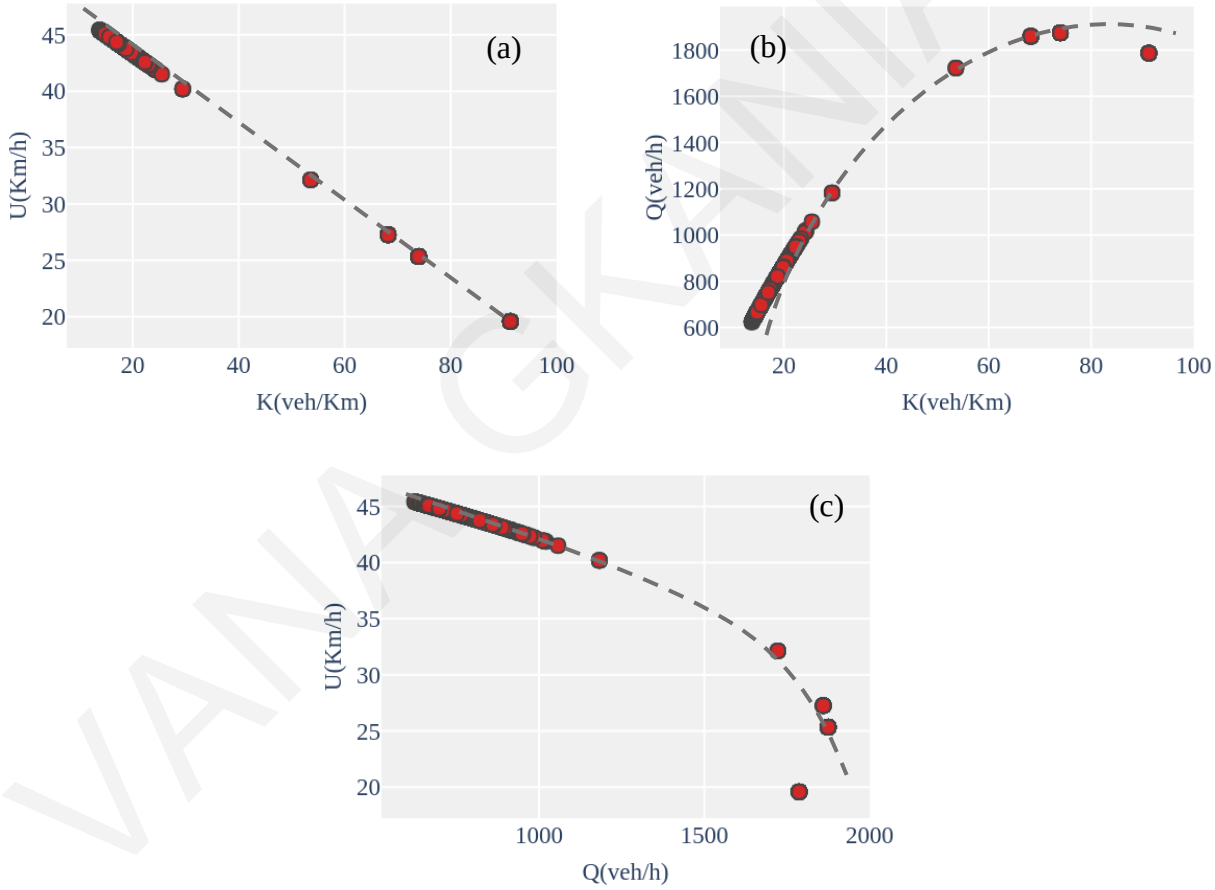


Figure 5.4: Nicosia's MFDs, for the selected week, (5-minute intervals): (a) average speed-average density diagram; (b) average flow-average density diagram; (c) average speed-average flow diagram.

The results of the Nicosia network shown in Figure 5.4, follow the theoretical shape of MFDs, although, the congested portion of the curve is not fully developed. Consequently, it can be concluded that the selected network remains practically uncongested for the examined time period, an element that is valid for the city of Nicosia.

Next, for the same city, the investigation of the relationship between the total vehicles A_t in the network and the average network speed U_t has been performed and presented in Figure 5.5 (a). Accordingly, Figure 5.5 (b) presents the relationship between accumulation and the average network flow. A connection among selected data (a-h) on the diagrams and the network's traffic state for the same intervals is highlighted on the same figures, illustrating the different network operating conditions as these are reflected in typical color-coding schemes.

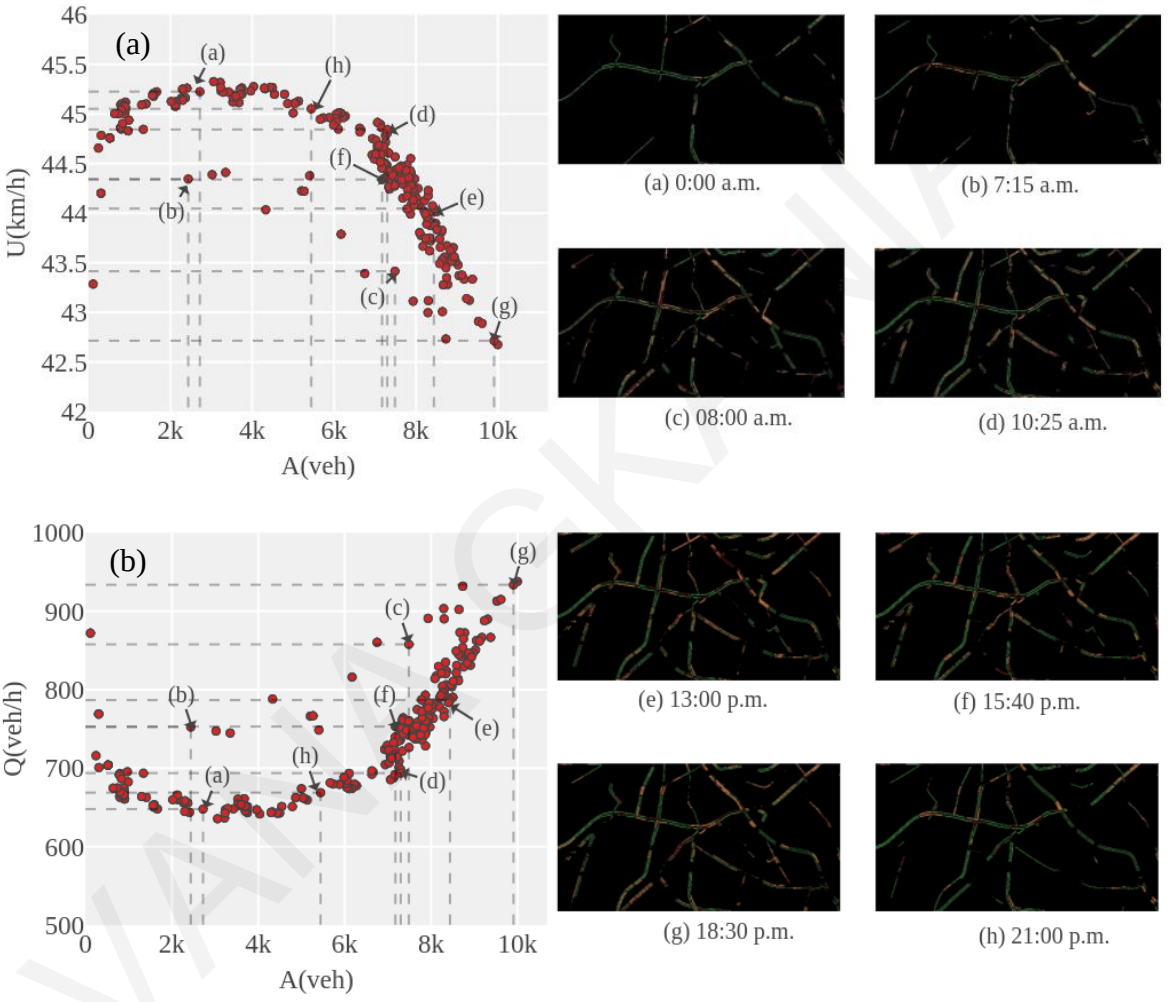


Figure 5.5: The relationship between average network speed(a); flow(b) versus the total number of vehicles, and network's traffic state for selected values of average accumulation during February 11th, 2019 (5-minute intervals).

Starting with point (a) in Figure 5.5 (a) that corresponds to 00:00 a.m. (time interval), higher average speed is observed for low accumulation, as probably drivers tend to drive faster in the almost empty network. During the morning peak hours, points (b) and (c), the number of vehicles that enter the network start to increase and as a result average speed start to decrease. At 10:25 a.m. (d) the network starts to improve the average network speed while the situation deteriorates again during the evening peak hour (g) at 18:30 p.m. After 21:00

p.m. (h) the number of the vehicles in the network present a significant reduction, and average speed is increased as it was expected.

Looking at the same points in Figure 5.5 (b), we can see that till 08:00 a.m. average flow of the network increases as more vehicles enter in the previous empty network while a steep incline appears at 13:00 a.m. point (e), where some links in the network start to phase congestion. During the evening peak hour at point (g), the city center seems to reach a maximum level of concentration where average traffic flow reaches 950 vehicles per hour and then phenomena of hysteresis loops seem to emerge.

5.4 Case Studies for the Different Urban Networks and Insights for Generalization

The previous section revealed that the fundamental macroscopic properties still hold even in cases of significant information reduction like the representation scheme adopted in online traffic maps. Thus, in this section, an extension of the analysis made so far is made for the rest cities of the selected study area, starting with the city of Paris, Istanbul, and Moscow using raster images from Yandex maps (see Figure 4.14) and then expanding the application to the rest dataset. Following the proposed analytical framework, the macroscopic phenomena that can be observed from the class-type information broadcasted by online traffic maps will be based on the MFD estimation and macroscopic traffic flow interpretation that can be performed based on it.

5.4.1 Application based on traffic information from Yandex maps

Following the proposed methodological framework, the MFDs (shown in Figure 5.6) has been prepared for the city of Paris, Moscow and Istanbul based on the following assumptions regarding the application of the Eq. (5.2):

- Jam density K_j= 150 veh/km and average free-flow speed U_f= 80Km/h
- Representative values for each color code are U_g=55Km/h, U_y=30Km/h, U_r=10Km/h, U_{dr}=5Km/h, and
- S=0.04 according to Table 4.3

Compared to Google maps, Yandex maps used to provide numerical values of average speed per colored link with a click on the link on the map when accessed. As, so representative values were chosen based on these values provided from Yandex maps.

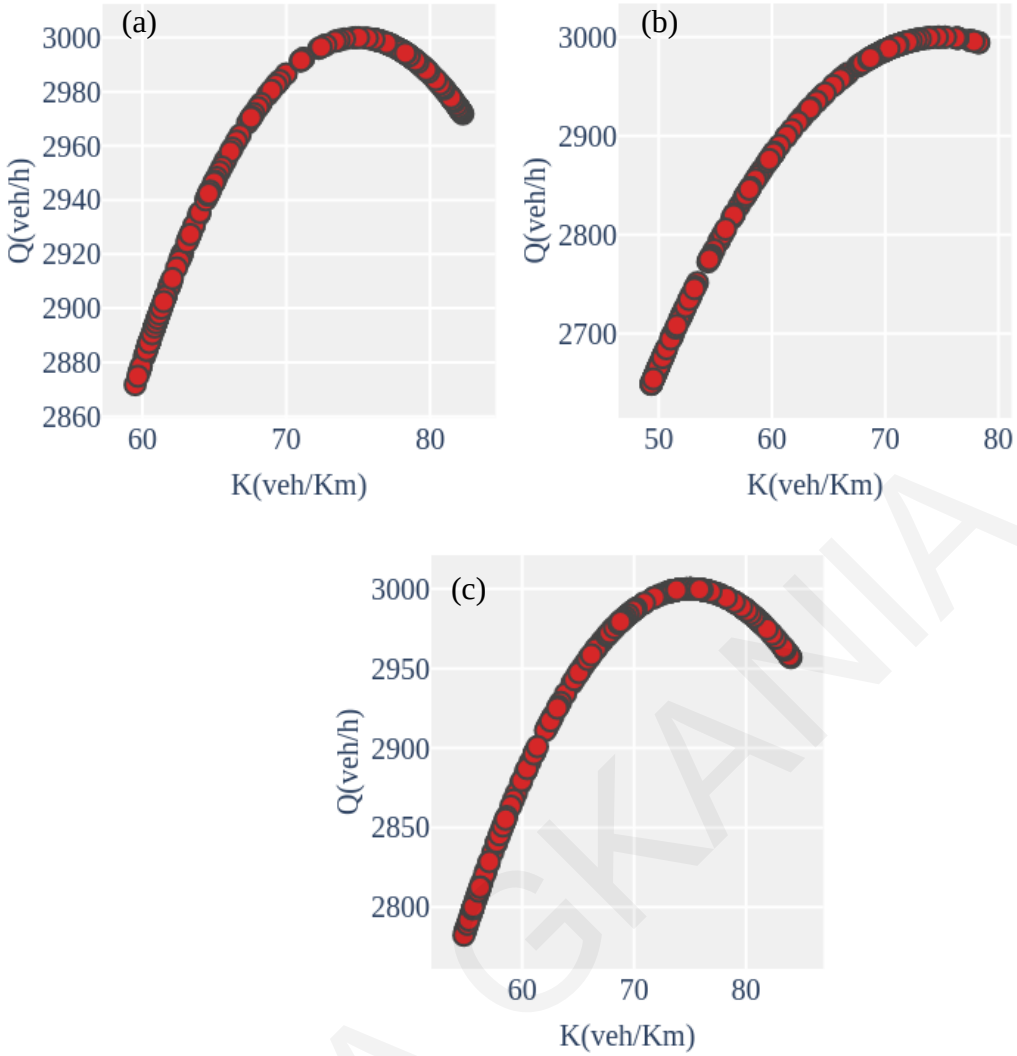


Figure 5.6: MFDs (average flow-average density diagram) for one typical weekday (5-minute intervals): (a) Paris; (b) Moscow; (c) Istanbul.

As it can be seen in Figure 5.6 the hourly weighted average capacity of the network for all the three cities is almost 3000 vehicles. It is also apparent that the capacity point at which the network operates at the maximum value is in the range of 73 to 75 vehicles per kilometer. The congested portion of the curve is more developed for the city of Paris and Istanbul (see Figure 5.6 (a, c).

At that point, it should be highlighted the importance of the assumptions used regarding the representative/average network values have in the form of the MFD. For gaining some insights on the sensitivity of the MFD with respect to these average values used, in Figure 5.7 two alternative MFDs have been estimated for the city of Moscow. In Figure 5.7 (a), the values for jam density is taken as K_j= 150 veh/km and for average free-flow as U_f= 80Km/h, while at Figure 5.7 (b) the respective values are K_j= 150 veh/km and U_f= 90Km/h.

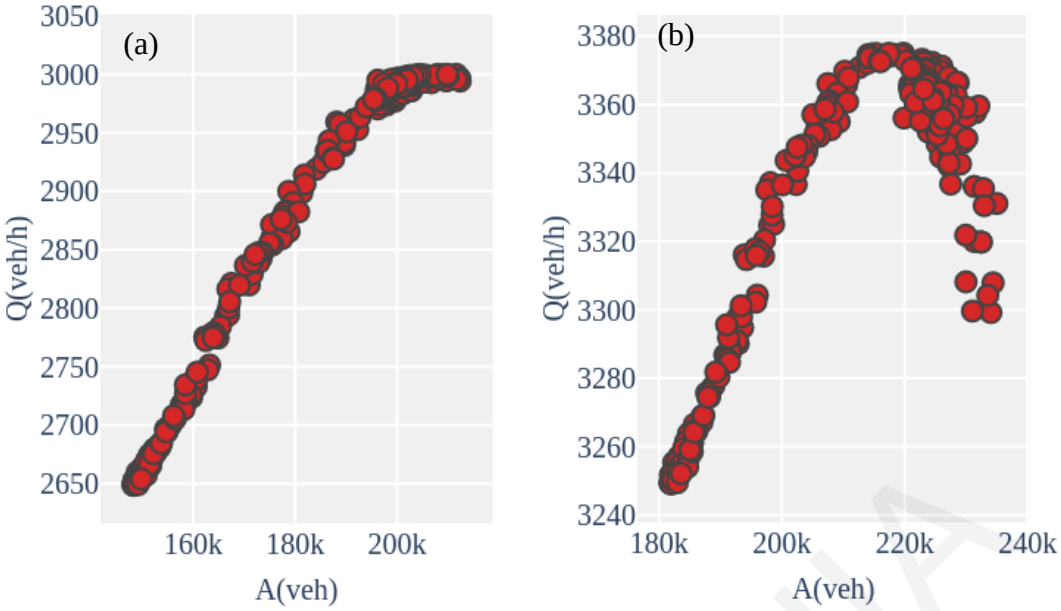


Figure 5.7: The relationship between average network flow and accumulation for (a), K_j= 150 veh/km and U_f= 80Km/h, and (b) K_j= 150 veh/km and U_f= 90Km/h (5-minute intervals) for the city of Moscow.

This slight modification of these gross values results to completely different shapes of the network MFD, an element exhibiting the sensitivity (and importance) of the MFD with respect to these network coefficients.

5.4.2 Network Operations and Performance – Phenomena of Traffic Hysteresis

In this section, we analyze the relationships between the total vehicles (accumulation) in the network and the average network flow and speed for the three cities (Figure 5.8 till Figure 5.10). A connection among selected data (*a*-*h*) on the diagrams and the network's traffic state for the same intervals is highlighted on the same figures, where it can be observed the different network operating conditions as these are reflected in typical color-coding schemes. Starting with the city of Paris (see Figure 5.8), points *(a)* and *(b)* correspond to low values of average network flow accumulation at 4:00 a.m. and 6:00 a.m., respectively, where the network's load is still rather empty. A steep incline occurs in the morning peak-hours (*c*-*d*) where many users have entered the network. Then at 8:35 a.m. (point e) the system seems to reach a maximum level of concentration (or maybe total network capacity) where average network traffic flow reaches 3,000 vehicles per hour and then phenomena of hysteresis loops seem to emerge (see Figure 5.8 (b)).

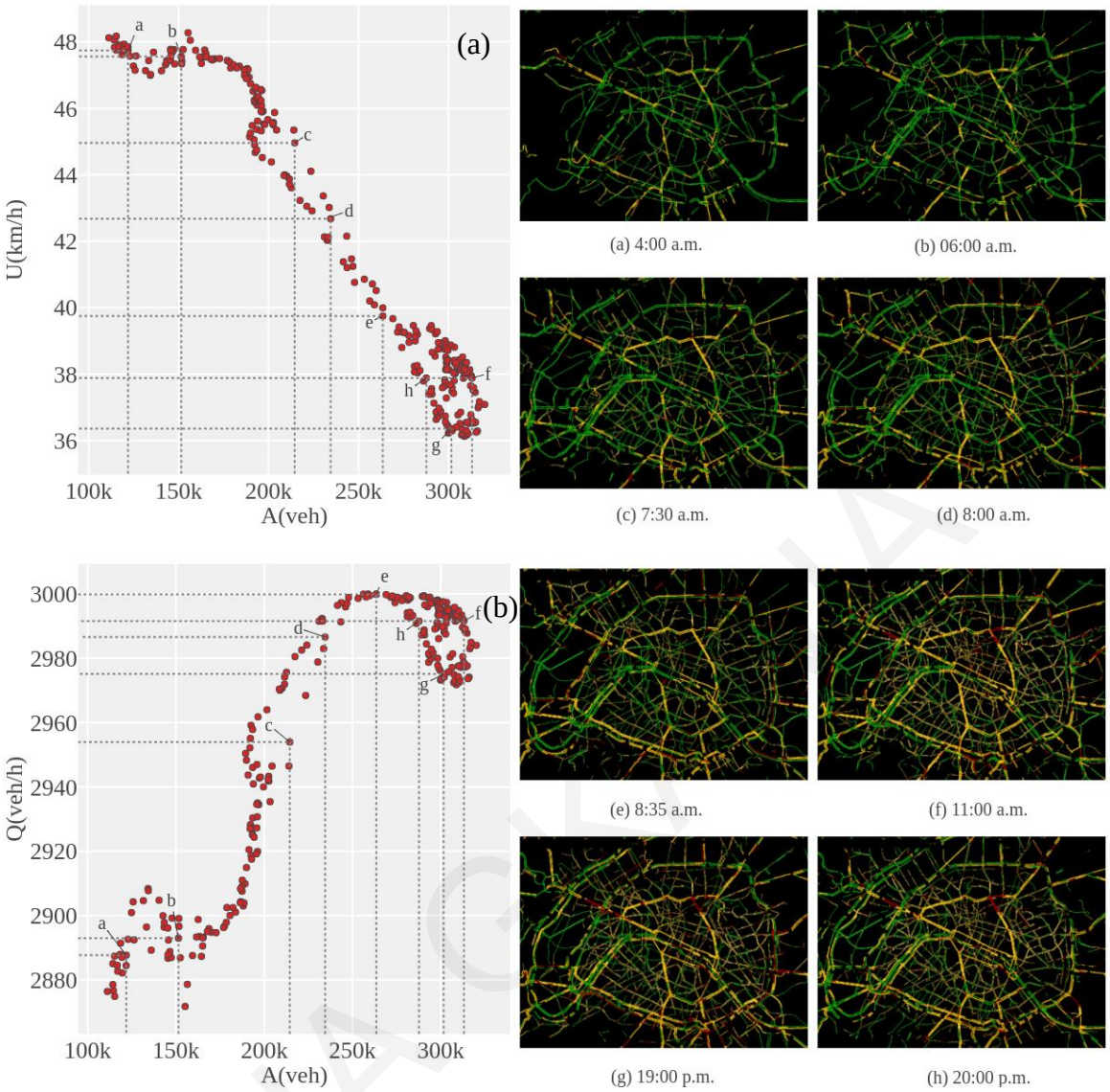


Figure 5.8: The relationship between average network flow (a), speed (b) versus the total number of vehicles, and network's traffic state for selected values of average accumulation during a weekday (5-minute intervals) for the city of Paris.

Moving to the city of Moscow (see Figure 5.9), a different shape of MFD appears while phenomena of hysteresis loops does not seem to emerge for this type of network. Here the system seems to reach a maximum level of concentration (or maybe total network capacity) where average network traffic flow reaches 3,000 vehicles per hour during the evening peak-hour (point g). That is also visible through the map image at the same interval where several links of the network are colored red that indicates congestion (see Figure 5.9 (b)).

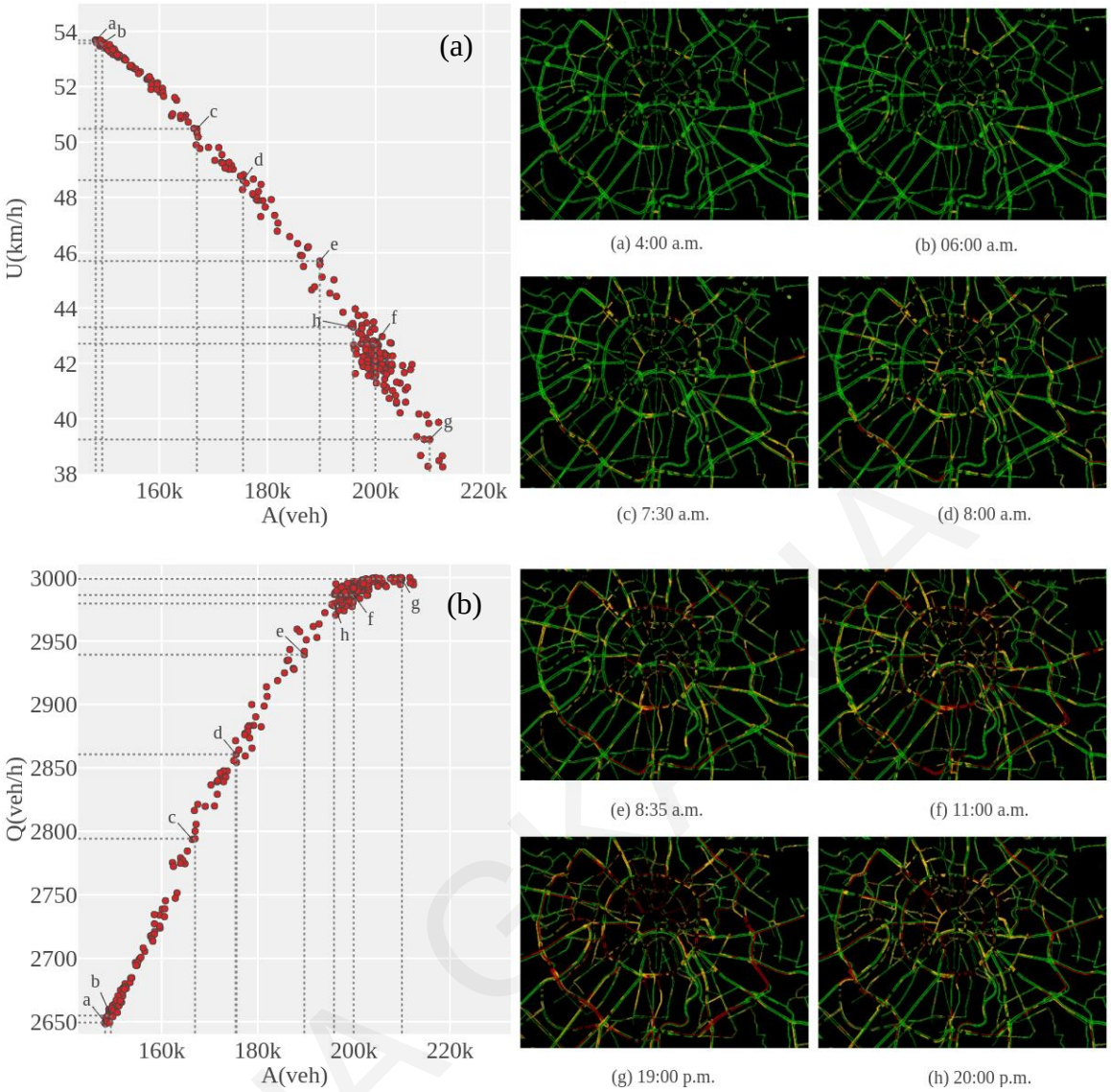


Figure 5.9: The relationship between average network flow (a), speed (b) versus the total number of vehicles, and network's traffic state for selected values of average accumulation during a weekday (5-minute intervals) for the city of Moscow.

In the next figure we see the results for the city of Istanbul. Again, point (b) corresponds to low values of average network flow accumulation at 6:00 a.m., respectively, where the network's load is still rather empty. A steep incline occurs in the morning peak-hours (c-d) where many users have entered the network. Then at 11:00 a.m. the system seems to reach a maximum level of concentration (or maybe total network capacity) where average network traffic flow reaches 3,000 vehicles per hour and then phenomena of hysteresis loops seem to emerge for the city of Istanbul, too.

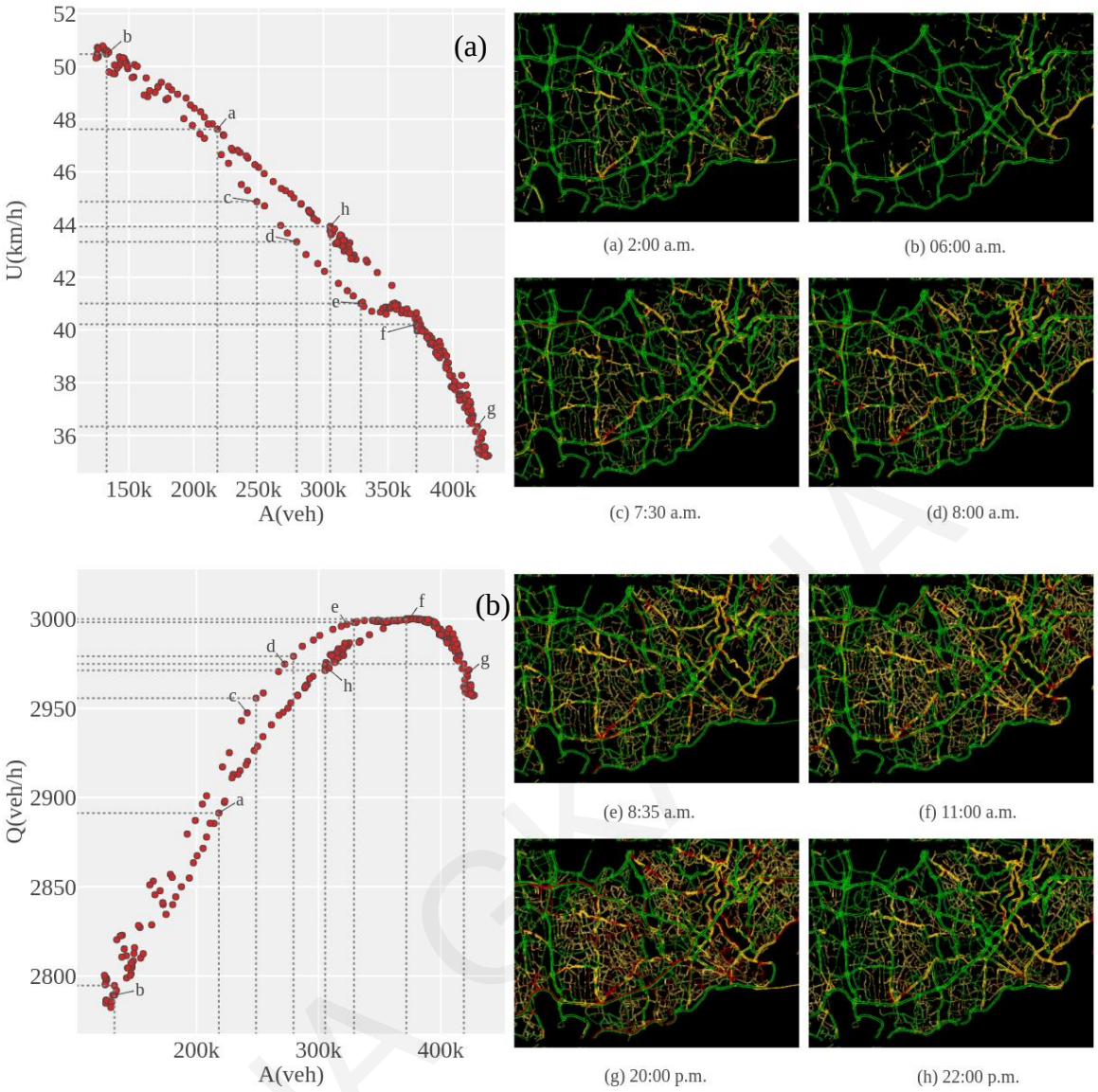


Figure 5.10: The relationship between average network flow (a), speed (b) versus the total number of vehicles, and network's traffic state for selected values of average accumulation during a weekday (5-minute intervals) for the city of Istanbul.

Focusing on the hysteresis phenomena observed in Figure 5.8, and Figure 5.10, it would be valuable to further investigate the temporal dimension of this phenomenon. In the case of Paris two distinctive hysteresis patterns can be observed. The first is an anti-clockwise between 8:35 a.m. and 13:00 p.m., where vehicles accumulation increases, but average flow drops at first and then improves gradually from 10:00 until 13:00 (see Figure 5.11(a)). This type of network recovery may reflect the response of the network users and operators towards a dynamic equilibrium (Geroliminis and Daganzo, 2008) that optimize network performance as this is related to the average traffic flow for the same level of vehicles accumulation in the network. The dominant phenomenon here is that network is improving its performance by improving average network flow. Then, a different type of

network hysteresis phenomenon has been captured, between 16:00 and 20:30. In this instance, average traffic flow deteriorates from 16:00 until 18:30 and recovers again at 20:30 p.m. (Figure 5.11 (b)), but this time in a clockwise manner. This means that the dominant phenomenon in this period is that the vehicles accumulation in the network is reduced (vehicles are exiting the network, which is reasonable for this time of the day), improving its performance. These macroscopic phenomena, although delicate and difficult to observe can be captured by online traffic maps, providing very important information valuable for managing large-scale urban road systems.

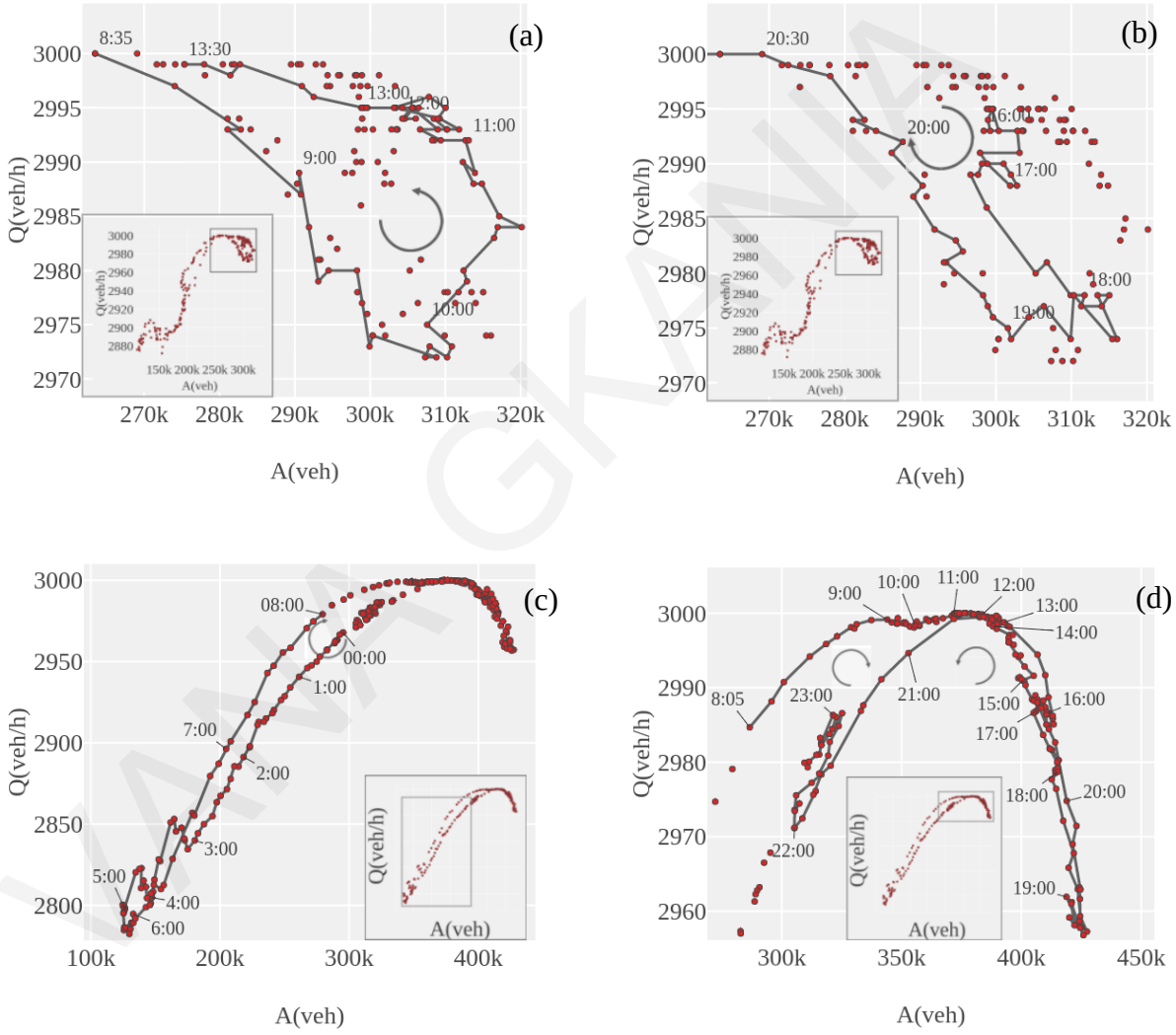


Figure 5.11: Average traffic flow vs. the total number of vehicles diagram during a weekday (5-minute intervals) for Paris: (a) anti-clockwise hysteresis; (b) clockwise hysteresis and for Istanbul; (c) clockwise hysteresis; (d) anti-clockwise hysteresis.

On the other hand, for the case of Istanbul (see Figure 5.11 (c), (d)) the shape of the estimated MFD suggests different network operating characteristics. As it can be observed by Figure 5.11 (c), in the hyper-congested (right) part of the MFD no hysteresis phenomenon occurs, which may reflect the rapid, severe and extensive (in time and space) network deterioration typically occurring in Istanbul road network. Though, the network loading and unloading process in Istanbul's network occurs in different manner; the loading phase (starting from empty network conditions) occurs rapidly and with better average network flow (upper branch of MFD's left side), while the traffic dissipation occurs with a slightly reduced network performance (lower average traffic flow) since it starts from congested network conditions (lower branch MFD's left side). Such characteristics require a more thorough investigation, since may provide important information to traffic management authorities.

5.4.3 Application based on traffic information from Google, Bing, and Here maps

In this section, the proposed methodological framework, is applied to the rest cities of the study area. Furthermore, the MFDs are estimated based on traffic information from Google, Bing and Here maps, for comparative reasons. The assumptions regarding the application of the Eq. (5.2) follow:

- Jam density K_j= 150 veh/km and average free-flow speed U_f= 80Km/h
- Representative values for each colour code are:

  U_g=55Km/h, U_o=30Km/h, U_r=10Km/h, U_{dr}=5Km/h, for Google, and Baidu, U_g=55Km/h, U_y=30Km/h, U_o=10Km/h, U_r=5Km/h for Bing, U_g=55Km/h, U_y=30Km/h, U_r=10Km/h for Here maps, and
- S according to Table 4.3

The representative values for each colour code was selected based on representative values of average speed that are provided from Yandex and Here maps for each traffic layer when accessed. A calibration process as described in Section 5.2 is proposed for accurate results for the exact shape of MFD. Although, by using logical assumptions, a clear view of the shape of the MFDs for the selected cities is feasible. As it can be seen from Figure 5.12 till Figure 5.14, that provides results for the city of Paris, Moscow and Istanbul, the hourly weighted average capacity of the network for all the cities is around 3000 vehicles. It is also apparent that the capacity point at which the network operates at the maximum value is in the range of 70 to 80 vehicles per kilometer.

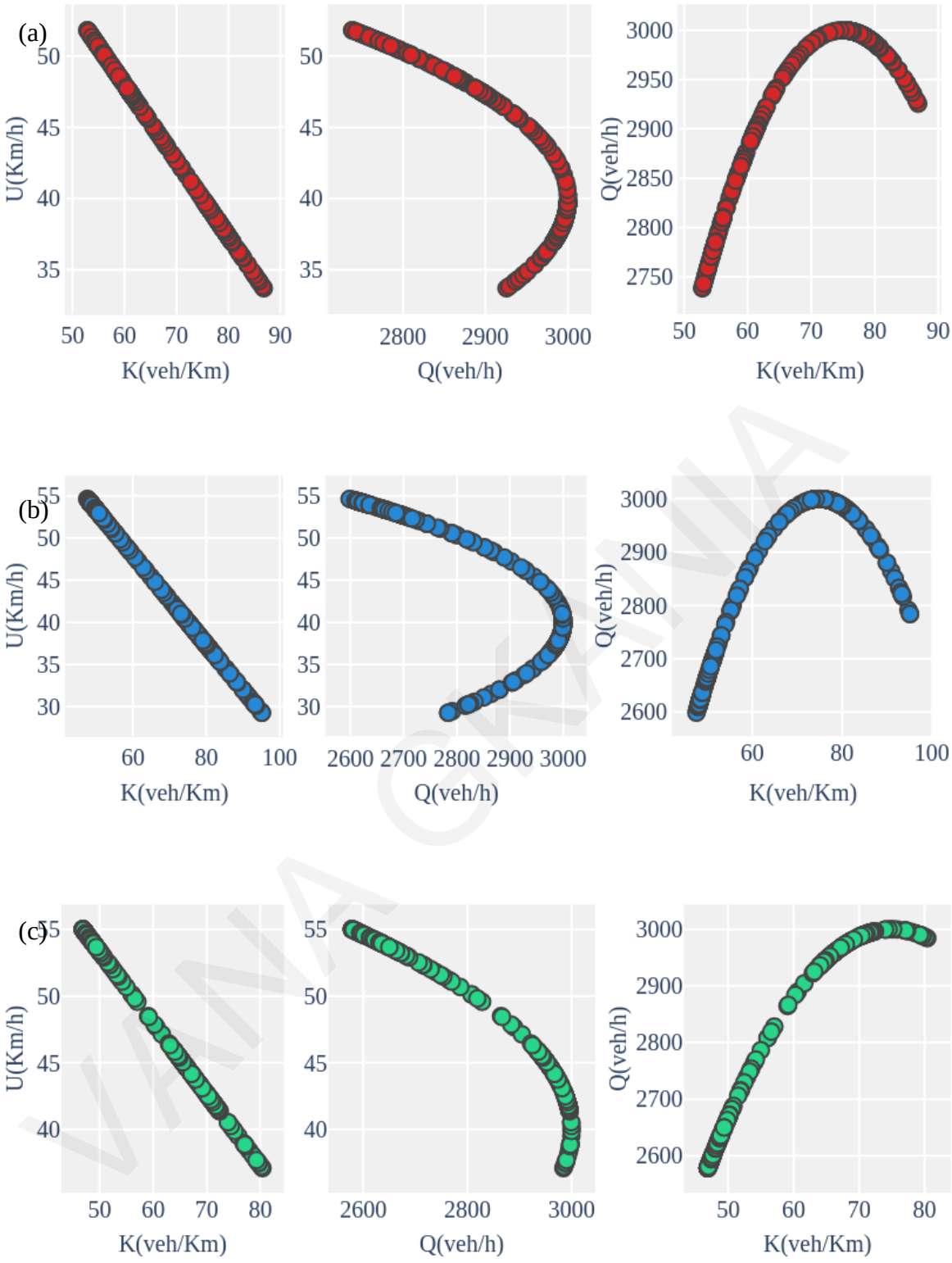


Figure 5.12: Paris's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

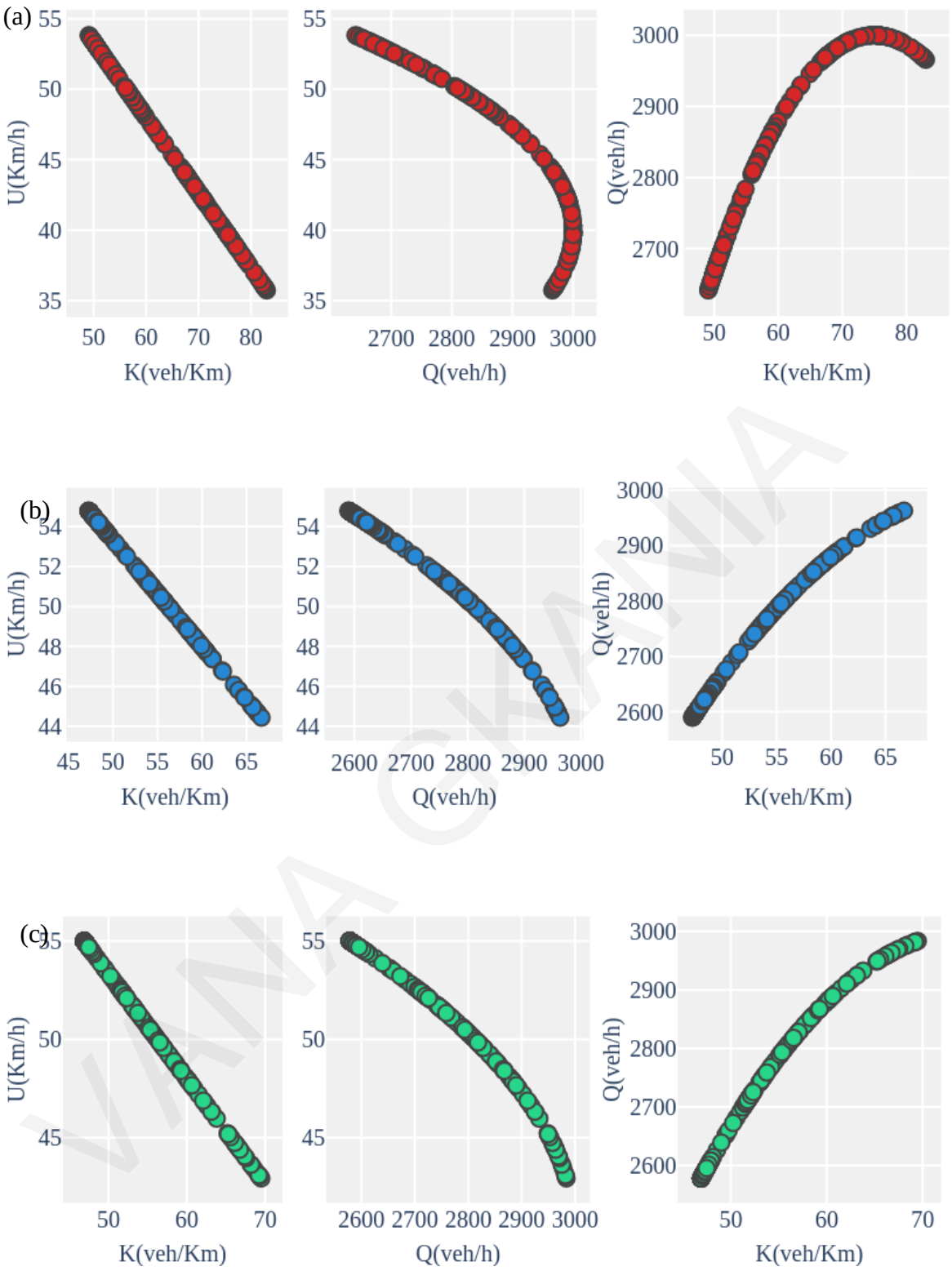


Figure 5.13: Istanbul's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

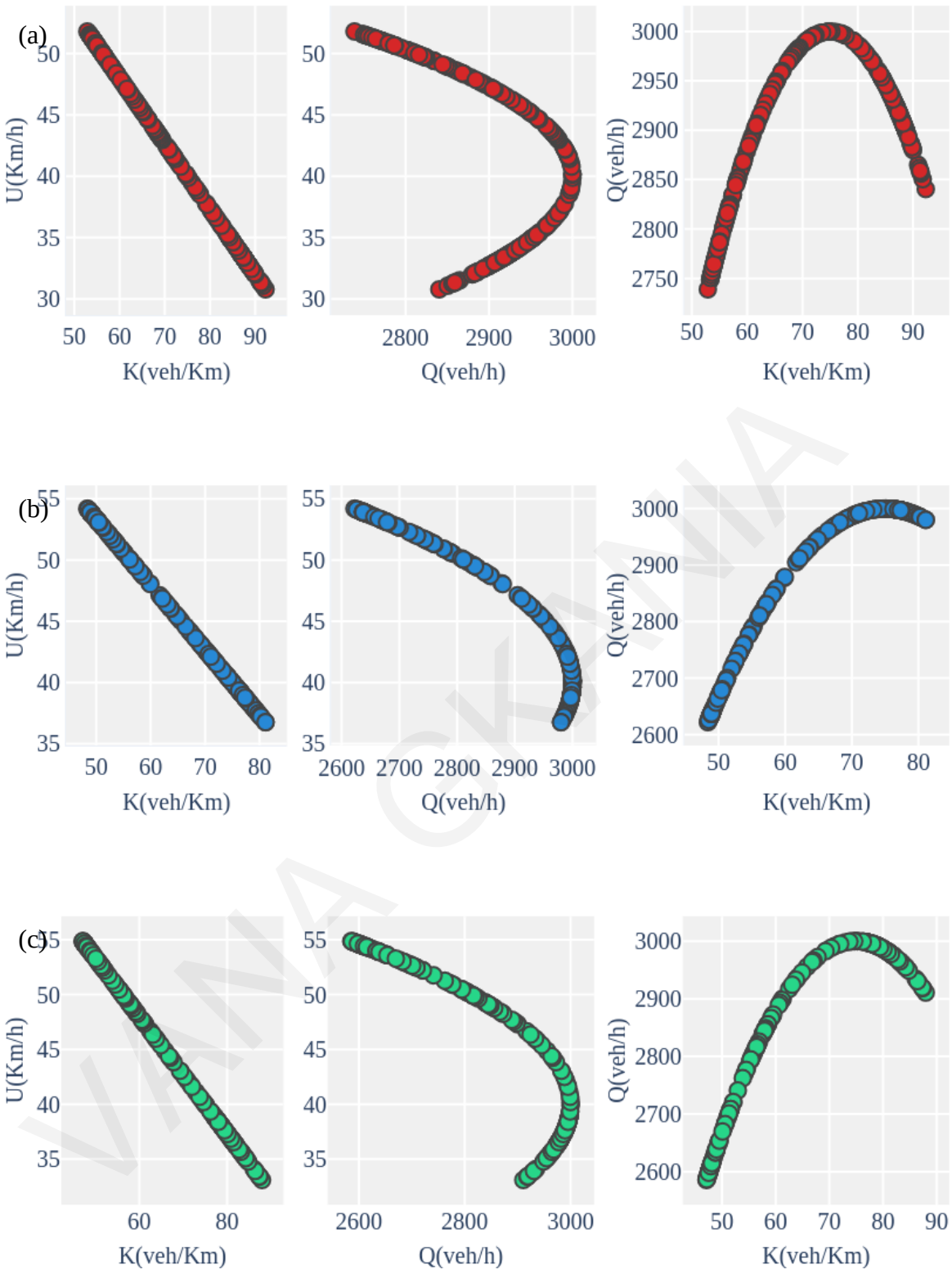


Figure 5.14: Moscow's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

In the previous sections, the MFDs for the city of Nicosia based on traffic information from Google maps and the MFDs for the city of Paris, Moscow, and Istanbul from Yandex

maps were presented. For consistency reasons, in Figures 5.12-5.14 the same cities were selected to illustrate the MFDs from the rest online traffic providers such as Google (red), Bing (blue) and Here maps (green). Results for the rest cities of the study area can be found in Appendix D.

Regarding the columns in these figures, the left one refers to the average speed-average density diagram, the middle column refers to the average speed-average flow diagram and last the right column corresponds to the average flow-average density diagram. The average speed-average-density diagram reflects the linear relationship between speed and density, as the Greenshields model was utilized in Eq. (5.4). The maximum value of average density ranges between 55-100 veh/Km. Note that, Paris and New Delhi presented the highest values almost 100 veh/Km while Beijing and Berlin the smallest less than 65 veh/Km. Regarding the average speed values, the range is between 25-55 Km/h for most of the cities. This range is strictly related to the assumptions of average speed for each traffic layer (e.g. 55Km/h for the green traffic layer). For instance, during the late-night hours where the green colour dominates in the online traffic maps, the average speed for the whole network is equal to the value of the average speed that corresponds to the green traffic layer.

Moving to the average speed-average flow diagram and to the average flow-average density diagram, a parabolic and a semi-parabolic curve is observed for most of the cities. Although, the congestion part of the curve in some cases is more developed for the same city depending on the online traffic provider. This can be observed for the city of Paris where the congestion part is more developed on the diagrams for Bing maps compare to Google and Here maps. The same applies for the city of New Delhi and Sao Paulo. On the contrary for the city of Moscow, Istanbul, New York, Los Angeles and Riyadh the congestion part of the curve is more developed in the red diagrams estimated based on data from Google maps. For the city of Sydney and Johannesburg, similar curves developed regardless the type of traffic maps used for estimation. For the city of Beijing the pattern of the diagrams is quite different form the rest of the cities, as the network remains uncongested and does not reach the hourly weighted average capacity (at that point the congested part of the curve starts to shape).

As a next step, the relationships between the total vehicles (accumulation) in the network and the average network flow for the rest cities were investigated in Figure 5.15 till Figure 5.20. Each raw on these figures refers to a different city of the dataset, while the left column corresponds to data from Google maps (red colour), the middle one to Bing maps

(blue colour) and last the right column to Here maps(green colour). The same colors per provider were used for consistency reasons. In the last Figure 5.20, the left column refers to the city of Tokyo using data form Google maps (red colour), while the rest two corresponds to the city of Beijing utilizing data from Bing (blue colour) and Baidu maps (green colour), respectively. Interestingly, mainly three different shapes can be observed from these figures, a parabolic one (e.g. Figure 5.16b(left)), a semi-parabolic (e.g. Figure 5.16a(right)) and an S-shape (e.g. Figure 5.16c (middle). Note that, similar patterns occurred for the city of Paris, Moscow, and Istanbul from Yandex maps (see Figure 5.8b, Figure 5.9b, and Figure 5.10b).

Furthermore, a clear view, of the network's operating conditions for the selected cities is provided from these figures. Starting with the city of Paris (see Figure 5.15a), low values of average network flow accumulation are observed when the network's load is still rather empty. A steep incline occurs where many users have entered the network. Then the system seems to reach a maximum level of concentration (or maybe total network capacity) where average network traffic flow reaches 3,000 vehicles per hour and then phenomena of hysteresis loops seem to emerge. For the same city (Figure 5.15a), but based on Bing maps(middle column) the maximum level of accumulation is quite lower (equal to 35000 veh) compare to 50000 veh from Google maps and even lower from Here maps (right column) that is equal to 25000veh. The same applies for most of the cities as the maximum level of accumulation from Here maps is quite lower compared to the other two providers. This can be explained due to the low urban coverage from Here maps according to Table 4.4.

Regarding the phenomena of traffic hysteresis, these seem to emerge for most of the cities based on information retrieved from Google maps (left column: Figure 5.15(a ,b, c), Figure 5.16(b, c), Figure 5.17(a ,b), Figure 5.18(a ,b, c), Figure 5.19(a ,b, c)). For the same cities but from Here maps a semi parabolic shape is observed without traffic hysteresis loops. Interestingly, that does not apply for the city of Sao Paulo that appears an S-shape curve (Figure 5.17b(right)). The S-shape curve is frequently observed for most of the cities based on estimation from Bing maps. This can be explained due to the urban coverage that Bing maps provide during the late night hours where the network is almost empty but still many links remained colored with the green traffic layer.

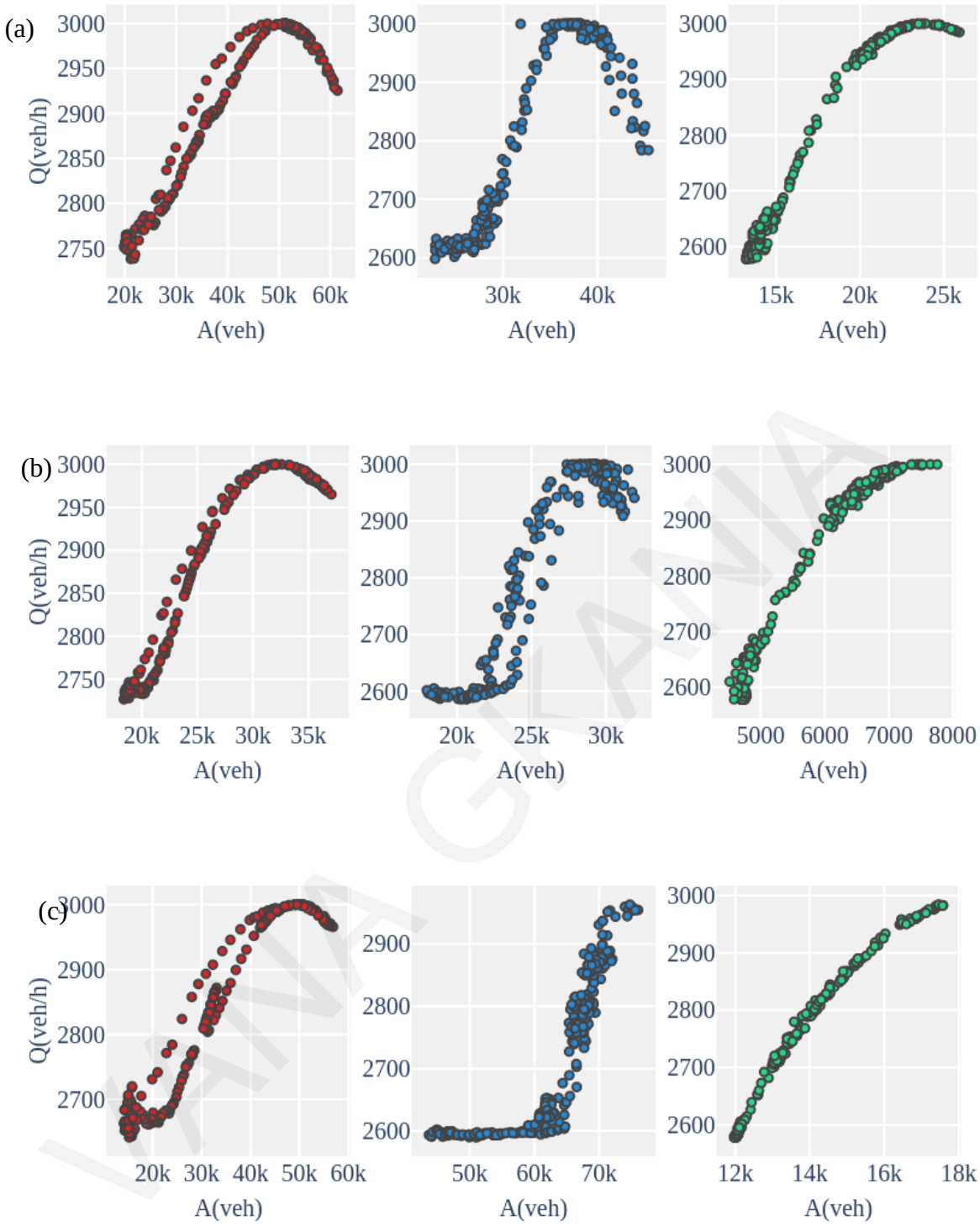


Figure 5.15: Average network flow versus accumulation(total number of vehicles) for the city of (a) Paris; (b) London; (c) Istanbul (5-minute intervals): (left column) from Google; (middle column) from Bing; and (right column) from Here maps.

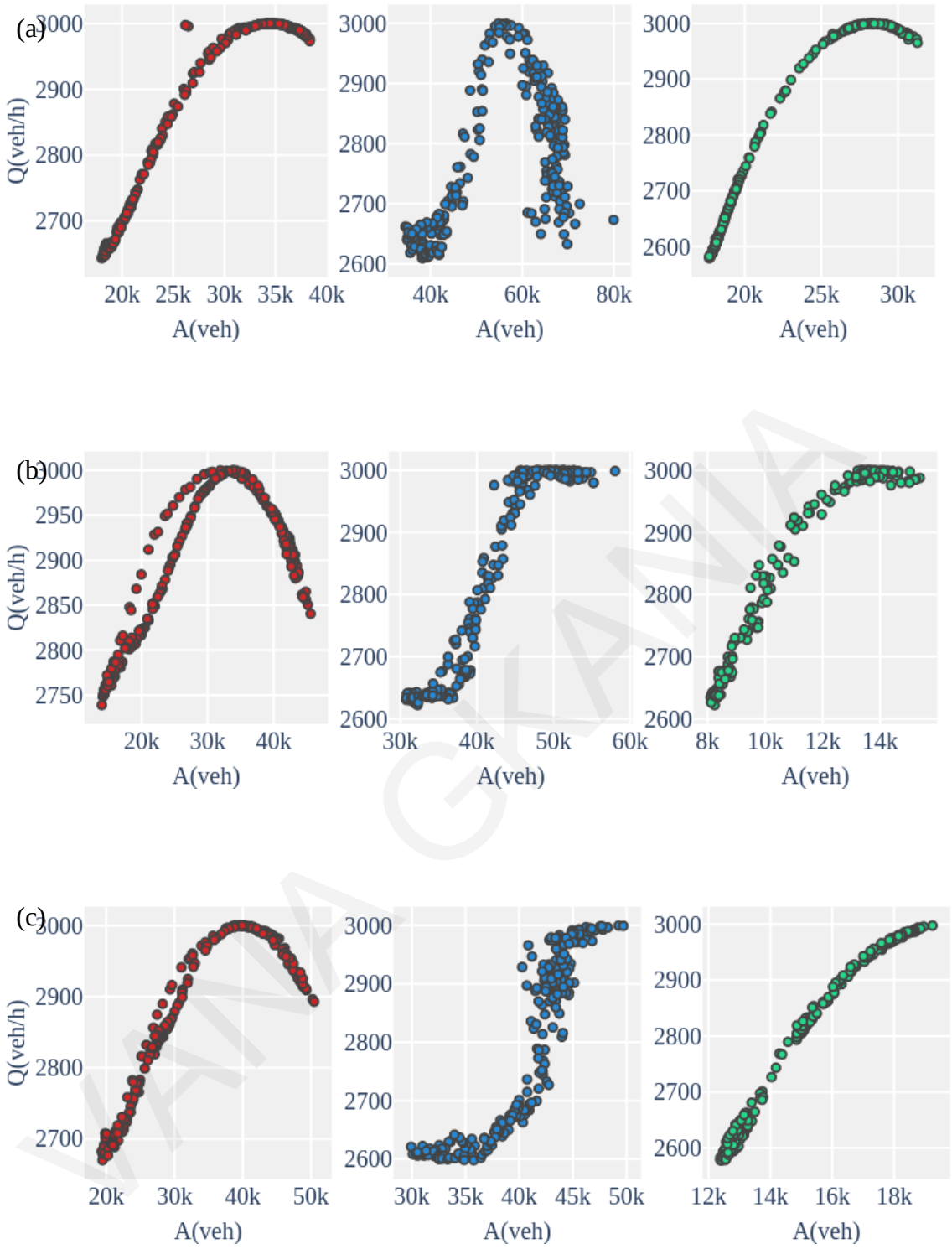


Figure 5.16: Average network flow versus accumulation(total number of vehicles) for the city of (a) New Delhi; (b) Moscow; (c) New York (5-minute intervals): (left column) from Google; (middle column) from Bing; and (right column) from Here maps.

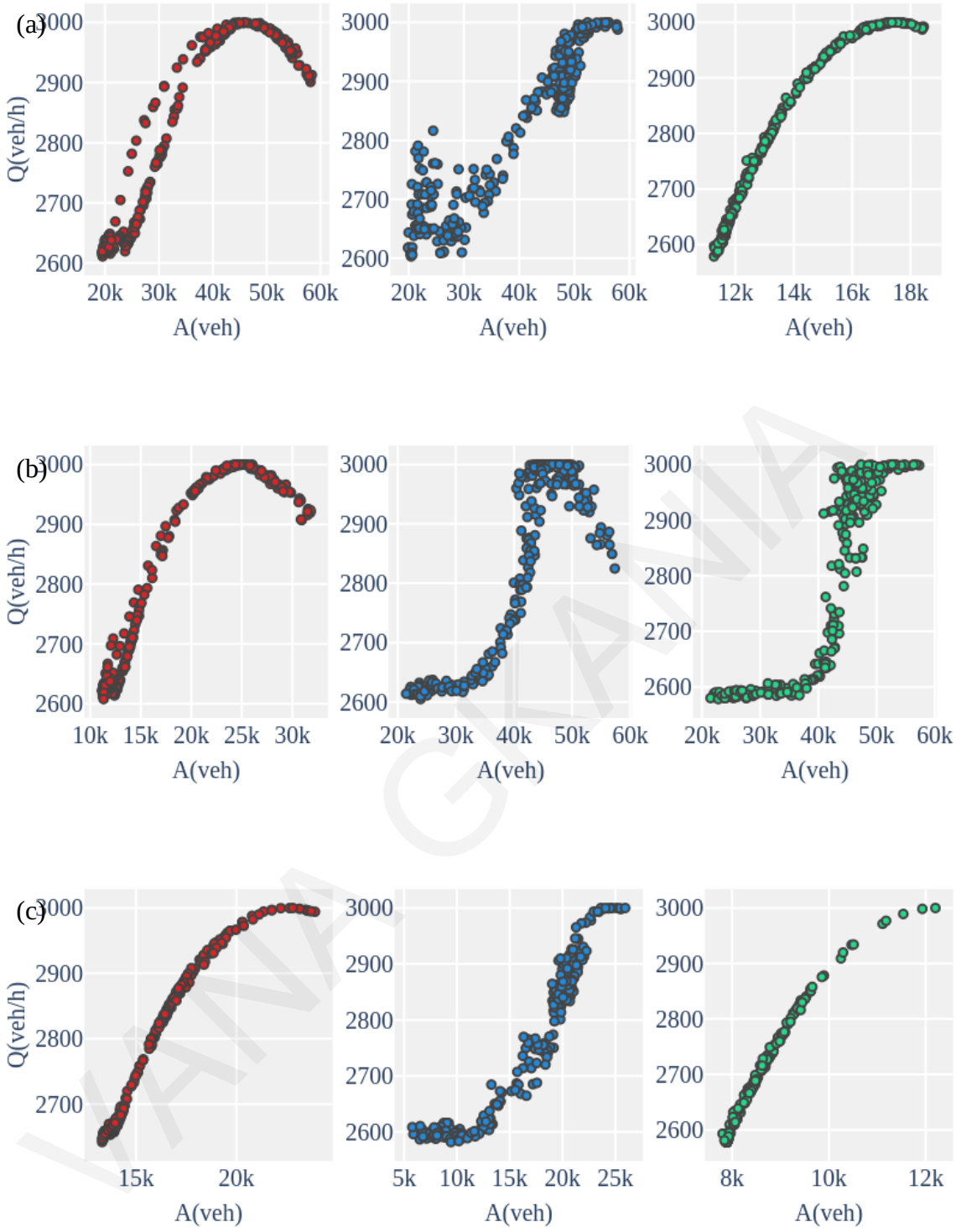


Figure 5.17: Average network flow versus accumulation(total number of vehicles) for the city of Los Angeles; (b) Sao Paulo; (c) Singapore (5-minute intervals): (left column) from Google; (middle column) from Bing; and (right column) from Here maps.

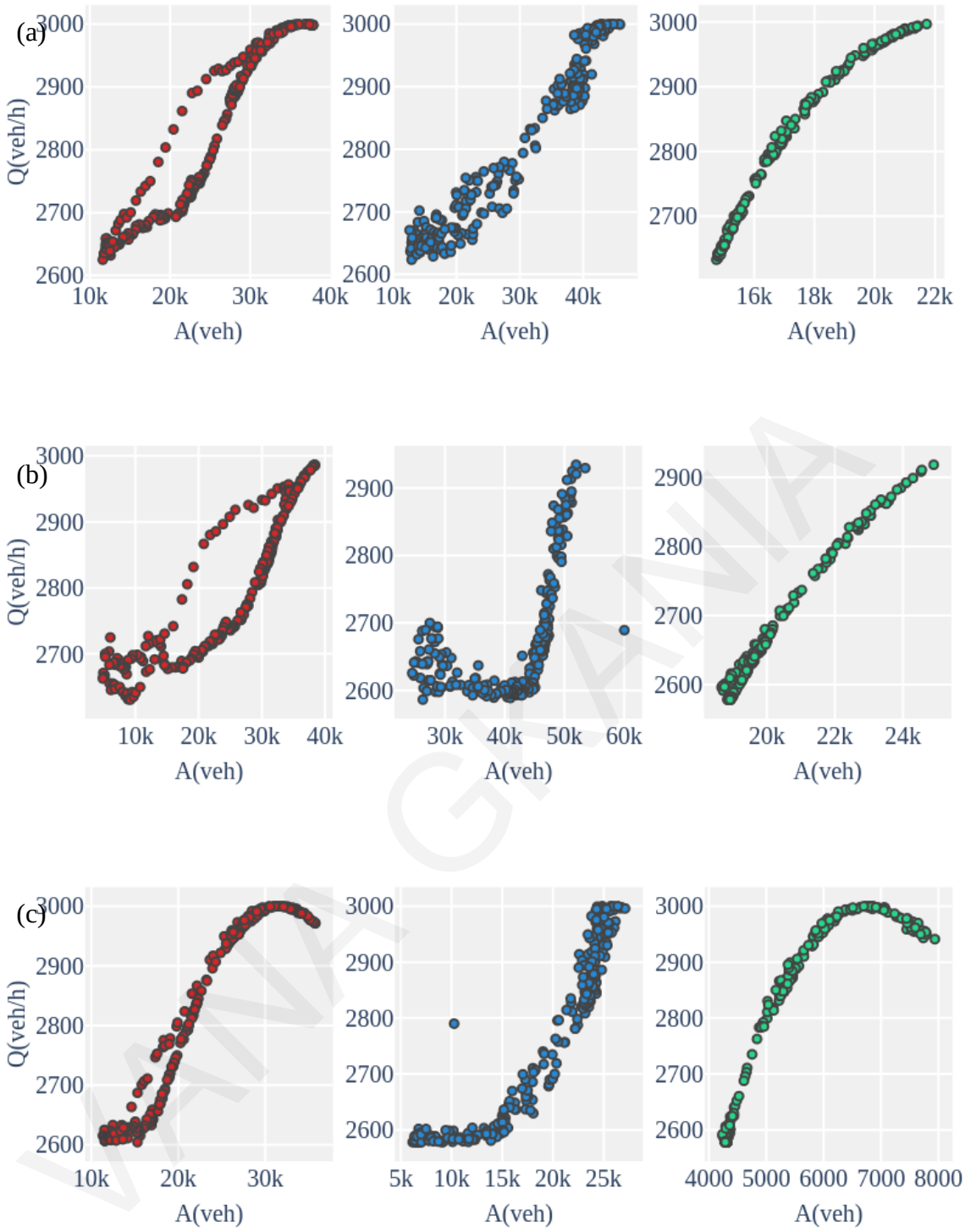


Figure 5.18: Average network flow versus accumulation(total number of vehicles) for the city of (a) Sydney; (b) Johannesburg; (c) Toronto (5-minute intervals): (left column) from Google; (middle column) from Bing; and (right column) from Here maps.

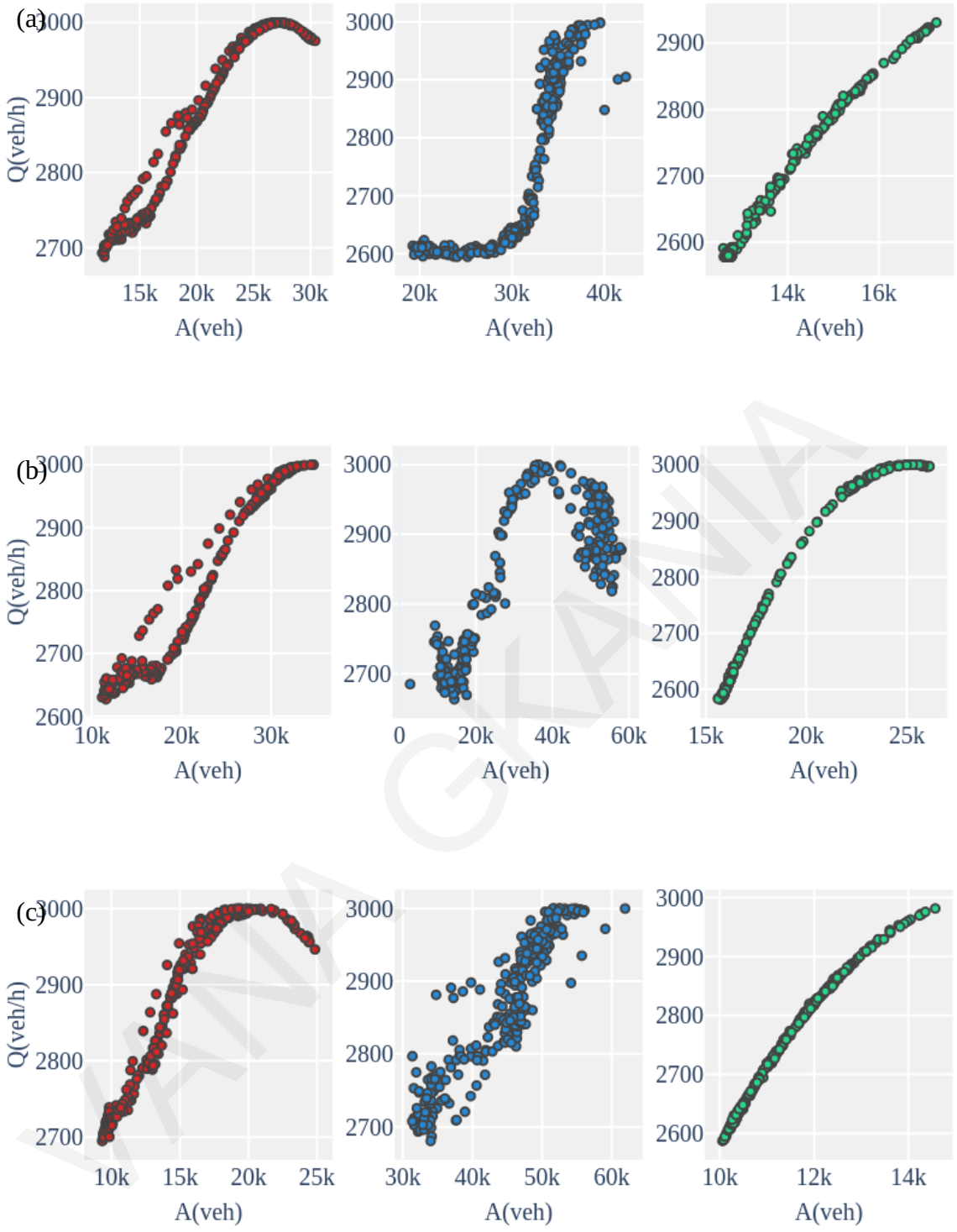


Figure 5.19: Average network flow versus accumulation(total number of vehicles) for the city of (a) Berlin; (b) Buenos Aires; (c) Riyadh (5-minute intervals): (left column) from Google; (middle column) from Bing; and (right column) from Here maps.

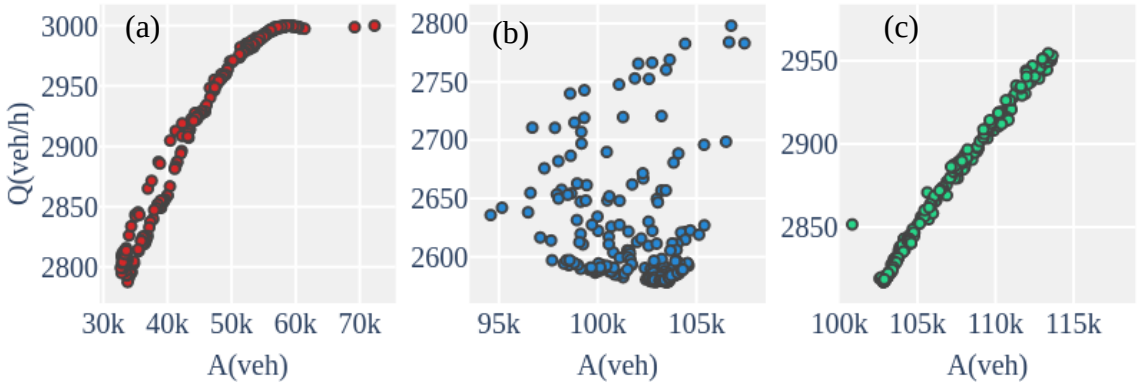


Figure 5.20: Average network flow versus accumulation(total number of vehicles) for the city of (a) Tokyo from Google; (b) Beijing from Bing and (c) Beijing from Baidu maps.

Last, in Figure 5.20 for the city of Tokyo a similar pattern with other cities is observed, where average network traffic flow reaches 3,000 vehicles/hour when the maximum level of concentration is equal to 60000veh in the network. On the contrary, for the city of Beijing in the same figure, a completely different pattern is observed from Bing and Baidu maps between average network flow and accumulation.

5.5 MFD estimation based on the Greenberg Model

In this section an alternative model is used to estimate the MFDs based on traffic maps information. Specifically, the same Eq. (5.2) is used to estimate the network average speed U^t at time interval t, while a nonlinear speed-density relationship based on Greenberg model (see Eq. (3.18)) is used to estimate the average network density K^t as:

$$K^t = K_{jam} \exp\left(-\frac{U^t}{U_0}\right) \qquad (5.7)$$

where U_o stands for the optimum average speed, that is equal to the half of the free speed (see Eq. (3.15)) and K_{jam} is the jam density of the network.

The equation of average flow-density relationship can be derived based on the fundamental traffic relationship, see Eq. (3.22) as:

$$Q^t = K^t U^t \qquad (5.8)$$

The total number of vehicles (accumulation) on the network A^t at time interval t, can be estimated by multiplying the average network density with the total length of the links in the network using Eq. (5.6).

In the next figures, the MFDs for the three cities (Paris, Moscow, and Istanbul) from Yandex maps are estimated based on a linear speed-density relationship (Greenshields

model) and on a logarithmic one (Greenberg model) for comparative reasons. One main difference between the two models is that the first requires knowledge of the free-flow speed and jam density parameters while the second requires the optimum speed and jam density parameters. The assumptions regarding the estimation of MFDs in this section follow:

- Jam density K_j= 150 veh/km, average free-flow speed U_f= 80Km/h, and optimum average speed U_o= 40Km/h
- Representative values for each colour code are:

  U_g=55Km/h, U_o=30Km/h, U_r=10Km/h, U_{dr}=5Km/h, for Google, and Baidu, U_g=55Km/h, U_y=30Km/h, U_o=10Km/h, U_r=5Km/h for Bing and Yandex, U_g=55Km/h, U_y=30Km/h, U_r=10Km/h for Here maps, and
- S according to Table 4.3

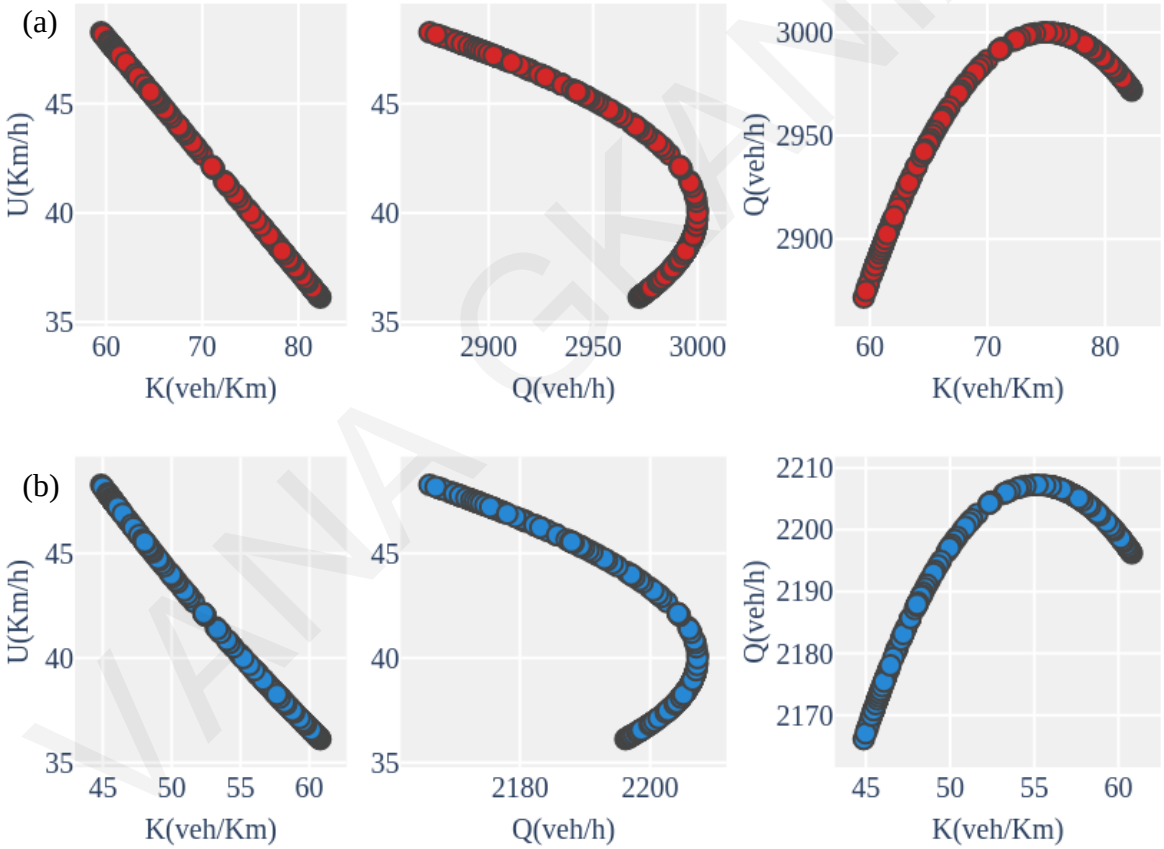


Figure 5.21: Paris's MFDs, for a typical weekday, (5-minute intervals) from Yandex maps based on (a) Greenshields model and (b) Greenberg model; (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

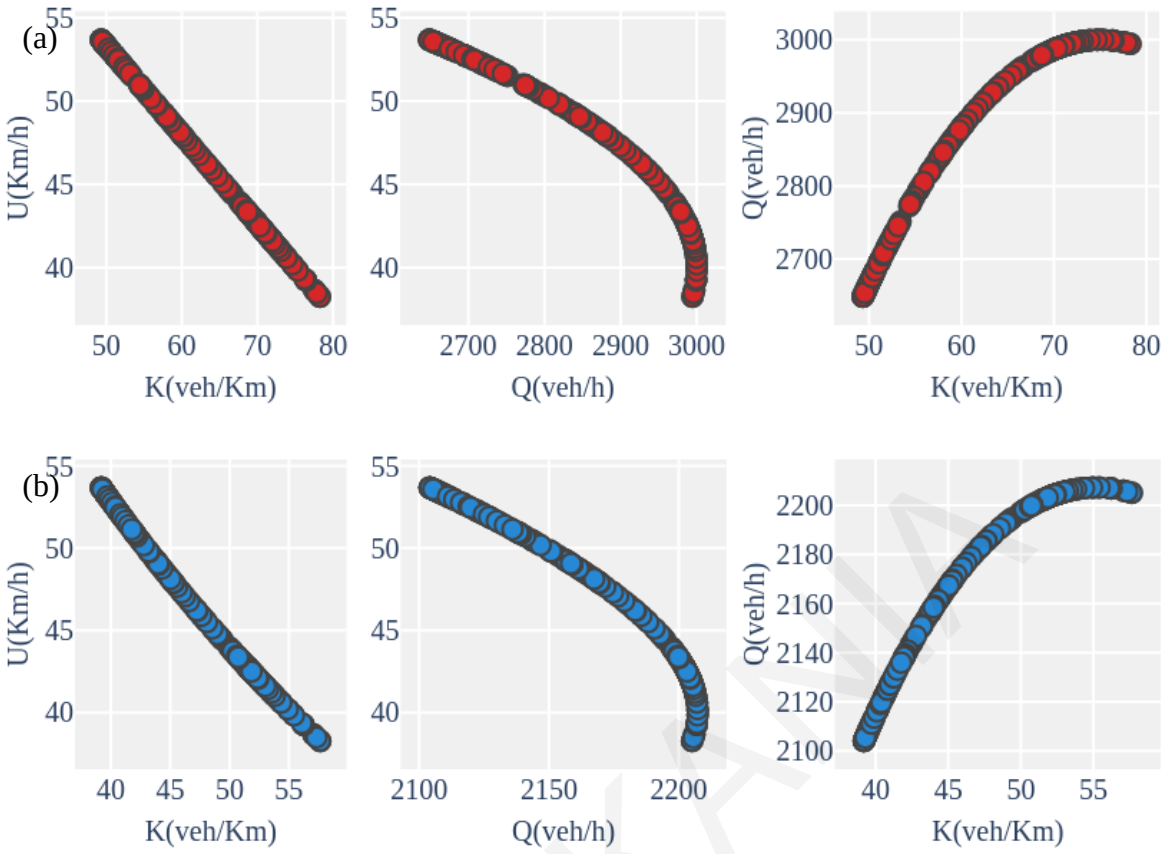


Figure 5.22: Moscow's MFDs, for a typical weekday, (5-minute intervals) from Yandex maps based on (a) Greenshields model and (b) Greenberg model; (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

Starting with the MFDs for the city of Paris (see Figure 5.21) similar patterns are observed for the average speed-average flow diagram and the average flow-average density diagram. Although, higher values of average flow and average density are observed based on the Greenshields model (Figure 5.21a). This can be explained due to the optimum speed parameter used that is equal to the half of the free speed. The same applies for the rest two cities (Figure 5.22, Figure 5.23). Regarding the average speed-average density diagrams based on Greenberg model, a non-linear pattern is observed as expected.

The last two figures of the section illustrate the relationship between average network flow versus accumulation based on the two models. Obviously, Figure 5.24a, Figure 5.24b, and Figure 5.24c are the same as Figure 5.8b, Figure 5.9b, Figure 5.10b, respectively. Concerning the same diagrams estimated based on the Greenberg model (Figure 5.25) the shame curves (S-shape and semi-parabolic) appeared but with significantly lower values of accumulation.

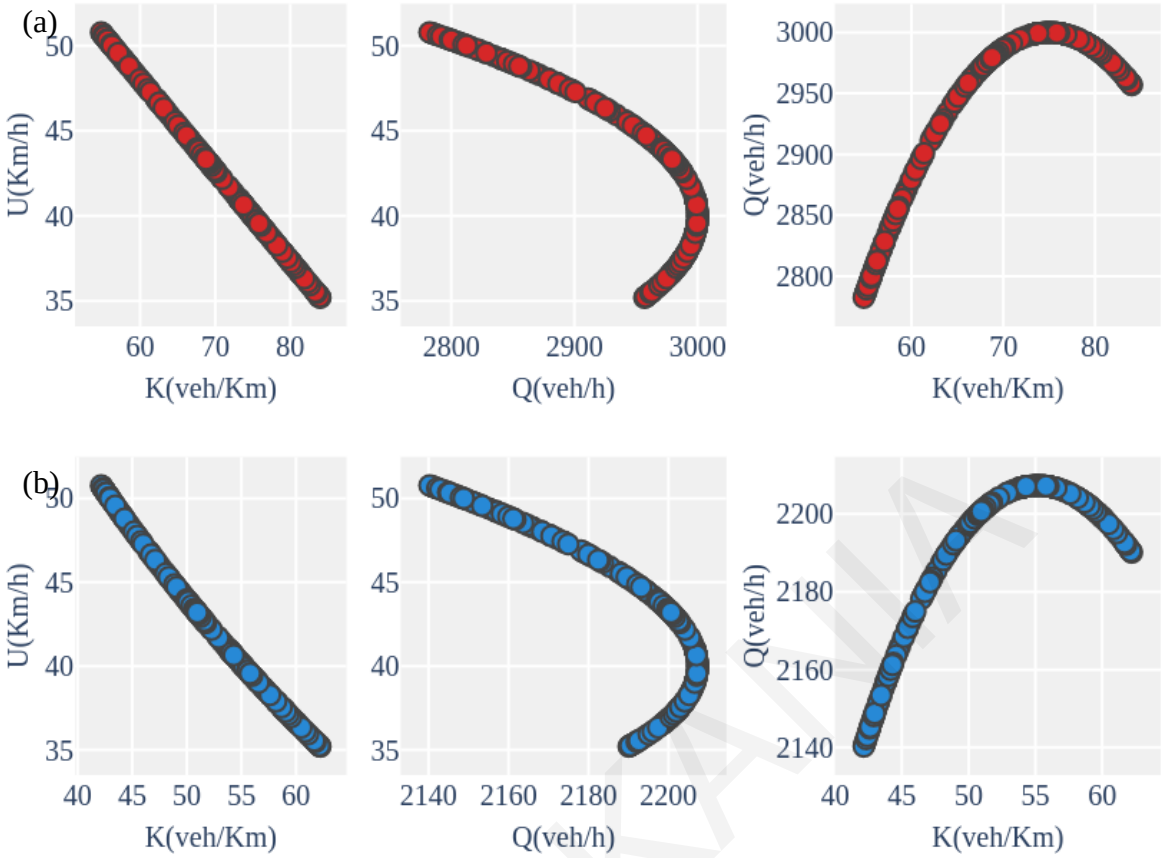


Figure 5.23: Istanbul's MFDs, for a typical weekday, (5-minute intervals) from Yandex maps based on (a) Greenshields model and (b) Greenberg model; (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

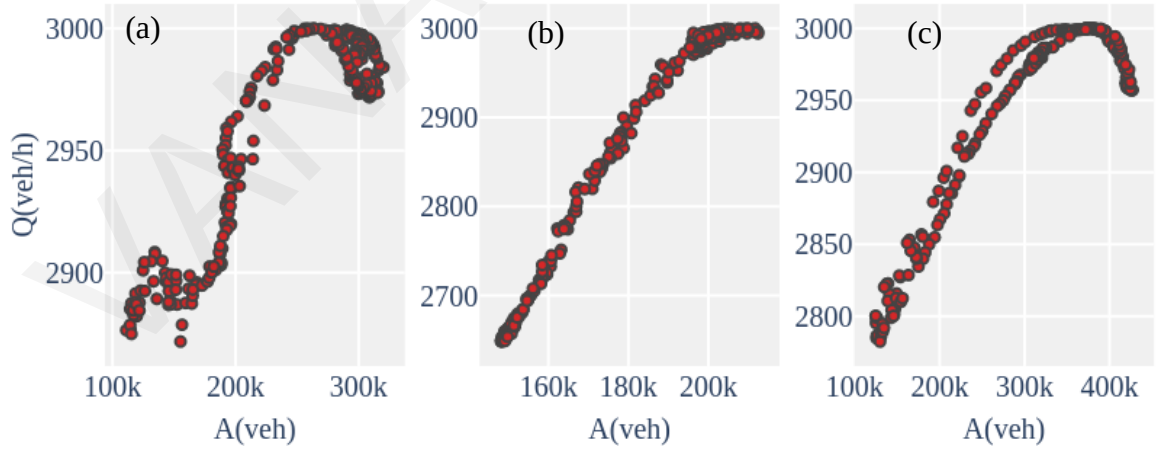


Figure 5.24: Average network flow versus accumulation(total number of vehicles) for the city of (a) Paris; (b) Moscow and (c) Istanbul (5-minute intervals) from Yandex maps based on Greenshields model.

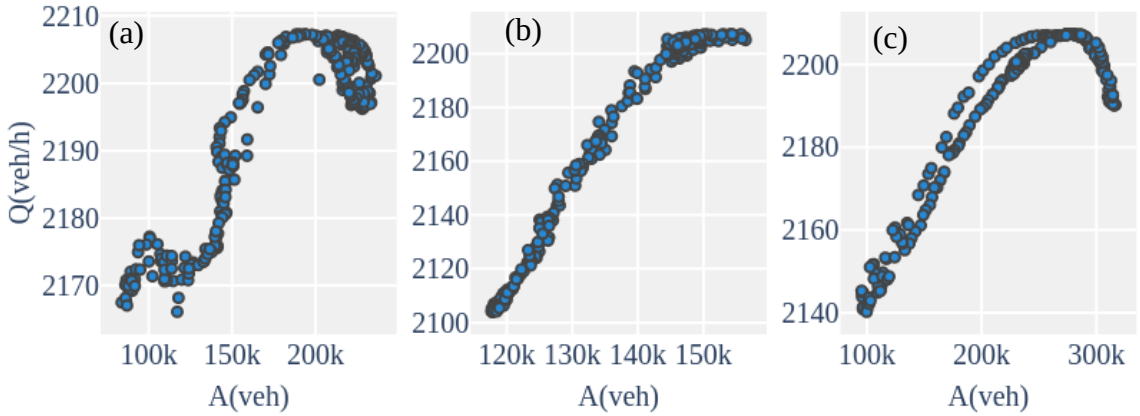


Figure 5.25: Average network flow versus accumulation(total number of vehicles) for the city of (a) Paris; (b) Moscow and (c) Istanbul (5-minute intervals) from Yandex maps based on Greenberg model.

5.6 Chapter Summary

In the current chapter the properties of traffic depiction and the feasibility of using online traffic maps information for capturing network-level traffic phenomena have been investigated. The proposed analysis is based on a traffic information extraction module for realistic networks, where traffic layers are captured by straightforward image processing techniques and processed as categorical type of data. Fundamental traffic relationships are then used, connecting microscopic information to macroscopic phenomena. To investigate the quality of traffic information depicted in online dynamic traffic maps, the estimation and the properties of MFD were used. The application of the proposed methodological framework is tested on commercial freely available online maps for different urban road networks. A thorough investigation about the value of spatiotemporal information that it is retained in the abstracted depiction of traffic conditions has been conducted, revealing that the fundamental macroscopic properties still hold even in cases of significant information reduction like the representation scheme adopted in online traffic maps. Not only the shapes and the magnitude of scale among the traffic variables are consistent but also valuable additional information can be extracted from traffic maps (like vehicles concentration, network heterogeneity, etc.) and delicate phenomena can be captured (like traffic hysteresis). In the next chapter this valuable information in the form of structured database after the proposed method is applied to estimate human mobility patterns across multiple cities.

PART II

Chapter 6 : Human Mobility Patterns based on Information from Online Traffic Maps

The current chapter introduces alternative and innovative methods in understanding movement within the urban fabric at the macroscopic level based on suitably developed quantified indices that may be applied to any city that has online traffic coverage. For illustration purposes, results are presented for multiple cities with different network structures and traffic characteristics (introduced in chapter 4), to reveal regularities/patterns and useful insights regarding the temporal and spatial dimension of urban mobility patterns.

6.1 Urban Morphology

Prior to the estimation of mobility patterns for the cities of the study area, a closer look to the urban morphology of these cities is provided. In Section 4.1, the spatial and cultural distribution along with the basic characteristics (land, population, and Gross Domestic Product) of the cities were presented. Here, emphasis is placed on the spatial correlation between the street connectivity and urban density that contributes to the modeling of a city's spatial structure. Starting with the European cities of the study area, we introduce the city outline for London.

6.1.1 London

London's complicated topography can be made simple by means of three basic patterns. First, there is the undulating line of the Thames separating northern from southern London. For historical reasons, most important destinations lie north of the river. The south is essentially an intricate patchwork of residential districts joined by miles of conventional through streets. At the centre is the area familiar to visitors—the City of London, the municipal corporation and borough of London, with its offices, shops, and public buildings. The first ring surrounding that area, the suburban belt—known for statistical purposes as Inner London—developed from the late 18th century until the beginning of World War I. The third zone—Outer London—consists of 20th-century suburban housing, chiefly created in a short, intensive building boom in 1925–39. The Metropolitan Green Belt forms a final concentric ring, defining the shape of the whole capital (Clout, 2019).

6.1.2 Paris

Moving to the city of Paris, that is positioned at the centre of the Île-de-France region, which is crossed by the Seine, Oise, and Marne rivers. The city proper is small; no corner is farther than about 6 miles (10 km) from the square in front of Notre-Dame Cathedral. It occupies a depression hollowed out by the Seine, and the surrounding heights have been respected as the limits of the city. Paris comprises 20 arrondissements (municipal districts), each of which has its own mayor and town hall. The numbering begins in the heart of Paris and continues in the spiralling shape of a snail shell, ending to the far east. The Seine flows for about 8 miles (13 km) through the centre of the city and 10 of the 20 arrondissements. It enters the city at the southeast corner, flows northwest, and turns gradually southwest, eventually leaving Paris at the southwest corner. As a result, what starts out as the stream's east bank becomes its north bank and ends as the west bank, and the Parisians therefore adopted the simple, unchanging designation of Right Bank and Left Bank (when facing downstream) (Ehrlich, 2019).

6.1.3 Berlin

Berlin is the capital and chief urban centre of Germany. The city lies at the heart of the North German Plain and is by far the largest city in Germany. Divided into several distinct areas, Berlin is a sprawling city, with the main districts being Charlottenburg, Kreuzberg, Mitte and Tiergarten. Berlin has always been a surprisingly green city, with luxuriant trees softening the effect of the stone apartment blocks in many streets. Water is even more prevalent, with the Spree River running through the city's centre, a broad belt of lakes spreading out east and west, and canals running through much of the city.

Until the "peaceful revolution" of 1989, the most notorious feature of the city's topography was the Berlin Wall, erected by the East German communist government in 1961 to stop free movement between East Berlin and West Berlin. The boundary between East and West Berlin and the boundary between West Berlin and East Germany, for a combined length of 103 miles (166 km), were closed until 1989 by a solid ring of barriers, consisting mostly of prefabricated concrete slabs. The political and physical division of Berlin had a profound and pervasive influence on urban planning. During the 1990s, massive construction projects transformed central Berlin. High-rise commercial construction in the Potsdamer Platz, on the site of the former wall, restored its role as a bustling urban centre. The central area of Berlin acquired broad north-south avenues, such as Wilhelm Strasse and

Friedrichstrasse, and its characteristic east-west road axis. Supplementing this main axis are several exit roads that now serve as major traffic arteries (Reuter, 2020).

6.1.4 Moscow

Moscow city is the capital of Russia, located in the far western part of the country. A map of Moscow presents a pattern of concentric rings that circle the rough triangle of the Kremlin and its rectangular extension, the Kitay-gorod, with outwardly radiating spokes connecting the rings; the whole pattern is modified by the twisting, northwest–southeast-trending Moscow River. These rings and radials mark the historical stages of the city's growth: successive epochs of development are traced by the Boulevard Ring and the Garden Ring, the Moscow Little Ring Railway (built in part along the line of the former Kamer-Kollezhsky customs barrier), and the Moscow Ring Road. Inner Moscow functions like a typical central business district. In this area are concentrated most of the government offices and administrative headquarters of state bodies, most of the hotels and larger shops, and the principal theatres, museums, and art galleries (French, 2019).

6.1.5 Nicosia

Nicosia, known locally as "Lefkosia" in Greek, is the centrally located capital of Cyprus. As a result of the Turkish intervention in 1974, part of the northern section of Nicosia, including the former international airport, has remained within the United Nations Forces in Cyprus buffer zone that separates the Republic of Cyprus (south) from the occupied areas (north). Nowadays, the city centre can be divided into two distinct areas: (a) the old city within the Venetian Walls and (b) the new city centre that includes the Central Business District. The circular core of the city centre has imposed a radial development of road systems over the years. Furthermore, Nicosia has the burden of the dividing line, halting its network to dead ends towards the north.

6.1.6 New York

New York City is a collection of many neighbourhoods scattered among the city's five boroughs—Manhattan, Brooklyn, the Bronx, Queens, and Staten Island—each exhibiting its own lifestyle. New York is the most populous and the most international city in the country. Its urban area extends into adjoining parts of New York, New Jersey, and Connecticut. Located where the Hudson and East rivers empty into one of the world's premier harbours, New York is both the gateway to the North American continent and its preferred exit to the oceans of the globe.

The city's ancient bedrock provides the immovable foundation for hundreds of modern skyscrapers. New York has more of these awesome structures than any other world city. Although often modified in specific cases, the rectilinear patterns imposed upon Manhattan in its infancy have determined its developmental patterns. In the 20th century, parkways were incorporated into the traffic patterns of all boroughs, as Eastern and Ocean parkways in Brooklyn, Riverside Drive (Manhattan), the Grand Concourse (the Bronx), and Queens Boulevard attest. Regardless of all its efforts, the modern city is infamous for the volume of traffic that clogs its well-laid-out street system (Lankevich, 2020).

6.1.7 Los Angeles

Los Angeles is the second most populous city and metropolitan area (after New York City) in the United States. The city sprawls across a broad coastal plain situated between mountains and the Pacific Ocean. The huge, sprawling, and tortuously shaped city of Los Angeles occupies a sizable portion of the southern part of the county. The city of Los Angeles is composed of a series of widely dispersed settlements loosely connected to downtown. It certainly does not conform to the popular Chicago school of urban theory of the 1920s and later, which held that a downtown was the main focus of community life, with its influence unfolding in a series of concentric circles out into the hinterlands. Among the outlying districts that lie within the city limits are Hollywood, located northwest of downtown; Encino, Van Nuys, and North Hollywood in the San Fernando Valley; Century City, Westwood, and Venice on the West Side; San Pedro and Wilmington in the harbour area; and Boyle Heights just east of the river. The main links connecting downtown and the suburbs are the famed Los Angeles freeways, which spread throughout the region in a vast network of concrete ribbons (Pitt, 2020).

6.1.8 Toronto

Toronto, city, is the capital of the province of Ontario, south-eastern Canada. It is the most populous city in Canada, a multicultural city, and the country's financial and commercial centre. Its location on the north-western shore of Lake Ontario, which forms part of the border between Canada and the United States, and its access to Atlantic shipping via the St. Lawrence Seaway and to major U.S. industrial centres via the Great Lakes, have enabled Toronto to become an important international trading centre. The city also features an extensive system of underground tunnels and concourses lined with shops, restaurants, and theatres. Through the construction of new housing and mixed-use projects, together with

the restoration and rehabilitation of heritage buildings, an extraordinary vitality has been brought to the urban core (McGillivray, 2020).

6.1.9 Sao Paulo

Sao Paulo, city, capital of Sao Paulo estate (state), south eastern Brazil, is the foremost industrial centre in Latin America. The city is located on a plateau of the Brazilian Highlands extending inland from the Serra do Mar, which rises as part of the Great Escarpment only a short distance inland from the Atlantic Ocean. In sharp contrast to Rio de Janeiro, where the ocean and mountains determine the city's configuration, dauntingly vast Sao Paulo sprawls virtually unrestrained in all directions. Although its nicer residential areas are to the southwest and west of the commercial centre, where elevations are generally higher, the spatial distribution of socioeconomic differences is not as distinct as it is in Rio. The lack of mountains or even significant hillsides means that Sao Paulo's favelas, or shantytowns, are horizontal, not vertical. The new favelas that are rising are well removed from the centre and disconnected from city services. As they grow, only some are integrated into the city; most still suffer substandard conditions. Sao Paulo's centre is marked by squares and parks, all within walking distance of one another. The vibrant commercial centre of Sao Paulo revolves around the famous "Triangle", the city's original centre. (Leite, 2020).

6.1.10 Buenos Aires

Buenos Aires, city is the capital of Argentina. Buenos Aires is one of Latin America's most important ports and most populous cities, as well as the national centre of commerce, industry, politics, culture, and technology. The metropolitan area is divided into the Federal District, established in 1880, and the surrounding suburbs. The Federal District contains less than one-fourth of the population of the metropolitan area, a proportion that shrinks as the suburbs continue to attract industry and residential communities. The city is divided into sections that coincide mostly with the traditional barrios (neighbourhoods). The city centre is built on the original colonial foundation. It has narrow streets laid out at right angles to form a grid pattern. This rectilinear pattern holds for more than 20 square blocks, an area that defined the limits of the city until the late 19th century. Since that time, expansion has been less planned, and the pattern of streets is less regular. The centre is the site of most major financial institutions and corporate headquarters (Bonilla, 2019)

6.1.11 New Delhi

The city of Delhi consists of two components: Old Delhi, in the north, the historic city; and New Delhi, in the south, since 1947 the capital of India, built in the first part of the 20th century as the capital of British India. One of the country's largest urban agglomerations, Delhi sits astride (but primarily on the west bank of) the Yamuna River, a tributary of the Ganges (Ganga) River, about 100 miles (160 km) south of the Himalayas. The national capital territory embraces Old and New Delhi and the surrounding metropolitan region, as well as adjacent rural areas. The city plan of Delhi is a mixture of old and new road patterns. The street network of Old Delhi reflects the defence needs of an earlier era, with a few transverse streets leading from one major gate to another. Occasionally a street from a subsidiary gate leads directly to the main axes, but most Old Delhi streets tend to be irregular in direction, length, and width. Narrow and winding paths, culs-de-sac, alleys, and byways form an intricate matrix that renders much of Old Delhi accessible only to pedestrian traffic. Conversely, the Civil Lines (residential areas originally built by the British for senior officers) in the north and New Delhi in the south embody an element of relative openness, characterized by green grass, trees, and a sense of order (Rao, 2018).

6.1.12 Riyadh

Riyadh itself is an amorphous expanse of neighbourhoods and subdivisions bounded by wide roads lined with commercial strip development. There are thousands of miles of paved roads in Riyadh, including the King Fahd (running north-south) and Mecca (Makkah; running east-west) highways, which constitute the two main axes of the city. With its grid system of wide thoroughfares and expressways, modern Riyadh was designed as an automobile-oriented city. The form and structure of the city has been reinforced by a number of large-scale construction projects undertaken in the latter half of the 20th century and in the early 21st century, including the establishment of the Diplomatic Quarter, where embassies and the offices of international organizations are located, and the redevelopment of the Qaṣr Al-Ḥukm ("Justice Palace") district, which houses most of the central shops (Kim, 2018).

6.1.13 Singapore

The city of Singapore is situated in the southern portion of the main island. Over time, urbanization has blurred the differences between city and country. Built-up areas now cover a large part of the city-state. The older parts of the city have been substantially refurbished, especially along the Singapore River but elsewhere as well. The once-common

Chinese shop-house, consisting of living quarters above a commercial establishment, gradually has been disappearing from the city. Instead, the government's Housing and Development Board (HDB) has relocated commerce into separate districts and has created integrated residential communities inhabited by people with a mixture of incomes. Singapore has one of the world's busiest ports in terms of shipping tonnage. The Port of Singapore Authority oversees all shipping activity and operates several terminals on the island. The island has a well-developed network of roads and highways, but traffic congestion frequently is a serious problem (Leinbach, 2020).

6.1.14 Beijing

The traditional core of Beijing essentially consisted of two walled cities (the walls no longer stand), the northern inner city and the southern outer city. Beijing has an increasingly dense network of highways radiating from the city, which are used by a growing number of privately-owned automobiles as well as by trucks and long-distance bus services. Beijing's road transport system, though improving rapidly, is still inadequate and cannot keep up with the rapid increase in vehicles, and traffic congestion is often severe. The once-ubiquitous bicycles and three-wheeled cycle carts continue to be heavily used for short-distance transport, despite the proliferation of automobiles (Bonavia, 2020).

6.1.15 Tokyo

Tokyo formerly (until 1868) Edo, is a metropolitan complex along the northern and western shores of Tokyo Bay, on the Pacific coast of the island of Honshu, central Japan. At its centre is the metropolitan prefecture, or metropolis of Tokyo, Japan's capital and largest city. Three prefectures (ken) bordering it—Saitama on the north, Chiba on the east, and Kanagawa on the south—may be said to make up the remainder of the complex, but there is more than one definition of Greater Tokyo, and large numbers of people live beyond the four prefectures and commute to work in the region.

Tokyo metropolis is one of three prefecture-level urban administrative units in Japan, the other two being Ōsaka and Kyōto. In addition to Tokyo city, the metropolis includes numerous industrial and residential suburbs and a large mountainous rural area to the west. Tokyo metropolis is a major component of the Tokyo-Yokohama Metropolitan Area as well as of the Keihin Industrial Zone. Tokyo is the chief transportation hub for Japan, as well as an important international traffic centre. It is served by a dense network of electric railways, subways, bus lines, and highways. Tokyo station is the central railroad terminal for all of

Japan, including the high-speed Shinkansen bullet trains from western Japan (Seidensticker, 2016).

6.1.16 Johannesburg

One of the youngest of the world's major cities, Johannesburg was founded in 1886, following the discovery of gold. Central Johannesburg, the commercial and financial heart of South Africa, is laid out in a rectangular grid pattern that is unchanged from the first city survey in 1886. Streets are narrow and cast into shadow by high-rise concrete blocks, creating an almost tunnel like effect. Architecturally, the city is a hodgepodge, reflecting decades of rapid growth and a singular indifference to historic preservation. Johannesburg is a hub for local, national, and international travel. Railroads and multilane freeways crisscross the metropolitan area, carrying hundreds of thousands of daily commuters to and from outlying suburbs and townships (Campbell, 2020).

6.1.17 Sydney

Located on Australia's south-eastern coast, Sydney is the country's largest city and, with its magnificent harbour and strategic position, is one of the most important ports in the South Pacific. Greater Sydney is spread over a vast area that stretches from the Blue Mountains in the west to the Pacific Ocean in the east and from the southern shore of Lake Macquarie in the north to south of Botany Bay. Only about one-third of this region is classified as urban, but the great bulk of the region's population lives in the urban area. A pattern of suburban sprawl, caused partly by the cheapness of land in earlier days and by the determination of ordinary Australians to own their own houses and gardens, has caused problems for the authorities responsible for sewerage and transport. The sprawl is also in marked contrast to the comparatively small and compact central business district, which is crammed into a narrow rocky peninsula between two arms of the harbour within the City of Sydney proper (Pringle, 2020)

6.1.18 Istanbul

Last, Istanbul, is the largest city and principal seaport of Turkey. The old walled city of Istanbul stands on a triangular peninsula between Europe and Asia. The old city contains about 9 square miles (23 square km), but the present municipal boundaries stretch a great deal beyond. By long tradition, the waters washing the peninsula are called "the three seas": they are the Golden Horn, the Bosporus, and the Sea of Marmara. The Bosporus is the channel connecting the Black Sea to the Mediterranean by way of the Sea of Marmara and

the straits of the Dardanelles. The narrow Golden Horn separates old Istanbul to the south from the "new" city of Beyoğlu to the north; the broader Bosporus divides European Istanbul from the city's districts on the Asian shore. The Galata and Atatürk bridges cross the Golden Horn to Beyoğlu. The separated geographical condition of Istanbul together with insufficient road network capacity and narrow access roads have generated constraints of transportation system (Ehrlich, 2020).

6.2 Methodological Approach and Outcomes

This section introduces the methodological framework, developed to estimate human mobility patterns for large urban areas, based on traffic maps. Starting with the three fundamental constituents of movement (space, time, and objects), three different foci are possible in analysing movement: the first focus on objects (movers, trajectories), the second on space (characteristics of location) and the third on time (characteristics of time units in terms of objects and space (Andrienko *et al.*, 2011). Here, the analysis concentrates not only on objects (human movements depicted through traffic maps) but also on space by pointing out locations within the city that appear congestion and further on time through the traffic state variation during the day. Specifically, two indicators named as Structural Similarity Index (SSIM), SSIM1 and SSIM2, that quantify traffic state alternation for the whole urban area were used to produce temporal mobilities patterns while Kernel Density estimation was applied for spatio-temporal patterns. Another clustering algorithm is utilized for the estimation of congestion patterns. Further elaboration on the terms and their use is provided in the dedicated sub-sections following next.

6.2.1 Quantification of variability in human mobility

To capture and monitor the daily variations across cities, image processing techniques were applied, as mentioned above. Specifically, a method for measuring the similarity between two images, known as the Structural Similarity Index (SSIM), was implemented. The SSIM is based on the computation of three terms, namely the luminance term *l*, the contrast term *c* and the structural term *s* between two nonnegative image signals x, y as follows (Wang *et al.*, 2004).

$$SSIM(x,y) = [l(x,y)]^{a} \cdot [c(x,y)]^{\beta} \cdot [s(x,y)]^{\gamma} \quad (6.1)$$

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (6.2)$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (6.3)$$

$$s(x,y) = \frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3} \quad (6.4)$$

where α>0, β>0 and γ>0 are parameters used to adjust the relative importance of the three components, where $\mu_x, \mu_y, \sigma_x, \sigma_y$ and σ_{xy} are the local means, standard deviations, and the cross-covariance for images x, y. The parameters $C_1 C_2, C_3$ are included to avoid instability when the denominator is very close to zero. The default value for exponents (α, β, γ) is equal to one.

In the case of map images, the similarity measure between two sequential images can capture and quantify the variation on mobility for the next time interval, for large urban areas. Furthermore, if an initial image is chosen as a baseline (i.e. off-peak hour state of the network) the network-wide propagation of traffic states can be monitored and compared within the day and across different cities. Thus, two different SSIM indices were used. The first named SSIM1 compares image x to image y, according to Eq. (6.5). In that case, x corresponds to the image taken on time interval t_k and y correspond to the image taken on time interval t_0 .The k variable lay in the range [1 288] for a day's images while, time interval t_0 refers to off-peak hour state of the network.

$$\mathrm{SSIM1}\left(\mathrm{img}_{t_k}, \mathrm{img}_{t_o}\right) = \left[\mathrm{l}\left(\mathrm{img}_{t_k}, \mathrm{img}_{t_o}\right)\right] \cdot \left[\mathrm{c}\left(\mathrm{img}_{t_k}, \mathrm{img}_{t_o}\right)\right] \cdot \left[\mathrm{s}\left(\mathrm{img}_{t_k}, \mathrm{img}_{t_o}\right)\right] \quad (6.5)$$

Calculating the SSIM1 index (see Algorithm 4, Appendix A) for all the selected cities the dynamic monitor of mobility patterns' variation, compared to the off-peak hour state of the network, can be captured and compared in a quantified way. In this respect, urban planners and traffic operators can explore the mobility profile of a city and measure the impacts on human mobility due to traffic congestion. By using, an initial image during off-peak hour as a baseline, the SSIM1 index at each time interval measures the difference on the level of mobility, compared to a stable traffic state in a city. Obviously, if image x is identical to image y the index is equal to 1, while if image x differs from y the index is closer to 0.

Figure 6.1 shows the daily variation of the SSIM1 index across the first nine cities of the study area, based on images from Google maps. By definition, a sharp drop in SSIM1 indicates a completely different traffic state of the network compared to off-peak hours, such as in case of a traffic jam, or during peak-hour traffic. Thus, initially, the index values are equal to 1 or close to 1, as activity drops at night and the sequential images resemble the

baseline image. A decline is observed in the morning peak hours for all the cities ranging from 0.95 to 0.7. Paris, Moscow and Istanbul exhibit a more abrupt drop (from 1 to 0.7-0.85) compared to the rest cities. Exceptionally, Singapore records the smallest variation at the morning peak (SSIM1 equal to 0.94) and a more stable pattern overall. The same stable pattern applies to Tokyo, London and Sao Paulo. During the rest of the day, variation becomes more stable. A second drop is observed in the evening peak-hours for all the cities followed by a rise, while activity drops again during the night hours.

Figure 6.2 shows the daily variation of the SSIM1 index across the rest cities of the study area. Again, images from Google maps were utilized for the same weekday. For the city of Beijing, images form Baidu maps were used, as Google does not provide online traffic coverage for this city. A more stable pattern is observed for the city of Beijing and for Singapore where a significant drop is noticed only during the morning and evening peak hours. Johannesburg, Buenos Aires and Sydney appear a steeper drop (from 1 to 0.85) compared to the other six. A steep incline after the morning peak is observed for the city of Toronto (from 0.87 to 0.95, that implies a fast transition of the network to the free-flow conditions. Sharp inclines after the morning peak is also detected for the city of Berlin, Sydney, Johannesburg, and Buenos Aires. Morning peak for the city of Riyadh is smoother compared to other cities but persists more hours.

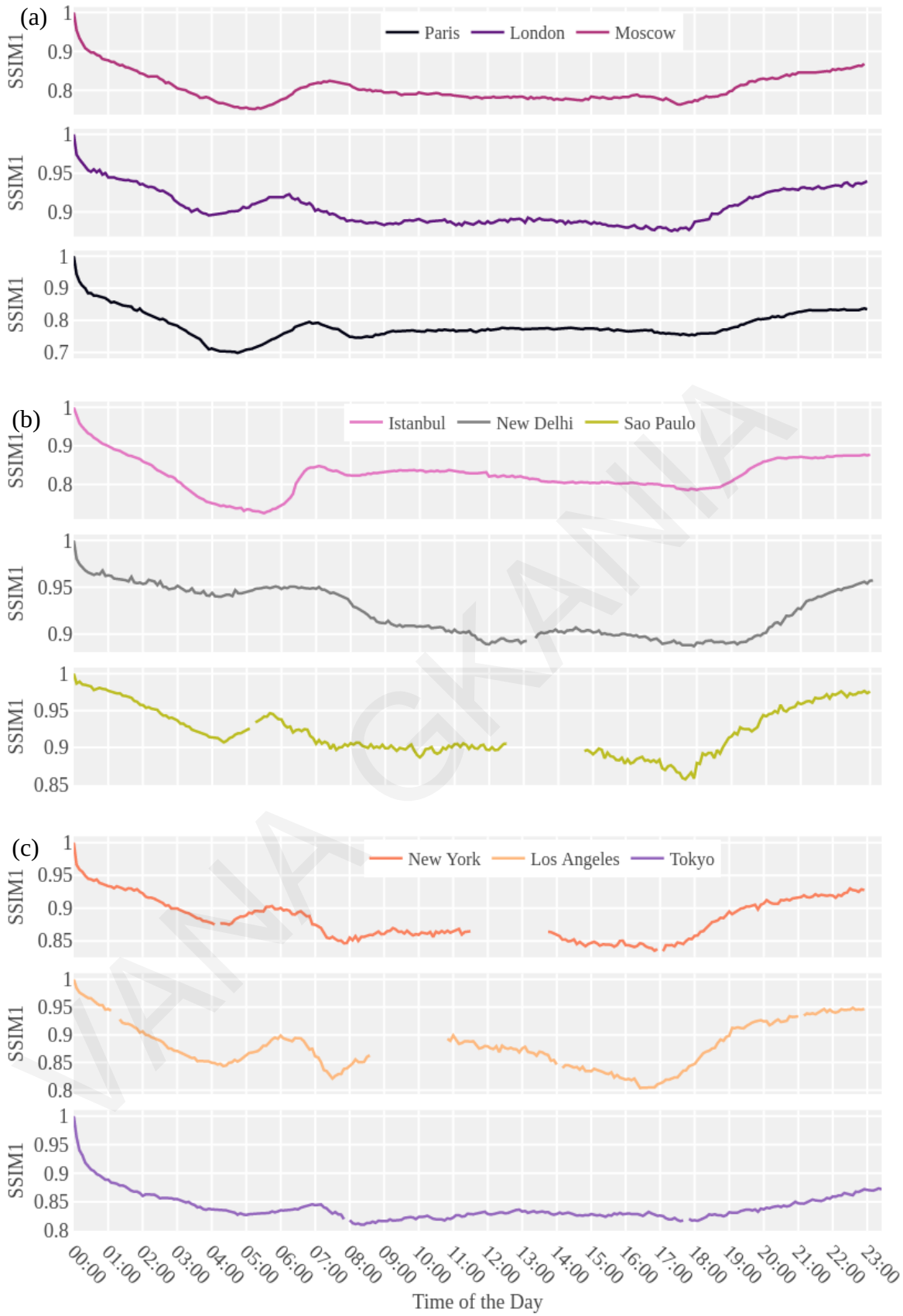


Figure 6.1: SSIM index for the city of (a) Paris, London, Istanbul; (b) Istanbul, New Delhi, Sao Paulo; (c) New York, Los Angeles, Tokyo.

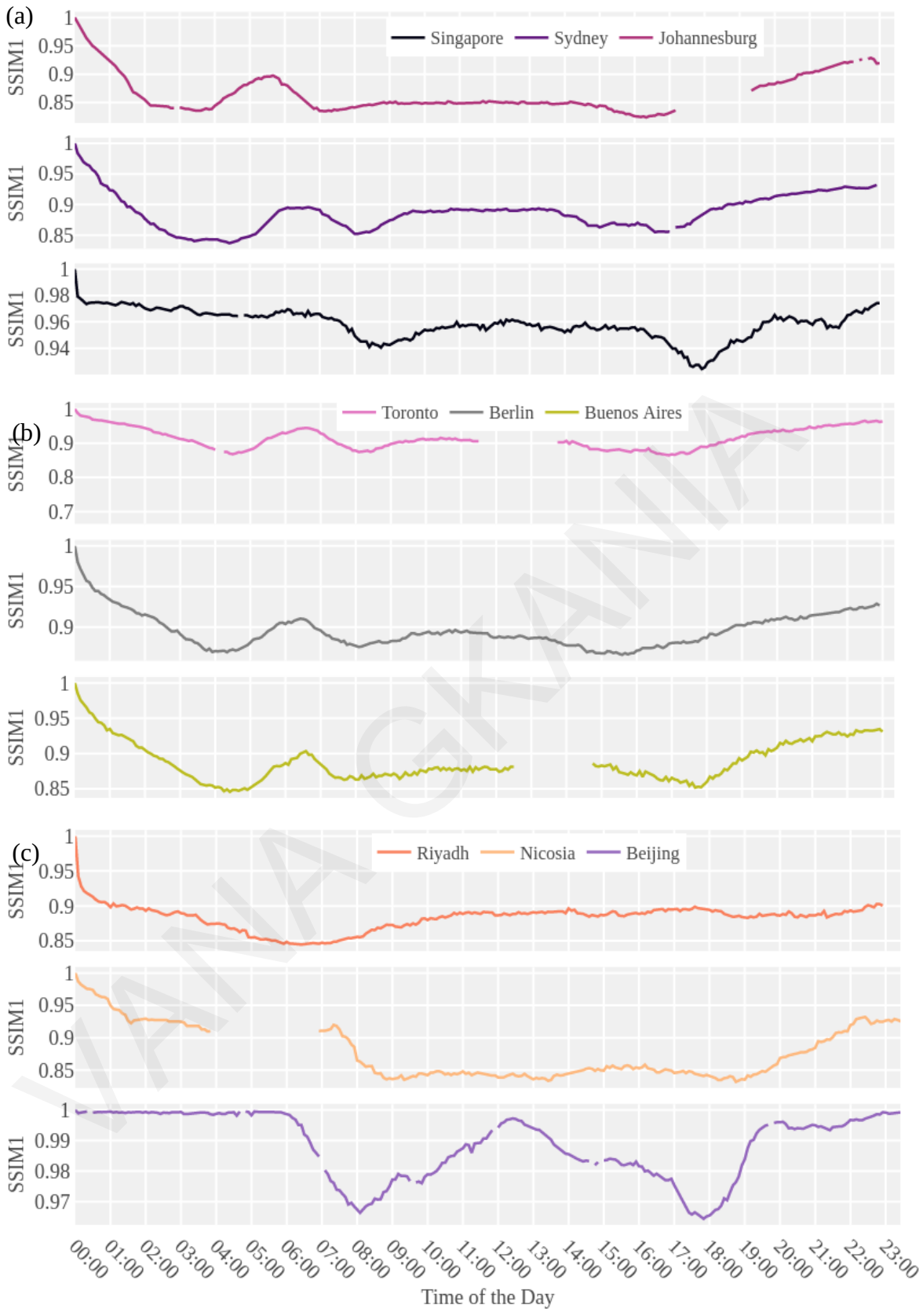


Figure 6.2: SSIM1 index for the city of (a) Singapore, Sydney, Johannesburg ; (b) Toronto, Berlin, Buenos Aires ; (c) Riyadh, Nicosia, Beijing.

For comparative reasons, Figure 6.3 and Figure 6.4 show the SSIM1 index for Paris, Moscow and Istanbul, estimated based on images from four different online traffic providers (Google, Bing, Here, and Yandex), for the same day. A quite different pattern for SSIM1 is observed for the same city across the four providers. This can be explained due to the different traffic coverage of the network for each provider. Starting with the estimation of the SSIM1 based on images from Google maps (Figure 6.3) a similar pattern can be seen for the three cities. A sharp drop during the morning peak (from 1 to 0.7) is noticed and then the network presents a more stable state for the rest of the day. A second drop is observed in the evening peak-hours for all the cities followed by a rise, while activity drops again during the night hours.

On the contrary, for the rest providers, SSIM1 values for each city presents a different rhythm per time interval, especially for Bing maps (Figure 6.4a). Except form morning and evening peak hours, smaller peak and drops can be further seen between 09:00 to 18:00. On the other hand, Here maps (Figure 6.4b) presents a stable pattern till 5:00 that implies that the sequential images are like the base image. Then a steeper drop is observed for the city of Paris between 8:00 and 9:00 in the morning. In contrast, Moscow and Istanbul present a smoother pattern during the day. Last, according to Yandex maps (Figure 6.4c), congestion starts earlier for the city of Paris and Istanbul as a steep drop (from 1 to 0.6) is observed till 5:00 in the morning.

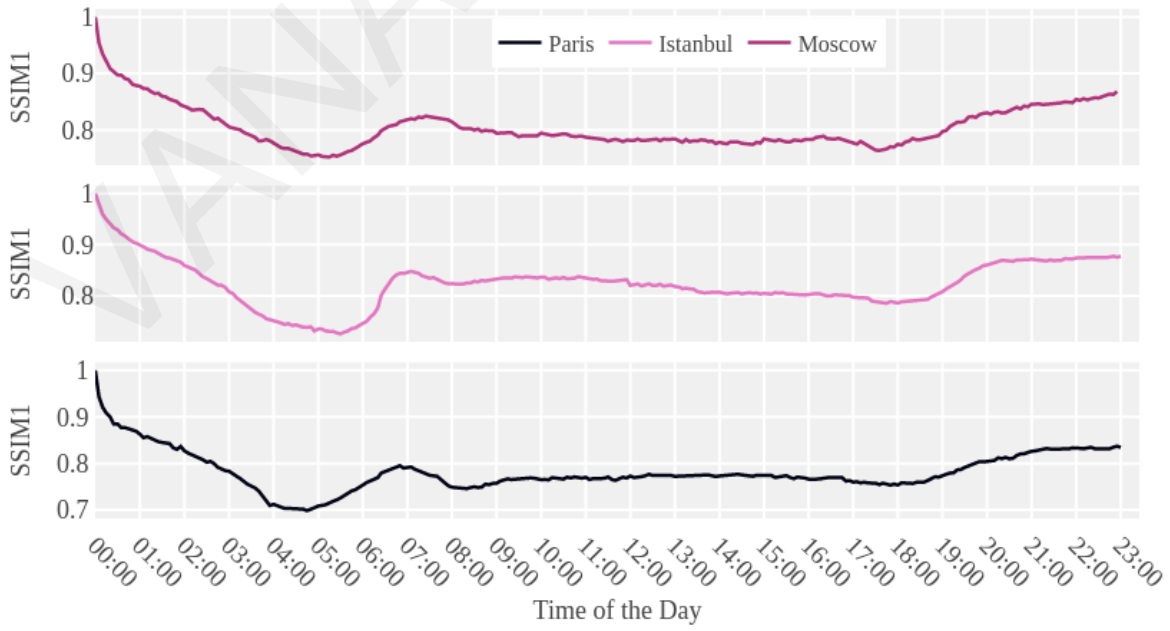


Figure 6.3: SSIM1 index for the city of Paris, Moscow and Istanbul from Google maps.

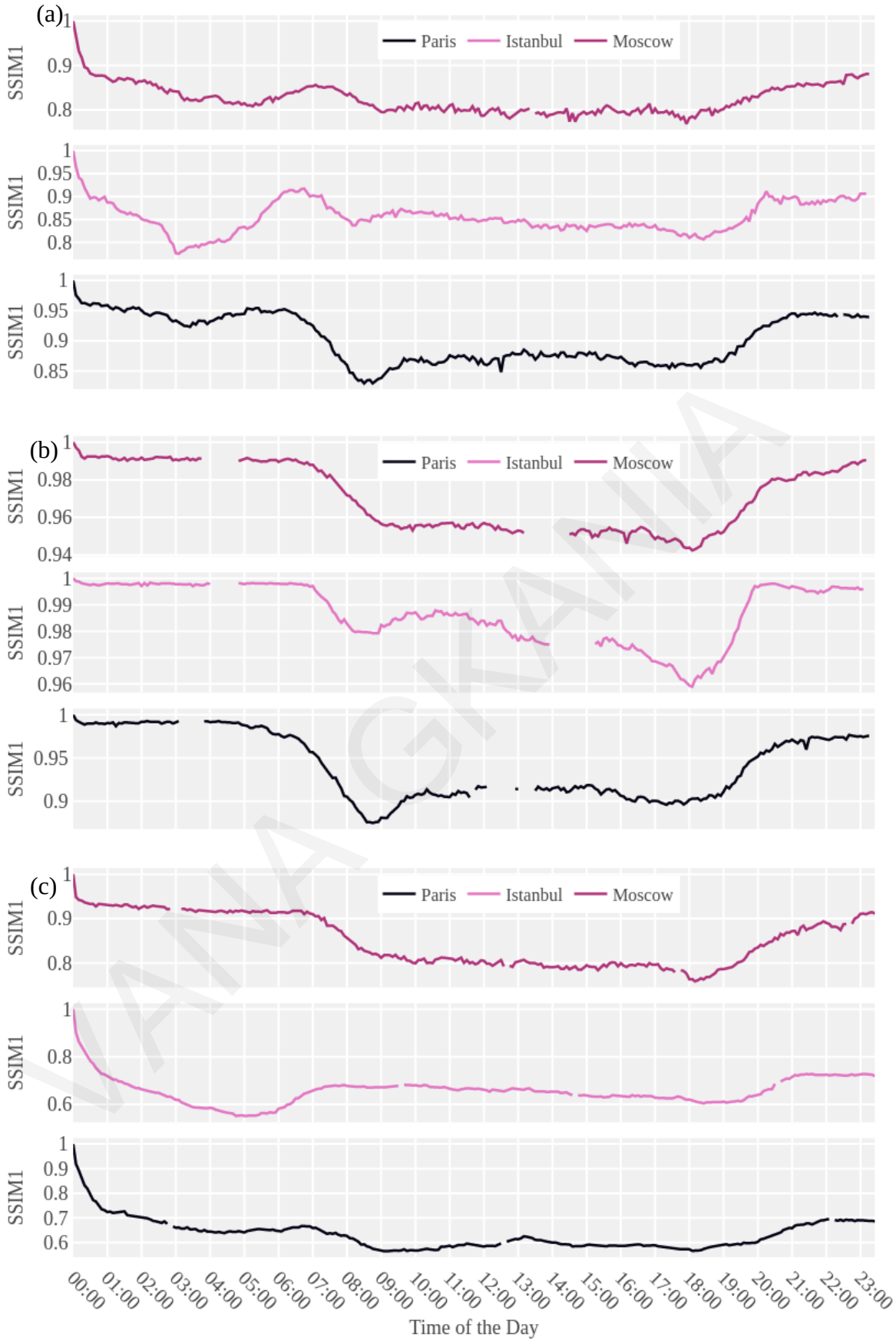


Figure 6.4: SSIM1 index for the city of Paris, Moscow and Istanbul from (a) Bing; (b) Here and (c)Yandex maps.

To directly compare the selected cities, we plotted the variability of SSIM1 in Figure 6.5.

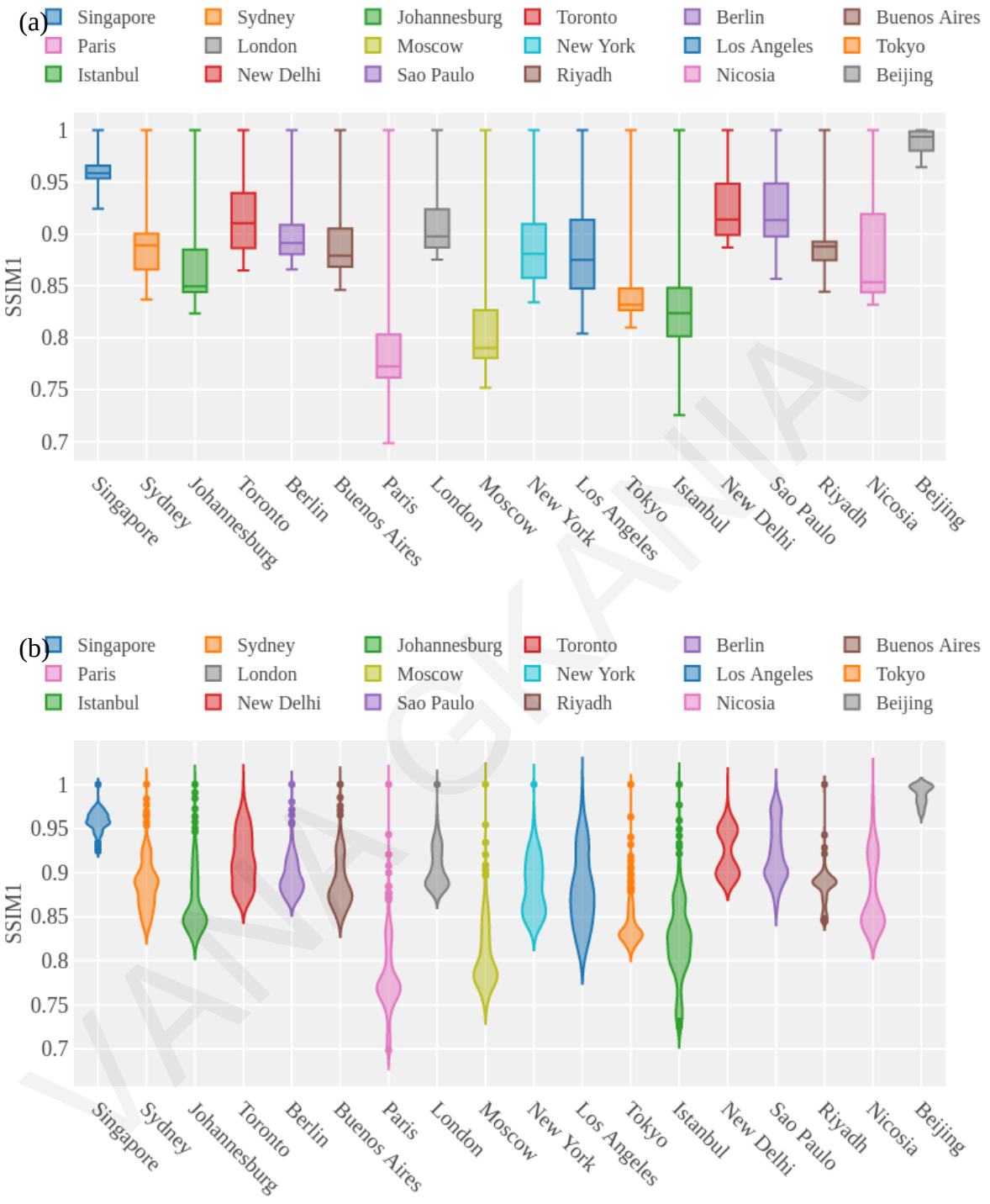


Figure 6.5: Variability of SSIM1 for the selected cities a) Boxplot graph; b) Violin graph

Starting with the boxplot (Figure 6.5a) cities can be grouped into three different clusters based on the median values. First cluster includes Singapore with the higher median value (0.94) while in the second cluster median values range between 0.85-0.95. In the third cluster, that contains Paris, Moscow, Tokyo, and Istanbul, median values are lower than 0.85. As far as the minimum values is concerned Paris, Moscow and Istanbul presented the

lower values. Besides boxplots, we can see a violin graph for the same data (Figure 6.5b). The unquestionable advantage of the violin plot over the box plot is that aside from showing the basic metrics (median, first and third quartile etc.) it also shows the entire distribution of the data. Wider sections of the violin plot represent a higher probability of observations taking a given value while the thinner sections correspond to a lower probability.

Based on the shape of the distributions, we can distinguish two main clusters of cities. First cluster includes cities where only one wider symmetric section is observed such as in Johannesburg, Moscow, Paris and Tokyo. The second cluster contains cities that appeared two or more wider sections such as the city of Nicosia, New Delhi and Sao Paulo. Regarding the number of outliers, there are more in the case of Paris, Istanbul and Moscow.

Next, to quantify human mobility alternation in the diurnal cycle, as reflected in the sequence of traffic maps images, a second SSIM2 index was estimated, as follows:

$$\mathrm{SSIM2}(\mathrm{img}_{t_{k+1}}, \mathrm{img}_{t_k}) = [\mathrm{l}(\mathrm{img}_{t_{k+1}}, \mathrm{img}_{t_k})] \cdot [\mathrm{c}(\mathrm{img}_{t_{k+1}}, \mathrm{img}_{t_k})] \cdot [\mathrm{s}(\mathrm{img}_{t_{k+1}}, \mathrm{img}_{t_k})] \quad (6.6)$$

The SSIM2 index compares image x to image y, where x corresponds to the image taken on time interval t_{k+1} and y to the image taken on time interval t_k, while k ranges between [1 288]. This index can be developed into an enhanced tool for monitoring and understanding the mobility profile of any city with online traffic coverage. Historical profiles of this index can be used to compare traffic anomalies (steep drops or peaks of the index) and their impact on human mobility for the entire city. Figure 6.6 and Figure 6.7 plots the SSIM2 index (Algorithm 5, Appendix A) for a typical weekday and for the eighteen cities of the study area. It can be said that all cities display a broadly comparable rhythm, common to all components of activity, based on the second index. Interestingly, Tokyo seems to differ, as a systematic alternation is observed in the sequence of images. Steep drops or peaks in these time-series suggest a sharp transition between the mobility patterns at the next time interval, thus SSIM2 could detect irregular operations or traffic anomalies. Different timeframes (hour, week, month) can be used to compare these temporal mobility patterns among cities.

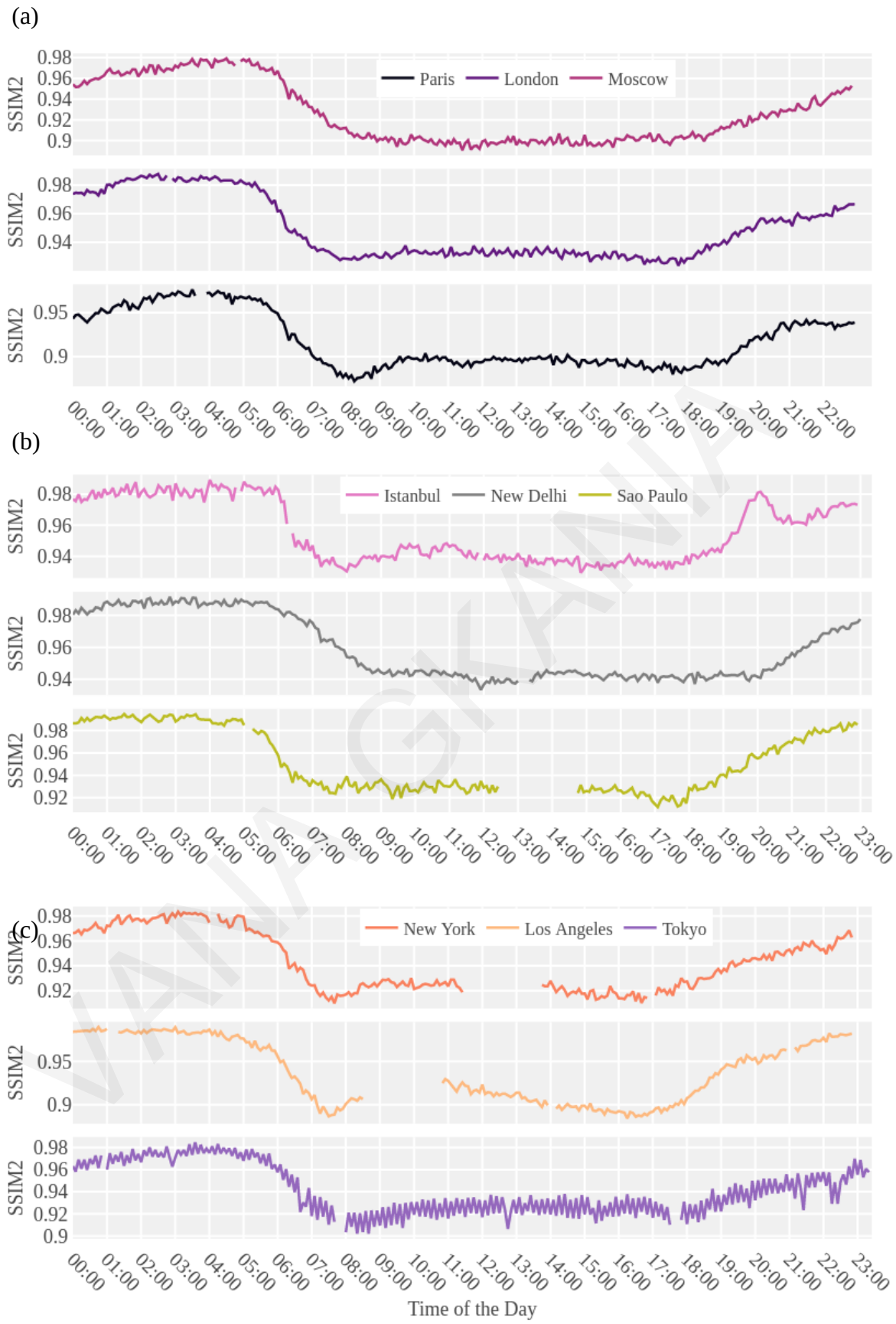


Figure 6.6: SSIM2 index for the city of (a) Paris, London, Istanbul; (b) Istanbul, New Delhi Sao Paulo; (c) New York, Los Angeles, Tokyo.

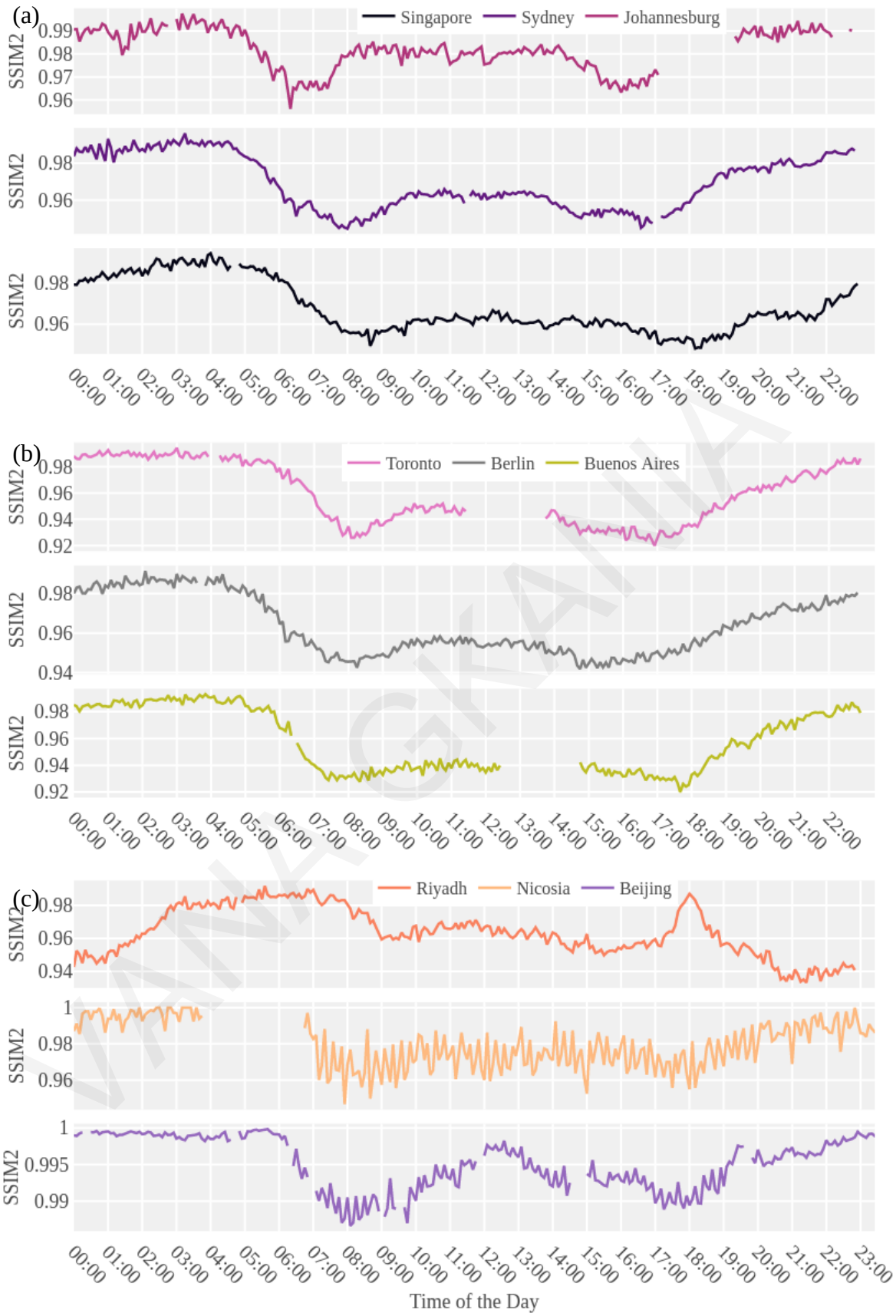


Figure 6.7: SSIM2 index for the city of (a) Singapore, Sydney, Johannesburg ; (b) Toronto, Berlin, Buenos Aires ; (c) Riyadh, Nicosia, Beijing.

To compare the values of SSIM2 estimated from map images from different providers, again three cities that have online traffic coverage from Google, Bing, Here, and Yandex were utilized, as shown Figure 6.8 and Figure 6.9. Greater alternation between sequential images can be seen for Bing and Yandex maps.

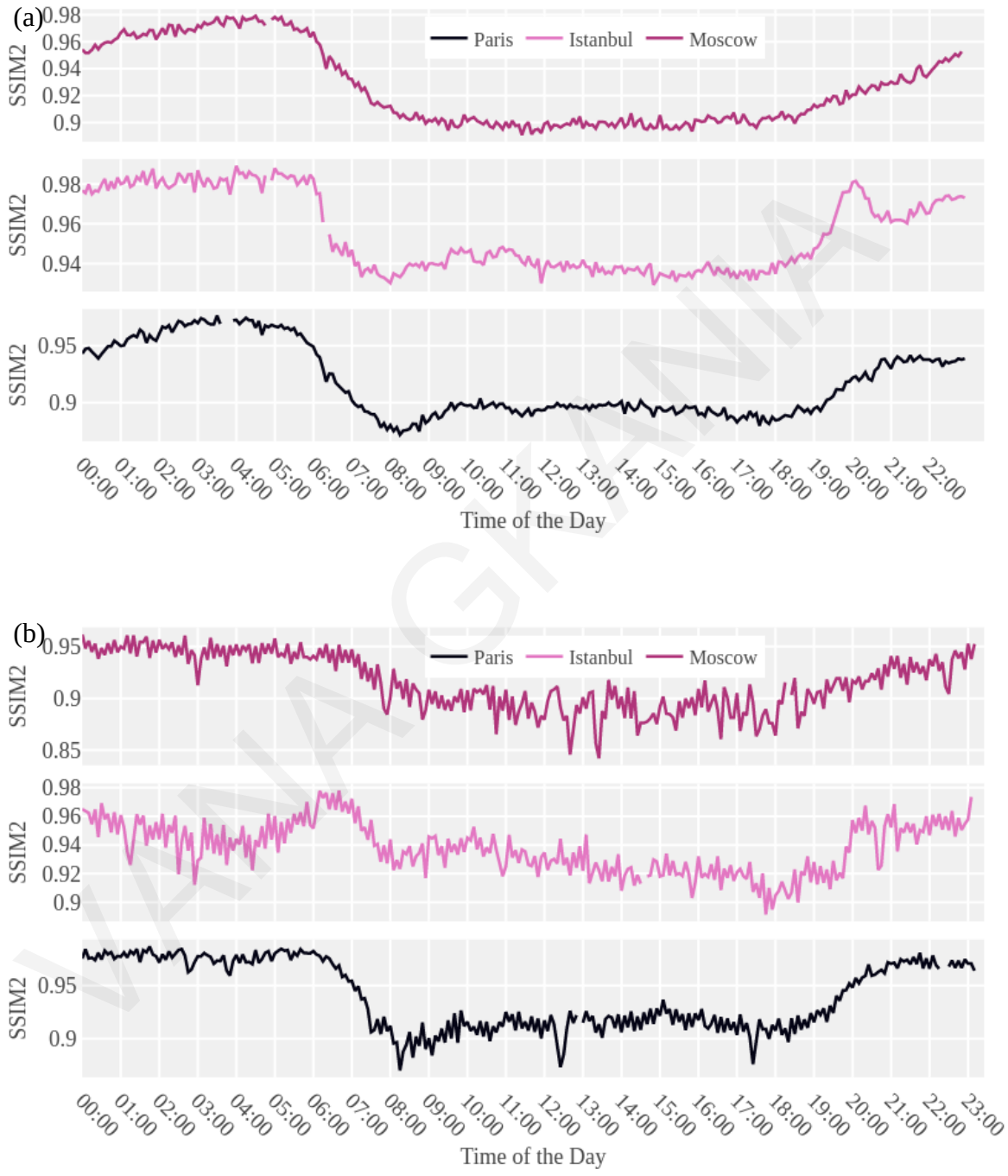


Figure 6.8: SSIM2 index for the city of Paris, Moscow and Istanbul from (a) Google and (b) Bing maps.

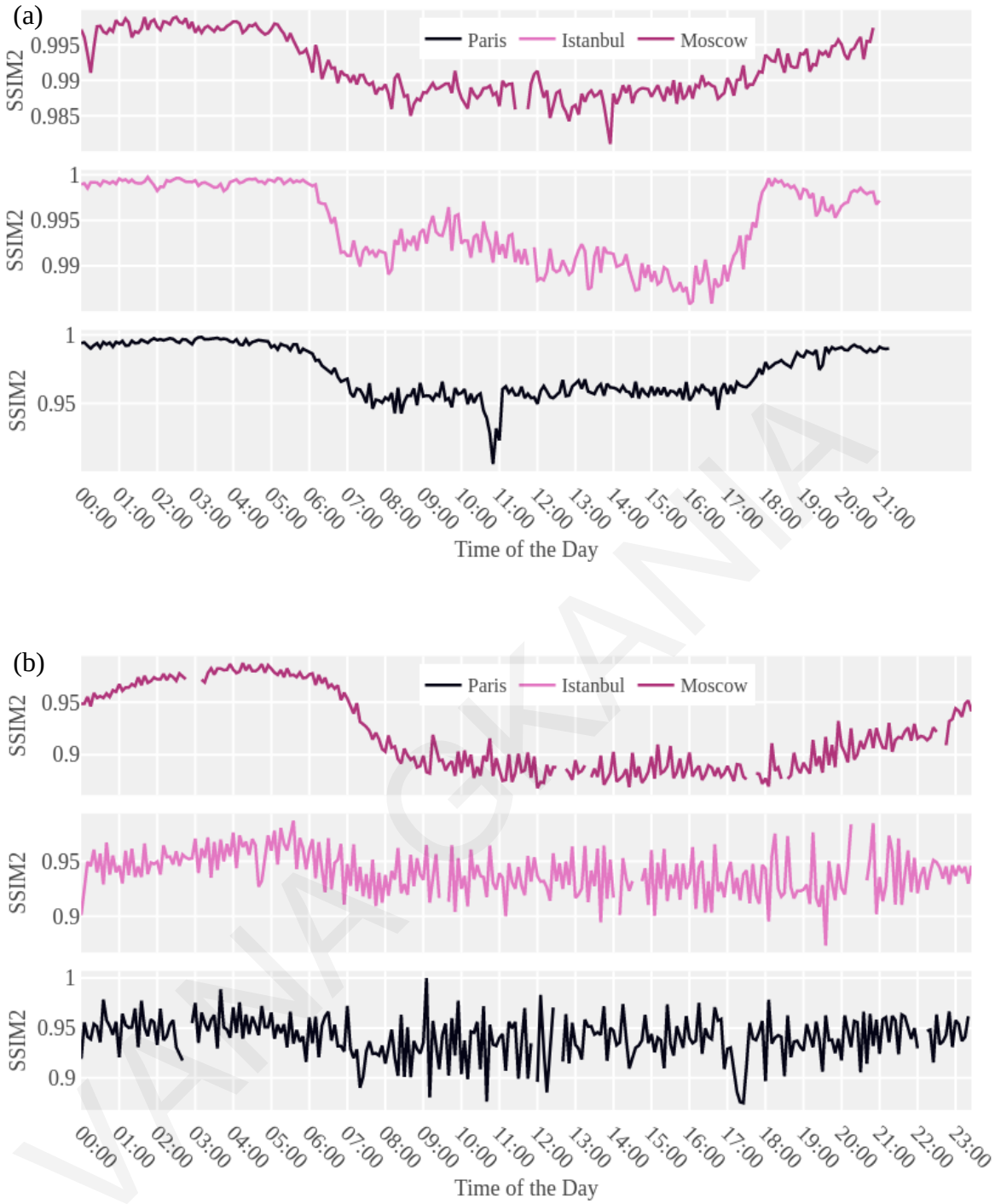


Figure 6.9: SSIM2 index for the city of Paris, Moscow and Istanbul from (a) Here and (b) Yandex maps.

For comparative analysis, the variability of SSIM2 data for the selected cities is illustrated in Figure 6.10 using boxplot and violin plots. Based on the first plot (Figure 6.10a) Singapore, Johannesburg and Sydney presented higher median values (over 0.96) while Paris the lowest (0.9). Regarding the shape of distribution in Figure 6.10b, regularities are observed for all cities, except from the city of Johannesburg. All cities appeared to have two

wider symmetric sections, that contains observations of the SSIM2 during the morning and evening peak-hours, accordingly. Paris and Los Angeles presented greater variability for the SSIM2, indicating multiple different traffic states of the network during the day.

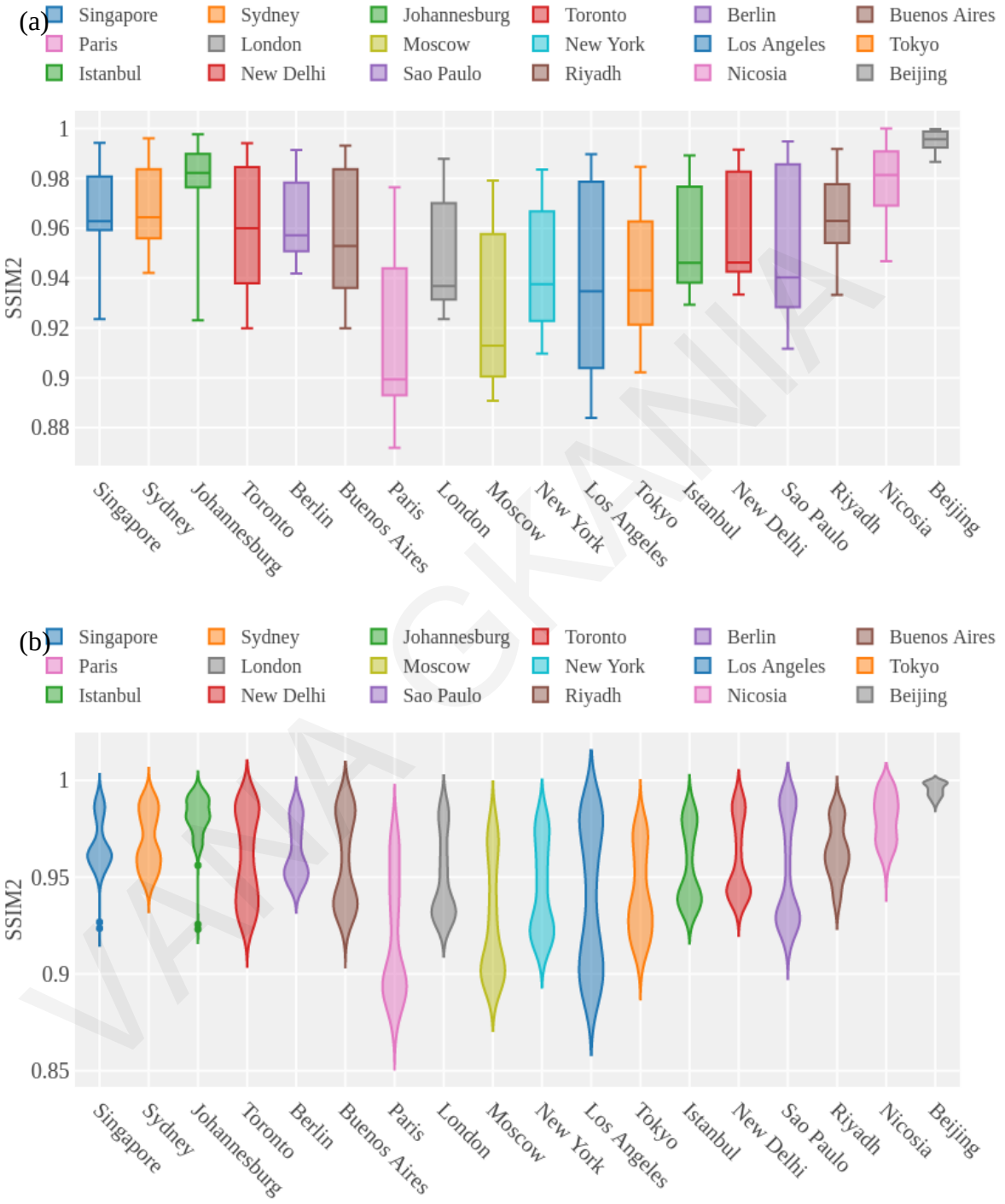


Figure 6.10: Variability of SSIM2 for the selected cities a) Boxplot graph; b) Violin graph

So far, the temporal dimension of mobility patterns was investigated across cities by monitoring the variation of pixels' percentages per traffic layer and by utilizing the SSIM indices. A common rhythm of mobility was observed across the cities during the day while

a closer look at the SSIM2 index revealed a dissimilar pattern for the city of Tokyo. In the next sub-section, emphasis will be given on the spatial dimension of mobility.

6.2.2 Spatio-temporal visualization of human mobility

Trying to explore the spatial dimension of human mobility, a method based on Kernel Density estimation function was applied. The particular method allows one to estimate the intensity of a point pattern and to represent it by means of a smoothed continuous surface that represents the variation of density of point events across the study region (Borruso, 2008).

The general form of a kernel estimator is given as

$$\widehat{f_h}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{6.7}$$

where $x_1, x_2 \ldots\ldots x_n$ are random samples from an unknown distribution, n is the sample size, $K(\cdot)$ is the kernel smoothing function, and h is the bandwidth (Bowman and Azzalini, 1997).

For each city, Kernel Density estimation was used to plot the intensity of each traffic layer, using five-minute time intervals, leading to mesmerized visualization of human mobility in the following figures. For easy comparison the corresponding colours of traffic layers were also used in the Kernel Density plots while two divergent time intervals were chosen, peak and off-peak hour to demonstrate the variation of mobility patterns on the urban grid. On the same figures, the peaks of the distributions reveal the inhomogeneity of activity in the urban space for a specific time interval while the number of peaks both in x and y axis indicates the homogenous centers of activity/mobility. A clear correlation between the four Kernel density plots for each city can be identified that resembles the relationship among traffic layers as shown in Figure 6.11 till Figure 6.16. For instance, when the percentage of green pixels (Light traffic) drops rapidly during the peak hours, a rise of the percentages of the rest layers appears at the same time. The vital information that the following figures provide stands for the exact location where this phenomenon takes place within the urban area.

Starting with the spatio-temporal patterns for the city of Paris, London and Moscow in Figure 6.11, the shape and the number of peaks for each traffic layer differs between the selected intervals and among cities. Although Paris and Moscow have similar road network geometry the total peaks of the distributions related to congestion (red and dark red colour)

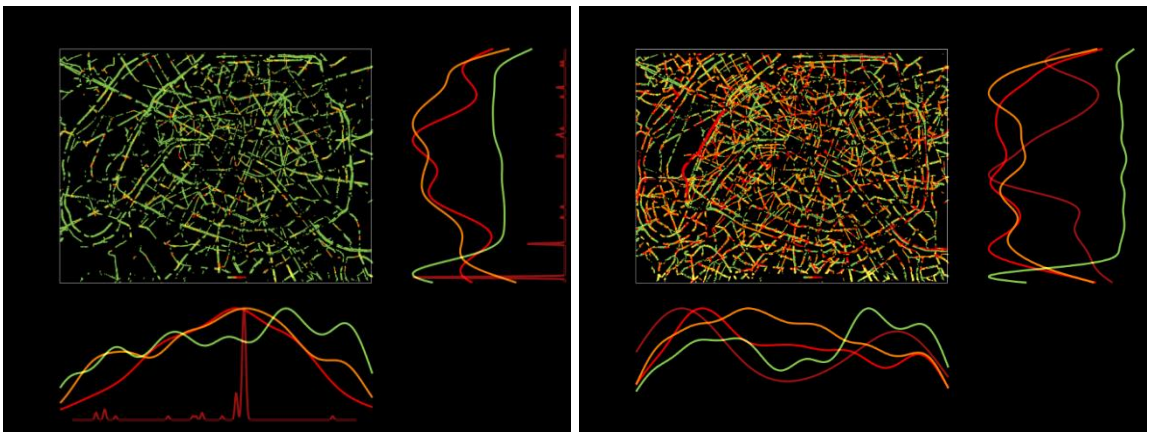
differs, as in Moscow we observe more congested locations. On the other hand, in London, we can see one big center of activity (peak of dark red distribution on y-axis) the City of London while secondary centers can be found as one moves away from the City.

Moving to Figure 6.12 it is easy to locate the congested areas within the city of New York, Los Angeles, and Tokyo during the peak hour (images on the right). In New York, the Kernel Density plots(red-dark) indicate the Business District Center in Manhattan as the most congested one and the same applies for the center district of Los Angeles. From the other hand, in Tokyo, the area near the port appears a strong activity during the peak hour that spreads towards the inner urban area while the shape of peaks for each traffic layer is smoothest compared to the rest of the cities.

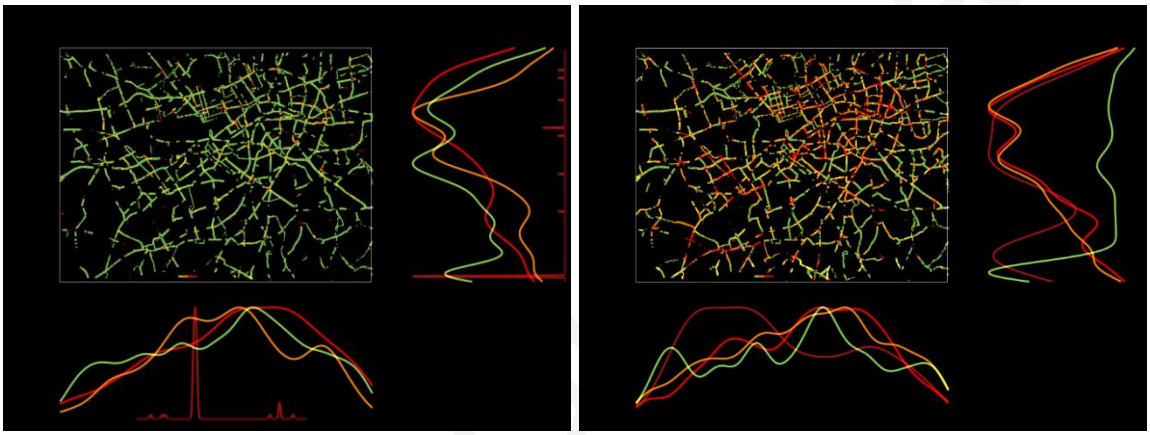
Next, results for the cities of Istanbul, New Delhi and Sao Paulo are presented in Figure 6.13. Here two distinguished congested locations are observed for the city of Istanbul and one for New Delhi. In Sao Paulo, multiple peaks of the Kernel distribution indicate a much more congested location apart from the city center. At that point, we should mention that the sharp peak of the dark-red Kernel Density plot, on the left images (off-peak interval) of the figures appears due to the map's legend.

Afterwards, the spatio-temporal patterns for the city of Singapore, Sydney and Johannesburg are presented in Figure 6.14. Starting with the city of Singapore, the peaks of the distributions related to congestion (red and dark red colour) on y-axis indicate that the north part of the city is the most congested one during the peak hours. On the other hand, in Sydney we can see two big centers of activity (peak of dark red distribution on y-axis), one on the south part of the city and the other in the centre of the map image. In Johannesburg, one distinguished congested location is observed based on the peak of red and dark red distribution on x-axis.

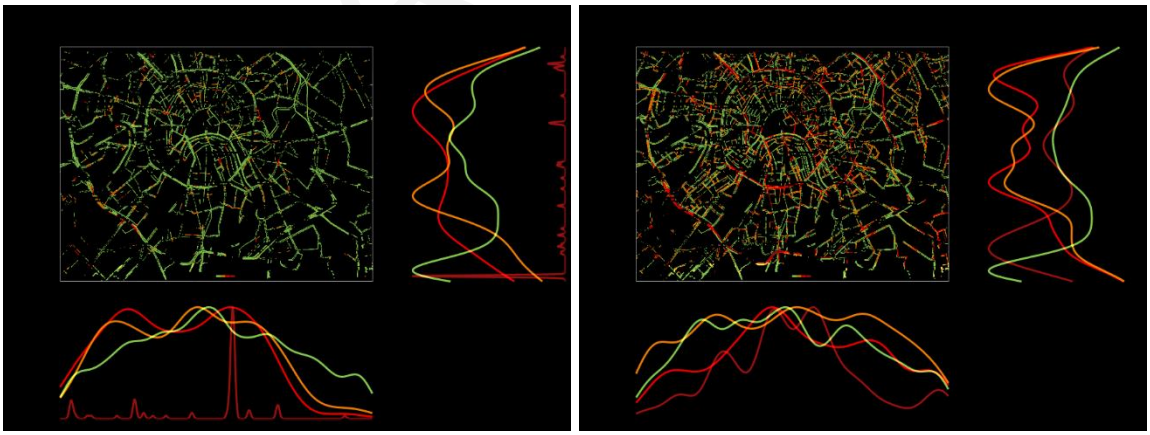
Moving to Figure 6.15, the shape and the number of peaks for each traffic layer differs between the selected intervals and among cities. For the city of Toronto, different parts of the rectangular urban grid face congestion problems during the peak hours. In Berlin, the Kernel Density plots (red-dark) indicate the Business District Center as the most congested one. On the other hand, in Buenos Aires, multiple peaks of the Kernel distribution indicate a much more congested location apart from the city center.



(a)

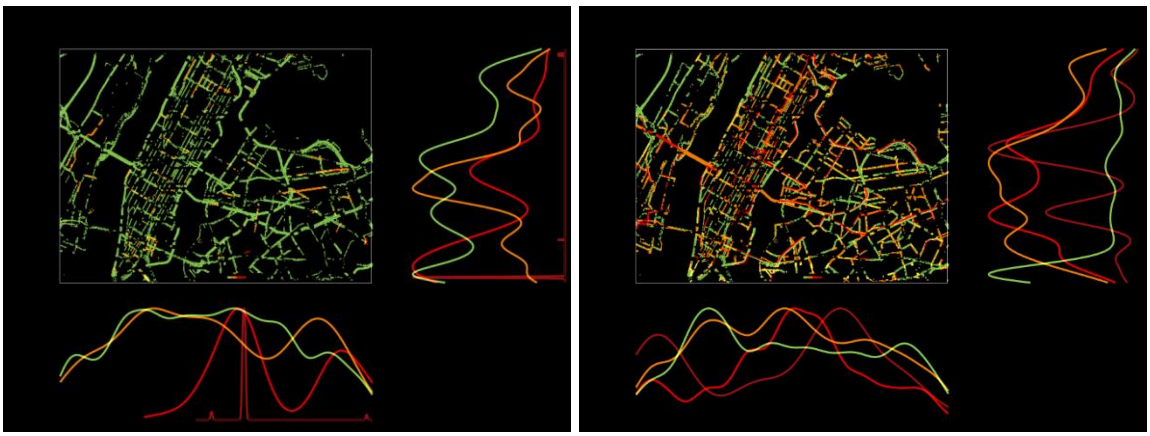


(b)

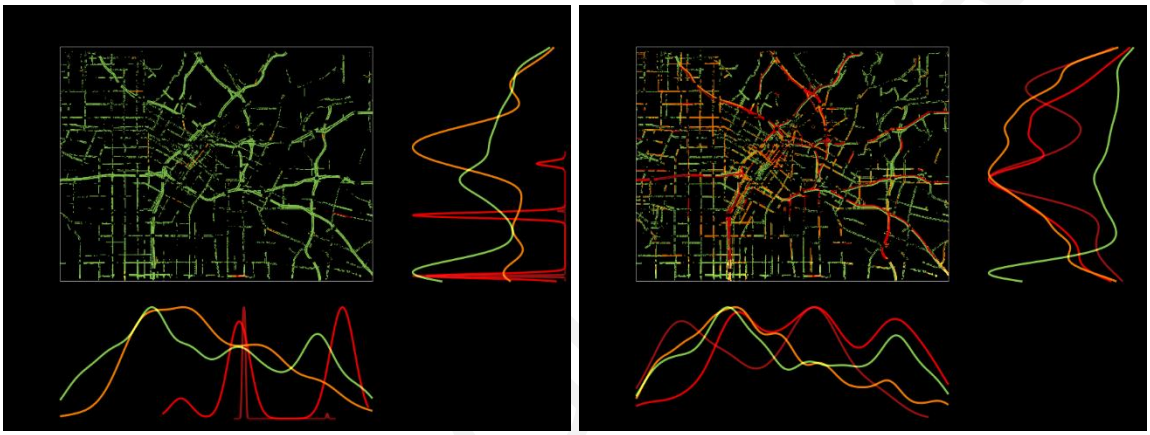


(c)

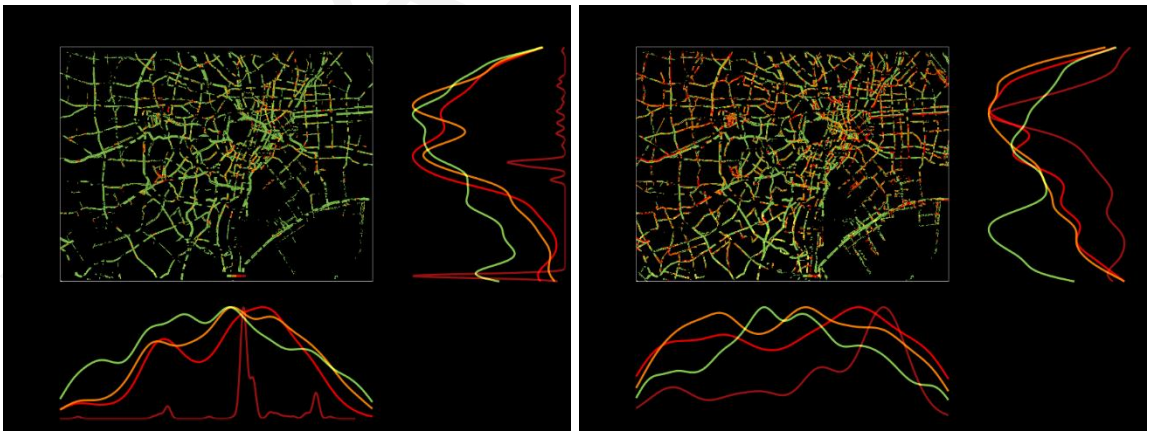
Figure 6.11:Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) Paris; (b) London; (c) Moscow.



(a)

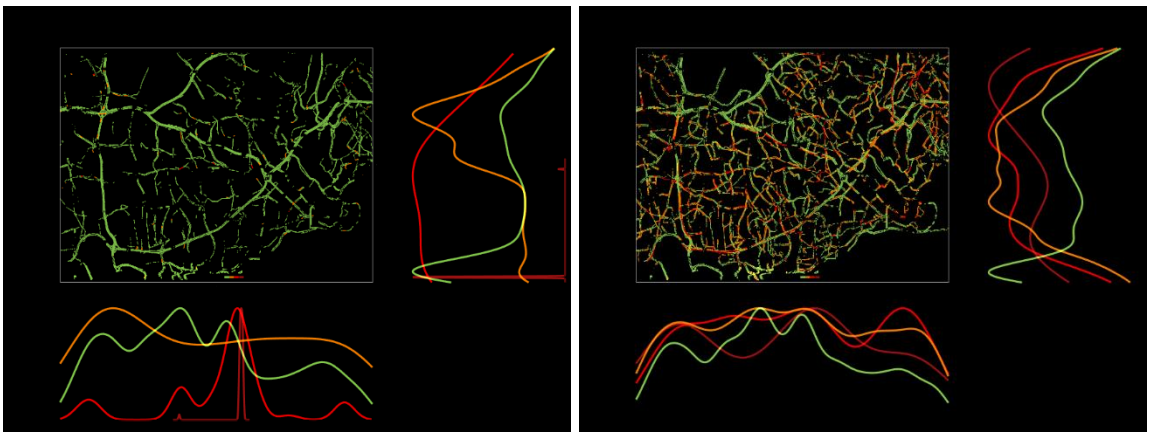


(b)

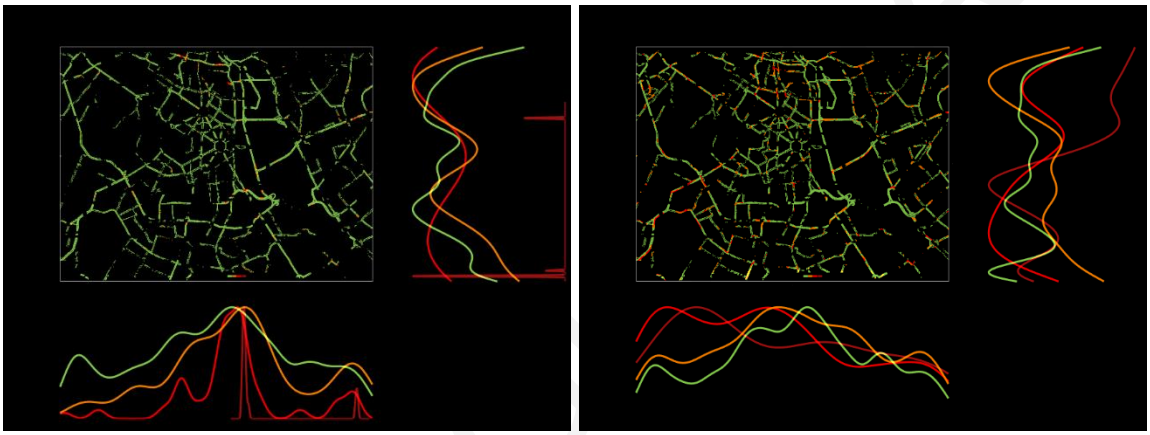


(c)

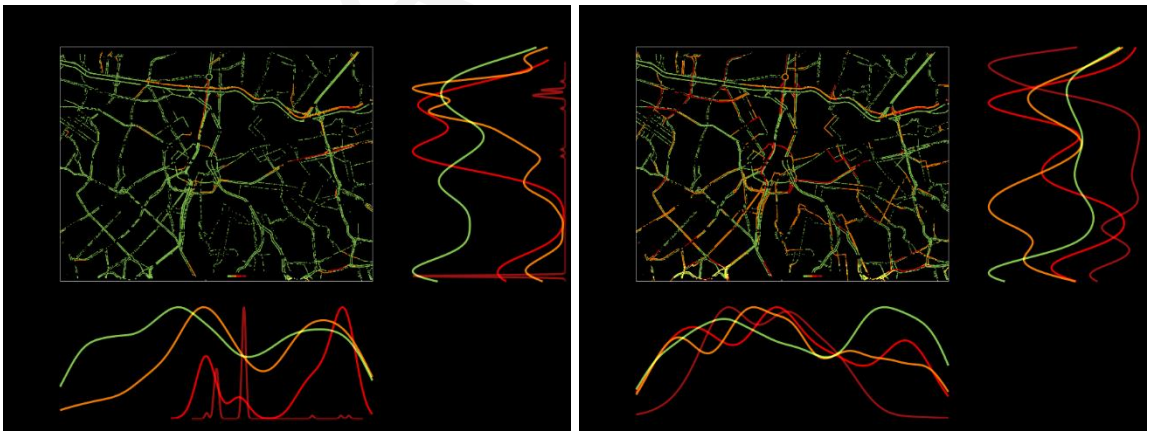
Figure 6.12:Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) New York; (b) Los Angeles; (c) Tokyo.



(a)

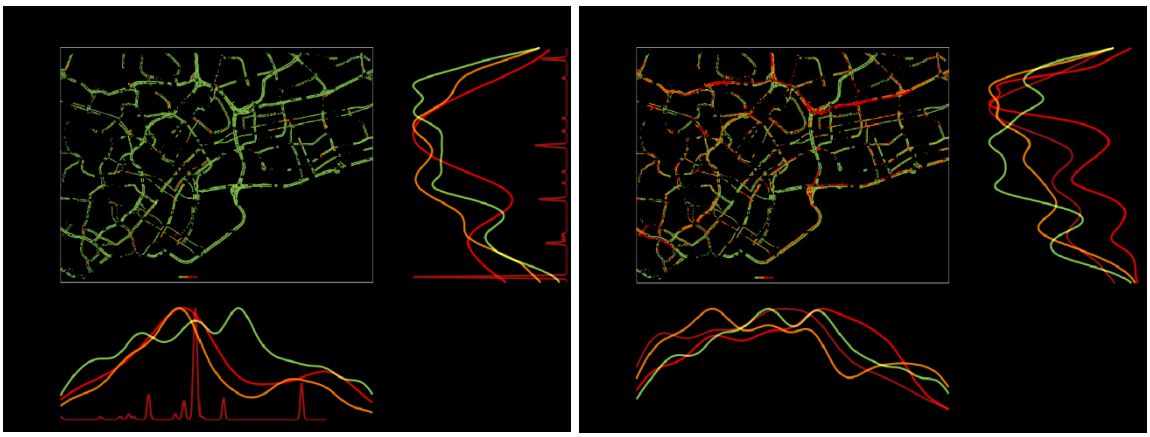


(b)

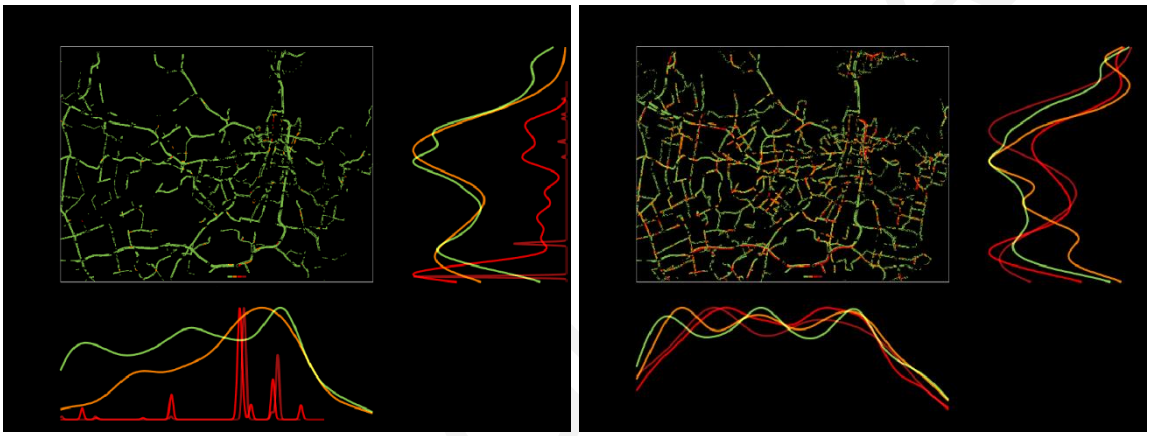


(c)

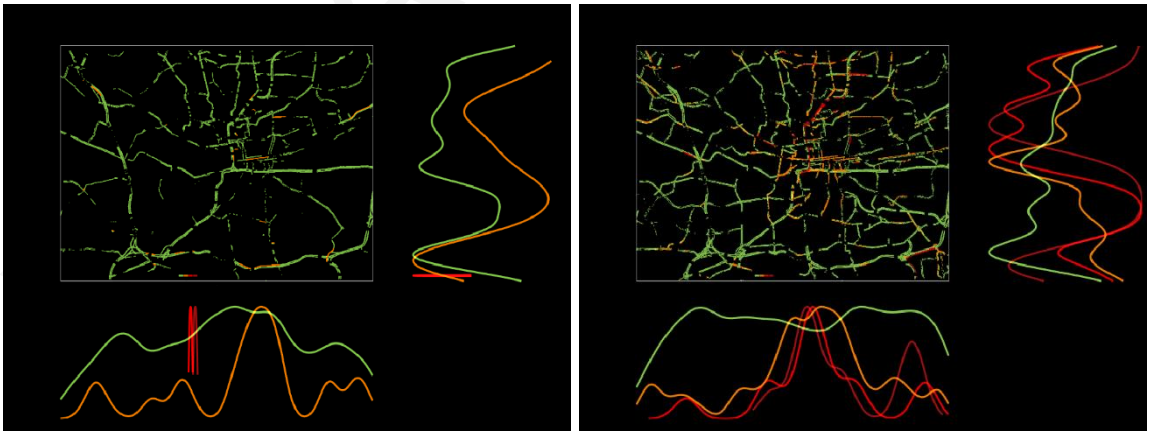
Figure 6.13: Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) Istanbul; (b) New Delhi; (c) Sao Paulo.



(a)

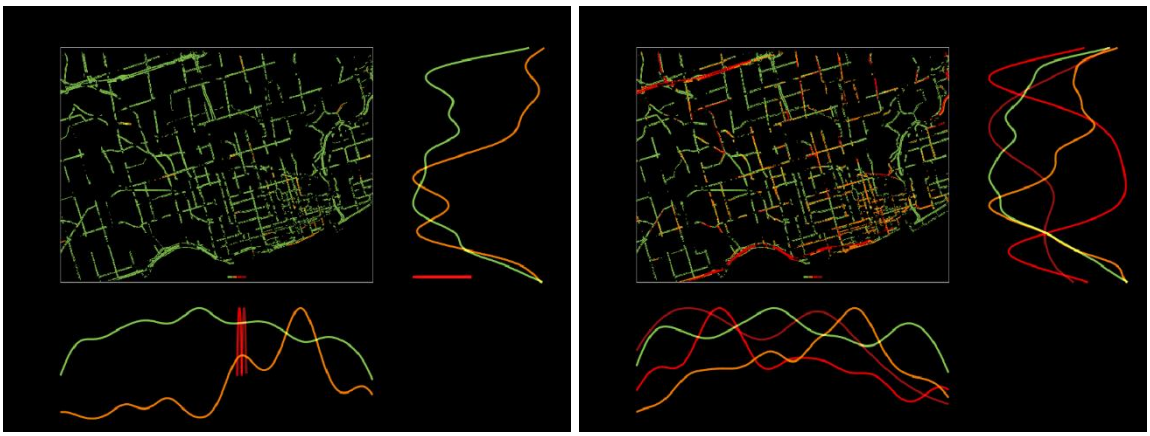


(b)

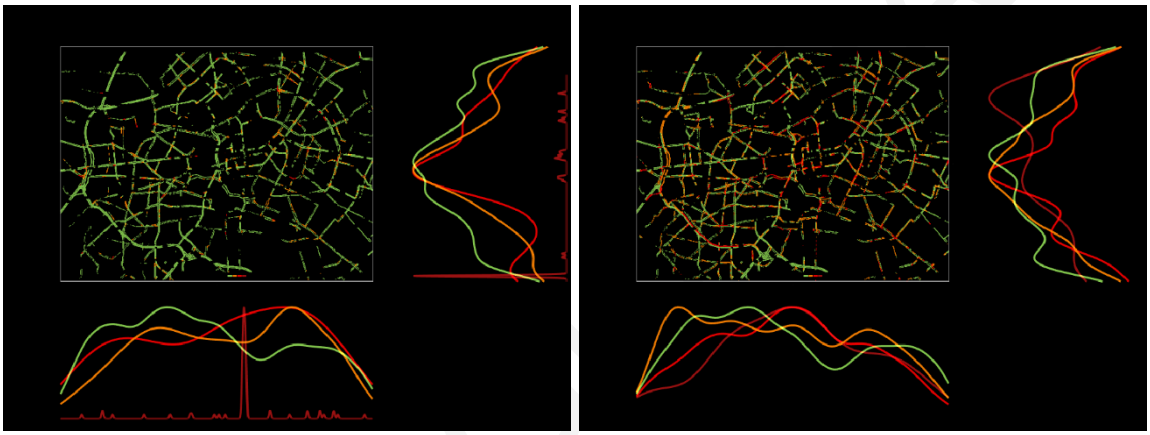


(c)

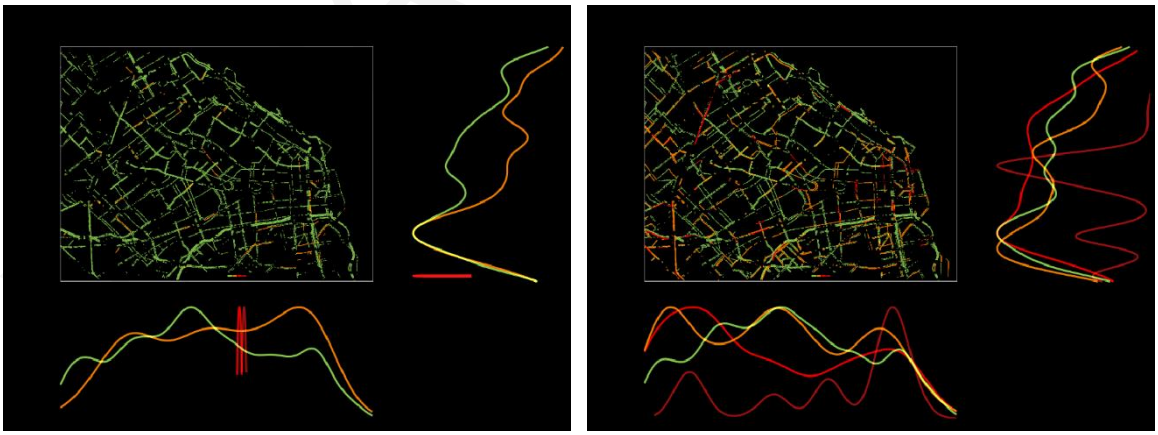
Figure 6.14: Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) Singapore; (b) Sydney; (c) Johannesburg



(a)

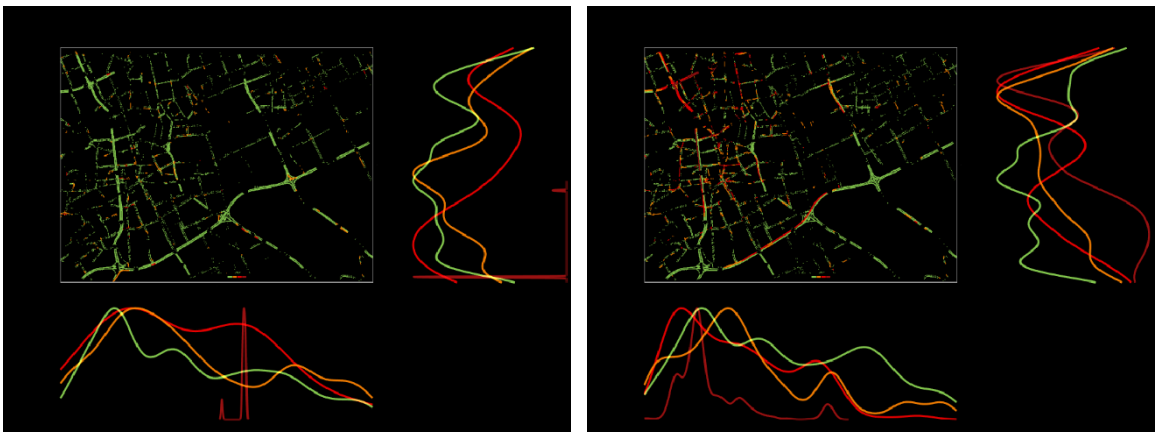


(b)

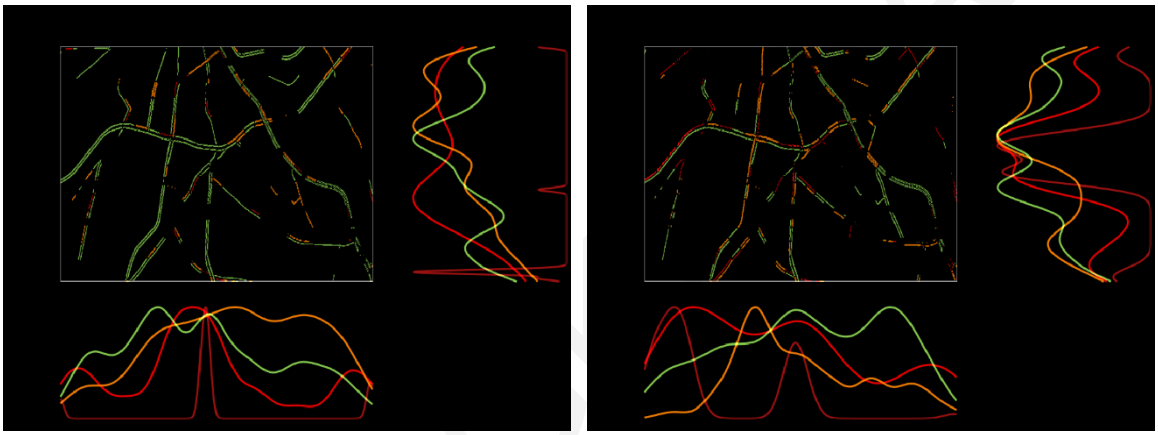


(c)

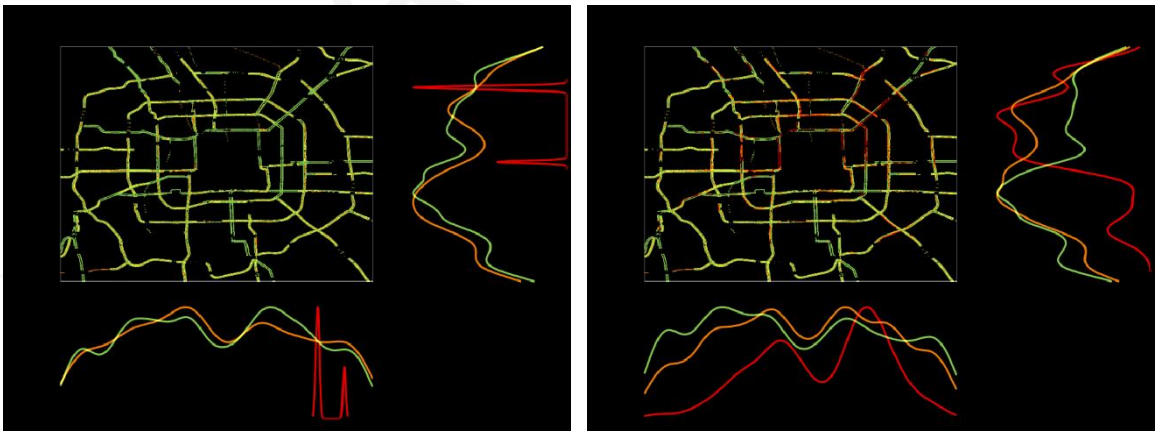
Figure 6.15: Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) Toronto; (b) Berlin; (c) Buenos Aires



(a)



(b)



(c)

Figure 6.16: Traffic layers' spatial distribution during off-peak hours (left) and peak hours (right) for (a) Riyadh; (b) Nicosia; (c) Beijing

Closing the Kernel Density plots, results for the cities of Riyadh, Nicosia and Beijing are presented in Figure 5.16. Here, several distinguished congested locations are observed for the city of Riyadh and two for the city of Beijing, respectively. Regarding the city of Nicosia, the selected road network represents the Business District Center outside of the Venetian Walls, where both horizontal and vertical streets exceeding their capacity during the morning and evening peak hours.

Previously, from the traffic information analysis (Section 4.4), we retrieved the urban space that remained uncongested across the cities through the estimation of the percentage of green pixels (light traffic). In this section, we further investigated the spatio-temporal dimension of mobility, by applying a spatial separation of the total requested activity and using the total pixels' marginal distributions per traffic layer based on Kernel Density estimation. The dynamic monitoring of these marginal distributions enabled a deeper understanding of the urban mobility patterns both in time and space, besides the two quantified indices (SSIM1 and SSIM2).

6.3 Dynamic Clustering and Propagation of Congestion

As a next step, clustering analysis is used to identify propagation of congestion and areas with similar congestion patterns within the city. Based on our approach these clusters emerge through image segmentation into perceptually meaningful atomic regions, known as superpixels (Achanta *et al.*, 2012). The superpixels algorithm applied in the dataset is the Simple Linear Iterative Clustering (SLIC). By default, the only parameter of the algorithm is the desired number of approximately equally sized superpixels. In our case, this number was set to 3500, resulting in a 14x14 pixels grid that corresponds to an urban area equal to 150m by 150m.

Figure 6.17 shows the steps followed to produce congestion patterns. The first step is the extraction of traffic information as previously described. Here only red and dark red traffic layers were used as they indicate heavy traffic and congestion. In the second step, the SLIC algorithm (see Algorithm 7, Appendix A) is applied and the initial image (step1) is divided into several regions. Although, two main clusters can be easily observed. The first one contains grid shape regions while the second consists of uniform regions in both size and shape. The grid shape area emerges due to the lack of traffic information (black pixels) while the shape changes in areas that pixels or their adjacent pixels are colored. In the next step, the color of each superpixel region is chosen according to pixel's maximum RGB value. Thus, the existence of dark red and red color dominates over black pixels and colors the

superpixel respectively. Last, in step 4, a combination of images in steps 2 and 3 gives a better visualization of the congestion patterns.

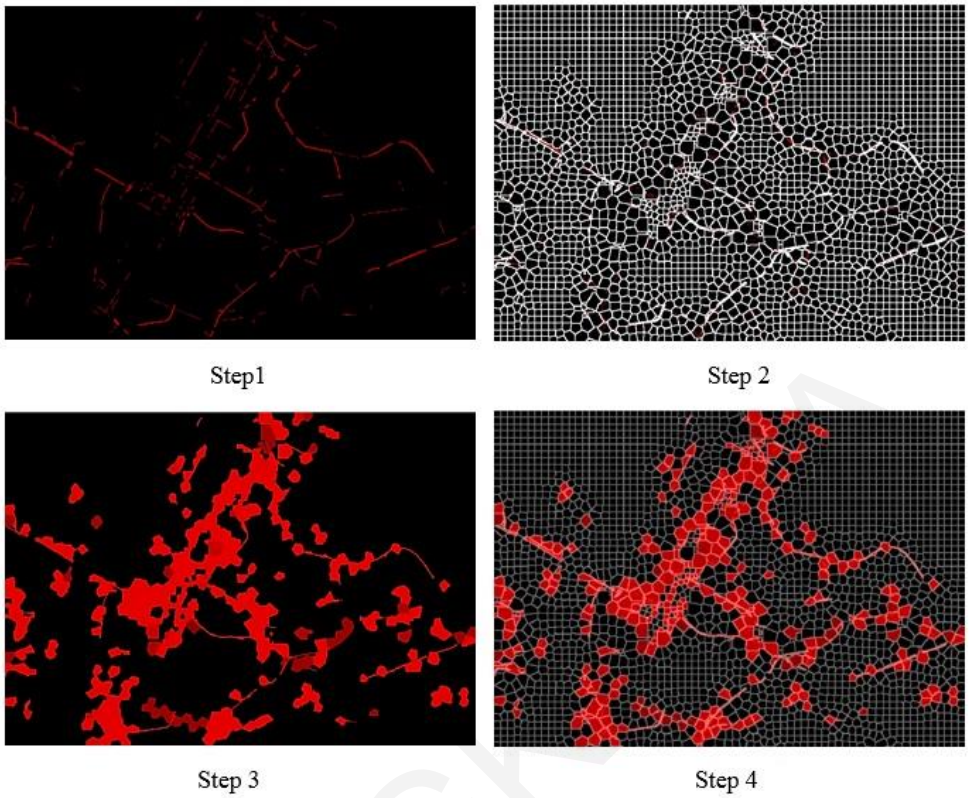


Figure 6.17: Steps followed to create congestion patterns: (Step 1) Heavy traffic layers isolation; (Step 2) Image segmentation to meaningful atomic regions by clustering pixels; (Step 3) Setting the color of each superpixel region; (Step 4) Congestion patterns visualization.

This transition from the linear perspective to regions/clusters with a similar level of congestion provides useful insights regarding the propagation of congestion both in space and time. In the following figures, we can see the variation of these patterns during the morning peak hours. Figure 6.18 and Figure 6.19 show the spatiotemporal variation of congestion for the city of Moscow(left) and Los Angeles(right), respectively. By comparing the two cities we can see the differences in propagation of congestion in space. In Moscow, the red clusters are scattered, and their number rapidly increases as we move to the next time interval whereas in Los Angeles congestion starts from the center and spread to the periphery during the morning peak hour.

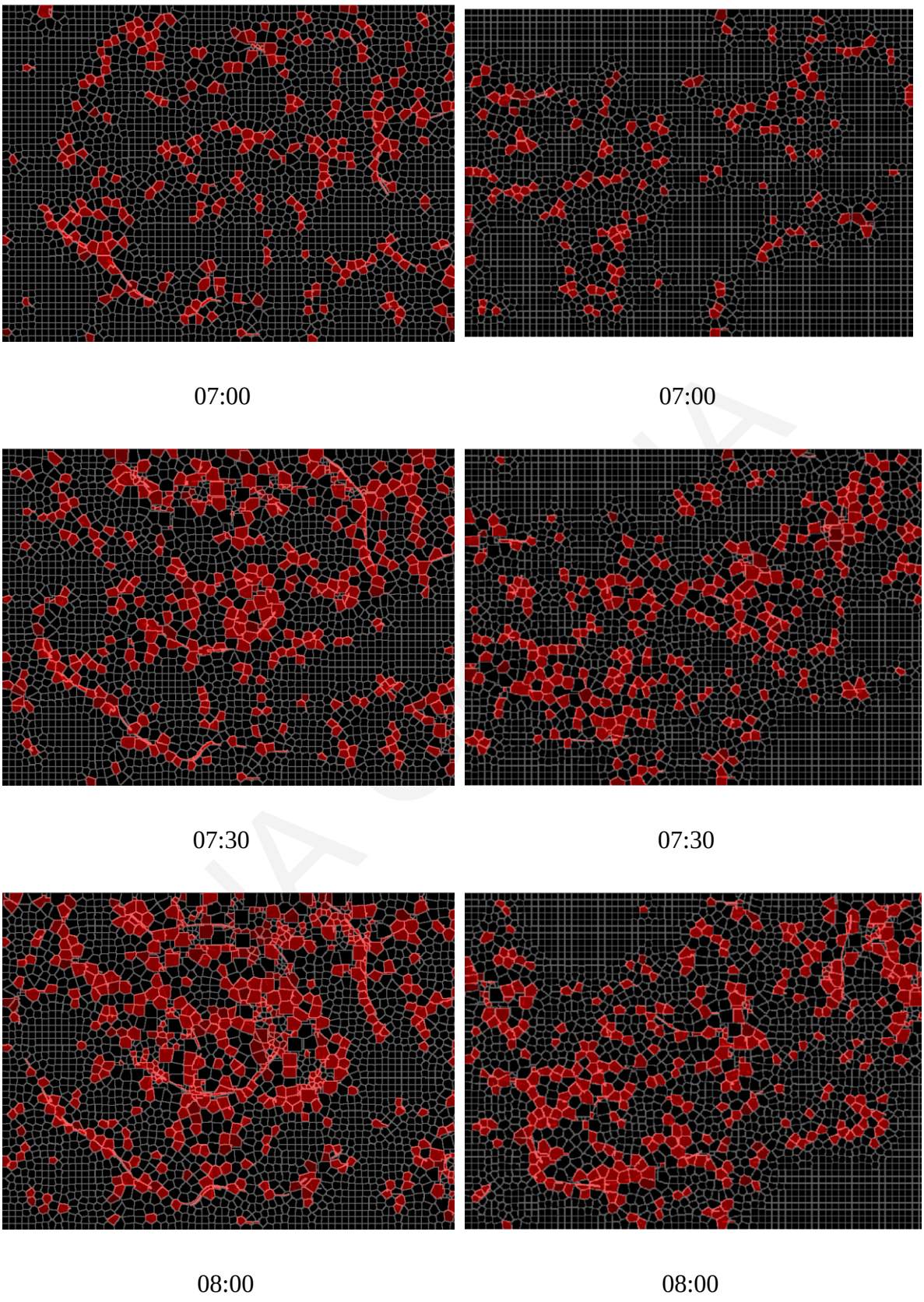
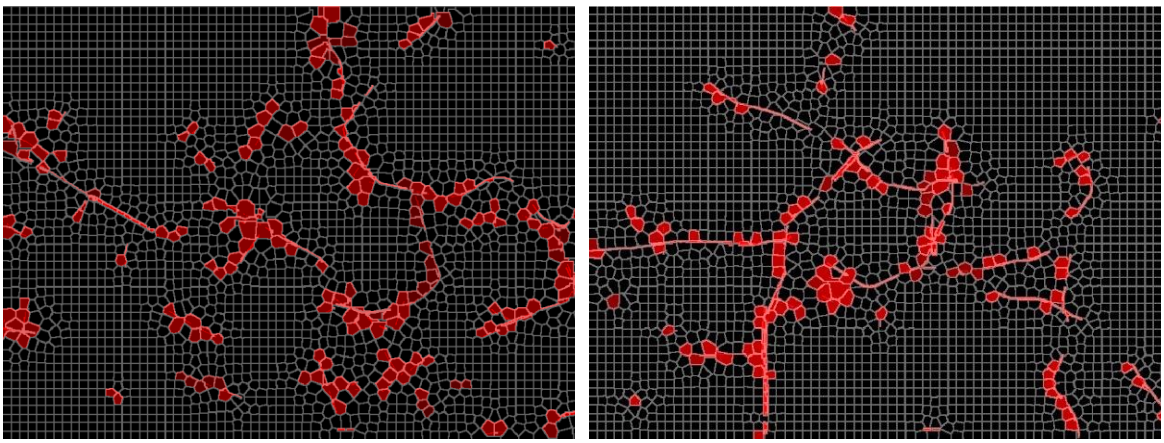
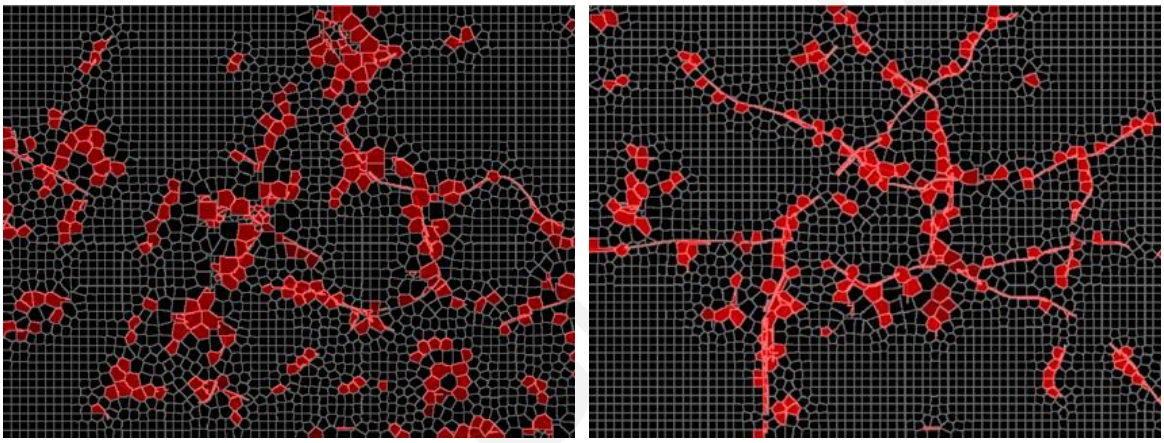


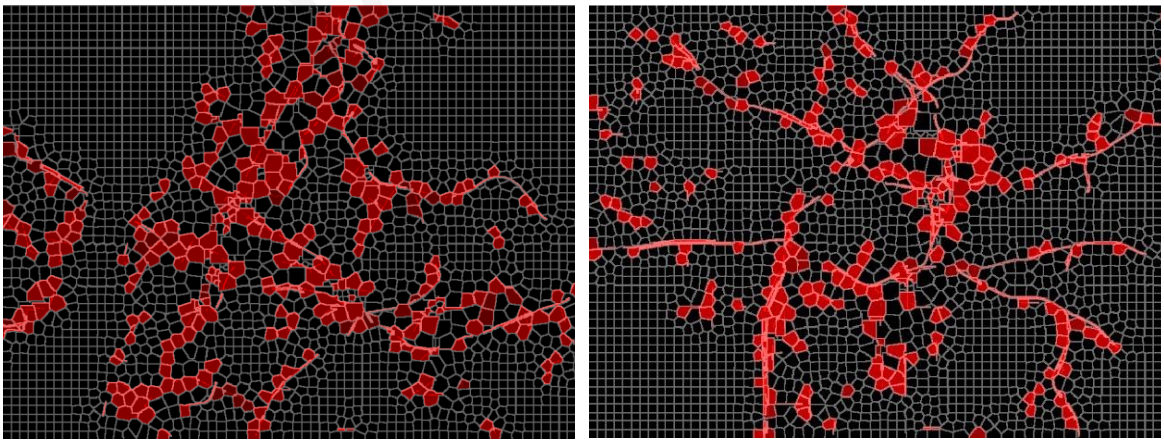
Figure 6.18: Spatio-temporal congestion patterns for Moscow (left) and Istanbul (right) during morning peak hours (07:00-08:00 a.m.).



07:00 07:00



07:30 07:30



08:00 08:00

Figure 6.19: Spatio-temporal congestion patterns for New York (left) and Los Angeles (right) during morning peak hours (07:00-08:00 a.m.).

Regarding the rest of the cities (Appendix E), the monitoring of the sequential images (Step 4) revealed similar congestion patterns to Los Angeles for the city of London and Tokyo where congestion starts from the city of London and Tokyo port, respectively and moves in the inner of the city. Paris' congestion pattern resembles Moscow's as congestion clusters were scattered in space. On the contrary, in New York, the congestion spread from the borough of Queens, Bronx, and Brooklyn towards the core of New York, Manhattan for the same time intervals.

6.4 Chapter Summary

In the current chapter the temporal variation of mobility patterns was firstly analysed revealing a broadly comparable rhythm across the cities. As a next step, emphasis was given to the spatiotemporal patterns. Kernel density plots were applied to investigate the homogeneity of activity within the city. Results showed that the selected cities share a common rhythm of mobility during the day while a closer look at the proposed indices revealed a dissimilar pattern for the city of Tokyo. The spatial repartition of mobility among the eighteen cities revealed a concentric organization, with strong activity level in the city center for most of the selected cities. Furthermore, clustering analysis was employed to identify propagation of congestion and areas with similar congestion patterns within the city. The transition from the linear perspective to regions/clusters with a similar level of congestion provided useful insights regarding the propagation of congestion both in space and time. Comparing the cities, Moscow and Paris congestion patterns were scattered and their number rapidly increased as moving to the next time interval whereas in Los Angeles, London and Tokyo the congestion started from the center and spread to the periphery during the morning peak hour. On the other hand, New York appeared a differential congestion pattern as strong activity was moving from the periphery boroughs to the core of New York, Manhattan.

Chapter 7 : Conclusions and Future Directions

7.1 Conclusions

This thesis has presented a new method to analyse and visualize urban dynamics, utilizing aggregated information from online traffic maps. As, it was mentioned in the literature overview in **Chapter 2**, while researchers have explored a wide variety of different mobility data, most of the studies focus on disaggregated data from traffic surveillance systems or mobile phones that are not available or free to all interested parties. In the light of other data sources limitation (access, cost, bias, coverage) to provide a network-level perspective, the capturing of macroscopic phenomena utilizing online traffic maps' information worldwide has great value especially in the era of Big Data, both for traffic monitoring, control and management of urban resources. The significance of this method can be summarized in the synthetic nature of mobility data that online maps depict, and in the extension of the application from one limited spatial entity (in general a city) to global coverage, given its widespread availability.

As, online traffic maps display condensed traffic information for the sake of readability, open research questions remain to be answered regarding the reduction of information that is performed in this display process. As it was demonstrated in **Chapter 5**, valid traffic characteristics, such as the MFD properties, still possess, that means that the data retrieved from online maps can be further utilised in urban mobility modeling especially in the scale of large and complex urban areas.

Specifically, the proposed method was based on a traffic information extraction module for realistic networks, where traffic layers are captured by straightforward image processing techniques and processed as categorical type of data, through a back-engineering approach. Fundamental traffic relationships are then used, connecting microscopic information to macroscopic phenomena. To investigate the quality of traffic information depicted in online dynamic traffic maps, the estimation and the properties of MFD were used. The application of the proposed methodological framework was tested on commercial freely available online maps from different providers and for different urban road networks. As a next step, in **Chapter 6** data retained from maps were used to analyse and visualize human mobility patterns (temporal, Spatio-temporal) for the selected cities.

Ultimately, this work makes the following specific contributions related to the research objectives formulated in **Chapter 1:**

a. *To gain a deep understanding of online traffic maps by investigating fundamental traffic flow characteristics and relationships that can be extracted by the aggregated type of traffic information used in them*

Initially, a deeper understanding of aggregated network traffic displayed on online dynamic traffic maps was achieved by investigating fundamental traffic flow characteristics and relationships that can be extracted by the aggregated type of traffic information used in them. To accomplish that, a representative study area and a vast dataset of numerous raster images of maps was utilised, presented in **Chapter 4** while fundamental characteristics of traffic flow and the relationships between these parameters were introduced in **Chapter 3**. In **Chapter 5** a thorough investigation about the value of spatiotemporal information that it is retained in the abstracted depiction of traffic conditions has been conducted, revealing that the fundamental macroscopic properties still hold even in cases of significant information reduction like the representation scheme adopted in online traffic maps. Not only the shapes and the magnitude of scale among the traffic variables are consistent but also valuable additional information can be extracted from traffic maps (like vehicles concentration, network heterogeneity, etc.) and delicate phenomena can be captured (like traffic hysteresis).

b. *To form the mathematical connection between the discretized/digitized map images on seamless pixels and the aggregated traffic variables across entire urban areas,*

Initially, in **Chapter 4** a connection between the number of pixels depicted in a raster image for the same zoom level, and the corresponding real linear distance on the road network, was achieved by forming an equation among the scale of the map, the number of pixels corresponding to the total number of pixels of the scale bar of the map and the real distance. Then this equation was applied on the dataset, and the urban coverage through maps was quantified by estimating the total road length in kilometres (Km) for each road network. In **Chapter 5** a mathematical connection between the average network flow, the average network speed, the average network density, and the number of pixels that belong to distinctive operating class of online traffic maps was established. Considering that, typical online maps offer information about links' average speed per coloured segment (complete links or link's stretches), the estimation of the network average speed U^t at time interval t was feasible by calculating the average speed form all the coloured links in the network at each time interval. Regarding the representative value of average speed (U_{ci}) according to colour c, ground truth data form loop detectors were used to assign representative values of average speed per traffic layer through empirical calibration, as described in **Chapter 5.**

Then a parametric approach was followed based on the fundamental linear speed-density relation proposed by Greenshields (Greenshields, 1935), to estimate the network's average density. A nonlinear speed-density relationship based on Greenberg model was also used for comparison, resulting to significantly lower values of accumulation for the same networks.

c. *To describe urban mobility through quantified indices and models that can be applied to every city that has online traffic coverage,*

In order to capture and monitor the daily time variations of mobility across cities, image processing techniques were applied and demonstrated for the whole dataset as presented in **Chapter 6**. Specifically, a method for measuring the similarity between two images, known as the Structural Similarity Index (SSIM), was implemented utilizing two different indices SSIM1 and SSIM2. The SSIM1 index at each time interval measures the difference on the level of mobility, compared to a stable traffic state in a city and enables the dynamic monitor of mobility patterns' variation, compared to the off-peak hour state of the network. In this respect, urban planners and traffic operators can explore the mobility profile of a city and measure the impacts on human mobility due to traffic congestion. By using, an initial image during off-peak hour as a baseline, the SSIM1 index at each time interval measures the difference on the level of mobility, compared to a stable traffic state in a city. The second index, SSIM2 compares image x to image y, where x corresponds to the image taken on time interval t_{k+1} and y to the image taken on time interval t_k. This index can be developed into an enhanced tool for monitoring as historical profiles of this index can be used to compare traffic anomalies (steep drops or peaks of the index) and their impact on human mobility for the entire city.

Emphasis was also given to the spatiotemporal patterns. Kernel density plots were applied to investigate the homogeneity of activity within the city. Results showed that the selected cities share a common rhythm of mobility during the day while a closer look at the proposed indices revealed a dissimilar pattern for the city of Tokyo. Furthermore, clustering analysis was employed to identify propagation of congestion and areas with similar congestion patterns within the city. The transition from the linear perspective to regions/clusters with a similar level of congestion provided useful insights regarding the propagation of congestion both in space and time.

d. *To establish a concrete connection between human mobility research and real-time traffic information.*

Endeavours to analyse human mobility are mired by their volume, structure, completeness, noise and other pitfalls associated with the use of big data sets. As such, high levels of aggregation are generally required to support planning and decision-making. Thus, online traffic maps that wrangle raw data into meaningful information are a valuable source that should be further utilised and explored in human mobility research. The proposed method in this thesis paved the way towards this respect.

7.2 Future Directions

This work is just a first step towards a goal of creating a new set of tools for urban operators, transport planners and policy makers to monitor complex urban issues based on information retrieved from online traffic maps.

Regarding the validity of the proposed method, this was tested by comparing the estimated MFDs to ground truth MFD obtained using empirical data from loop detectors at a small scale and only for one city. Thus, it is crucial, as future work to expand the validation testing both at different levels of scales and for different urban networks, utilizing available empirical data or measurements.

Overall, directions for further research can be organised into three main categories:

i. Acquisition and evaluation of the rest information that online maps offer

In **Chapter 4** we presented a method to extract dynamic traffic information from map images by distinguishing and then capturing the traffic layers from the rest of the map. Although, online traffic maps, provide further information regarding typical traffic state per day. The acquisition of this information for investigation is highly recommended for comparative reasons, as it can be used as a baseline to identify non-recurring incidents or shock-events that temporarily reduce road capacity in an urban area. Another valuable information that online traffic maps offers, stands for travel time between a predefined set of Origin-Destinations (O-D) locations. These travel times could be used as input data in order to determine travel time distributions (not individual values/predictions) on a rolling horizon, based on a method presented in (Dimitriou and Gkania, 2016) for large urban-scale networks.

ii. Connection of visual analytics of human mobility with traffic phenomena

In **Chapter 5** the validity of the simplified/coded information that dynamic traffic maps provide was investigated in terms of traffic operational characteristics. Particularly, the estimation and the investigation of the Macroscopic Fundamental Diagram-MFD properties

were performed utilizing the Greenshields and Greenberg model. Although, different macroscopic models could also be tested, such as the Two-fluid model, that quantifies traffic performance on a network by studying the interaction between moving and stopped vehicles in the traffic stream.

In **Chapter 6** the daily variation of mobility was quantified using computer vision and image processing techniques. The estimated indices and the spatio-temporal patterns of mobility could be further connected with specific traffic phenomena such as traffic breakdown, hysteresis, stop-and-go traffic and synchronized flow.

iii. Application of machine learning methods

As machine learning is a class of methods for automatically creating models from data, images from online traffic maps could be utilized for classification of mobility patterns in an urban area and for prediction of congested areas within the cities.

Last, it could be said that, online traffic maps constitute a new set of lenses, providing tremendous opportunities to examine urban mobility dynamics. Numerous visualizations tools and models could be used apart from Kernel Densities diagrams or the SSIM indices, and SLIC algorithm proposed in this thesis for classification, prediction or monitoring of urban dynamics.

Publications

Book Chapters

1. Gkania, V., & Dimitriou, L. (2019). A Back-Engineering Approach to Explore Human Mobility Patterns Across Megacities Using Online Traffic Maps. In Mobility Patterns, Big Data and Transport Analytics (pp. 345-363). Elsevier.

Journal Papers

1. Gkania, V., & Dimitriou, L. (2020). Linking the microscopic traffic flow mechanics with the macroscopic phenomena by exploiting class-type traffic information retrieved from online traffic maps. In 23rd EURO Working Group on Transportation Meeting, EWGT 2020, Paphos, Cyprus.
2. Dimitriou, L., & Gkania, V. (2016). Dynamic Short-Term Projections of Travel Time Distributions in Urban Signalized Networks utilizing Composite Information of Traffic Characteristics. IFAC-Papers Online, 49(3), 237-242.
3. Dimitriou, L., Kourti, E., Christodoulou. C., & Gkania. V. (2016). 'Dynamic Estimation of Optimal Dispatching Locations for Taxi Services in Mega-Cities based on Detailed GPS Information', *IFAC-PapersOnLine*. Elsevier B.V., 49(3), pp. 197–202. doi: 10.1016/j.ifacol.2016.07.033.
4. Gkania, V., & Dimitriou, L. (2020). Uncovering Human Mobility Patterns in Cities by applying Image Processing Techniques on Traffic Maps submitted to PLOS ONE (under review).
5. Dimitriou, L., & Gkania, V. (2020). Integrating Location-specific Flow Mechanics for Analyzing Network-wide Operational States on Online Traffic Maps submitted to Transportation Research Part C-Emerging Technologies (under review).

Peer Reviewed Conference Papers

1. Gkania, V., & Dimitriou, L. (2020). Estimation of Macroscopic Fundamental Diagrams Based on Information Retrieved from Online Traffic Maps: Calibration, Empirical Data and Validation. In 99th Annual Meeting of the Transportation Research Board; Washington, D.C.
2. Gkania, V., & Dimitriou, L. (2019). Uncovering human mobility patterns in cities through the lens of traffic maps. In 98th Annual Meeting of the Transportation Research Board; Washington, D.C.
3. Gkania, V., & Dimitriou, L. (2019). Identification of macroscopic network performance and phase transition phenomena in urban areas based on information retrieved by online traffic maps. In 98th Annual Meeting of the Transportation Research Board; Washington, D.C.
4. Gkania, V., & Dimitriou, L. (2017). Prediction of traffic characteristics in smart cities based on deep learning mechanisms. In eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2016: Proceedings of the 11th European

Conference on Product and Process Modeling (ECPPM 2016), Limassol, Cyprus, 7-9 September 2016 (p. 283). CRC Press.

5. Gkania, V., & Dimitriou, L. (2017). Dynamic Short-Term Projections of Travel Time Distributions Based on Heterogeneous Traffic Data (No. 17-03354). In 96th Annual Meeting of the Transportation Research Board; Washington, D.C.
6. Dimitriou, L., & Gkania, V. (2016). Adaptive short-term prediction of freeways travel-time distributions based on multivariate hazard models. In 2016 18th Mediterranean Electrotechnical Conference (MELECON) (pp. 1-6). IEEE.

Working Papers

1. Gkania, V., & Dimitriou, L. A Review on the Evolution of Human Mobility Mapping and Data

Online repository material

1. https://www.youtube.com/watch?v=U0o0WwuxaGw&t=3s
2. https://www.youtube.com/watch?v=wDr_haEBl1E
3. https://www.youtube.com/watch?v=OXq3PwNQ3Lw
4. https://www.youtube.com/watch?v=_f7QP3bPIUg
5. https://www.youtube.com/watch?v=hh6DCOqNebc
6. https://www.youtube.com/watch?v=HFbKih31w-Q
7. https://www.youtube.com/watch?v=JVhlwiFM1Xo
8. https://www.youtube.com/watch?v=7VmDlgVAYLE&t=3s
9. https://www.youtube.com/watch?v=czEVIHzmc-o&t=17s
10. https://www.youtube.com/watch?v=DHEtfdGG9mg
11. https://www.youtube.com/watch?v=G9O7CFaTyeY&t=26s
12. https://www.youtube.com/watch?v=76l5689wiR8
13. https://www.youtube.com/watch?v=Q-lkveHh5_U&feature=youtu.be
14. https://www.youtube.com/watch?v=Y3d39rdH6Oo&feature=youtu.be

References

Achanta, R. *et al.* (2012) 'SLIC superpixels compared to state-of-the-art superpixel methods', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11), pp. 2274–2281. doi: 10.1109/TPAMI.2012.120.

Alexander, L. *et al.* (2015) 'Origin–destination trips by purpose and time of day inferred from mobile phone data', *Transportation Research Part C: Emerging Technologies*. Pergamon, 58, pp. 240–250. doi: 10.1016/J.TRC.2015.02.018.

Ambühl, L. and Menendez, M. (2016) 'Data fusion algorithm for macroscopic fundamental diagram estimation', *Transportation Research Part C: Emerging Technologies*, 71, pp. 184–197. doi: 10.1016/j.trc.2016.07.013.

Amirgholy, M. and Gao, H. O. (2017) 'Modeling the dynamics of congestion in large urban networks using the macroscopic fundamental diagram: User equilibrium, system optimum, and pricing strategies', *Transportation Research Part B: Methodological*. Elsevier Ltd, 104, pp. 215–237. doi: 10.1016/j.trb.2017.07.006.

Ampountolas, K. and Kouvelas, A. (2015) 'Real-Time Estimation Of Critical Values Of The Macroscopic Fundamental Diagram For Maximum Network Throughput', *In Transportation Research Board 94th Annual Meeting*, No. 15-177, pp. 1–15.

Andrienko, G. *et al.* (2011) 'A conceptual framework and taxonomy of techniques for analyzing movement', *Journal of Visual Languages & Computing*. Academic Press, 22(3), pp. 213–232. doi: 10.1016/J.JVLC.2011.02.003.

Andrienko, G. *et al.* (2017) 'Visual analytics of mobility and transportation: State of the art and further research directions', *IEEE Transactions on Intelligent Transportation Systems*, 18(8), pp. 2232–2249. doi: 10.1109/TITS.2017.2683539.

Andrienko, N. and Andrienko, G. (2013) 'Visual analytics of movement: An overview of methods, tools and procedures', *Information Visualization*. SAGE PublicationsSage UK: London, England, 12(1), pp. 3–24. doi: 10.1177/1473871612457601.

Ardekani, S. and Herman, R. (1987) 'Urban Network-Wide Traffic Variables and Their Relations', *Transportation Science*. INFORMS, 21(1), pp. 1–16. doi: 10.1287/trsc.21.1.1.

Ardekani, S., Torres-Verdin, V. and Herman, R. (1985) 'The Two-Fluid Model and Traffic Quality in Mexico City (El Modelo Bifluido y La Calidad Del Transito En La Ciudad de Mexico)', *Revista Ingenieria Civil*.

Banks, J. H. (1991) 'Two-capacity phenomenon at freeway bottlenecks: A basis for ramp metering', *Transportation Research Record*, 1320, pp. 83–90.

Barth, D. (2009) *Official Google Blog: The bright side of sitting in traffic: Crowdsourcing road congestion data*, *Official Blog Google*. Available at: https://googleblog.blogspot.com/2009/08/bright-side-of-sitting-in-traffic.html (Accessed: 2 December 2020).

Bayir, M. A., Demirbas, M. and Eagle, N. (2010) 'Mobility profiler: A framework for discovering mobility profiles of cell phone users', *Pervasive and Mobile Computing*. Elsevier, 6(4), pp. 435–454. doi: 10.1016/J.PMCJ.2010.01.003.

Bazzani, A. *et al.* (2011) 'Towards Congestion Detection in Transportation Networks Using GPS Data', *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*. IEEE, (May 2010), pp. 1455–1459. doi: 10.1109/PASSAT/SocialCom.2011.249.

Beck, H. (2016) *Harry Beck's Tube map - Transport for London*. Available at: https://tfl.gov.uk/corporate/about-tfl/culture-and-heritage/art-and-design/harry-becks-tube-map (Accessed: 9 February 2018).

Beibei, J. Y., van Zuylen, H. J. and Shoufeng, L. (2016) 'Determining the Macroscopic Fundamental Diagram on the Basis of Mixed and Incomplete Traffic Data', *In Transportation Research Board 95th Annual Meeting*, No. 16-260.

Bell, M. and Ward, G. (1998) 'Patterns of temporary mobility in Australia: Evidence from the 1991 Census', *Australian Geographical Studies*. Blackwell Publishers Ltd, 36(1), pp. 58–81. doi: 10.1111/1467-8470.00039.

Bhattacharya, J. (2016) *RUDIMENTS OF MODERN COMPUTER APPLICATION: PART 1*. Academic Publishers.

Bonavia, D. M. (2020) *Beijing | Province, City, Capital, & History | Britannica*, *Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Beijing (Accessed: 17 May 2020).

Bonilla, J. (2019) *Buenos Aires | History, Population, Map, Culture, & Facts | Britannica*, *Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Buenos-Aires/Climate#ref59302 (Accessed: 16 May 2020).

Borruso, G. (2008) 'Network density estimation: A GIS approach for analysing point patterns in a network space', *Transactions in GIS*. Wiley/Blackwell (10.1111), 12(3), pp. 377–402. doi: 10.1111/j.1467-9671.2008.01107.x.

Bowman, A. W. and Azzalini, A. (1997) *Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations*. OUP Oxford.

Buisson, C. and Ladier, C. (2009) 'Exploring the Impact of Homogeneity of Traffic Measurements on the Existence of Macroscopic Fundamental Diagrams', *Transportation Research Record: Journal of the Transportation Research Board*. Transportation Research Board of the National Academies, 2124, pp. 127–136. doi: 10.3141/2124-12.

Calabrese, F. *et al.* (2011) 'Rome. Real-Time Urban Monitoring Using Cell Phones: A Case Study', *IEEE Transactions on Intelligent Transportation Systems*, 12(1), pp. 141–151. doi: 10.1109/TITS.2010.2074196.

Calabrese, F. *et al.* (2013) 'Understanding individual mobility patterns from urban sensing data: A mobile phone trace example', *Transportation Research Part C: Emerging Technologies*. Pergamon, 26, pp. 301–313. doi: 10.1016/j.trc.2012.09.009.

Campbell, J. T. (2020) *Johannesburg | City, History, & Points of Interest | Britannica*, *Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Johannesburg-South-Africa (Accessed: 16 May 2020).

Cao, G. *et al.* (2015) 'A scalable framework for spatiotemporal analysis of location-based social media data', *Computers, Environment and Urban Systems*. Pergamon, 51, pp. 70–82. doi: 10.1016/J.COMPENVURBSYS.2015.01.002.

Cheng, Z. *et al.* (2011) 'Exploring Millions of Footprints in Location Sharing Services', *Icwsm*, 2010(Cholera), pp. 81–88. doi: papers3://publication/uuid/0C46BD5D-4908-4A8A-BD06-5BCB2F1DE282.

Chiang, Y. Y. and Knoblock, C. A. (2009) 'A method for automatically extracting road layers from raster maps', *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, pp. 838–842. doi: 10.1109/ICDAR.2009.274.

Clark, J. (2012) *Movement in Manhattan, Neoformix*. Available at: https://neoformix.com/2012/MovementInManhattan.html (Accessed: 13 February 2018).

Claudel, M., Nagel, T. and Ratti, C. (2016) 'From origins to destinations: The past, present and future of visualizing flow maps', *Built Environment*, 42(3), pp. 338–355. doi: 10.2148/benv.42.3.338.

Clout, H. D. (2019) *London | History, Maps, Points of Interest, & Facts | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/London (Accessed: 14 October 2020).

Cox, W. (2019) *DEMOGRAPHIA WORLD URBAN AREAS 15 th ANNUAL EDITION*.

Csáji, B. C. *et al.* (2013) 'Exploring the mobility of mobile phone users', *Physica A: Statistical Mechanics and its Applications*. North-Holland, 392(6), pp. 1459–1473. doi: 10.1016/j.physa.2012.11.040.

CTPP (2018) *CTPP 2006-2010 Census Tract Flows Query Tutorial, Federal Highway Administration*. Available at: https://www.fhwa.dot.gov/planning/census_issues/ctpp/data_products/2006-2010_tract_flows/index.cfm (Accessed: 2 March 2019).

Daganzo, C. F. (2007) 'Urban gridlock: Macroscopic modeling and mitigation approaches', *Transportation Research Part B: Methodological*, 41(1), pp. 49–62. doi: 10.1016/j.trb.2006.03.001.

Dimitriou, L. *et al.* (2016) 'Dynamic Estimation of Optimal Dispatching Locations for Taxi Services in Mega-Cities based on Detailed GPS Information', *IFAC-PapersOnLine*. Elsevier, 49(3), pp. 197–202. doi: 10.1016/J.IFACOL.2016.07.033.

Dimitriou, L. and Gkania, V. (2016) 'Dynamic Short-Term Projections of Travel Time Distributions in Urban Signalized Networks utilizing Composite Information of Traffic Characteristics', *IFAC-PapersOnLine*, 49(3), pp. 237–242. doi: 10.1016/j.ifacol.2016.07.040.

Edie, L. C. (1961) 'Car-Following and Steady-State Theory for Noncongested Traffic', *Operations Research*. Institute for Operations Research and the Management Sciences (INFORMS), 9(1), pp. 66–76. doi: 10.1287/opre.9.1.66.

Ehrlich, B. (2019) *Paris | Definition, Points of Interest, Facts, & History | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Paris#ref59646 (Accessed: 14 May 2020).

Ehrlich, B. (2020) *Istanbul | History, Points of Interest, & Map | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Istanbul (Accessed: 16 May 2020).

Ferreira, N. *et al.* (2013) 'Visual exploration of big spatio-temporal urban data: A study of New York city taxi trips', *IEEE Transactions on Visualization and Computer Graphics*, 19(12), pp. 2149–2158. doi: 10.1109/TVCG.2013.226.

French, R. A. (2019) *Moscow | History, Geography, & Map | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Moscow (Accessed: 16 May 2020).

Freundlich, B. (2011) 'The History of Cartography Project', *Cartography and Geographic Information Science*. University of Chicago Press, 38(3), pp. 341–343. doi: 10.1559/15230406382341.

Friendly, M. (1975) ‘A Brief History of Data Visualization’, *Handbook of Data Visualization*, pp. 15–56. doi: 10.1007/978-3-540-33037-0_2.

García-Ramírez, Y. (2020) ‘Developing a traffic congestion model based on google traffic data: A case study in Ecuador’, *VEHITS 2020 - Proceedings of the 6th International Conference on Vehicle Technology and Intelligent Transport Systems*, pp. 137–144. doi: 10.5220/0009594501370144.

Gartner N., Messer C., Rathi A., (Editors) (2001) *Traffic Flow Theory A State-of-the-Art Report Revised, TRB’s Committee on Traffic Flow Theory and Characteristics*.

Gayah, V. V., Dixit, V. V. and Guler, S. I. (2014) ‘Relationship between mean and day-to-day variation in travel time in urban networks’, *EURO Journal on Transportation and Logistics*. Springer Berlin Heidelberg, 3(3–4), pp. 227–243. doi: 10.1007/s13676-013-0032-2.

Gayah, V. V., Gao, X. and Nagle, A. S. (2014) ‘On the impacts of locally adaptive signal control on urban network stability and the Macroscopic Fundamental Diagram’, *Transportation Research Part B: Methodological*. Elsevier Ltd, 70(January), pp. 255–268. doi: 10.1016/j.trb.2014.09.010.

Geroliminis, N. and Daganzo, C. F. (2008) ‘Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings’, *Transportation Research Part B: Methodological*, 42(9), pp. 759–770. doi: 10.1016/j.trb.2008.02.002.

Geroliminis, N. and Levinson, D. M. (2009) ‘Cordon Pricing Consistent with the Physics of Overcrowding’, *Transportation and Traffic Theory 2009: Golden Jubilee*. Boston, MA: Springer US, (1995), pp. 219-240\r740. doi: 10.1007/978-1-4419-0820-9_11.

Geroliminis, N. and Sun, J. (2011) ‘Properties of a well-defined macroscopic fundamental diagram for urban traffic’, *Transportation Research Part B: Methodological*, 45(3), pp. 605–617. doi: 10.1016/j.trb.2010.11.004.

Geroliminis, N., Zheng, N. and Ampountolas, K. (2014) ‘A three-dimensional macroscopic fundamental diagram for mixed bi-modal urban networks’, *Transportation Research Part C: Emerging Technologies*, 42, pp. 168–181. doi: 10.1016/j.trc.2014.03.004.

Giannotti, F. *et al.* (2011) ‘Unveiling the complexity of human mobility by querying and mining massive trajectory data’, *VLDB Journal*. Springer-Verlag New York, Inc., 20(5), pp. 695–719. doi: 10.1007/s00778-011-0244-8.

Godfrey, J. W. (1969) ‘The mechanism of a road network’, *Traffic Engineering and Control*, 11(7), pp. 323–327.

González, M. C., Hidalgo, C. A. and Barabási, A. L. (2008) ‘Understanding individual human mobility patterns’, *Nature*, 453(7196), pp. 779–782. doi: 10.1038/nature06958.

Greenshields, B. D. *et al.* (1934) ‘The Photographic Method Of Studying Traffic Behavior’, *Highway Research Board Proceedings*, pp. 382–399.

Greenshields, B. D. (1935) ‘A study of traffic capacity’, in *14 Annual Meeting of the Highway Research Board Proceedings*, pp. 448–477.

Grigg, D. B. (1977) ‘E. G. Ravenstein and the “laws of migration”’, *Journal of Historical Geography*, 3(1), pp. 41–54. doi: 10.1016/0305-7488(77)90143-8.

Guenan, L. (2014) ‘Empirical Macroscopic Fundamental Diagrams Insights from loop detector and floating car data’, *In 96th Annual Meeting of the Transportation Research Board, Washington, DC.*

Guo, D. and Zhu, X. (2014) 'Origin-destination flow data smoothing and mapping', *IEEE Transactions on Visualization and Computer Graphics*, 20(12), pp. 2043–2052. doi: 10.1109/TVCG.2014.2346271.

Guo, Q. L. and Karimi, H. A. (2017) 'A novel methodology for prediction of spatial-temporal activities using latent features', *Computers, Environment and Urban Systems*. Pergamon, 62, pp. 74–85. doi: 10.1016/j.compenvurbsys.2016.10.006.

Hai, T. and Bao, Y. L. (2008) 'Study on extracting the colors of road network based on the highway traffic map images', in *Proceedings - 2008 International Conference on MultiMedia and Information Technology, MMIT 2008*, pp. 385–388. doi: 10.1109/MMIT.2008.206.

Hall, F. L. (1992) 'Traffic stream characteristics', *Revised Monograph on Traffic Flow Theory*, 165, pp. 2.1-2.36.

Hall, F. L. and Agyemang-Duah, K. (1991) 'Freeway Capacity Drop and the Definition of Capacity', *Transportation Research Record: Journal of the Transportation Research Board*, 1320, pp. 91–98.

Hasan, S., Zhan, X. and Ukkusuri, S. V. (2013) 'Understanding urban human activity and mobility patterns using large-scale location-based data from online social media', in *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing - UrbComp '13*. New York, New York, USA: ACM Press, p. 1. doi: 10.1145/2505821.2505823.

Hawelka, B. *et al.* (2014) 'Geo-located Twitter as proxy for global mobility patterns', *Cartography and Geographic Information Science*. Taylor & Francis, 41(3), pp. 260–271. doi: 10.1080/15230406.2014.890072.

Herman, R. and Ardekani, S. (1984) 'Characterizing Traffic Conditions in Urban Areas', *Transportation Science*. INFORMS , 18(2), pp. 101–140. doi: 10.1287/trsc.18.2.101.

Herman, R. and Prigogine, I. (1979) 'A two-fluid approach to town traffic.', *Science (New York, N.Y.)*, 204(4389), pp. 148–151. doi: 10.1126/science.204.4389.148.

Hoogendoorn, S. and Knoop, V. (2012) 'Traffic flow theory and modelling', *The Transport System and Transport Policy: An Introduction*, pp. 125–159.

Hoogendoorn, S. P. *et al.* (2017) 'Macroscopic Fundamental Diagram for pedestrian networks: Theory and applications', in *Transportation Research Procedia*, pp. 480–496. doi: 10.1016/j.trpro.2017.05.027.

Hoteit, S. *et al.* (2014) 'Estimating human trajectories and hotspots through mobile phone data', *Computer Networks*. Elsevier, 64, pp. 296–307. doi: 10.1016/j.comnet.2014.02.011.

IMF: World Economic Outlook (WEO), October 2019 - knoema.com (2019). Available at: https://knoema.com/IMFWEO2019Oct/imf-world-economic-outlook-weo-october-2019 (Accessed: 1 March 2020).

Ji, Y. and Geroliminis, N. (2012) 'On the spatial partitioning of urban transportation networks', *Transportation Research Part B: Methodological*, 46(10), pp. 1639–1656. doi: 10.1016/j.trb.2012.08.005.

Ji, Y., Luo, J. and Geroliminis, N. (2014) 'Empirical observations of congestion propagation and dynamic partitioning with probe data for large scale systems', *Research Record*, 2422(2), pp. 1–11.

Jiang, B., Yin, J. and Zhao, S. (2009) 'Characterizing the human mobility pattern in a large

street network', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 80(2). doi: 10.1103/PhysRevE.80.021136.

Jiang, S. *et al.* (2017) 'Human mobility in space from three modes of public transportation', *Physica A: Statistical Mechanics and its Applications*. North-Holland, 483, pp. 227–238. doi: 10.1016/J.PHYSA.2017.04.182.

Jurdak, R. *et al.* (2015) 'Understanding Human Mobility from Twitter', *PLOS ONE*. Edited by Y. Wu. Public Library of Science, 10(7), p. e0131469. doi: 10.1371/journal.pone.0131469.

Kang, C. *et al.* (2012) 'Intra-urban human mobility patterns: An urban morphology perspective', *Physica A: Statistical Mechanics and its Applications*. North-Holland, 391(4), pp. 1702–1717. doi: 10.1016/j.physa.2011.11.005.

Kashiyama, T., Pang, Y. and Sekimoto, Y. (2017) 'Open PFLOW: Creation and evaluation of an open dataset for typical people mass movement in urban areas', *Transportation Research Part C: Emerging Technologies*. Pergamon, 85, pp. 249–267. doi: 10.1016/J.TRC.2017.09.016.

Keyvan-Ekbatani, M. *et al.* (2012) 'Exploiting the fundamental diagram of urban networks for feedback-based gating', *Transportation Research Part B: Methodological*, 46(10), pp. 1393–1403. doi: 10.1016/j.trb.2012.06.008.

Kim, T. J. (2018) *Riyadh | Geography, History, & Population | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Riyadh#ref276658 (Accessed: 16 May 2020).

Knoop, V., Hoogendoorn, S. and Van Lint, J. (2012) 'Routing strategies based on macroscopic fundamental diagram', *Transportation Research Record*. SAGE PublicationsSage CA: Los Angeles, CA, 2315(2315), pp. 1–10. doi: 10.3141/2315-01.

Knoop, V. L. *et al.* (2018) 'Empirical MFDs using Google Traffic Data', in *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*. Institute of Electrical and Electronics Engineers Inc., pp. 3832–3839. doi: 10.1109/ITSC.2018.8570005.

Kondor, D. *et al.* (2015) 'Visualizing signatures of human activity in cities across the globe', *arXiv preprint*, 3(3), pp. 1–6. doi: 10.1007/slaf-0054-15008.

Lankevich, G. (2020) *New York City | Layout, People, Economy, Culture, & History | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/New-York-City (Accessed: 17 May 2020).

Laurila, J. K. *et al.* (2013) 'From big smartphone data to worldwide research: The Mobile Data Challenge', in *Pervasive and Mobile Computing*. Elsevier, pp. 752–771. doi: 10.1016/j.pmcj.2013.07.014.

Laval, J. A. (2011) 'Hysteresis in traffic flow revisited: An improved measurement method', *Transportation Research Part B: Methodological*. Elsevier Ltd, 45(2), pp. 385–391. doi: 10.1016/j.trb.2010.07.006.

Leclercq, L., Chiabaut, N. and Trinquier, B. (2014) 'Macroscopic Fundamental Diagrams: A cross-comparison of estimation methods', *Transportation Research Part B: Methodological*, 62, pp. 1–12. doi: 10.1016/j.trb.2014.01.007.

Leinbach, T. R. (2020) *Singapore | Facts, Geography, History, & Points of Interest | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Singapore (Accessed: 16 May 2020).

Leite, A. (2020) *Sao Paulo | Points of Interest, History, & Facts | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Sao-Paulo-Brazil (Accessed: 17 May 2020).

Lighthill, M. J. and Whitham, G. B. (1955) 'On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads', *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 229(1178), pp. 317–345. doi: 10.1098/rspa.1955.0089.

Lina J., K. (2009) 'Lossless Image Compression', in *The Essential Guide to Image Processing*. Elsevier Inc., pp. 385–419. doi: 10.1016/B978-0-12-374457-9.00016-0.

Liu, X. *et al.* (2015) 'Revealing travel patterns and city structure with taxi trip data', *Journal of Transport Geography*. Pergamon, 43, pp. 78–90. doi: 10.1016/J.JTRANGEO.2015.01.016.

Loder, A. *et al.* (2017) 'Empirics of multi-modal traffic networks – Using the 3D macroscopic fundamental diagram', *Transportation Research Part C: Emerging Technologies*, 82, pp. 88–101. doi: 10.1016/j.trc.2017.06.009.

Luo, F. *et al.* (2016) 'Explore spatiotemporal and demographic characteristics of human mobility via Twitter: A case study of Chicago', *Applied Geography*. Pergamon, 70, pp. 11–25. doi: 10.1016/j.apgeog.2016.03.001.

Ma, X. *et al.* (2017) 'Understanding commuting patterns using transit smart card data', *Journal of Transport Geography*. Pergamon, 58, pp. 135–145. doi: 10.1016/J.JTRANGEO.2016.12.001.

Mahmassani, H., Hou, T. and Saberi, M. (2013) 'Connecting Networkwide Travel Time Reliability and the Network Fundamental Diagram of Traffic Flow', *Transportation Research Record: Journal of the Transportation Research Board*, 2391, pp. 80–91. doi: 10.3141/2391-08.

Mahmassani, H. S. ., Williams, J. C. and Herman, R. (1984) 'Investigation of Network-Level Traffic Relationships: Some Simulation Results', *Transportation Research Record*, 971, pp. 121–140.

Manca, M. *et al.* (2017) 'Using social media to characterize urban mobility patterns: State-of-the-art survey and case-study', *Online Social Networks and Media*. Elsevier, 1, pp. 56–69. doi: 10.1016/j.osnem.2017.04.002.

MapQuest (2020) 'Official MapQuest - Maps, Driving Directions, Live Traffic'.

Mathew, T. V. and Rao, K. V. K. (2006) 'Fundamental relations of traffic flow', *Introduction to Transportation Engineering*, 1, pp. 1–8.

May, A. (1990) *Traffic Flow Fundamentals*.

McGillivray, B. (2020) *Toronto | History, Points of Interest, & Facts | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Toronto (Accessed: 17 May 2020).

MIT Senseable City Lab (2020) *MIT Senseable City Lab, Senseable City Laboratory*. Available at: http://senseable.mit.edu/ (Accessed: 22 February 2018).

Newell, G. F. (1962) 'Theories of Instability in Dense Highway Traffic', *The Operations Research Society of Japan*, 1, pp. 9–54.

Ni, L., Wang, X. and Chen, X. (2018) 'A spatial econometric model for travel flow analysis and real-world applications with massive mobile phone data', *Transportation Research Part C*. Pergamon, 86, pp. 510–526. doi: 10.1016/j.trc.2017.12.002.

Noulas, A. *et al.* (2012) 'A Tale of Many Cities: Universal Patterns in Human Urban Mobility', *PLoS ONE*. Edited by J. A. Añel. Public Library of Science, 7(5), p. e37027. doi: 10.1371/journal.pone.0037027.

Outram, C., Ratti, C. and Biderman, A. (2010) *The Copenhagen Wheel:! An innovative electric bicycle system that harnesses the power of real-time information and crowd sourcing*.

Pappalardo, L. *et al.* (2015) 'Returners and explorers dichotomy in human mobility', *Nature Communications*. Nature Publishing Group, 6(1), p. 8166. doi: 10.1038/ncomms9166.

Petrovska, N. and Stevanovic, A. (2015) 'Traffic Congestion Analysis Visualisation Tool', in *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*. IEEE, pp. 1489–1494. doi: 10.1109/ITSC.2015.243.

Phan, D. *et al.* (2005) 'Flow map layout', in *Proceedings - IEEE Symposium on Information Visualization, INFO VIS*. IEEE, pp. 219–224. doi: 10.1109/INFVIS.2005.1532150.

Piaget, J. and Inhelder, B. (1967) *Th Child's Conception of Space*. Routledge & Kegan Paul.

Pitt, L. M. (2020) *Los Angeles - City layout | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Los-Angeles-California/City-layout (Accessed: 15 May 2020).

Pringle, J. D. (2020) *Sydney | History & Points of Interest | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Sydney-New-South-Wales#ref60757 (Accessed: 16 May 2020).

Qu, X., Wang, S. and Zhang, J. (2015) 'On the fundamental diagram for freeway traffic: A novel calibration approach for single-regime models', *Transportation Research Part B: Methodological*. doi: 10.1016/j.trb.2015.01.001.

Rao, K. V. K. and Mathew, T. V. (2007) 'Traffic stream models', *Introduction to Transportation Engineering*, pp. 1–5.

Rao, V. L. S. P. (2018) *Delhi - City layout | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Delhi/City-layout (Accessed: 16 May 2020).

Ratti, C. *et al.* (2006) *Real Time Rome, MIT*. Available at: http://senseable.mit.edu/realtimerome/ (Accessed: 13 February 2018).

Ratti, C. *et al.* (2007) 'Mobile Landscapes: Graz in Real Time', in *Location Based Services and TeleCartography*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 433–444. doi: 10.1007/978-3-540-36728-4_31.

Ravenstein, E. G. (1885) 'The Laws of Migration', *Journal of the Statistical Society of London*, 48(2), p. 167. doi: 10.2307/2979181.

Reuter, L. R. (2020) *Berlin | History, Map, Population, Attractions, & Facts | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Berlin/The-city-layout (Accessed: 14 May 2020).

Richards, P. (1956) 'Shock Waves on the Highway', *Operations Research*, 4(1), pp. 42–51. doi: 10.1287/opre.4.1.42.

Robinson, A. H. (1955) 'The 1837 Maps of Henry Drury Harness', *The Geographical Journal*, pp. 440–450. doi: 10.2307/1791753.

Robinson, A. H. (1967) 'The thematic maps of Charles Joseph Minard', *Imago Mundi*, 21(1), pp. 95–108. doi: 10.1080/03085696708592302.

Saberi, M., Mahmassani, H. S. and Zockaie, A. (2014) 'Network capacity, traffic instability, and adaptive driving: findings from simulated urban network experiments', *EURO Journal on Transportation and Logistics*, 3(3–4), pp. 289–308. doi: 10.1007/s13676-013-0040-2.

Seidensticker, E. G. (2016) *Tokyo-Yokohama Metropolitan Area | Japan | Britannica, Encyclopædia Britannica*. Available at: https://www.britannica.com/place/Tokyo-Yokohama-Metropolitan-Area (Accessed: 16 May 2020).

Shen, L. and Stopher, P. R. (2014) 'Review of GPS Travel Survey and GPS Data-Processing Methods', *Transport Reviews*. Taylor & Francis, 34(3), pp. 316–334. doi: 10.1080/01441647.2014.903530.

Smeed, R. J. (1966) 'Road Capacity of City Centres', *Traffic Engineering and Control*, 8(7), pp. 455–458.

Smeed, R. J. (1968) 'Traffic Studies and Urban Congestion', *Journal of Transport Economics and Policy*, 2(1), pp. 33–70. doi: 10.1177/03063127067078012.

Steiger, E. *et al.* (2015) 'Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data', *Computers, Environment and Urban Systems*. Pergamon, 54, pp. 255–265. doi: 10.1016/j.compenvurbsys.2015.09.007.

Stragier, J. and Mechant, P. (2013) 'Mobile fitness apps for promoting physical activity on Twitter: the RunKeeper case', *Etmaal van de Communicatiewetenschap*, (2010).

Sun, J. B. *et al.* (2011) 'Exploring spacetime structure of human mobility in urban space', *Physica A: Statistical Mechanics and its Applications*. North-Holland, 390(5), pp. 929–942. doi: 10.1016/j.physa.2010.10.033.

Sun, L. and Axhausen, K. W. (2016) 'Understanding urban mobility patterns with a probabilistic tensor factorization framework', *Transportation Research Part B: Methodological*. Pergamon, 91, pp. 511–524. doi: 10.1016/J.TRB.2016.06.011.

Tang, J. *et al.* (2015) 'Uncovering urban human mobility from large scale taxi GPS data', *Physica A: Statistical Mechanics and its Applications*. North-Holland, 438, pp. 140–153. doi: 10.1016/J.PHYSA.2015.06.032.

Thomson J.M (1967) 'Speeds and flows of traffic in central London: Speed-flow relations', *Traffic Engineering and Control*, 8(12), pp. 721–725.

Tsou, M.-H. (2011) 'Revisiting Web Cartography in the United States: the Rise of User-Centered Design', *Cartography and Geographic Information Science*. Taylor & Francis Group, 38(3), pp. 250–257. doi: 10.1559/15230406382250.

Tsubota, T., Bhaskar, A. and Chung, E. (2014) 'Macroscopic Fundamental Diagram for Brisbane, Australia', *Transportation Research Record: Journal of the Transportation Research Board*, 2421, pp. 12–21. doi: 10.3141/2421-02.

Veenendaal, B. (2016) 'Eras of web mapping developments: Past, present and future', *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 41(July), pp. 247–252. doi: 10.5194/isprsarchives-XLI-B4-247-2016.

Verhetsel, A. and Vanelslander, T. (2010) 'What location policy can bring to sustainable commuting: An empirical study in Brussels and Flanders, Belgium', *Journal of Transport Geography*. Pergamon, 18(6), pp. 691–701. doi: 10.1016/j.jtrangeo.2009.11.003.

Wachowicz, M. (2010) *Movement-aware applications for sustainable mobility: Technologies and approaches, Movement-Aware Applications for Sustainable Mobility:*

Technologies and Approaches. doi: 10.4018/978-1-61520-769-5.

Wang, P. F. *et al.* (2015) 'An Empirical Analysis of Macroscopic Fundamental Diagrams for Sendai Road Networks', *Interdisciplinary Information Sciences*. The Editorial Committee of the Interdisciplinary Information Sciences, 21(1), pp. 49–61. doi: 10.4036/iis.2015.49.

Wang, Z. *et al.* (2004) 'Image quality assessment: From error visibility to structural similarity', *IEEE Transactions on Image Processing*, 13(4), pp. 600–612. doi: 10.1109/TIP.2003.819861.

Wardrop J.G (1968) 'Journey speed and flow in central urban areas', *Traffic Engineering and Control*, 9(11), pp. 528–532.

Weller, C. (2017) *Mesmerizing maps show the flow of refugees over the last 15 years - Business Insider*, *Business Insider*. Available at: http://www.businessinsider.com/maps-flow-refugees-last-15-years-2017-5/#2001-saw-roughly-500000-refugees-fleeing-primarily-middle-eastern-countries-such-as-afghanistan-and-african-countries-such-as-sudan-and-the-democratic-republic-of-congo-1 (Accessed: 13 February 2018).

Williams, J. C., Mahmassani, H. S. and Herman, R. (1985) 'Analysis of Traffic Network Flow Relations and Two-Fluid Model Parameter Sensitivity', *Transportation Research Record*. Transportation Research Board, (1005), pp. 95–106.

Wilson, E. (2011) 'From Inductance Loops to Vehicle Trajectories', *75 Years of the Fundamental Diagram for Traffic Flow Theory*, E-C149(June), p. 134.

Xu, Y. *et al.* (2017) 'How friends share urban space: An exploratory spatiotemporal analysis using mobile phone data', *Transactions in GIS*, 21(3), pp. 468–487. doi: 10.1111/tgis.12285.

Yandex (2014) *Yandex Traffic Jam Technology Overview*. Available at: https://yandex.com/company/technologies/traffic_jams_technology/ (Accessed: 2 December 2020).

Yau, N. (2007) *Where People Run in Major Cities | FlowingData*, *FlowingData*. Available at: https://flowingdata.com/2014/02/05/where-people-run/ (Accessed: 17 February 2019).

Yildirimoglu, M., Limniati, Y. and Geroliminis, N. (2015) 'Investigating empirical implications of hysteresis in day-to-day travel time variability', *Transportation Research Part C: Emerging Technologies*. Elsevier Ltd, 55, pp. 340–350. doi: 10.1016/j.trc.2015.03.012.

Yildirimoglu, M., Ramezani, M. and Geroliminis, N. (2015) 'Equilibrium analysis and route guidance in large-scale networks with MFD dynamics', in *Transportation Research Procedia*, pp. 185–204. doi: 10.1016/j.trpro.2015.07.011.

Yuan, H. *et al.* (2014) 'Human mobility discovering and movement intention detection with GPS trajectories', *Decision Support Systems*. North-Holland, 63, pp. 39–51. doi: 10.1016/j.dss.2013.09.010.

Zahavi, Y. (1972) 'Traffic Performance Evaluation of Road Networks by the α-Relationship. PartsI & II', *Traffic Engineering & Control*, Vol. 14(No. 6), pp. 228–293.

Zhong, C. *et al.* (2016) 'Variability in Regularity: Mining Temporal Mobility Patterns in London, Singapore and Beijing Using Smart-Card Data', *PLOS ONE*. Edited by T. Preis. Public Library of Science, 11(2), p. e0149222. doi: 10.1371/journal.pone.0149222.

Appendix A

Algorithm 1: Screen capture

```
% % Google maps -screen-capture for one city


for k=1:288
% %  1.Paris
% Define a particular map in your browser (copy the link):
url                                                                    =
'https://www.google.com/maps/place/Paris,+France/@48.8529646,2.3024943,11
316m/data=!3m1!1e3!4m5!3m4!1s0x47e66e1f06e2b70f:0x40b82c3688c9460!8m2!3d4
8.856614!4d2.3522219!5m1!1e1';
% Open that map from MATLAB
web(url,'-browser');
% Hold on 20 seconds, in order to be sure that it has been loaded properly
pause(20)
% Make a screen-capture as an image (img)
img = screencapture(0, 'Position', [0 0 1920 1080]);
filename='C:\Users\vagkania01\Documents\MATLAB\screen_capture_google\Pari
s_google';
NewImage1 = sprintf('Paris_G_%03d.png', k);
fullFileName = fullfile(filename, NewImage1);
imwrite(img,fullFileName);


end
```

Algorithm 2: Crop image

```
% % Crop image
% Specify the folder where the files live.
myFolder = 'E:\MATLAB\TRB_2017_Paper2\Google\01.Paris\2017-05-31';
% Check to make sure that folder actually exists.  Warn user if it doesn't.
if ~isdir(myFolder)
  errorMessage = sprintf('Error: The following folder does not exist:\n%s',
myFolder);
  uiwait(warndlg(errorMessage));
  return;
end

% Get a list of all files in the folder with the desired file name pattern.
filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
you need.
theFiles = dir(filePattern);
for k = 1 :length(theFiles)
```

```
  baseFileName = theFiles(k).name;
  fullFileName = fullfile(myFolder, baseFileName);
  fprintf(1, 'Now reading %s\n', fullFileName);
  % Now do whatever you want with this file name,
  % such as reading it in as an image array with imread()

imageArray= imread(fullFileName);
croppedImage= imcrop(imageArray,[448 70 1756-448 1027-70]);
imshow(croppedImage);
% drawnow; % Force display to update immediately.
filename2='E:\MATLAB\TRB_2017_Paper2\Google\01.Paris\C_G';
croppedImage= sprintf('P_G_%03d.png', k);
fullFileName2 = fullfile(filename2, croppedImage);

set(gcf,'PaperPositionMode','auto')
print(fullFileName2,'-dpng','-r0')

end
```

Algorithm 3: Separate network and count pixels per traffic layer per imag

```
%% Seperate road network
% % Specify the folder where the files live.
myFolder = 'E:\MATLAB\TRB_2017_Paper2\Google\01.Paris\C_G';
% Check to make sure that folder actually exists.  Warn user if it doesn't.
if ~isdir(myFolder)
  errorMessage = sprintf('Error: The following folder does not exist:\n%s',
myFolder);
  uiwait(warndlg(errorMessage));
  return;
end
% Get a list of all files in the folder with the desired file name pattern.
filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
you need.
theFiles = dir(filePattern);
for k = 1:length(theFiles)
  baseFileName = theFiles(k).name;
  fullFileName = fullfile(myFolder, baseFileName);
  fprintf(1, 'Now reading %s\n', fullFileName);
  % Read an image
pic1= imread(fullFileName);
```

```
% % green
% Define thresholds for channel 1 based on histogram settings
channel1Min = 118.000;
channel1Max = 132.000;

% Define thresholds for channel 2 based on histogram settings
channel2Min = 153.000;
channel2Max = 202.000;

% Define thresholds for channel 3 based on histogram settings
channel3Min = 80.000;
channel3Max = 86.000;



for mm=1:size(pic1,1)

    for nn=1:size(pic1,2)
        if  pic1(mm,nn,1)<channel1Min  ||  pic1(mm,nn,1)>channel1Max  ||
pic1(mm,nn,2)<channel2Min      ||      pic1(mm,nn,2)>channel2Max      ||
pic1(mm,nn,3)<channel3Min  || pic1(mm,nn,3)>channel3Max
            gsc=0*pic1(mm,nn,1)+0*pic1(mm,nn,2)+0*pic1(mm,nn,3);
            pic1(mm,nn,:)=[gsc gsc gsc];

        end
    end
end
% % Count green pixels
A=pic1;
[sz1 sz2 sz3]=size(A)
N=256;
ColorList={'Green' 'Yellow' 'Red'};
gr=0:1/(N-1):1;
% checking how many different values pixels take in this picture
P=unique(impixel(A,sz2,sz1));
% filtering reds
cMap=zeros(N,3);cMap(:,1)=gr;
figure(2);hr=imshow(ind2rgb(A(:,:,1),cMap));title(ColorList{1});
% accessing pixel values in field CData of image handle hr, for black
R=hr.CData;
R1=R(:,:,1);R2=R(:,:,2);R3=R(:,:,3);
amount_black_pixels=numel(find(R1==0 & R2==0 & R3==0));
fprintf('\n amount black pixels: %s \n', num2str(amount_black_pixels));
```

```
amount_green_pixels=(sz1*sz2)-amount_black_pixels;
fprintf('\n amount green pixels: %s \n', num2str(amount_green_pixels));
amount_green_pixels_all(k,1)=amount_green_pixels(1,1);
% % yellow
pic2= imread(fullFileName);
% Define thresholds for channel 1 based on histogram settings
channel1Min = 187.000;
channel1Max = 240.000;

% Define thresholds for channel 2 based on histogram settings
channel2Min = 110.000;
channel2Max = 143.000;

% Define thresholds for channel 3 based on histogram settings
channel3Min = 0.000;
channel3Max = 67.000;

for mm=1:size(pic2,1)

    for nn=1:size(pic2,2)
        if  pic2(mm,nn,1)<channel1Min  ||  pic2(mm,nn,1)>channel1Max  ||
pic2(mm,nn,2)<channel2Min     ||     pic2(mm,nn,2)>channel2Max      ||
pic2(mm,nn,3)<channel3Min  || pic2(mm,nn,3)>channel3Max
            gsc=0*pic1(mm,nn,1)+0*pic1(mm,nn,2)+0*pic1(mm,nn,3);
            pic2(mm,nn,:)=[gsc gsc gsc];

        end
    end
end
% % Count yellow pixels
A=pic2;
[sz1 sz2 sz3]=size(A)
N=256;
ColorList={'Green' 'Yellow' 'Red'};
gr=0:1/(N-1):1;
% checking how many different values pixels take in this picture
P=unique(impixel(A,sz2,sz1));
% filtering reds
cMap=zeros(N,3);cMap(:,1)=gr;
figure(2);hr=imshow(ind2rgb(A(:,:,1),cMap));title(ColorList{2});
% accessing pixel values in field CData of image handle hr, for black
R=hr.CData;
```

```
R1=R(:,:,1);R2=R(:,:,2);R3=R(:,:,3);
amount_black_pixels=numel(find(R1==0 & R2==0 & R3==0));
fprintf('\n amount black pixels: %s \n', num2str(amount_black_pixels));
amount_yellow_pixels=(sz1*sz2)-amount_black_pixels;
fprintf('\n amount yellow pixels: %s \n', num2str(amount_yellow_pixels));
amount_yellow_pixels_all(k,1)=amount_yellow_pixels(1,1);
% % orange
pic3= imread(fullFileName);
% Define thresholds for channel 1 based on histogram settings
channel1Min = 163.000;
channel1Max = 255.000;

% Define thresholds for channel 2 based on histogram settings
channel2Min = 0.000;
channel2Max = 45.000;

% Define thresholds for channel 3 based on histogram settings
channel3Min = 0.000;
channel3Max = 255.000;

for mm=1:size(pic3,1)

    for nn=1:size(pic3,2)
        if  pic3(mm,nn,1)<channel1Min  ||  pic3(mm,nn,1)>channel1Max  ||
pic3(mm,nn,2)<channel2Min    ||    pic3(mm,nn,2)>channel2Max    ||
pic3(mm,nn,3)<channel3Min  || pic3(mm,nn,3)>channel3Max
           gsc=0*pic1(mm,nn,1)+0*pic1(mm,nn,2)+0*pic1(mm,nn,3);
            pic3(mm,nn,:)=[gsc gsc gsc];

        end
    end
end
% % Count orange pixels
A=pic3;
[sz1 sz2 sz3]=size(A)
N=256;
ColorList={'Red' 'Green' 'Blue'};
gr=0:1/(N-1):1;
% checking how many different values pixels take in this picture
P=unique(impixel(A,sz2,sz1));
% filtering reds
cMap=zeros(N,3);cMap(:,1)=gr;
```

```
figure(2);hr=imshow(ind2rgb(A(:,:,1),cMap));title(ColorList{1});
% accessing pixel values in field CData of image handle hr, for black
R=hr.CData;
R1=R(:,:,1);R2=R(:,:,2);R3=R(:,:,3);
amount_black_pixels=numel(find(R1==0 & R2==0 & R3==0));
fprintf('\n amount black pixels: %s \n', num2str(amount_black_pixels));
amount_orange_pixels=(sz1*sz2)-amount_black_pixels;
fprintf('\n amount orange pixels: %s \n', num2str(amount_orange_pixels));
amount_orange_pixels_all(k,1)=amount_orange_pixels(1,1);
% % red
pic4= imread(fullFileName);
% Define thresholds for channel 1 based on histogram settings
channel1Min = 136.000;
channel1Max = 158.000;

% Define thresholds for channel 2 based on histogram settings
channel2Min = 19.000;
channel2Max = 35.000;

% Define thresholds for channel 3 based on histogram settings
channel3Min = 19.000;
channel3Max = 36.000;

for mm=1:size(pic4,1)

    for nn=1:size(pic4,2)
        if  pic4(mm,nn,1)<channel1Min  ||  pic4(mm,nn,1)>channel1Max  ||
pic4(mm,nn,2)<channel2Min      ||      pic4(mm,nn,2)>channel2Max      ||
pic4(mm,nn,3)<channel3Min  || pic4(mm,nn,3)>channel3Max
          gsc=255*pic1(mm,nn,1)+255*pic1(mm,nn,2)+255*pic1(mm,nn,3);
           pic4(mm,nn,:)=[gsc gsc gsc];

        end
    end
end
% % Count red pixels
A=pic4;
[sz1 sz2 sz3]=size(A)
N=256;
ColorList={'Green' 'Yellow' 'Red'};
gr=0:1/(N-1):1;
% checking how many different values pixels take in this picture
```

```
P=unique(impixel(A,sz2,sz1));
% filtering reds
cMap=zeros(N,3);cMap(:,1)=gr;
figure(2);hr=imshow(ind2rgb(A(:,:,1),cMap));title(ColorList{2});
% accessing pixel values in field CData of image handle hr, for black
R=hr.CData;
R1=R(:,:,1);R2=R(:,:,2);R3=R(:,:,3);
amount_black_pixels=numel(find(R1==0 & R2==0 & R3==0));
fprintf('\n amount black pixels: %s \n', num2str(amount_black_pixels));
amount_red_pixels=(sz1*sz2)-amount_black_pixels;
fprintf('\n amount red pixels: %s \n', num2str(amount_red_pixels));
amount_red_pixels_all(k,1)=amount_red_pixels(1,1);

pic=[pic1+pic2+pic3+pic4];
imshow(pic);
drawnow; % Force display to update immediately.
filename2='E:\MATLAB\TRB_2017_Paper2\Google\01.Paris\N_G';
pic = sprintf('N_G_01_%03d.png', k);
fullFileName2 = fullfile(filename2,pic);
set(gcf,'PaperPositionMode','auto')
print(fullFileName2,'-dpng','-r0')
end
Pixels_01=[amount_green_pixels_all                amount_yellow_pixels_all
amount_orange_pixels_all amount_red_pixels_all];
save('Pixels_01.mat', 'Pixels_01');
close all
clear all
```

Algorithm 4: SSIM1 for one city

```
%% Specify the folder where the files live.
myFolder = 'E:\Google\01.Paris\N_G';
% Check to make sure that folder actually exists.  Warn user if it doesn't.
if ~isdir(myFolder)
  errorMessage = sprintf('Error: The following folder does not exist:\n%s',
myFolder);
  uiwait(warndlg(errorMessage));
  return;
end
% Get a list of all files in the folder with the desired file name pattern.
filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
you need.
theFiles = dir(filePattern);
pic0=N_G_01_001;
Ast01= zeros(300,1);
for k = 1: length(theFiles)
  baseFileName = theFiles(k).name;
  fullFileName = fullfile(myFolder, baseFileName);
```

```
  fprintf(1, 'Now reading %s\n', fullFileName);
  % Read an image
pic1= imread(fullFileName);
% Me ref tin prwti eikona

Ast01(k,1) = ssim(pic1,pic0);

 end
```

Algorithm 5: SSIM2 for one city

```
%% Specify the folder where the files live.
myFolder = 'E:\Google\01.Paris\N_G';
% Check to make sure that folder actually exists.  Warn user if it
doesn't.
if ~isdir(myFolder)
  errorMessage = sprintf('Error: The following folder does not
exist:\n%s', myFolder);
  uiwait(warndlg(errorMessage));
  return;
end
% Get a list of all files in the folder with the desired file name
pattern.
filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
you need.
theFiles = dir(filePattern);
Ast01=zeros(300,1);
for k = 2:length(theFiles)
  baseFileName = theFiles(k-1).name;
  fullFileName = fullfile(myFolder, baseFileName);
  fprintf(1, 'Now reading %s\n', fullFileName);
  pic0=imread(fullFileName);
  baseFileName = theFiles(k).name;
  fullFileName = fullfile(myFolder, baseFileName);
  fprintf(1, 'Now reading %s\n', fullFileName);
  pic1= imread(fullFileName);
% Me ref tin proigoumeni eikona

Ast01(k-1,1)=ssim(pic1,pic0);
 end
```

Algorithm 6: Kernel density estimation for one city

```
% Google Scatter Hist

%%  Green
% Specify the folder where the files live.
myFolder = 'E:\TRB_2017_Paper2\Google\01.Paris\N_G';
% Check to make sure that folder actually exists.  Warn user if it
doesn't.
if ~isdir(myFolder)
  errorMessage = sprintf('Error: The following folder does not
exist:\n%s', myFolder);
  uiwait(warndlg(errorMessage));
  return;
end
% Get a list of all files in the folder with the desired file name
pattern.
filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
you need.
theFiles = dir(filePattern);
```

```
358
359  for r = 1:length(theFiles)
360    baseFileName = theFiles(r).name;
361    fullFileName = fullfile(myFolder, baseFileName);
362    fprintf(1, 'Now reading %s\n', fullFileName);
363    % Now do whatever you want with this file name,
364    % such as reading it in as an image array with imread()
365
366  pic2= imread(fullFileName);
367  % Define thresholds for channel 1 based on histogram settings
368  channel1Min = 118.000;
369  channel1Max = 132.000;
370
371  % Define thresholds for channel 2 based on histogram settings
372  channel2Min = 153.000;
373  channel2Max = 202.000;
374
375  % Define thresholds for channel 3 based on histogram settings
376  channel3Min = 80.000;
377  channel3Max = 86.000;
378
379  Fsh=zeros(10,2);
380  k=0;
381
382  for i=1:size(pic2,1)
383      for j=1:size(pic2,2)
384
385
386  if pic2(i,j,1)>=channel1Min && pic2(i,j,1)<=channel1Max &&
387  pic2(i,j,2)>=channel2Min && pic2(i,j,2)<=channel2Max &&
388  pic2(i,j,3)>=channel3Min  && pic2(i,j,3)<=channel3Max
389              k=k+1;
390              Fsh(k,:)=[i,j];
391
392  end
393      end
394  end
395
396
397
398  % Create scatterhist
399  scatterhist(Fsh(:,2),Fsh(:,1),'Location','NorthEast',
400  'LineWidth',[1.5,1.5,1.5],'Kernel','on','Color',[0.52 0.79
401  0.31],'Marker','.','MarkerSize',[0.01,0.01,0.01])
402
403  xlim([0 873])
404  ylim([0 639])
405
406  set(gca,'FontName', 'Times New Roman')
407  set(gca,'FontSize',1)
408  set(gcf,'color','k'); %%% change
409  set(gca,'color','none')
410  set(gca,'xtick',[],'ytick',[])
411
412  % Save new image
413  filename2='E:\TRB_2017_Paper2\Google\01.Paris\N_G_SG';
414  Image= sprintf('N_G_SG_01_%03d.png',r);
415  fullFileName2 = fullfile(filename2, Image);
416  set(gcf,'inverthardcopy','off'); %%% change
417  set(gcf,'color','k')  %%% change
418  saveas(gcf,fullFileName2); %%% change
419
```

```
end

%% Orange

% Specify the folder where the files live.
myFolder = 'E:\TRB_2017_Paper2\Google\01.Paris\N_G';
% Check to make sure that folder actually exists.  Warn user if it
doesn't.
if ~isdir(myFolder)
  errorMessage = sprintf('Error: The following folder does not
exist:\n%s', myFolder);
  uiwait(warndlg(errorMessage));
  return;
end
% Get a list of all files in the folder with the desired file name
pattern.
filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
you need.
theFiles = dir(filePattern);

for r = 1:length(theFiles)
  baseFileName = theFiles(r).name;
  fullFileName = fullfile(myFolder, baseFileName);
  fprintf(1, 'Now reading %s\n', fullFileName);
  % Now do whatever you want with this file name,
  % such as reading it in as an image array with imread()

pic2= imread(fullFileName);
% Define thresholds for channel 1 based on histogram settings
channel1Min = 187.000;
channel1Max = 240.000;

% Define thresholds for channel 2 based on histogram settings
channel2Min = 110.000;
channel2Max = 143.000;

% Define thresholds for channel 3 based on histogram settings
channel3Min = 0.000;
channel3Max = 67.000;
Fsh=zeros(10,2);
k=0;

for i=1:639
    for j=1:873


if pic2(i,j,1)>=channel1Min && pic2(i,j,1)<=channel1Max &&
pic2(i,j,2)>=channel2Min && pic2(i,j,2)<=channel2Max &&
pic2(i,j,3)>=channel3Min  && pic2(i,j,3)<=channel3Max
            k=k+1;
            Fsh(k,:)=[i,j];

end
    end
end

% Create scatterhist
scatterhist(Fsh(:,2),Fsh(:,1),'Location','NorthEast','LineWidth',[1.5,1.5
,1.5],'Kernel','on','Color',[0.94  0.49
0.01],'Marker','.','MarkerSize',[0.01,0.01,0.01])

xlim([0 873])
```

```
ylim([0 639])

set(gca,'FontName', 'Times New Roman')
set(gca,'FontSize', 1)
set(gcf,'color','k');
set(gca,'color','none')
set(gca,'xtick',[],'ytick',[])

% Save new image
filename2='E:\TRB_2017_Paper2\Google\01.Paris\N_G_SO';
Image= sprintf('N_G_SO_01_%03d.png',r);
fullFileName2 = fullfile(filename2, Image);
set(gcf,'inverthardcopy','off');
set(gcf,'color','k')
saveas(gcf,fullFileName2);
end

%% Red

% Specify the folder where the files live.
myFolder = 'E:\TRB_2017_Paper2\Google\01.Paris\N_G';
% Check to make sure that folder actually exists.  Warn user if it
doesn't.
if ~isdir(myFolder)
  errorMessage = sprintf('Error: The following folder does not
exist:\n%s', myFolder);
  uiwait(warndlg(errorMessage));
  return;
end
% Get a list of all files in the folder with the desired file name
pattern.
filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
you need.
theFiles = dir(filePattern);
for r = 1:length(theFiles)
  baseFileName = theFiles(r).name;
  fullFileName = fullfile(myFolder, baseFileName);
  fprintf(1, 'Now reading %s\n', fullFileName);
  % Now do whatever you want with this file name,
  % such as reading it in as an image array with imread()

pic2= imread(fullFileName);

% Define thresholds for channel 1 based on histogram settings
channel1Min = 163.000;
channel1Max = 255.000;

% Define thresholds for channel 2 based on histogram settings
channel2Min = 0.000;
channel2Max = 45.000;

% Define thresholds for channel 3 based on histogram settings
channel3Min = 0.000;
channel3Max = 255.000;


Fsh=zeros(10,2);
k=0;

for i=1:639
    for j=1:873

```

```
544
545 if pic2(i,j,1)>=channel1Min && pic2(i,j,1)<=channel1Max &&
546 pic2(i,j,2)>=channel2Min && pic2(i,j,2)<=channel2Max &&
547 pic2(i,j,3)>=channel3Min  && pic2(i,j,3)<=channel3Max
548             k=k+1;
549             Fsh(k,:)=[i,j];
550 end
551     end
552 end
553
554 % Create scatterhist
555 scatterhist(Fsh(:,2),Fsh(:,1),'Location','NorthEast','LineWidth',[1.5,1.5
556 ,1.5],'Kernel','on','Color',[0.90  0.00
557 0.00],'Marker','.','MarkerSize',[0.01,0.01,0.01])
558
559 xlim([0 873])
560 ylim([0 639])
561
562 set(gca,'FontName', 'Times New Roman')
563 set(gca,'FontSize', 1)
564 set(gcf,'color','k');
565 set(gca,'color','none')
566 set(gca,'xtick',[],'ytick',[])
567
568 % Save new image
569 filename2='E:\TRB_2017_Paper2\Google\%  orange\N_G_SR';
570 Image=sprintf('N_G_SR_01_%03d.png',r);
571 fullFileName2 = fullfile(filename2, Image);
572 set(gcf,'inverthardcopy','off');
573 set(gcf,'color','k')
574 saveas(gcf,fullFileName2);
575
576
577 end
578 %% Dark red
579
580 % Specify the folder where the files live.
581 myFolder = 'E:\TRB_2017_Paper2\Google\01.Paris\N_G';
582 % Check to make sure that folder actually exists.  Warn user if it
583 doesn't.
584 if ~isdir(myFolder)
585   errorMessage = sprintf('Error: The following folder does not
586 exist:\n%s', myFolder);
587   uiwait(warndlg(errorMessage));
588   return;
589 end
590 % Get a list of all files in the folder with the desired file name
591 pattern.
592 filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
593 you need.
594 theFiles = dir(filePattern);
595 for r =1:length(theFiles)
596   baseFileName = theFiles(r).name;
597   fullFileName = fullfile(myFolder, baseFileName);
598   fprintf(1, 'Now reading %s\n', fullFileName);
599   % Now do whatever you want with this file name,
600   % such as reading it in as an image array with imread()
601
602 pic2= imread(fullFileName);
603
604 % Define thresholds for channel 1 based on histogram settings
605 channel1Min = 136.000;
```

```
channel1Max = 158.000;

% Define thresholds for channel 2 based on histogram settings
channel2Min = 19.000;
channel2Max = 35.000;

% Define thresholds for channel 3 based on histogram settings
channel3Min = 19.000;
channel3Max = 36.000;


Fsh=zeros(10,2);
k=0;

for i=1:639
    for j=1:873


if pic2(i,j,1)>=channel1Min && pic2(i,j,1)<=channel1Max &&
pic2(i,j,2)>=channel2Min && pic2(i,j,2)<=channel2Max &&
pic2(i,j,3)>=channel3Min  && pic2(i,j,3)<=channel3Max
            k=k+1;
            Fsh(k,:)=[i,j];
end
    end
end

% Create scatterhist
scatterhist(Fsh(:,2),Fsh(:,1),'Location','NorthEast','LineWidth',[1.5,1.5
,1.5],'Kernel','on','Color',[0.60  0.07
0.07],'Marker','.','MarkerSize',[0.01,0.01,0.01])

xlim([0 873])
ylim([0 639])

set(gca,'FontName', 'Times New Roman')
set(gca,'FontSize', 1)
set(gcf,'color','k');
set(gca,'color','none')
set(gca,'xtick',[],'ytick',[])


% Save new image
filename2='E:\TRB_2017_Paper2\Google\01.Paris\N_G_SDR';
Image= sprintf('N_G_SDR_01_%03d.png',r);
fullFileName2 = fullfile(filename2, Image);
set(gcf,'inverthardcopy','off');
set(gcf,'color','k')
saveas(gcf,fullFileName2);
end

% Add images

for k = 1:288
    % % Specify the folder where the files live.
myFolder = 'E:\TRB_2017_Paper2\Google\01. Paris\N_G_SG';
% Check to make sure that folder actually exists.  Warn user if it
doesn't.
if ~isdir(myFolder)
  errorMessage = sprintf('Error: The following folder does not
exist:\n%s', myFolder);
  uiwait(warndlg(errorMessage));
```

```
668     return;
669   end
670   % Get a list of all files in the folder with the desired file name
671   pattern.
672   filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
673   you need.
674   theFiles = dir(filePattern);
675
676   %     length(theFiles)
677     baseFileName = theFiles(k).name;
678     fullFileName = fullfile(myFolder, baseFileName);
679     fprintf(1, 'Now reading %s\n', fullFileName);
680     % Read an image
681   pic1= imread(fullFileName);
682
683   % % Specify the folder where the files live.
684   myFolder = 'E:\TRB_2017_Paper2\Google\01. Paris\N_G_SO';
685   % Check to make sure that folder actually exists.  Warn user if it
686   doesn't.
687   if ~isdir(myFolder)
688     errorMessage = sprintf('Error: The following folder does not
689   exist:\n%s', myFolder);
690     uiwait(warndlg(errorMessage));
691     return;
692   end
693   % Get a list of all files in the folder with the desired file name
694   pattern.
695   filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
696   you need.
697   theFiles = dir(filePattern);
698   % for k = 1:length(theFiles)
699     baseFileName = theFiles(k).name;
700     fullFileName = fullfile(myFolder, baseFileName);
701     fprintf(1, 'Now reading %s\n', fullFileName);
702     % Read an image
703   pic2= imread(fullFileName);
704
705   % % Specify the folder where the files live.
706   myFolder = 'E:\TRB_2017_Paper2\Google\01. Paris\N_G_SR';
707   % Check to make sure that folder actually exists.  Warn user if it
708   doesn't.
709   if ~isdir(myFolder)
710     errorMessage = sprintf('Error: The following folder does not
711   exist:\n%s', myFolder);
712     uiwait(warndlg(errorMessage));
713     return;
714   end
715   % Get a list of all files in the folder with the desired file name
716   pattern.
717   filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
718   you need.
719   theFiles = dir(filePattern);
720   % for k = 1:length(theFiles)
721     baseFileName = theFiles(k).name;
722     fullFileName = fullfile(myFolder, baseFileName);
723     fprintf(1, 'Now reading %s\n', fullFileName);
724     % Read an image
725   pic3= imread(fullFileName);
726
727
728   % % Specify the folder where the files live.
729   myFolder = 'E:\TRB_2017_Paper2\Google\01. Paris\N_G_SDR';
```

```
% Check to make sure that folder actually exists.  Warn user if it
doesn't.
if ~isdir(myFolder)
  errorMessage = sprintf('Error: The following folder does not
exist:\n%s', myFolder);
  uiwait(warndlg(errorMessage));
  return;
end
% Get a list of all files in the folder with the desired file name
pattern.
filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
you need.
theFiles = dir(filePattern);
% for k = 1:length(theFiles)
  baseFileName = theFiles(k).name;
  fullFileName = fullfile(myFolder, baseFileName);
  fprintf(1, 'Now reading %s\n', fullFileName);
  % Read an image
pic4= imread(fullFileName);
% I3=pic1;
% I4=pic2;
% halphablend = vision.AlphaBlender;
% pic = step(halphablend,I3,I4);
pic=[pic1+pic2+pic3+pic4];
imshow(pic);
filename2='E:\TRB_2017_Paper2\Google\01. Paris\N_G_K';
pic = sprintf('N_G_K_%03d.png', k);
fullFileName2 = fullfile(filename2,pic);
set(gcf,'inverthardcopy','off');
set(gcf,'color','k')
saveas(gcf,fullFileName2);


end
```

Algorithm 7: SLIC Algorithm

```
% % Specify the folder where the files live.
myFolder = 'I:\Vana Ucy\Conference\Chapter\Google\06.Moscow\N_G_C';
% Check to make sure that folder actually exists.  Warn user if it
doesn't.
if ~isdir(myFolder)
  errorMessage = sprintf('Error: The following folder does not
exist:\n%s', myFolder);
  uiwait(warndlg(errorMessage));

  return;
end
% Get a list of all files in the folder with the desired file name
pattern.
filePattern = fullfile(myFolder, '*.png'); % Change to whatever pattern
you need.
theFiles = dir(filePattern);
for k = 1:length(theFiles)
  baseFileName = theFiles(k).name;
  fullFileName = fullfile(myFolder, baseFileName);
  fprintf(1, 'Now reading %s\n', fullFileName);
  % Read an image
A= imread(fullFileName);
% % Calculate superpixels of the image.
```

```
[L,N] = superpixels(A,3500);
% % Display the superpixel boundaries overlaid on the original image.

BW = boundarymask(L);

% % Set the color of each pixel in the output image to the max RGB color
of the superpixel region.
outputImage = zeros(size(A),'like',A);
idx = label2idx(L);
numRows = size(A,1);
numCols = size(A,2);
for labelVal = 1:N
    redIdx = idx{labelVal};
    greenIdx = idx{labelVal}+numRows*numCols;
    blueIdx = idx{labelVal}+2*numRows*numCols;
    outputImage(redIdx) = max(A(redIdx));
    outputImage(greenIdx) = min(A(greenIdx));
    outputImage(blueIdx) = min(A(blueIdx));
end


drawnow; % Force display to update immediately.drawnow;
iptsetpref('ImshowBorder','tight');
C=imshow(outputImage,'InitialMagnification',67)
filename2='I:\Vana Ucy\Conference\Chapter\Google\06.Moscow\N_G_S_V';
Image=sprintf('N_G_S_C_%03d.png',k);
fullFileName2 = fullfile(filename2, Image);
saveas(gcf,fullFileName2);

end
close all
clear al
```

Appendix B

Bing Maps

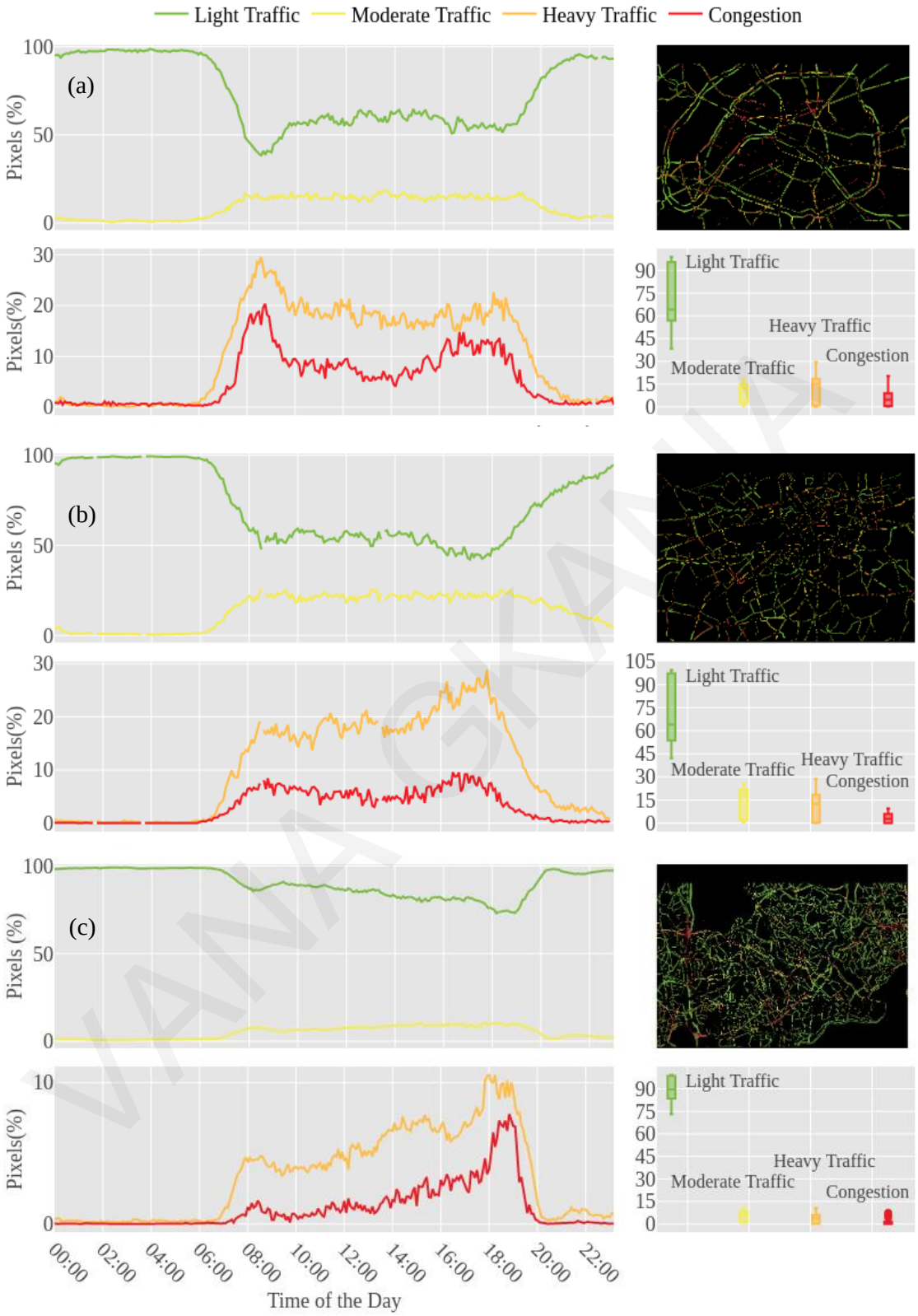


Figure B. 1: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Paris; (b) London; (c) Istanbul

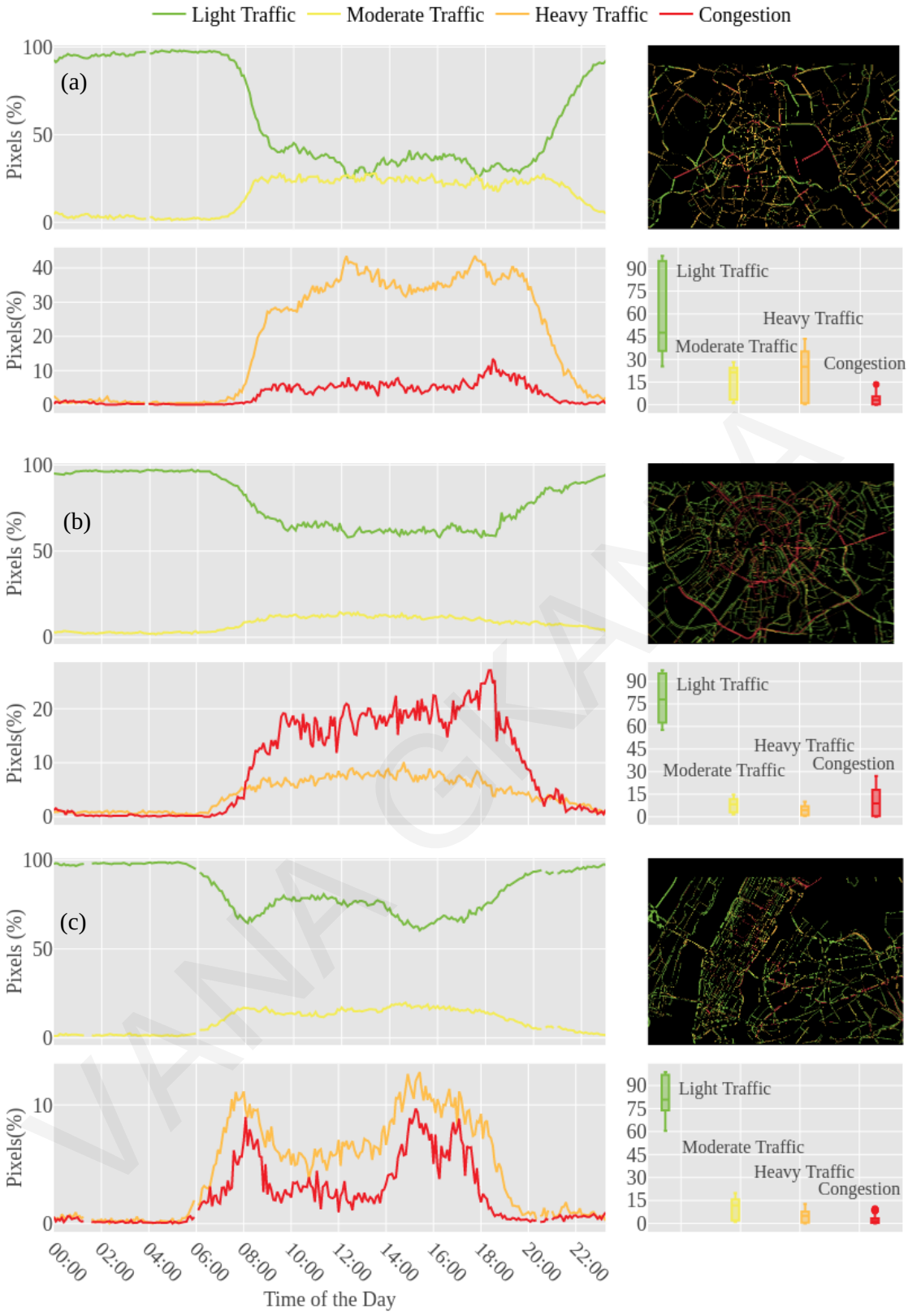


Figure B. 2: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) New Delhi; (b) Moscow; (c) New York

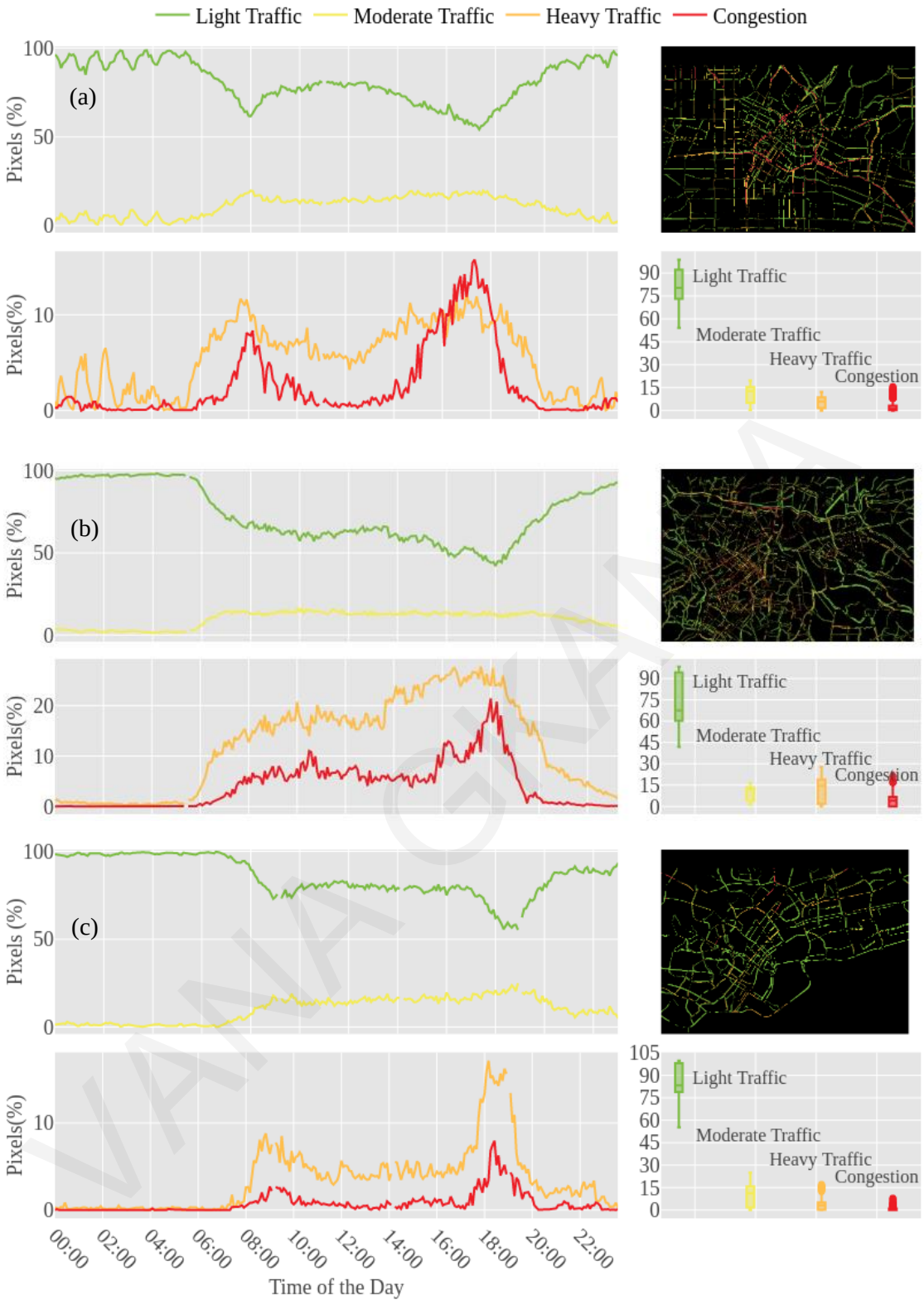


Figure B. 3: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Los Angeles; (b) Sao Paulo; (c) Singapore

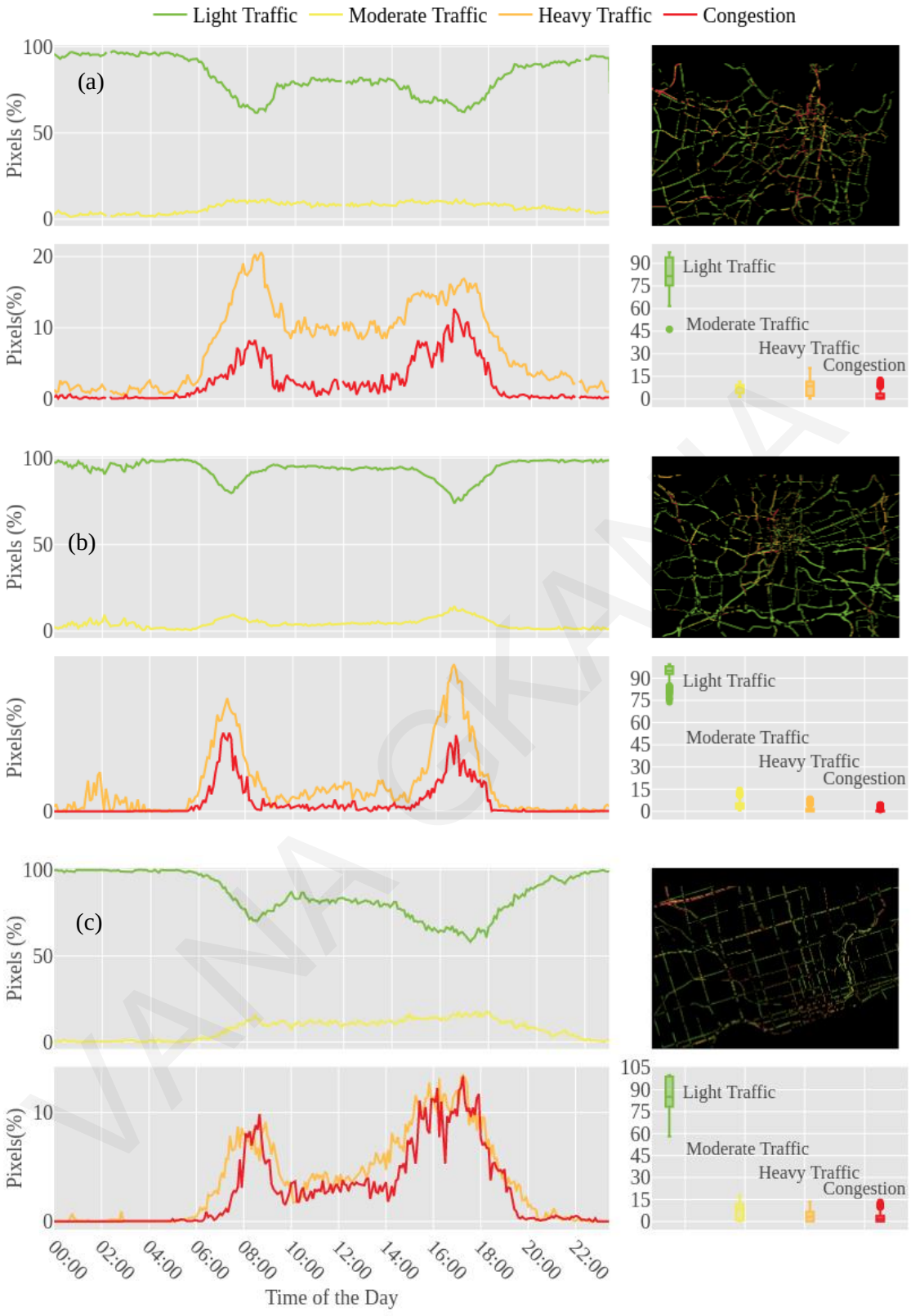


Figure B. 4: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Sydney; (b) Johannesburg; (c) Toronto

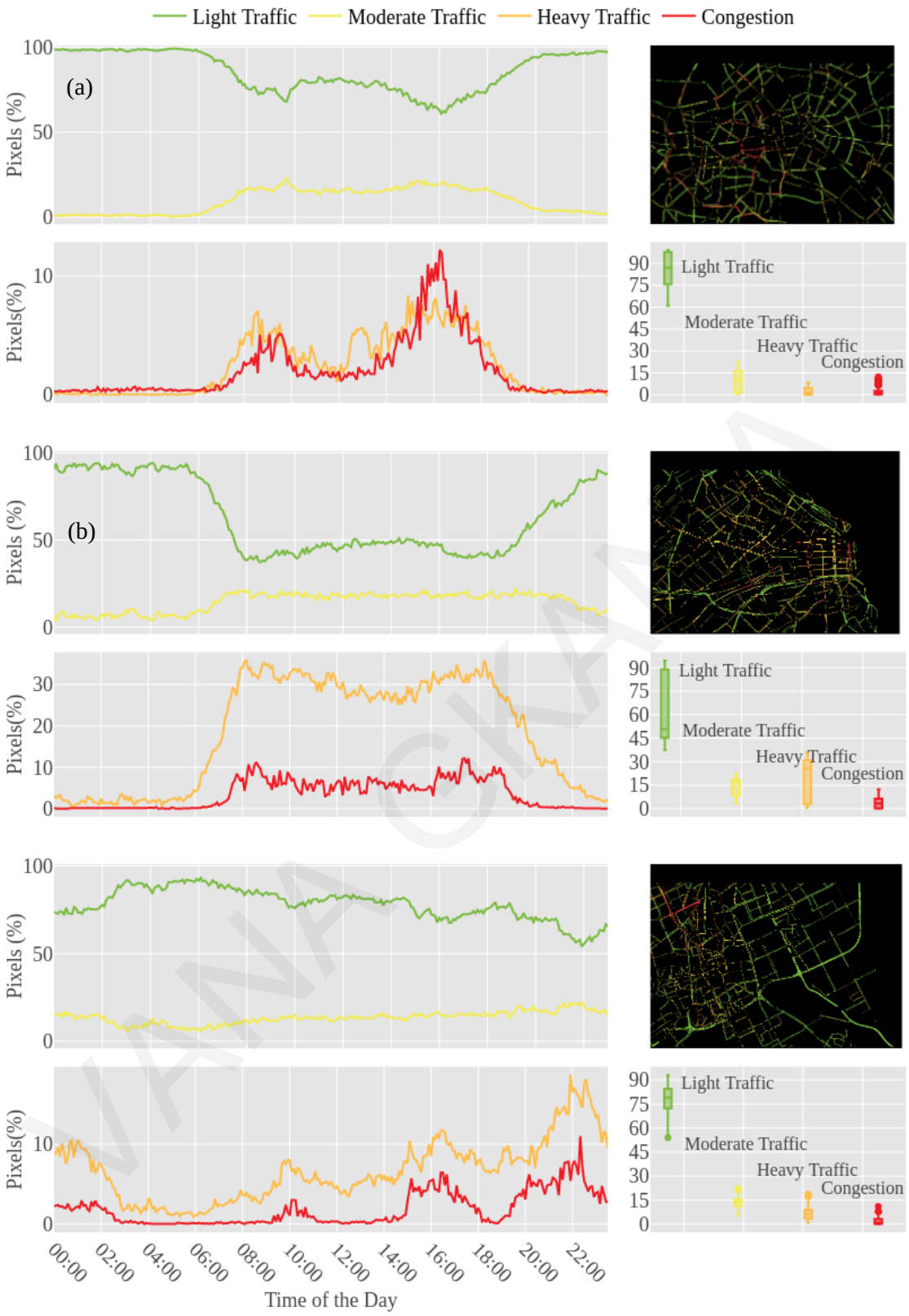


Figure B. 5: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Berlin; (b) Buenos Aires; (c) Riyadh.

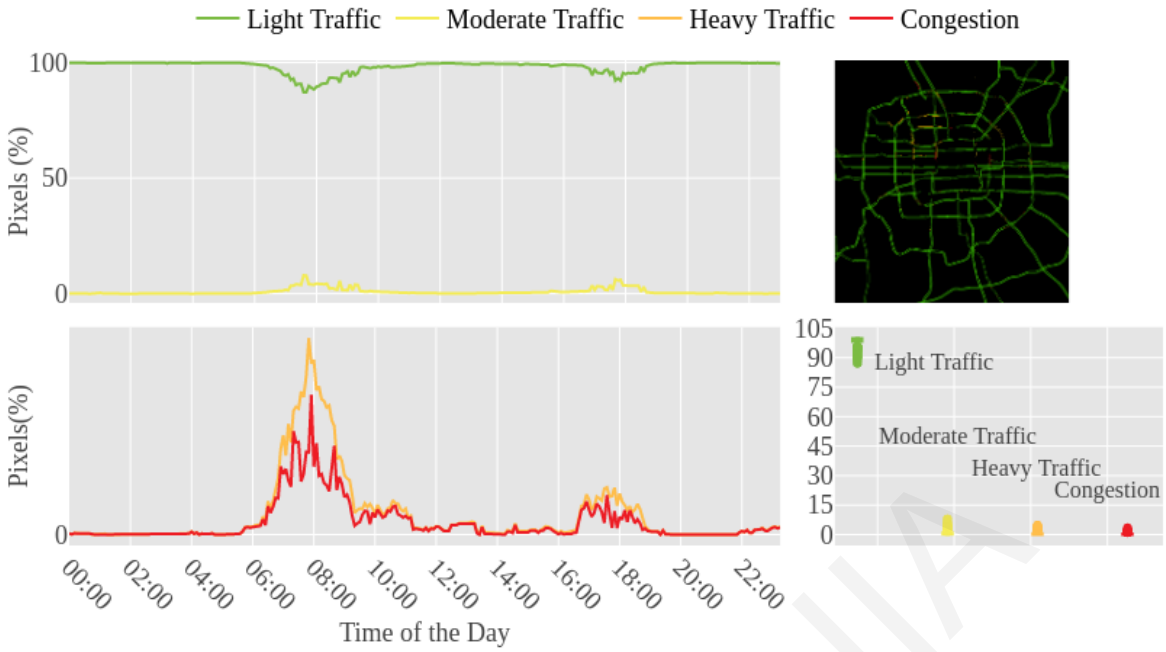


Figure B. 6: Timeseries of pixels and boxplot per traffic layer for one weekday for Beijing

Appendix C

Here Maps

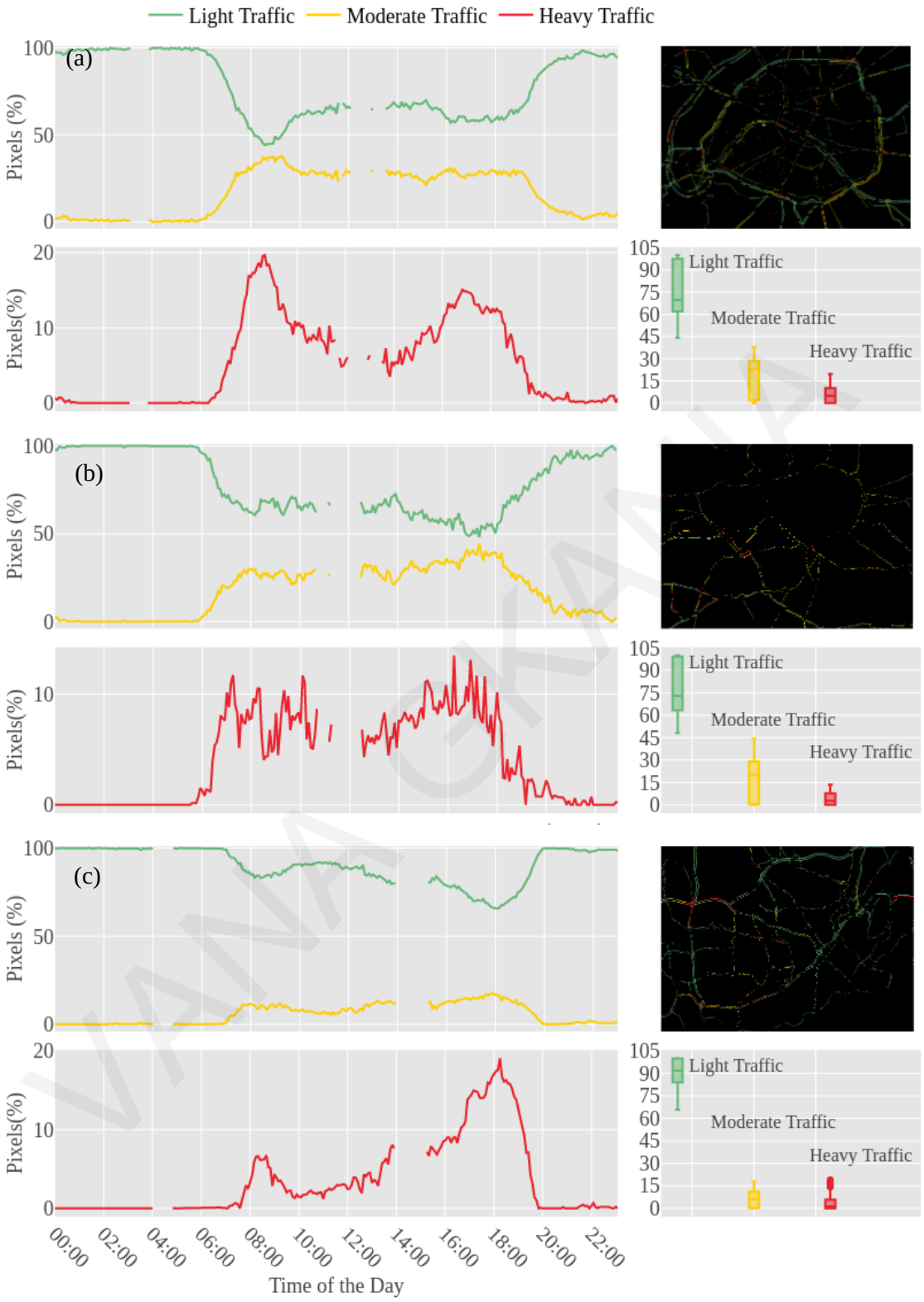


Figure C. 1: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Paris; (b) London; (c) Istanbul

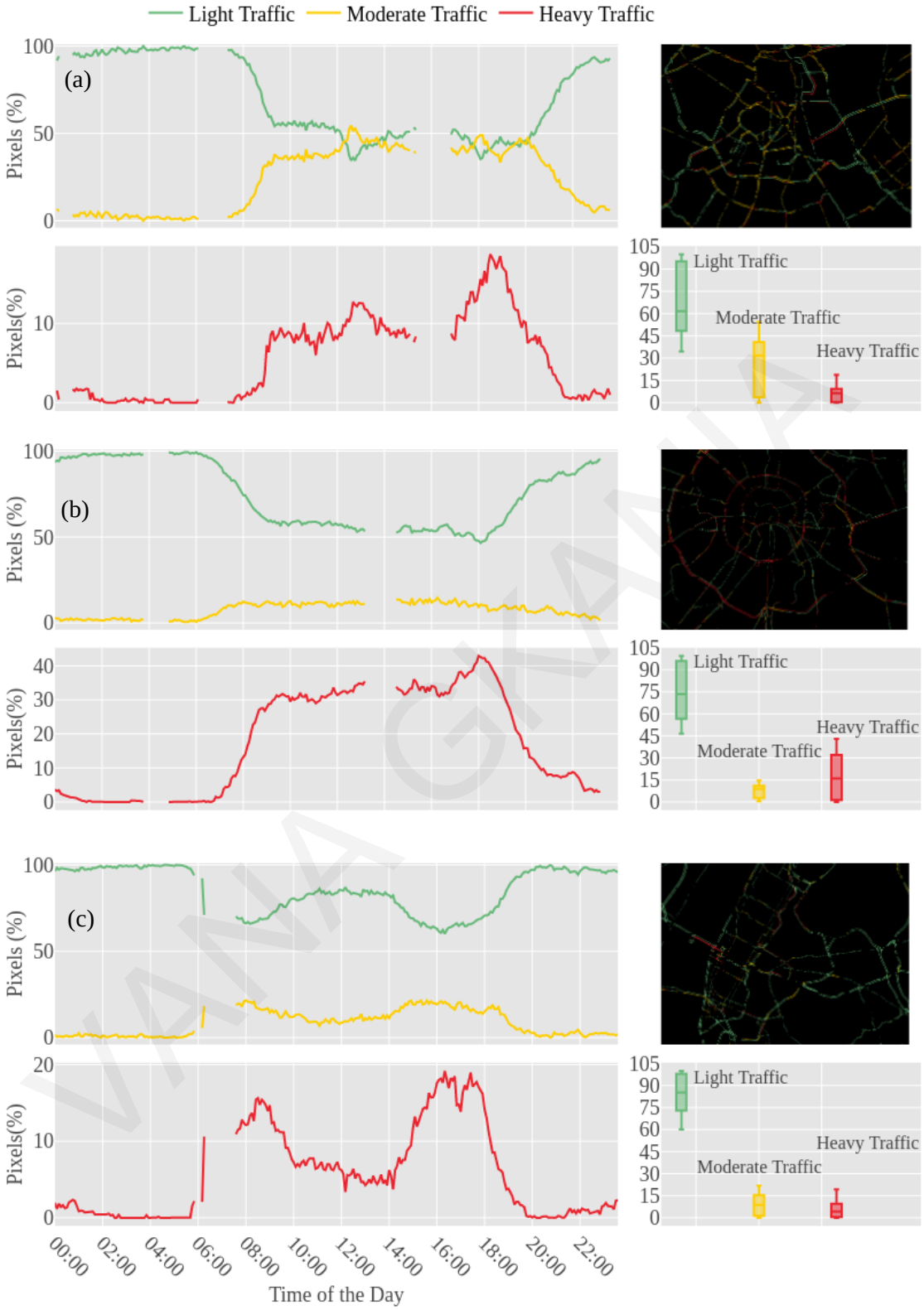


Figure C. 2: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) New Delhi; (b) Moscow; (c) New York

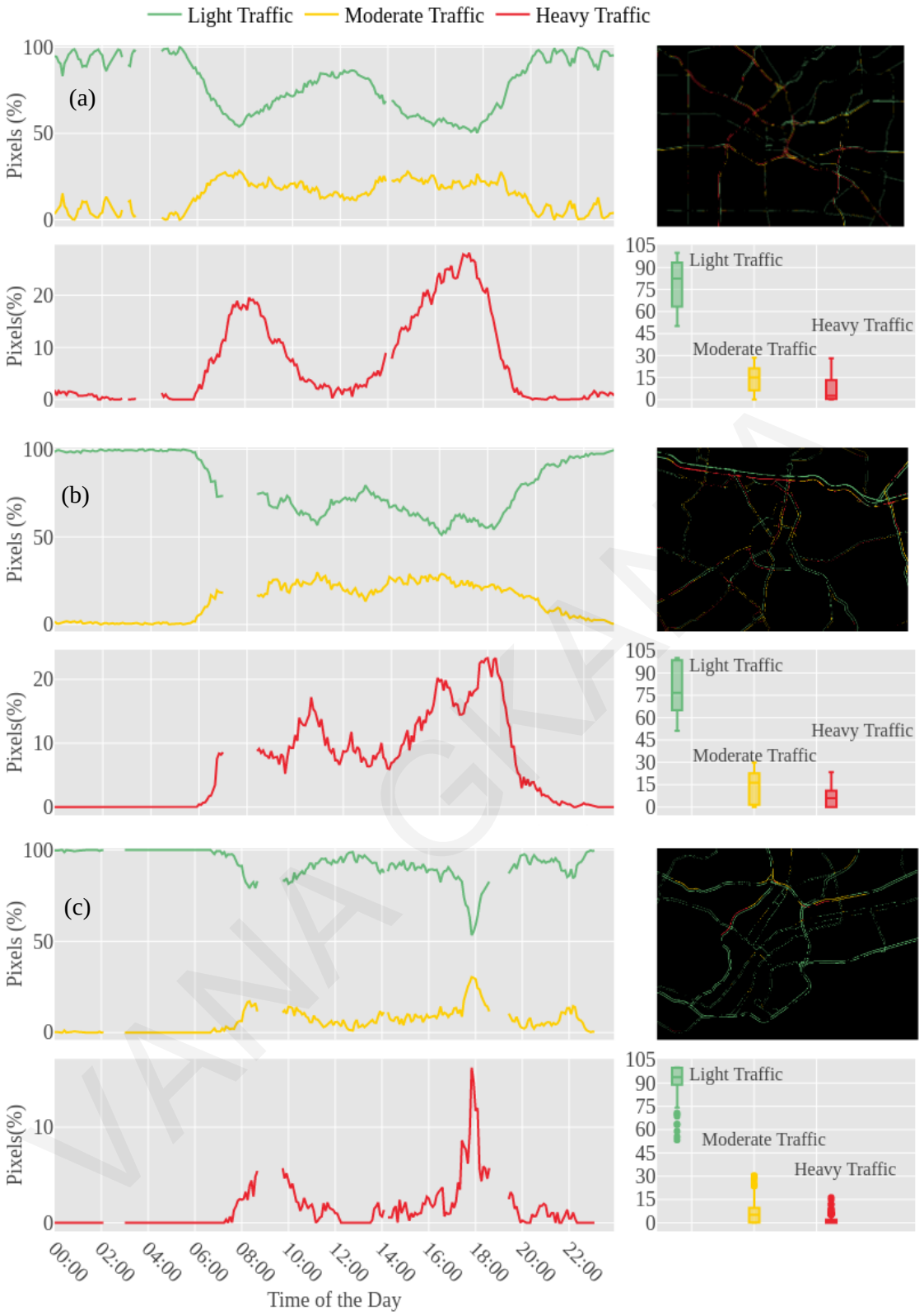


Figure C. 3: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Los Angeles; (b) Sao Paulo; (c) Singapore

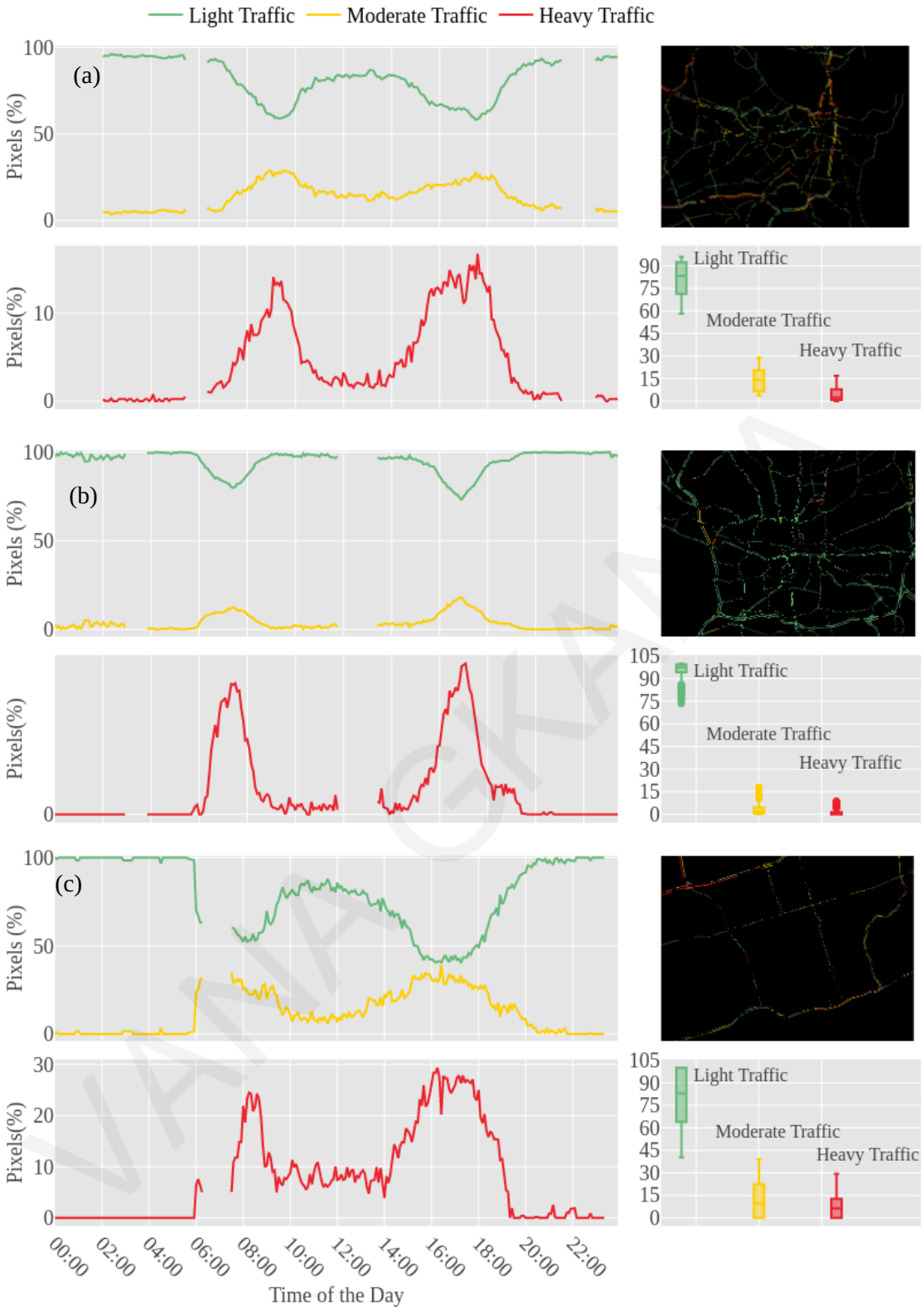


Figure C. 4: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Sydney; (b) Johannesburg; (c) Toronto

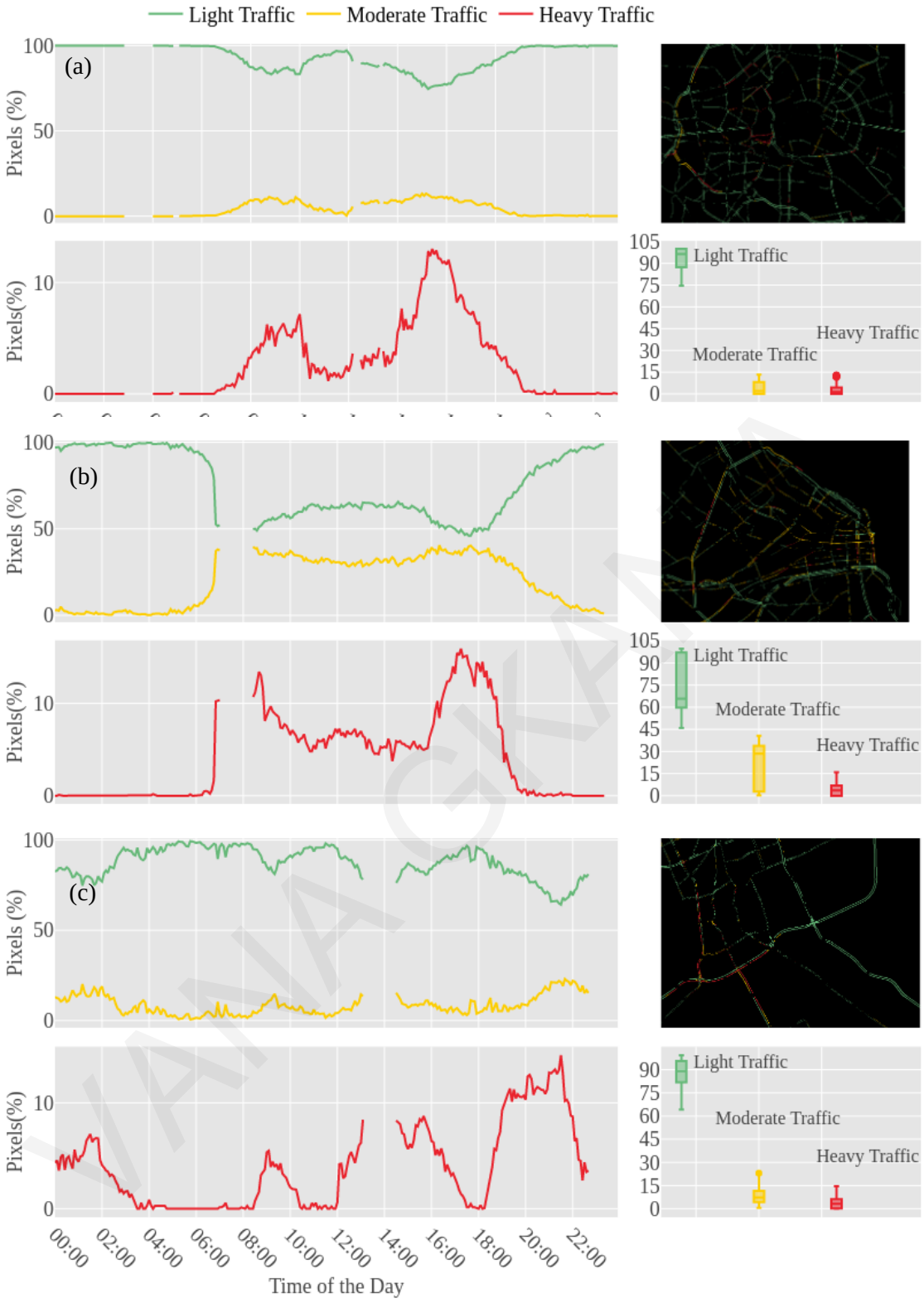


Figure C. 5: Timeseries of pixels and boxplot per traffic layer for one weekday for (a) Berlin; (b) Buenos Aires; (c) Riyadh.

Appendix D

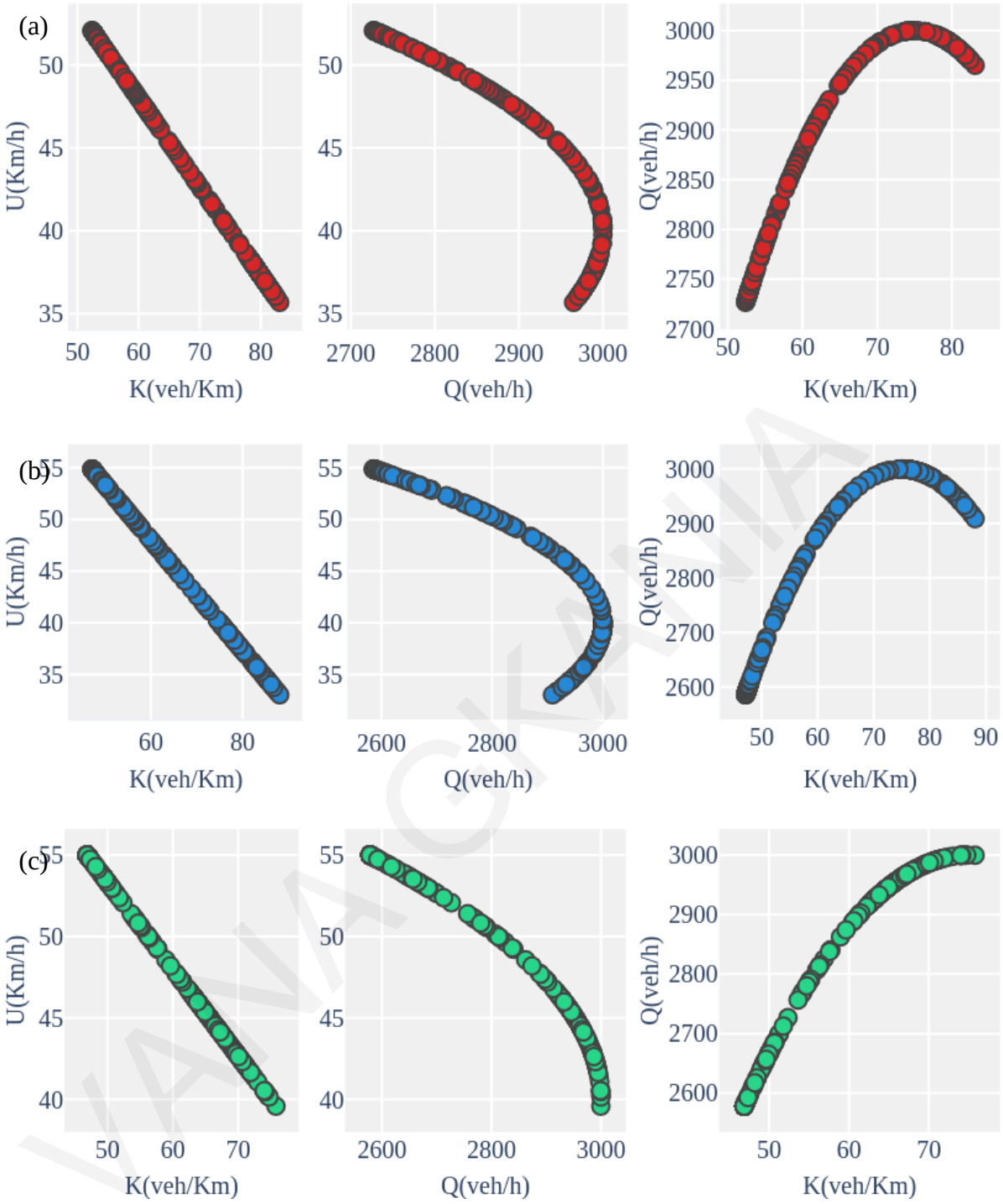


Figure D. 1: London's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

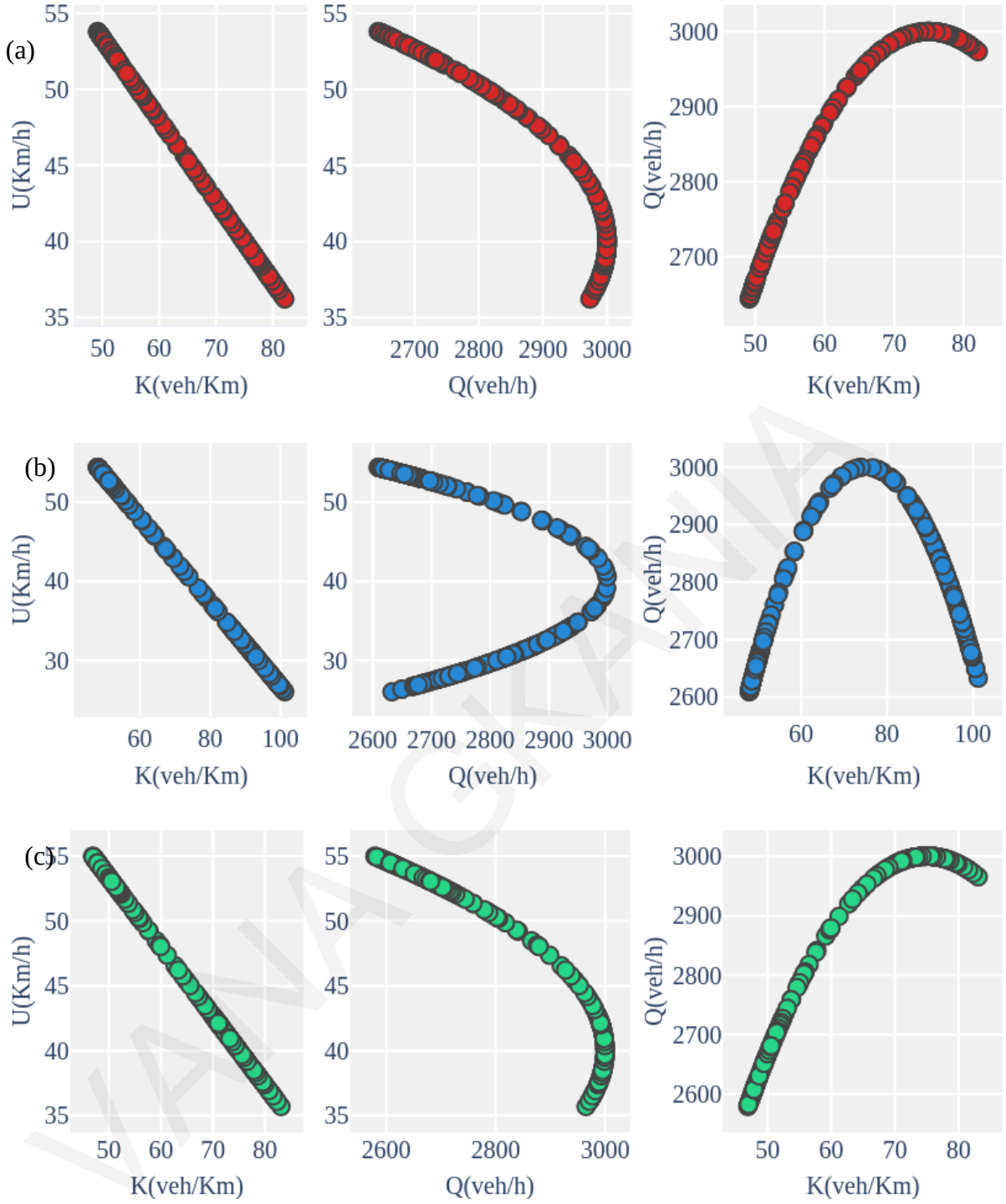


Figure D. 2: New Delhi's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

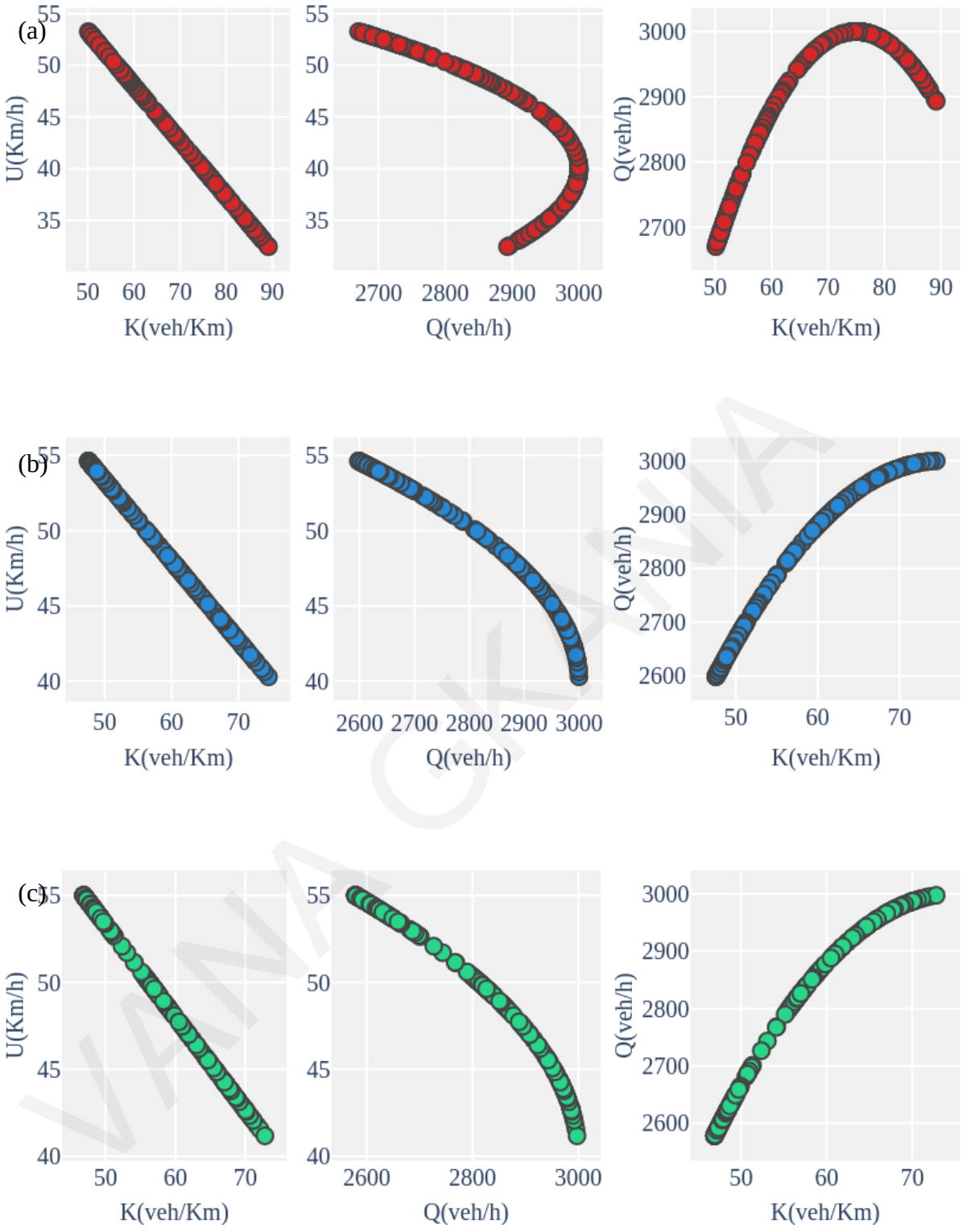


Figure D. 3: New York's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

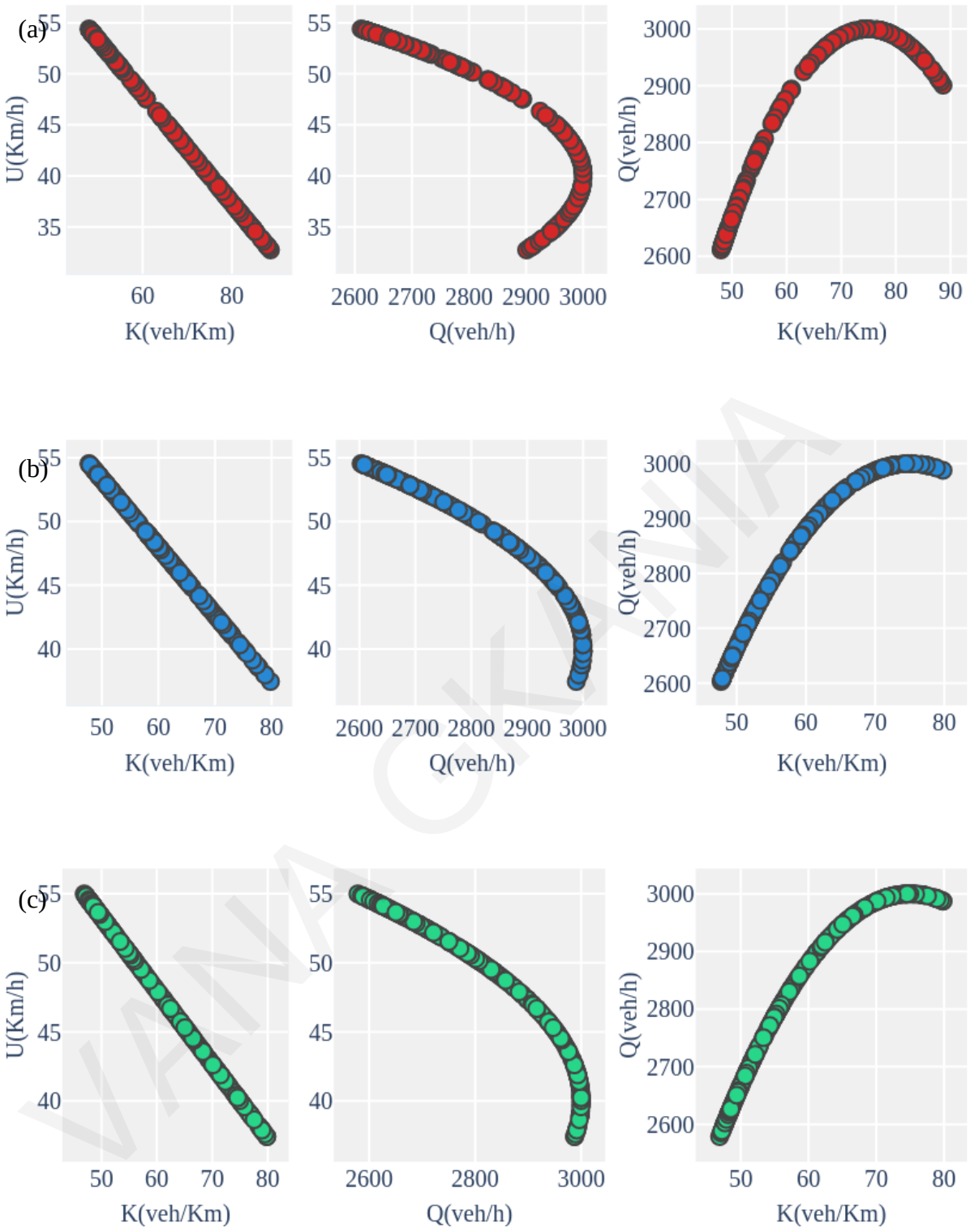


Figure D. 4: Los Angeles's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

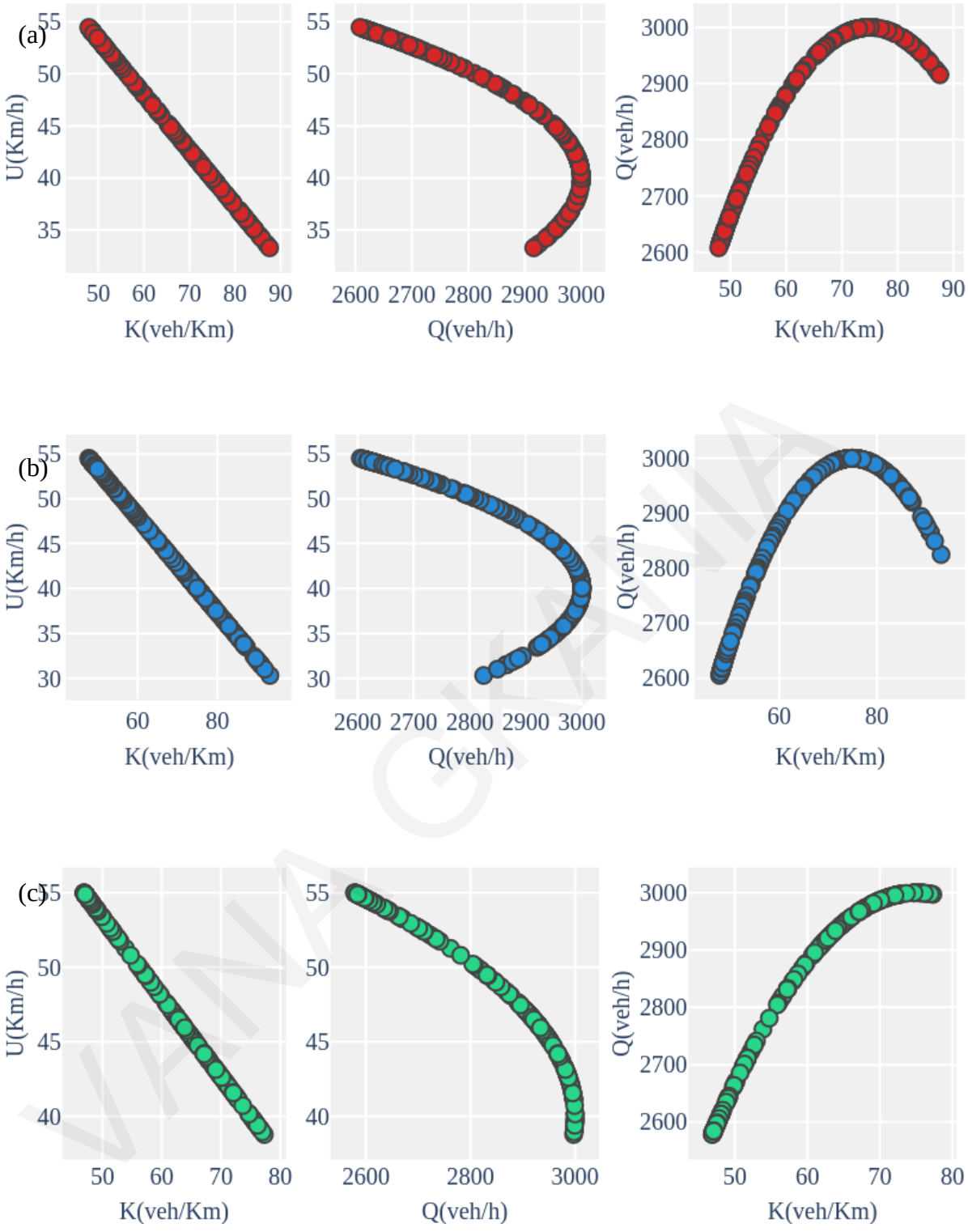


Figure D. 5: Sao Paulo's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

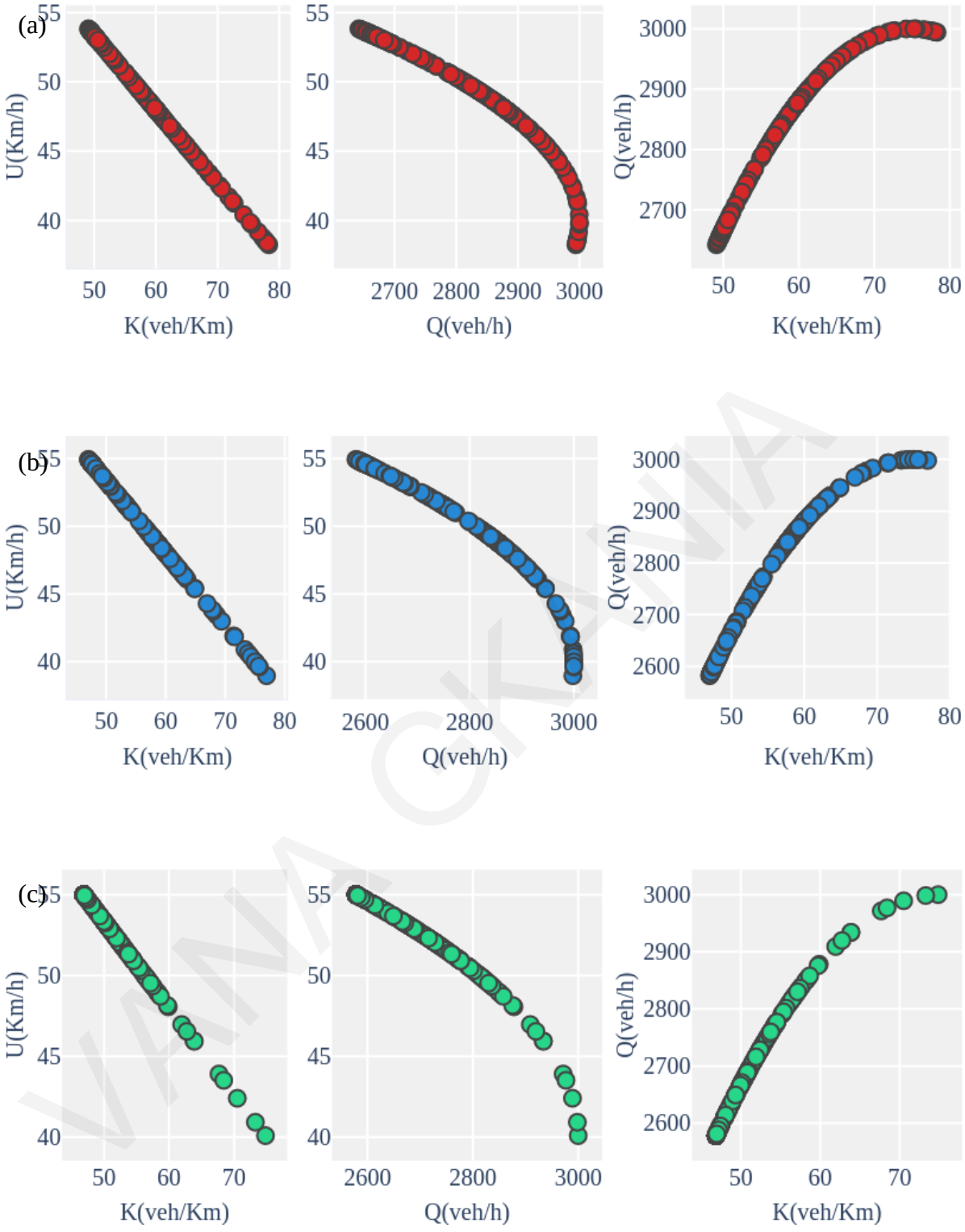


Figure D. 6: Singapore's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

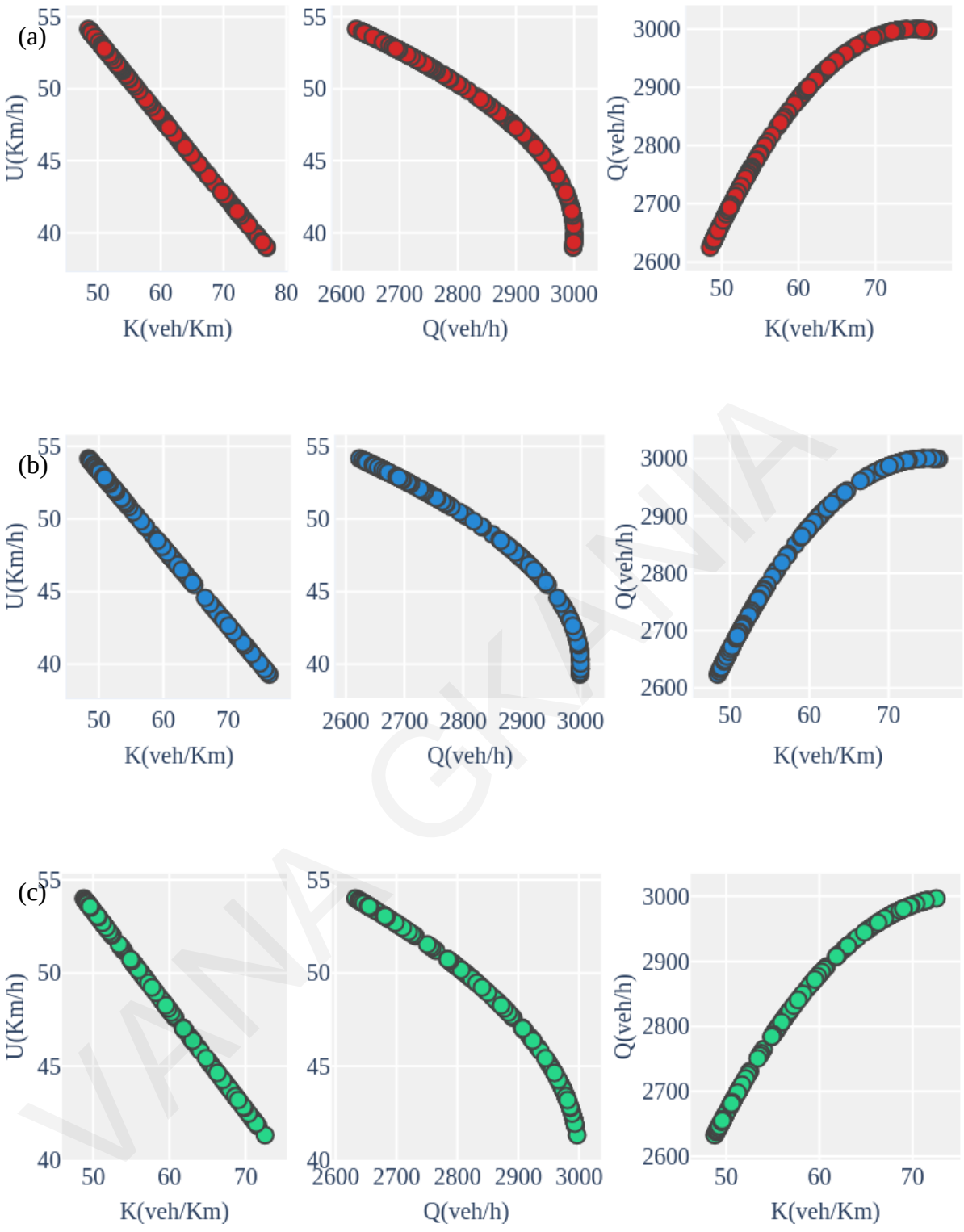


Figure D. 7: Sydney's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

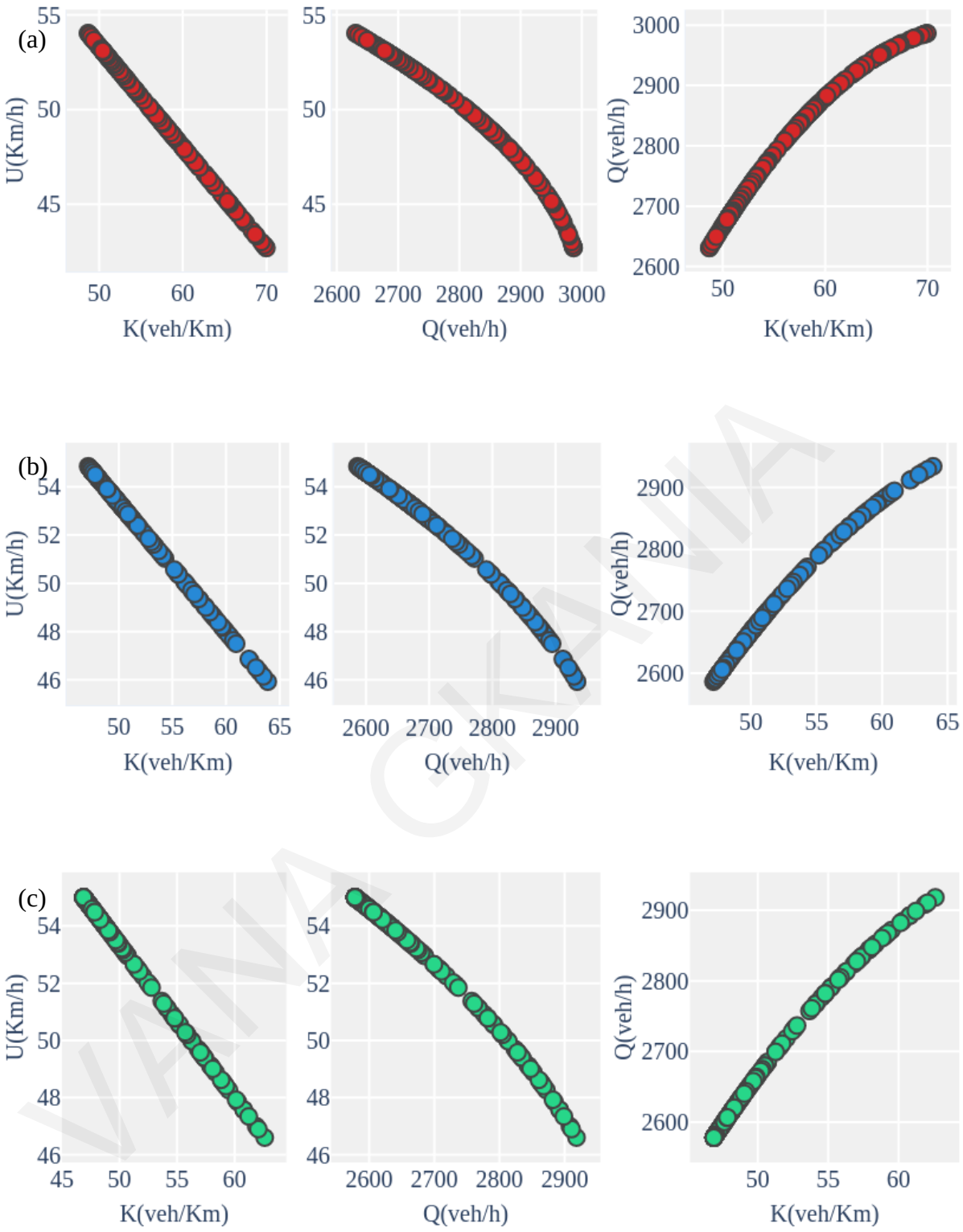


Figure D. 8: Johannesburg's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

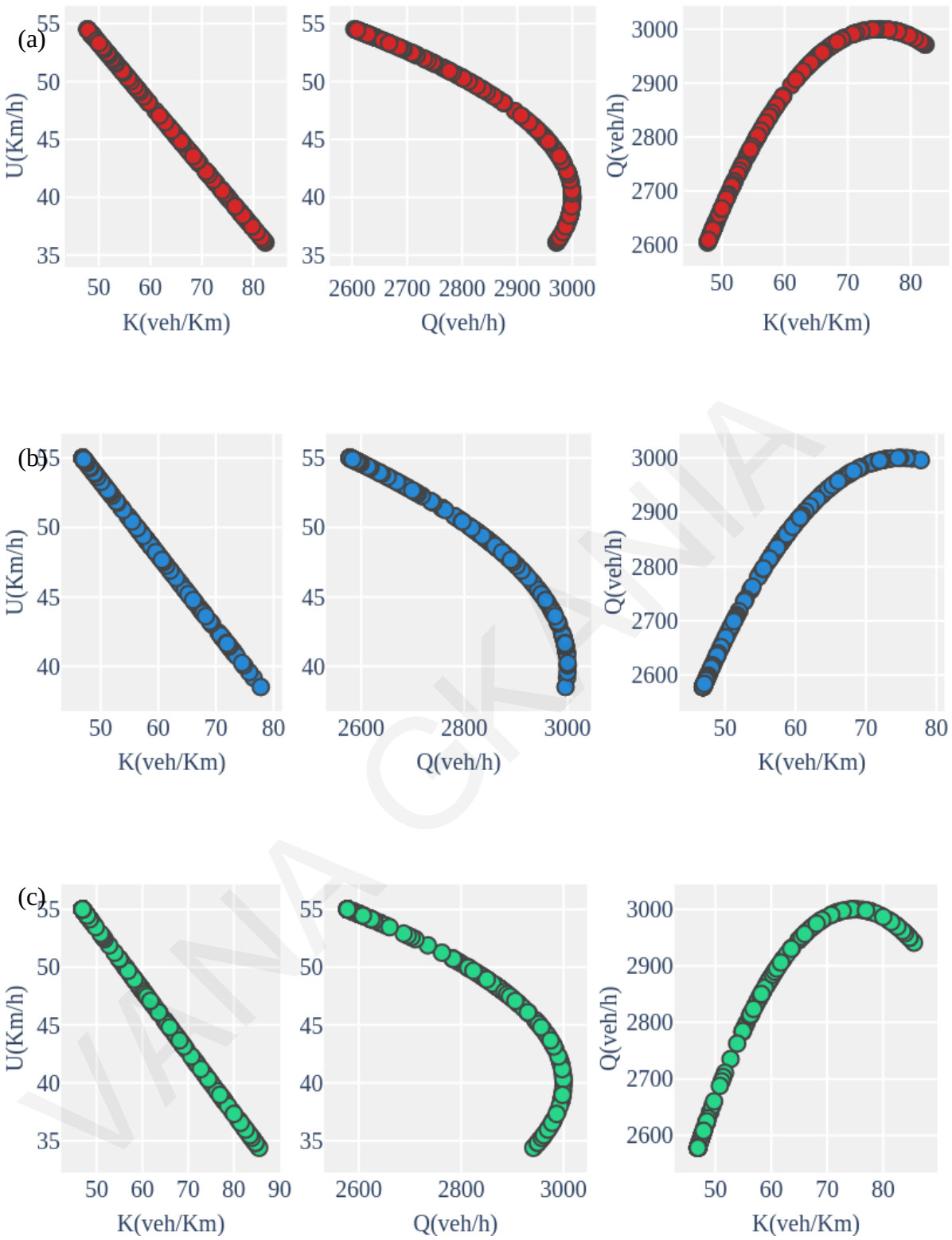


Figure D. 9: Toronto's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

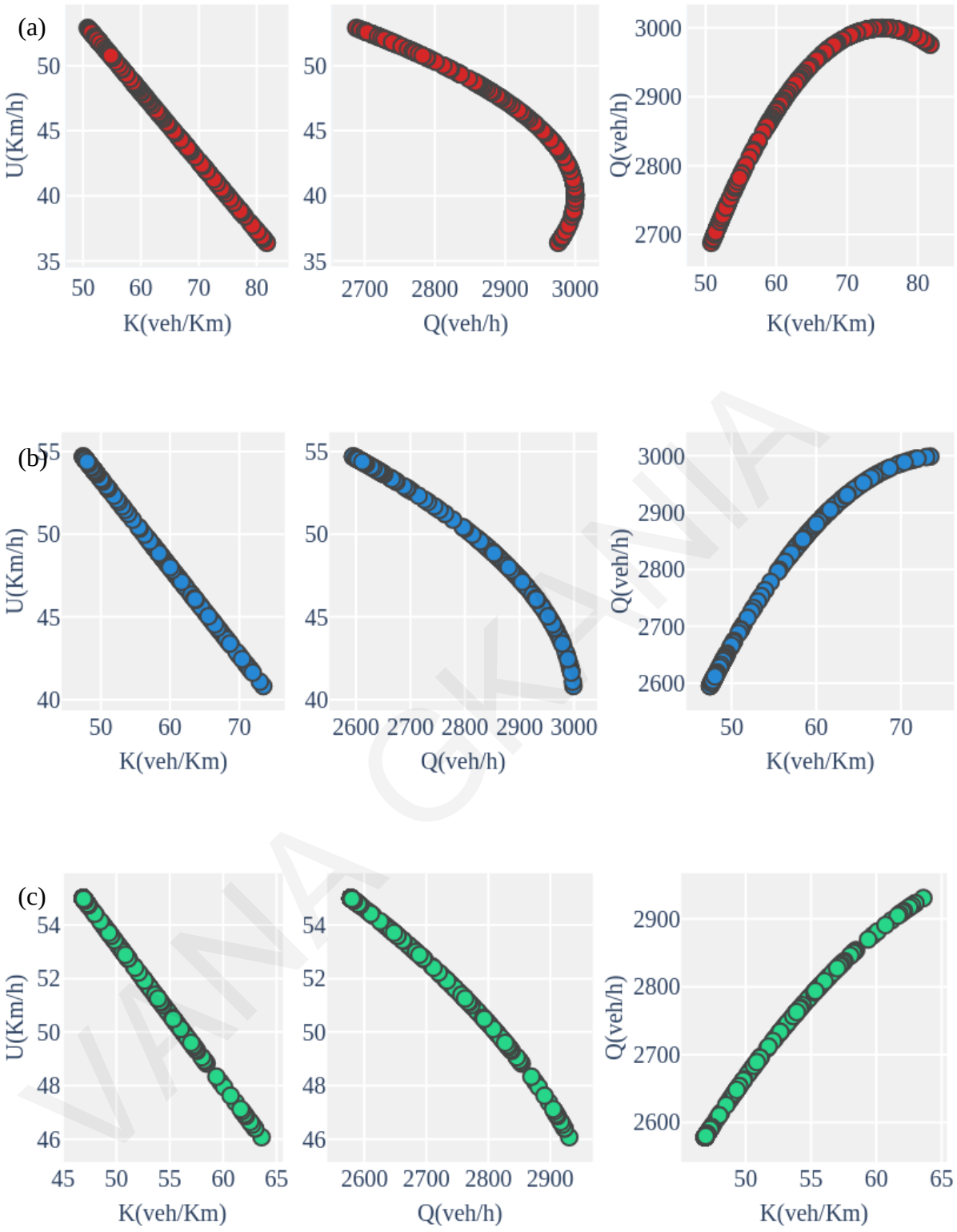


Figure D. 10: Berlin's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

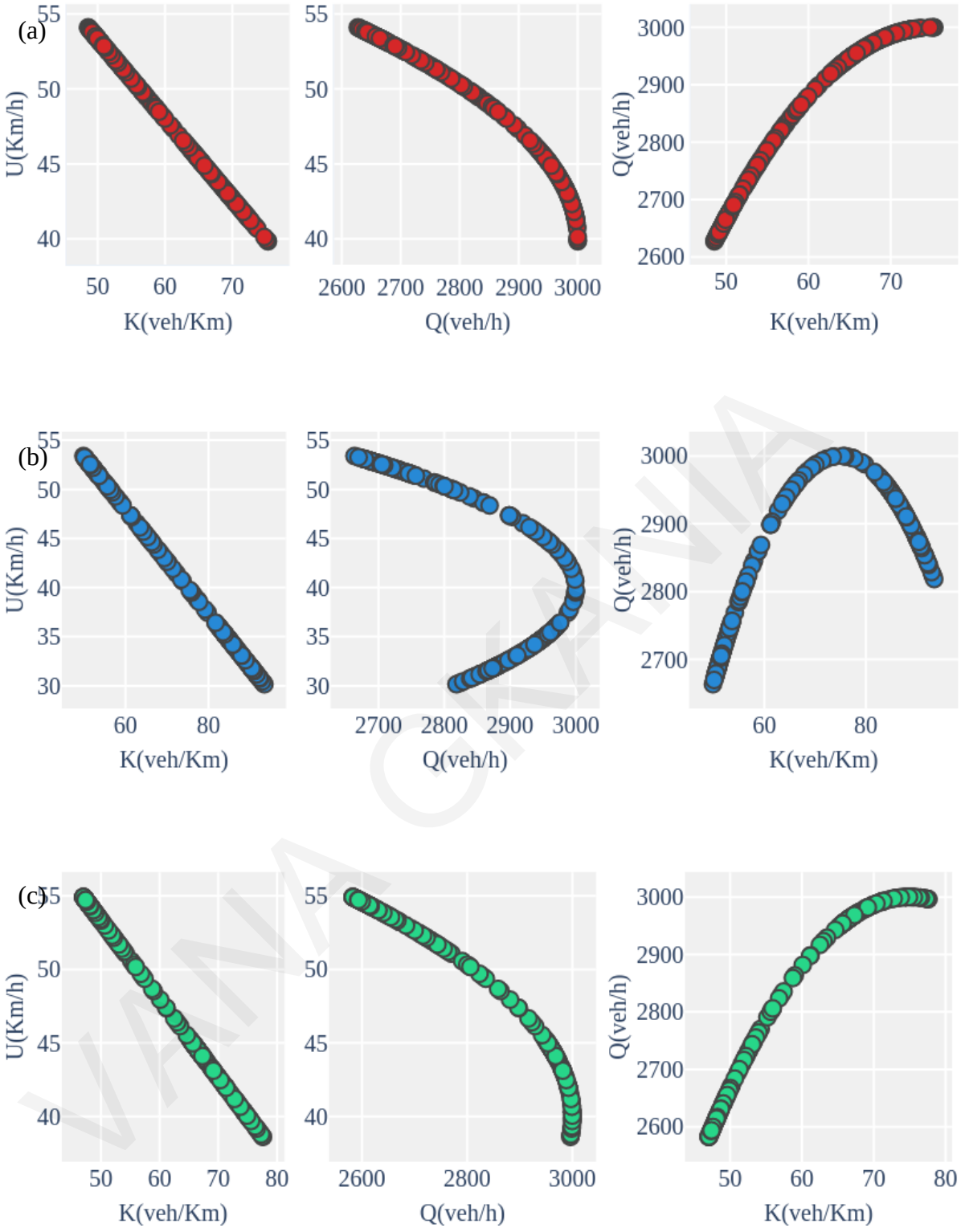


Figure D. 11: Buenos Aires's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

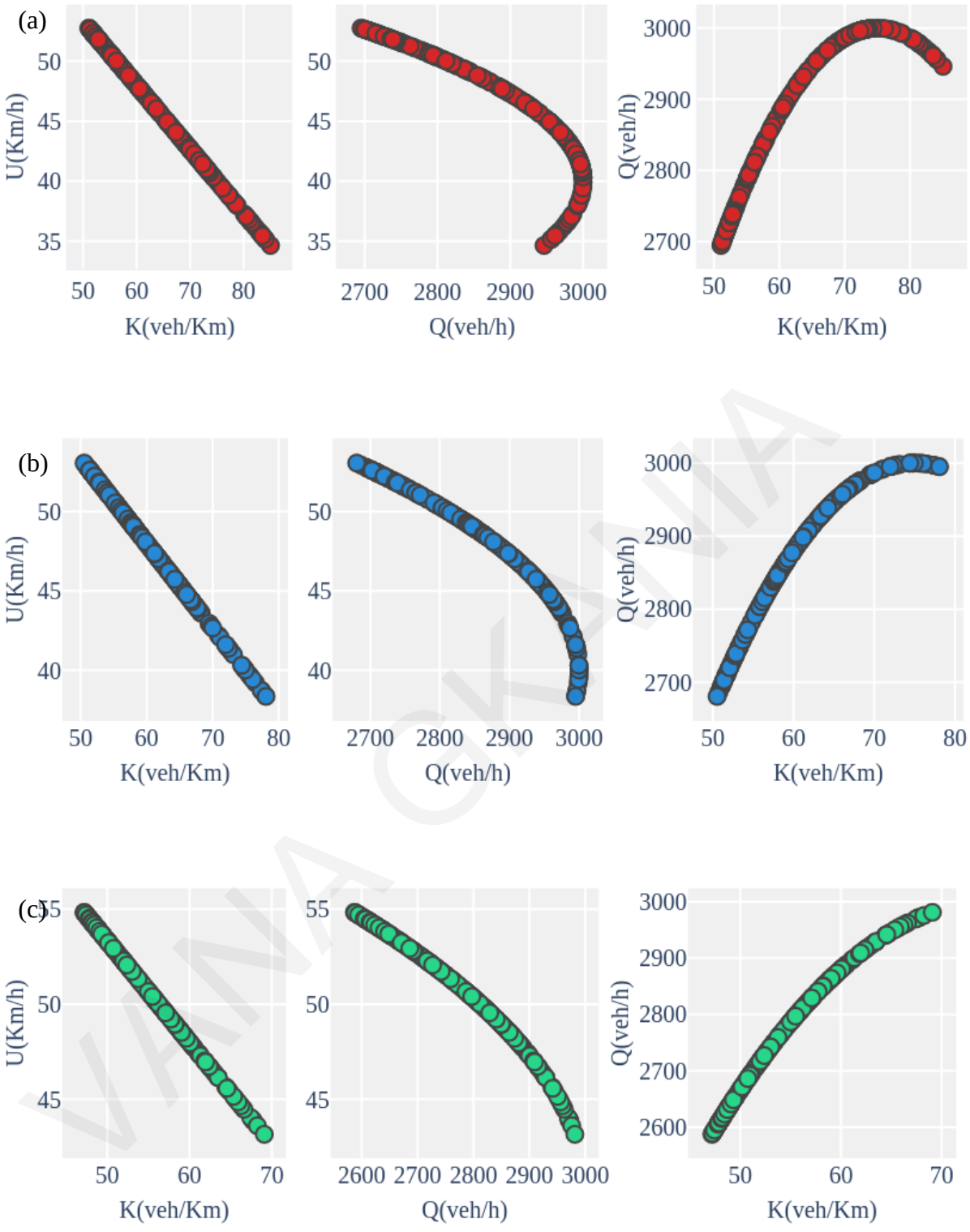


Figure D. 12: Riyadh's MFDs, for a typical weekday, (5-minute intervals) from (a) Google; (b) Bing; (c) Here maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

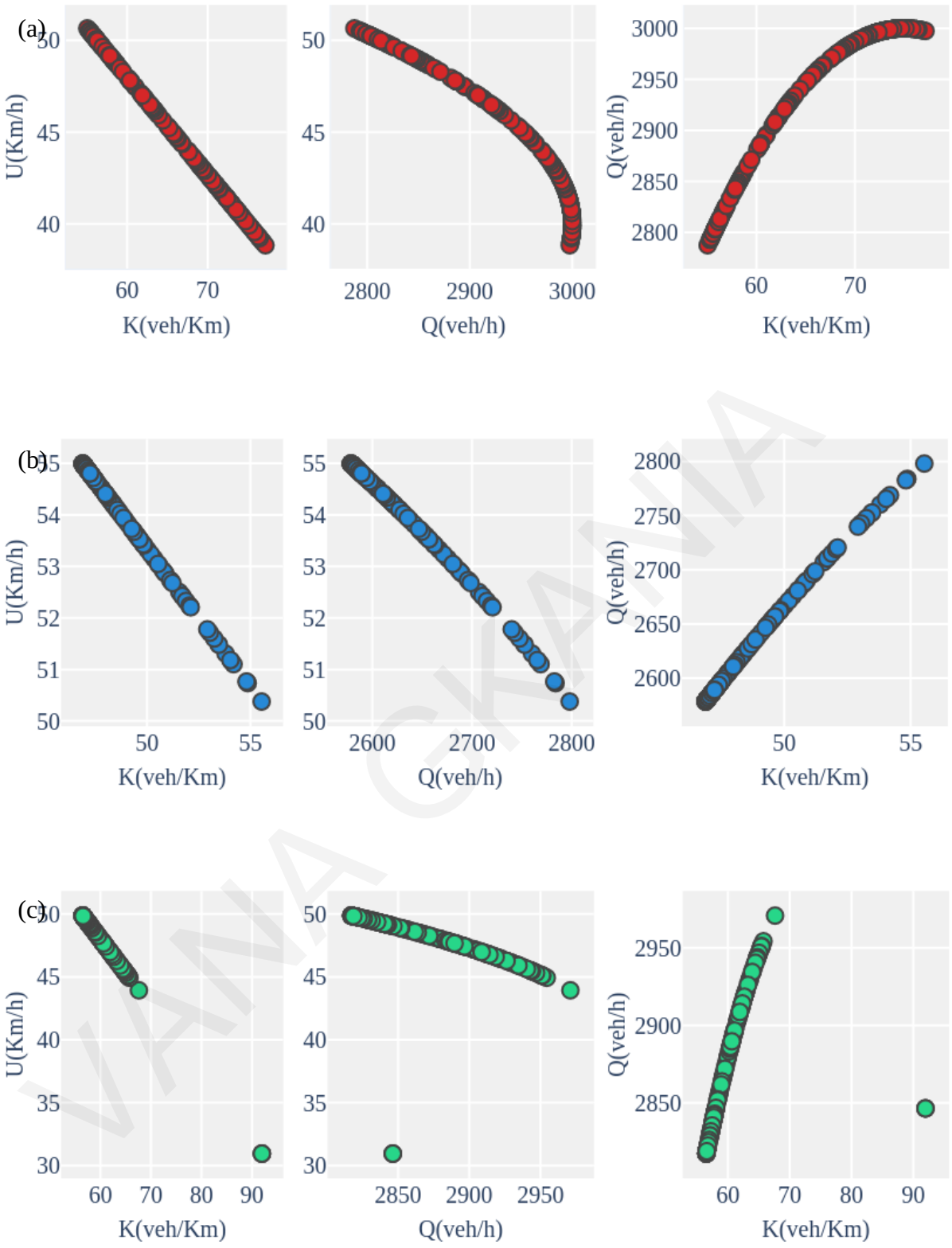
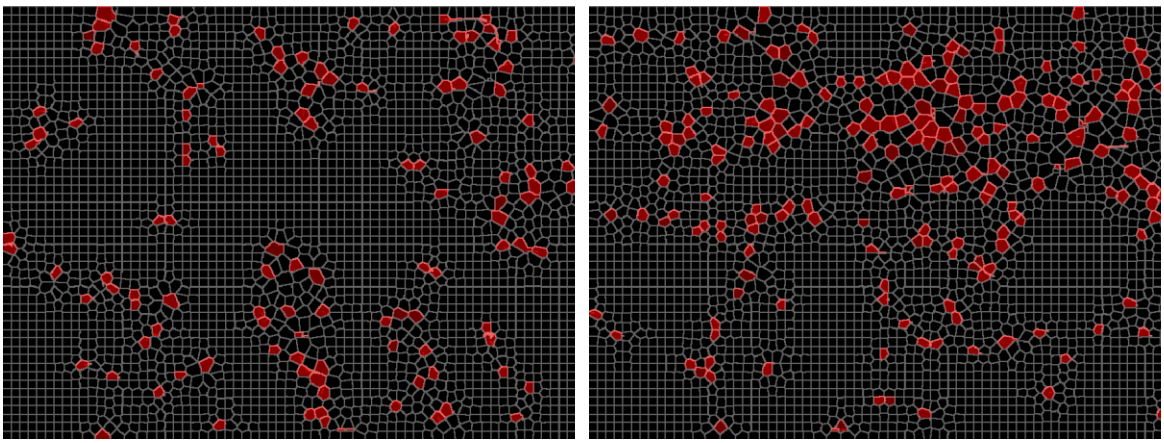
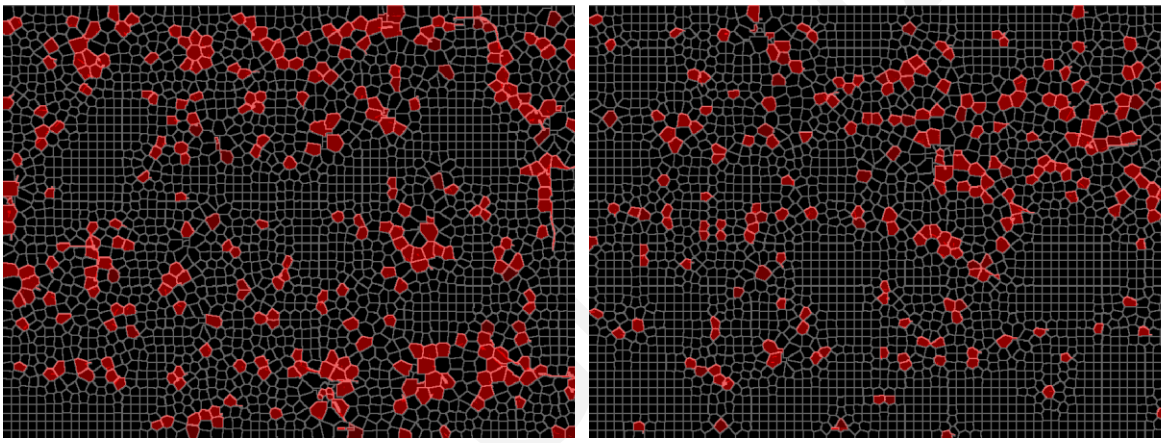


Figure D. 13: (a) Tokyo's MFDs, for a typical weekday, (5-minute intervals) from Google; (b) Beijing's MFDs, for a typical weekday, (5-minute intervals) from Bing; (c) Beijing's MFDs, for a typical weekday, (5-minute intervals) from Baidu maps: (left column) average speed-average density diagram; (middle column) average speed-average flow diagram; (right column) average flow-average density diagram.

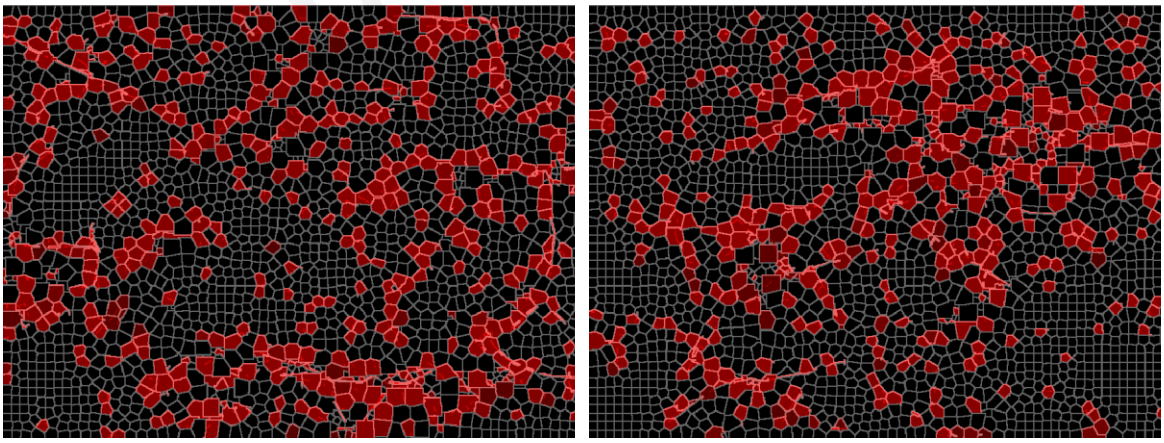
Appendix E



07:00 07:00

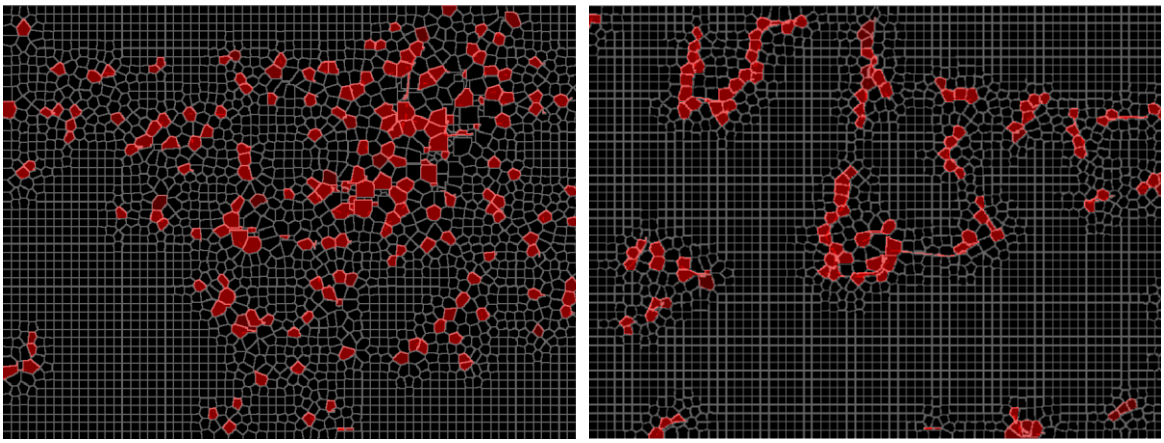


07:30 07:30

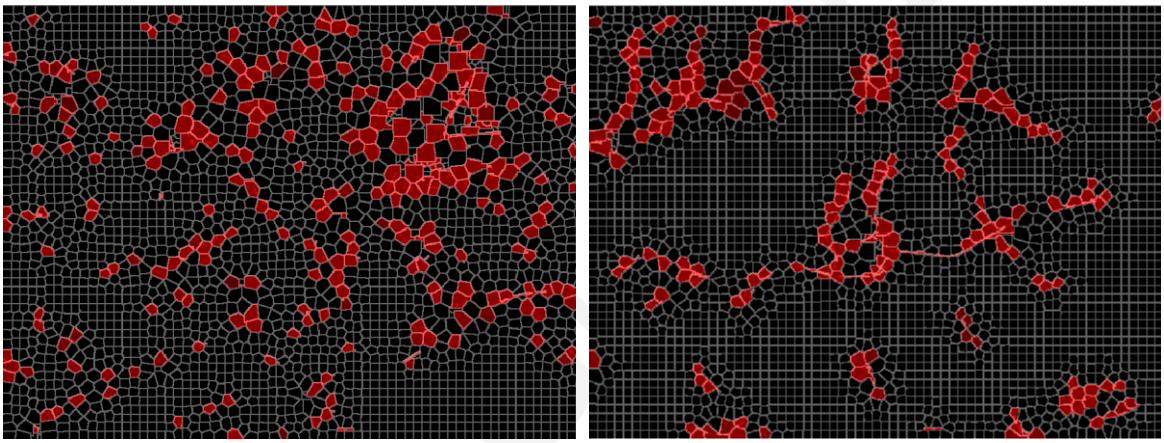


08:00 08:00

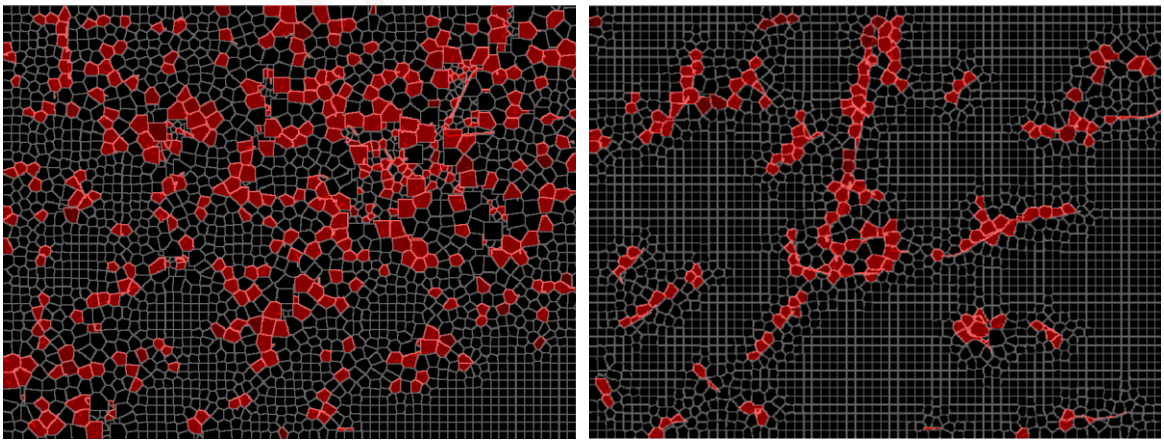
Figure E. 1: Spatio-temporal congestion patterns for Paris (left) and London (right) during morning peak hours (07:00-08:00 a.m.).



07:00 07:00

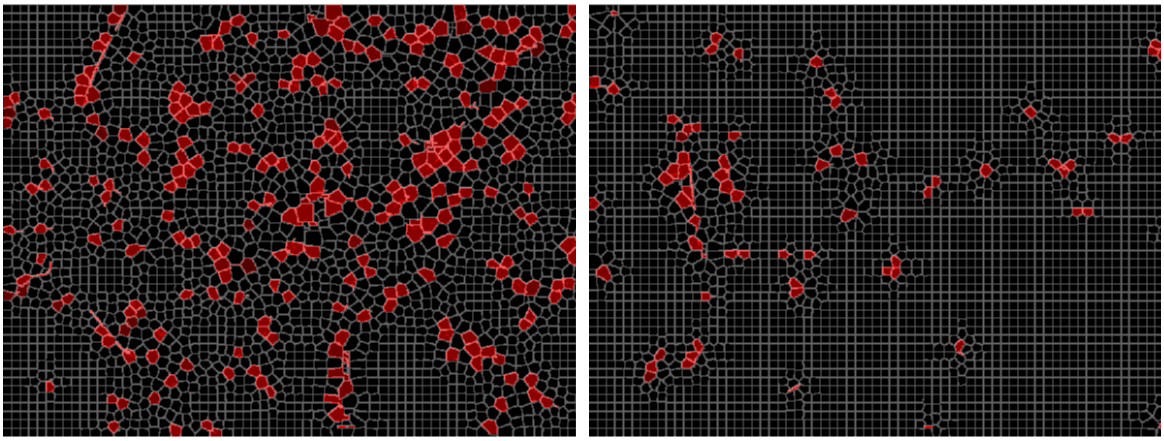


07:30 07:30

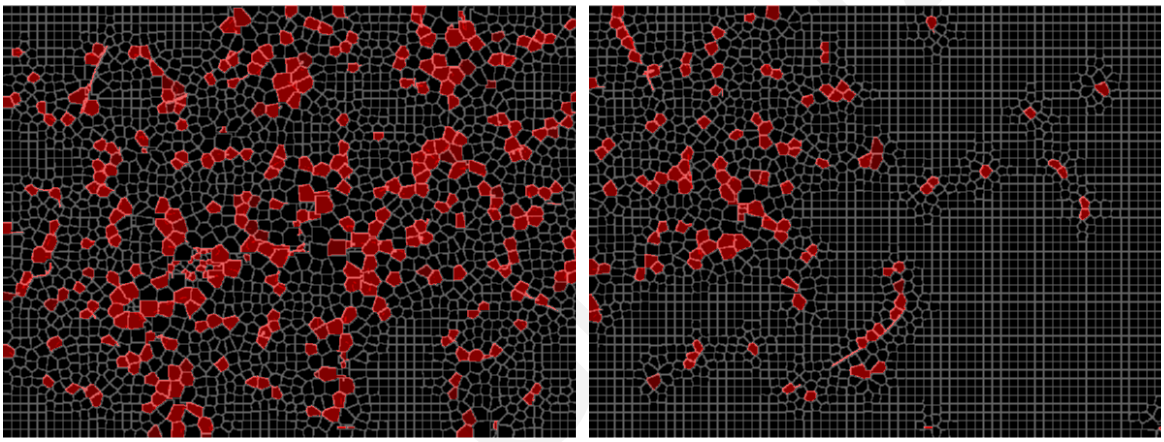


08:00 08:00

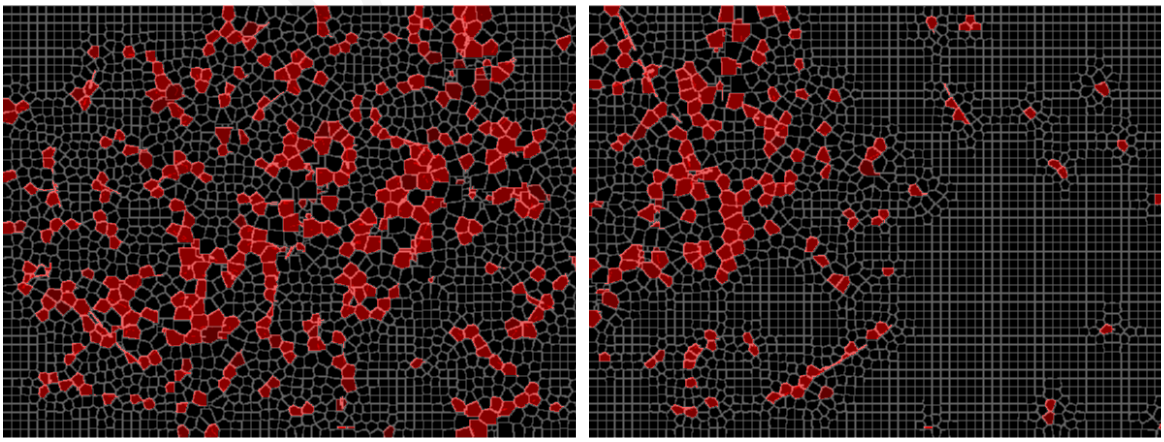
Figure E. 2: Spatio-temporal congestion patterns for Tokyo (left) and Sao Paulo (right) during morning peak hours (07:00-08:00 a.m.).



07:00 07:00

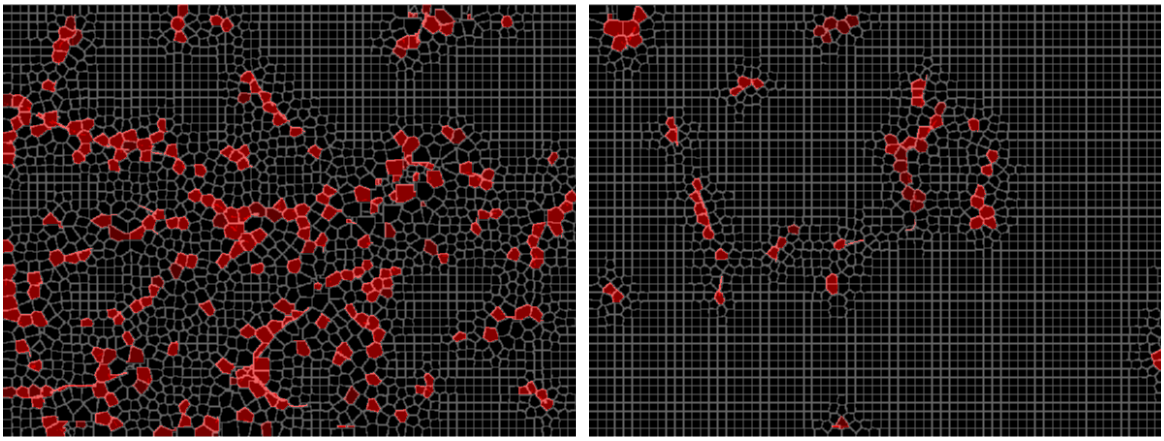


07:30 07:30

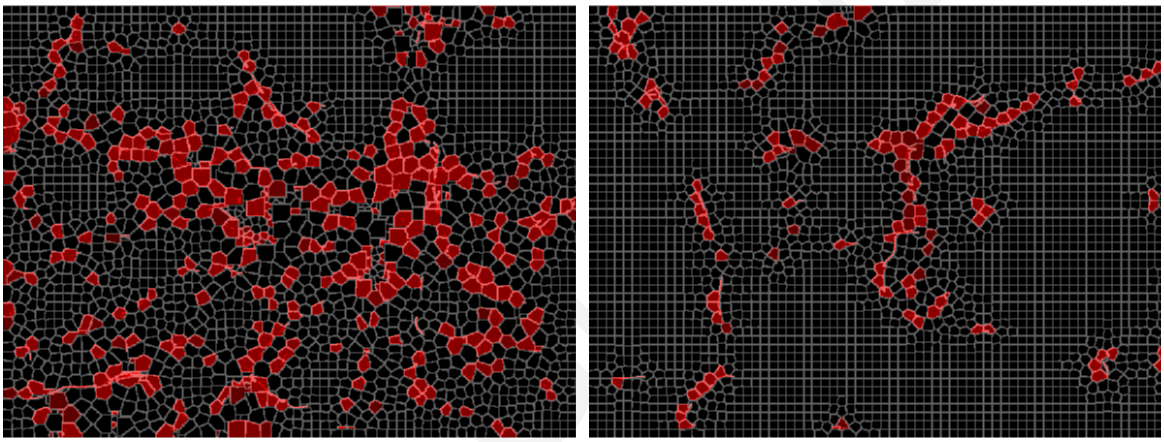


08:00 08:00

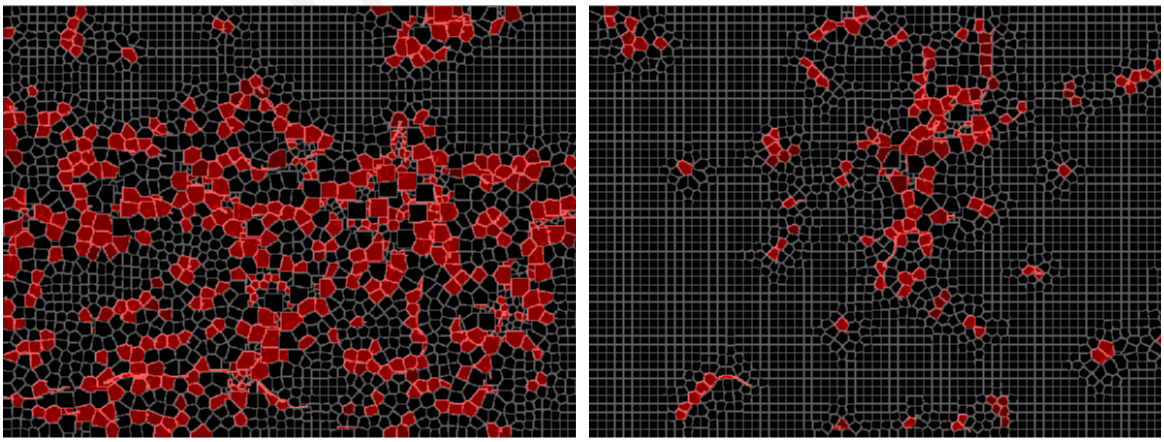
Figure E. 3: Spatio-temporal congestion patterns for Berlin (left) and Riyadh (right) during morning peak hours (07:00-08:00 a.m.).



07:00 07:00

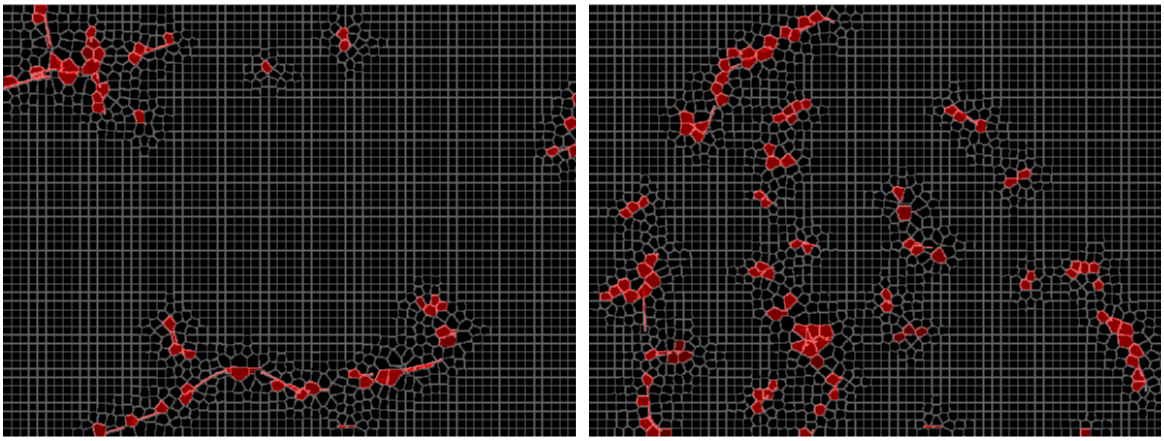


07:30 07:30

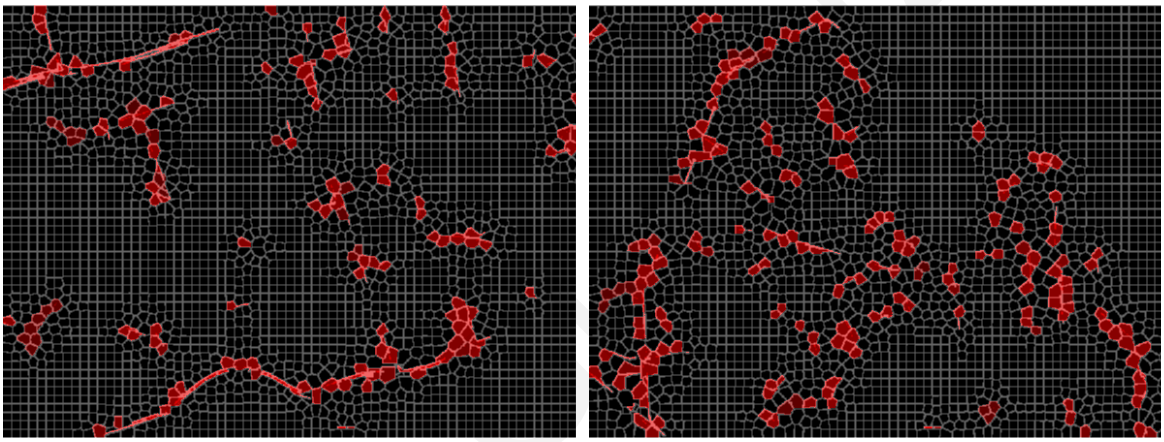


08:00 08:00

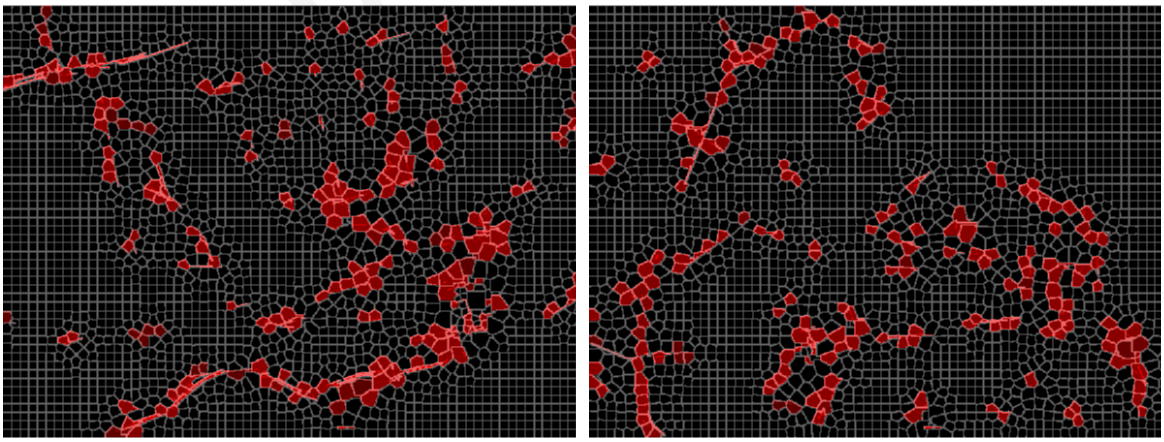
Figure E. 4: Spatio-temporal congestion patterns for Sydney (left) and Johannesburg (right) during morning peak hours (07:00-08:00 a.m.).



07:00 07:00

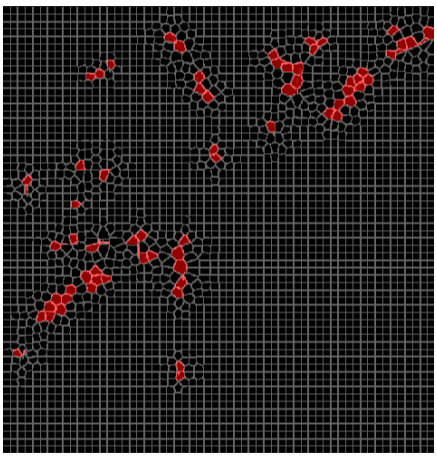


07:30 07:30

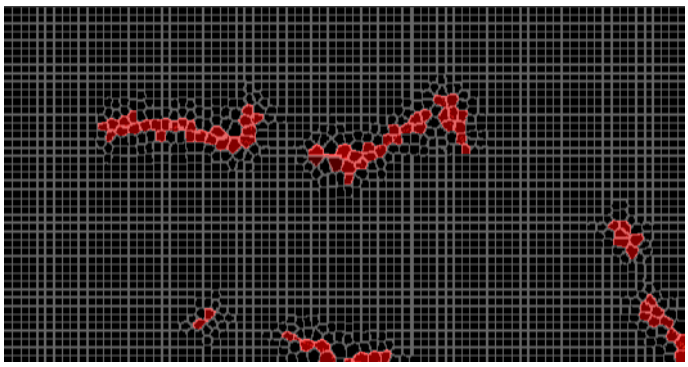


08:00 08:00

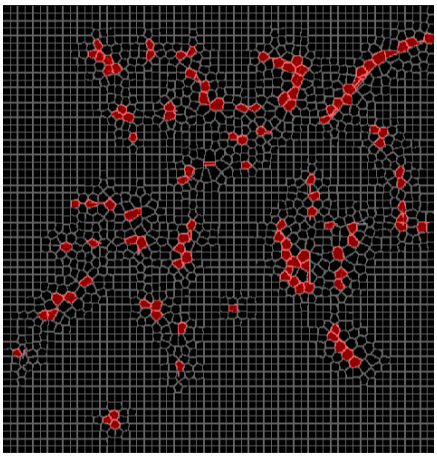
Figure E. 5: Spatio-temporal congestion patterns for Toronto (left) and Buenos Aires (right) during morning peak hours (07:00-08:00 a.m.).



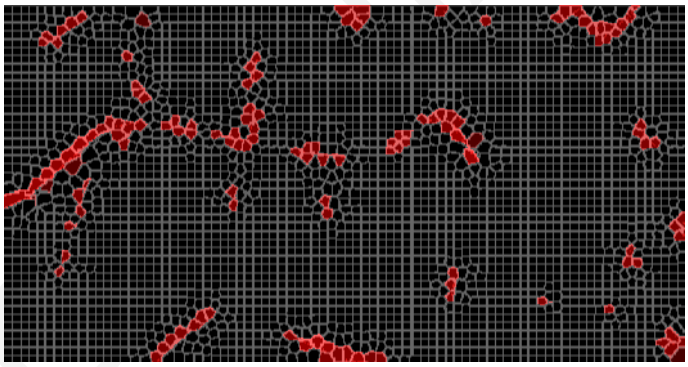
07:00



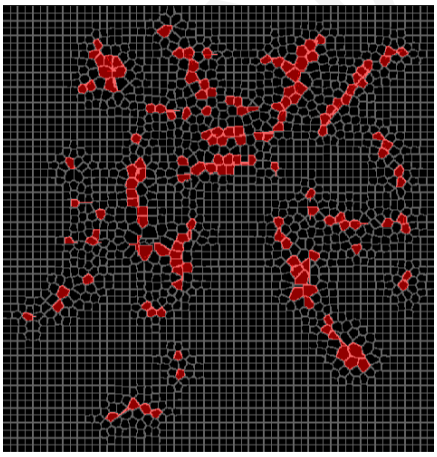
07:00



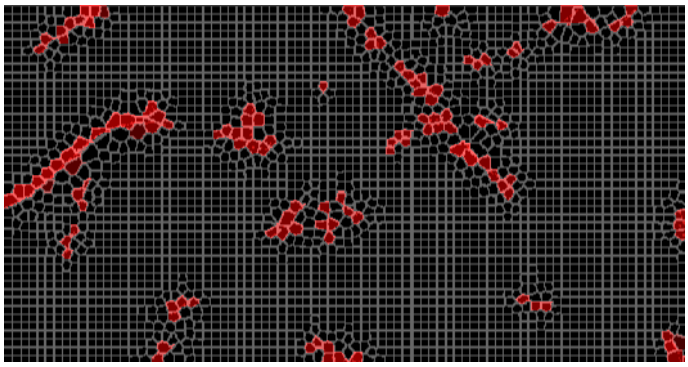
07:30



07:30



08:00



08:00

Figure E. 6: Spatio-temporal congestion patterns for Beijing (left) and Nicosia (right) during morning peak hours (07:00-08:00 a.m.)