ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ ΚΑΙ ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

Ανάπτυξη μεθοδολογίας εντοπισμού και εκτίμησης διάρκειας συμβάντων κορεσμού σε αστικά οδικά δίκτυα

Διπλωματική Εργασία

Στέφανος Δεμένεγας

Επιβλέπουσα: Βλαχογιάννη Ελένη, Καθηγήτρια Ε.Μ.Π.

ΑΘΗΝΑ, ΙΟΥΛΙΟΣ 2023

Copyright © Στέφανος Δεμένεγας, 2023
Με επιφύλαξη παντός δικαιώματος

Απαγορεύεται η αντιγραφή, αποθήκευση σε αρχείο πληροφοριών, διανομή, αναπαραγωγή, μετάφραση ή μετάδοση της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό, υπό οποιαδήποτε μορφή και με οποιοδήποτε μέσο επικοινωνίας, ηλεκτρονικό ή μηχανικό, χωρίς την προηγούμενη έγγραφη άδεια του συγγραφέα. Επιτρέπεται η αναπαραγωγή, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν στη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Η έγκριση της διπλωματικής εργασίας από τη Σχολή Πολιτικών Μηχανικών του Εθνικού Μετσοβίου Πολυτεχνείου δεν υποδηλώνει αποδοχή των απόψεων συγγραφέα (Ν. 5343/1932, Άρθρο 202).

Copyright © Stefanos Demenegas, 2023
All Rights Reserved

Neither the whole nor any part of this diploma thesis may be copied, stored in a retrieval system, distributed, reproduced, translated, or transmitted for commercial purposes, in any form or by any means now or hereafter known, electronic or mechanical, without the written permission from the author. Reproducing, storing and distributing this thesis for non-profitable, educational or research purposes is allowed, without prejudice to reference to its source and to inclusion of the present text. Any queries in relation to the use of the present thesis for commercial purposes must be addressed to its author.

Approval of this diploma thesis by the School of Civil Engineering of the National Technical University of Athens (NTUA) does not constitute in any way an acceptance of the views of the author contained herein by the said academic organization (L. 5343/1932, art. 202).

Στέφανος Δεμένεγας (2023)

Ανάπτυξη μεθοδολογίας εντοπισμού και εκτίμησης διάρκειας συμβάντων κορεσμού σε αστικά οδικά δίκτυα

Διπλωματική Εργασία

Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα.

Stefanos Demenegas (2023)

Diploma Thesis

Method development of detecting and estimating congestion duration in urban traffic networks

National Technical University of Athens, Greece

Ευχαριστίες

Η ολοκλήρωση της διπλωματικής μου εργασίας σηματοδοτεί την περάτωση του κύκλου προπτυχιακών σπουδών μου στη Σχολή Πολιτικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου. Θα ήθελα να ευχαριστήσω θερμά την επιβλέπουσα καθηγήτρια κ. Ελένη Βλαχογιάννη, για την ανάθεση του ενδιαφέροντος αντικειμένου μελέτης και την πολύτιμη καθοδήγηση καθ' όλη τη διάρκεια εκπόνησης της εργασίας.

Επίσης, ιδιαίτερα σημαντική ήταν και η συμβολή του υποψήφιου διδάκτορα Ε.Μ.Π. Παναγιώτη Φαφουτέλλη, ο οποίος ήταν πάντα διαθέσιμος ώστε να παρέχει καθοριστικές συμβουλές και η βοήθειά του κατά την ανάλυση και συγγραφή της εργασίας ήταν ανεκτίμητη. Τον ευχαριστώ για την άψογη συνεργασία.

Τέλος, θα ήθελα να ευχαριστήσω από καρδιάς την οικογένειά μου και τους φίλους μου, για την υπομονή, κατανόηση και αμέριστη υποστήριξη όλα τα χρόνια των σπουδών μου.

Στέφανος Δεμένεγας

Αθήνα, Ιούλιος 2023

Ανάπτυξη μεθοδολογίας εντοπισμού και εκτίμησης διάρκειας συμβάντων κορεσμού σε αστικά οδικά δίκτυα

Στέφανος Δεμένεγας

Επιβλέπουσα Καθηγήτρια: Ελένη Ι. Βλαχογιάννη, Καθηγήτρια ΕΜΠ

Σύνοψη

Το συνηθέστερο, ίσως, πρόβλημα στην κυκλοφορία σε αστικά οδικά δίκτυα, είναι η εμφάνιση κορεσμού. Η διάρκεια του φαινομένου ποικίλει από συμβάν σε συμβάν. Σκοπός της παρούσας έρευνας, είναι η ανάπτυξη μεθοδολογίας εντοπισμού και εκτίμησης της διάρκειας συμβάντων κορεσμού σε αστικά οδικά δίκτυα, με βάση ιστορικά στοιχεία κυκλοφοριακών μεγεθών. Αρχικά, εφαρμόζονται δύο μέθοδοι εντοπισμού κορεσμού, σε δύο πόλεις, την Αθήνα και το Μάντσεστερ, οι οποίες σαν μεταβλητή εξόδου δίνουν τον αριθμό των συμβάντων κορεσμού. Για την ανάπτυξη μοντέλων εκτίμησης της διάρκειας του κορεσμού, επιλέγονται οι κατάλληλες ανεξάρτητες επεξηγηματικές μεταβλητές, σύμφωνα με την υπόθεση ότι επηρεάζουν τη διάρκεια. Ύστερα, εφαρμόζονται τρία διαφορετικά μοντέλα Ανάλυσης Επιβίωσης (Survival Analysis), ώστε να εξεταστεί η επιρροή κάθε μεταβλητής. Από τα μοντέλα προκύπτει πως η ώρα αιχμής, δηλαδή το διάστημα κατά το οποίο παρουσιάζονται οι μεγαλύτεροι κυκλοφοριακοί φόρτοι, είναι αυτό που επηρεάζει περισσότερο την διάρκεια του κορεσμού.

Λέξεις-κλειδιά: κορεσμός, διάρκεια, Ανάλυση Επιβίωσης, εντοπισμός κορεσμού,

Method development of detecting and estimating congestion duration in urban traffic networks

Stefanos Demenegas

Supervising Professor: Eleni I. Vlahogianni, Professor NTUA

Abstract

Congestion occurrence might be the most common traffic problem in urban traffic networks. The duration of each incident varies from one another. The purpose of this research is the method development of detecting and estimating congestion duration in urban traffic networks, based on historical traffic data. Firstly, two congestion detection methods are applied, in two cities, Athens and Manchester, with the output being the number of congestion occurrences. For the development of the congestion duration estimating models, the most suitable independent explanatory variables are chosen according to the assumption that they have impact on duration. Then, three different Survival Analysis models are applied, to examine the impact of each variable. In line with the results, it turns out that peak hour, which represents the period of time when traffic flow reaches its peak, affects congestion duration the most.

Keywords: congestion, duration, Survival Analysis, congestion detection

Περίληψη

Η αστικοποίηση είναι ένα φαινόμενο που παρατηρείται να έχει αυξητική τάση τα τελευταία χρόνια. Η συνεχής μετακίνηση και εγκατάσταση πληθυσμών στα αστικά κέντρα έχει άμεσες επιπτώσεις στην κυκλοφορία. Οι αυξημένες ανάγκες του πληθυσμού για μετακινήσεις έχουν πολύ συχνά ως αποτέλεσμα τον κορεσμό της κυκλοφορίας στα αστικά οδικά δίκτυα. Η εκτίμηση της διάρκειας ενός συμβάντος κορεσμού θα μπορούσε να φανεί χρήσιμη στον τομέα της διαχείρισης κυκλοφορίας.

Ο κορεσμός εμφανίζεται σε μία οδική υποδομή, όταν ο εισερχόμενος σε αυτή φόρτος οχημάτων υπερβαίνει την κυκλοφοριακή της ικανότητα. Τα είδη του κορεσμού είναι δύο: επαναλαμβανόμενος και μη επαναλαμβανόμενος κορεσμός. Στην πρώτη περίπτωση πρόκειται για συμβάντα τα οποία παρατηρούνται σε τμήμα ενός δικτύου ή σε συγκεκριμένες χρονικές στιγμές με συγκεκριμένη συχνότητα, ενώ στη δεύτερη, για μη αναμενόμενα συμβάντα τα οποία προκαλούνται από αναπάντεχα γεγονότα, όπως ακραία καιρικά φαινόμενα ή οδικά ατυχήματα.

Σε βάσεις δεδομένων κυκλοφοριακών μεγεθών στην Αθήνα και το Μάντσεστερ, εφαρμόζονται δύο μέθοδοι εντοπισμού κορεσμού. Η πρώτη, η μέθοδος «Spillover», λαμβάνει υπόψη μία εξίσωση η οποία υπολογίζει το ανώτατο όριο της κατάληψης, πάνω από το οποίο υπάρχει κορεσμός, συναρτήσει του κυκλοφοριακού φόρτου. Στη δεύτερη μέθοδο λαμβάνονται υπόψη μόνο οι καταγεγραμμένες ταχύτητες. Τίθεται ένα στατιστικό όριο για την ταχύτητα, βασισμένο στη μέθοδο boxplot, κάτω από το οποίο θεωρείται πως οι ταχύτητες είναι ακραία χαμηλές και αποτελούν μετρήσεις κατά τη διάρκεια κορεσμού. Κάθε μέθοδος έχει ως τιμή εξόδου τα συμβάντα κορεσμού. Για τη συνέχεια της έρευνας, επιλέγεται η δεύτερη μέθοδος ως η καταλληλότερη για τις δύο βάσεις δεδομένων. Η διάρκεια των συμβάντων υπολογίζεται με την κατάλληλη επεξεργασία σε γλώσσα προγραμματισμού.

Για την ανάπτυξη μοντέλων πρόβλεψης της διάρκειας του κορεσμού, χρησιμοποιείται η μέθοδος της Ανάλυσης Επιβίωσης. Συγκεκριμένα, χρησιμοποιούνται παραμετρικά μοντέλα τα οποία βασιζόμενα σε ανεξάρτητες μεταβλητές κατασκευάζουν την καμπύλη επιβίωσης και εκφράζουν την πιθανότητα η διάρκεια ενός συμβάντος να έχει συγκεκριμένη τιμή. Ειδικότερα, εφαρμόζονται το μοντέλο κινδύνου Cox, η μέθοδος κατανομής Weibull και η μέθοδος Random Survival Forest. Οι ανεξάρτητες μεταβλητές που εισάχθηκαν στα μοντέλα ήταν η ώρα αιχμής, ο αριθμός των λωρίδων και ο κυκλοφοριακός φόρτος κατά την έναρξη του κορεσμού.

Τελικά, τη μεγαλύτερη επιρροή στη διάρκεια είχε η ώρα αιχμής, που στην Αθήνα εντός αιχμής οι διάρκειες είναι μεγαλύτερες, ενώ στο Μάντσεστερ εντός αιχμής είναι μικρότερες. Συνολικά, ενώ και στις δύο πόλεις η κατανομή των διαρκειών είναι παρόμοια, στο Μάντσεστερ οι διάρκειες είναι ελάχιστα μεγαλύτερες. Από την άλλη, τη μικρότερη επιρροή στα μοντέλα φάνηκε να έχει ο κυκλοφοριακός φόρτος κατά την έναρξη του συμβάντος. Ο αριθμός των λωρίδων, παρά το γεγονός ότι έδειχνε να επηρεάζει τα μοντέλα, δεν αποτελούσε στατιστικά σημαντική μεταβλητή, με βάση το στατιστικό έλεγχο που εφαρμόστηκε.

Τα αποτελέσματα της παρούσας έρευνας θα μπορούσαν να χρησιμοποιηθούν στον τομέα της διαχείρισης κυκλοφορίας. Με περισσότερες ανεξάρτητες μεταβλητές, όπως οι καιρικές συνθήκες ή μεγαλύτερο όγκο δεδομένων, είναι πιθανό με την εφαρμογή των μοντέλων να προκύψουν πιο ξεκάθαρα αποτελέσματα για την επιρροή κάθε μεταβλητής στη διάρκεια του κορεσμού.

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ 11

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ 11

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ 12

1. ΕΙΣΑΓΩΓΗ 13

1.1 Γενικά 13

1.1.1. Η έννοια του κορεσμού και οι συνέπειες στην κοινωνία 13

1.2. Η έννοια της πρόβλεψης στις μεταφορές 15

1.3. Σκοπός της εργασίας 15

1.4. Διάρθρωση της εργασίας 16

2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ 17

2.1. Εισαγωγή 17

2.2. Εντοπισμός φαινομένων κορεσμού 17

2.3. Έρευνες πρόβλεψης διάρκειας φαινομένων 19

2.4. Σύνοψη βιβλιογραφικής ανασκόπησης 21

3. ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ 23

3.1. Ανάπτυξη μεθοδολογίας 23

3.2. Βασικές έννοιες της Ανάλυσης Επιβίωσης 24

3.2.1. Είδη μοντέλων 26

3.2.2. Αξιολόγηση μοντέλων 27

4. ΕΦΑΡΜΟΓΗ ΜΕΘΟΔΟΛΟΓΙΑΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ 29

4.1. Περιγραφή δεδομένων 29

4.1.1. Βάση Δεδομένων Αθήνας 29

4.1.2. Βάση δεδομένων Μάντσεστερ 30

4.2. Επεξεργασία δεδομένων 31

4.2.1. Εντοπισμός κορεσμού 31

4.2.2. Υπολογισμός διάρκειας συμβάντων κορεσμού 37

4.2.3. Επιλογή ανεξάρτητων επεξηγηματικών μεταβλητών 38

4.3. Μοντέλα Survival Analysis 40

4.3.1. Μοντέλο Κινδύνου Cox 41

4.3.2. Μέθοδος Κατανομής Weibull 44

4.3.3. Μέθοδος Random Survival Forest 47

5. ΣΥΜΠΕΡΑΣΜΑΤΑ 49

5.1. Εισαγωγή 49
5.2. Βασικά συμπεράσματα 49
5.3. Προτάσεις για περαιτέρω έρευνα 50
ΒΙΒΛΙΟΓΡΑΦΙΑ 53
ΠΑΡΑΡΤΗΜΑ 57

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 1.1. Τυπική μορφή "Bottleneck" 14
Εικόνα 1.2. Αιτίες κορεσμού της κυκλοφορίας 14
Εικόνα 1.3. Δομή της διπλωματικής εργασίας 16
Εικόνα 3.1. Οπτικοποίηση των "δέντρων" απόφασης της Random Survival Forest (Datema et al., 2012) 27
Εικόνα 4.1. Σημεία μέτρησης κυκλοφοριακών μεγεθών στην Αθήνα 29
Εικόνα 4.2. Σημεία μέτρησης κυκλοφοριακών μεγεθών στο Μάντσεστερ 30
Εικόνα 4.3. Μορφή πίνακα state (0 ακόρεστη ροή, 1 κορεσμένη ροή) 37

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 2.1. . Διάγραμμα Κυκλοφοριακού Φόρτου-Κατάληψης - Μετάβαση ροής από ελεύθερη σε κορεσμένη (Ι. Μ. Φραντζεσκάκης, Ι.Κ. Γκόλιας, Μ.Χ. Πίτσιαβα-Λατινοπούλου, 2009) 17
Διάγραμμα 2.2. Διάγραμμα φόρτου-κατάληψης με επισήμανση του "spillover" 18
Διάγραμμα 3.1. Διάγραμμα ροής μεθοδολογίας 23
Διάγραμμα 3.2. Απεικόνιση βασικών συναρτήσεων Survival Analysis (Washington, Karlaftis, Mannering, Anastasopoulos, 2004) 24
Διάγραμμα 3.3. Εισαγωγή επεξηγηματικής μεταβλητής στη βασική συνάρτηση κινδύνου με αναλογικό μοντέλο(Washington, Karlaftis, Mannering, Anastasopoulos, 2004) 25
Διάγραμμα 4.1. Θέση μέτρησης MS249 και Διάγραμμα φόρτου-κατάληψης θέσης MS249 (Βασιλίσσης Σοφίας, ρεύμα προς Βασιλέως Κωνσταντίνου, πριν την Παπαδιαμαντοπούλου) 32
Διάγραμμα 4.2. Θέση μέτρησης MS503 και Διάγραμμα φόρτου-ταχύτητας στη θέση MS503 (Λεωφόρος Κηφισίας, ρεύμα προς Αθήνα, μετά από Πανόρμου) 32
Διάγραμμα 4.3. Θέση μέτρησης MS523 και Διάγραμμα φόρτου-κατάληψης θέσης MS523 (Λεωφόρος Βεΐκου, ρεύμα προς Αθήνα, ύψος Λεωφόρου Γαλατσίου) 32
Διάγραμμα 4.4. Θέση μέτρησης MS650 και Διάγραμμα φόρτου-κατάληψης θέσης MS650 (Ράμπα εισόδου στη Λεωφόρο Αθηνών από Θηβών με κατεύθυνση προς Αθήνα) 33
Διάγραμμα 4.5. Θέση μέτρησης MS704 και Διάγραμμα φόρτου-κατάληψης στη θέση MS704 (Λεωφόρος Πειραιώς, ρεύμα προς Αθήνα, ύψος Ιεράς Οδού) 33
Διάγραμμα 4.6. Διάγραμμα φόρτου-ταχύτητας στη θέση MS249 (Βασιλίσσης Σοφίας, ρεύμα προς Βασιλέως Κωνσταντίνου, πριν την Παπαδιαμαντοπούλου) 34
Διάγραμμα 4.7. Διάγραμμα φόρτου-ταχύτητας στη θέση MS503 (Λεωφόρος Κηφισίας, ρεύμα προς Αθήνα, μετά από Πανόρμου) 34
Διάγραμμα 4.8. Διάγραμμα φόρτου-ταχύτητας στη θέση MS523 (Λεωφόρος Βεΐκου, ρεύμα προς Αθήνα, ύψος Λεωφόρου Γαλατσίου) 34
Διάγραμμα 4.9. Διάγραμμα φόρτου-ταχύτητας στη θέση MS650 (Ράμπα εισόδου στη Λεωφόρο Αθηνών από Θηβών με κατεύθυνση προς Αθήνα) 35
Διάγραμμα 4.10. Διάγραμμα φόρτου-ταχύτητας στη θέση MS704 (Λεωφόρος Πειραιώς, ρεύμα προς Αθήνα, ύψος Ιεράς Οδού) 35
Διάγραμμα 4.11. Θέση μέτρησης 1004 και Διάγραμμα φόρτου-ταχύτητας στη θέση 1004 (Park Road, κόμβος με Manchester Road) 36
Διάγραμμα 4.12. Θέση μέτρησης 1048 και Διάγραμμα φόρτου-ταχύτητας στη θέση 1048 (Chester Road προς Bridgewater Way) 36
Διάγραμμα 4.13. Θέση μέτρησης 1081 και Διάγραμμα φόρτου-ταχύτητας στη θέση 1081 (Ardwick Green South, ρεύμα προς κόμβο Downing Street, μετά τον κυκλικό κόμβο με Hyde Road και Stockport Road) 36

Διάγραμμα 4.14. Θέση μέτρησης 1083 και Διάγραμμα φόρτου-ταχύτητας στη θέση 1083 (Stockport Road, κόμβος με Crowcroft Road) 37
Διάγραμμα 4.15. Συμβάντα κορεσμού ανά διάστημα αιχμής στην Αθήνα (0 εκτός αιχμής, 1 πρωινή αιχμή 7:00-11:00, 2 απογευματινή αιχμή 14:00-19:00) 38
Διάγραμμα 4.16. Συμβάντα κορεσμό ανά διάστημα αιχμής στο Μάντσεστερ (0 εκτός αιχμής, 1 εντός αιχμής 10:00-18:00) 39
Διάγραμμα 4.17. Συμβάντα κορεσμού ανά αριθμό λωρίδων διατομής στην Αθήνα 39
Διάγραμμα 4.18. Καμπύλη Survival πραγματικών συμβάντων στην Αθήνα 41
Διάγραμμα 4.19. Καμπύλη Survival πραγματικών συμβάντων στο Μάντσεστερ 41
Διάγραμμα 4.20. Επιρροή ανεξάρτητων μεταβλητών στο μοντέλο κινδύνου Cox στην Αθήνα 42
Διάγραμμα 4.21. Καμπύλη Survival μοντέλου κινδύνου Cox για ώρες εντός κι εκτός πρωινής αιχμής στην Αθήνα 42
Διάγραμμα 4.22. Καμπύλη Survival μοντέλου κινδύνου Cox για ώρες εντός κι εκτός απογευματινής αιχμής στην Αθήνα 43
Διάγραμμα 4.23. Επιρροή ανεξάρτητων μεταβλητών στο μοντέλο κινδύνου Cox στο Μάντσεστερ .. 43
Διάγραμμα 4.24. Καμπύλη Survival μοντέλου κινδύνου Cox για ώρες εντός κι εκτός αιχμής στο Μάντσεστερ 44
Διάγραμμα 4.25. Επιρροή ανεξάρτητων μεταβλητών στη μέθοδο κατανομής Weibull στην Αθήνα .. 45
Διάγραμμα 4.26. Καμπύλη survival κατανομής Weibull για ώρες εντός κι εκτός πρωινής αιχμής στην Αθήνα 45
Διάγραμμα 4.27. Καμπύλη survival κατανομής Weibull για ώρες εντός κι εκτός απογευματινής αιχμής στην Αθήνα 45
Διάγραμμα 4.28. Επιρροή ανεξάρτητων μεταβλητών στη μέθοδο κατανομής Weibull στο Μάντσεστερ 46
Διάγραμμα 4.29. Καμπύλη survival μεθόδου κατανομής Weibull για ώρες εντός κι εκτός αιχμής στο Μάντσεστερ 46
Διάγραμμα 4.30. Καμπύλες Random Survival Forest για μερικά τυχαία σημεία στην Αθήνα 47
Διάγραμμα 4.31. Καμπύλες Random Survival Forest για μερικά τυχαία σημεία στο Μάντσεστερ 47

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 4.1. Μορφή βάσης δεδομένων μετρήσεων Αθήνας 30
Πίνακας 4.2. Μορφή βάσης δεδομένων μετρήσεων Μάντσεστερ 31
Πίνακας 4.3. Πίνακας διαρκειών συμβάντων κορεσμού Αθήνας με ανεξάρτητες μεταβλητές 40
Πίνακας 4.4. Πίνακας διαρκειών συμβάντων κορεσμού Μάντσεστερ με ανεξάρτητες μεταβλητές ... 40
Πίνακας 4.5. Αποτελέσματα εφαρμογής μοντέλου κινδύνου Cox στην Αθήνα 42
Πίνακας 4.6. Αποτελέσματα εφαρμογής μοντέλου κινδύνου Cox στο Μάντσεστερ 43
Πίνακας 4.7. Αποτελέσματα εφαρμογής μεθόδου κατανομής Weibull στην Αθήνα 44
Πίνακας 4.8. Αποτελέσματα εφαρμογής μεθόδου κατανομής Weibull στο Μάντσεστερ 46
Πίνακας 4.9. Δείκτης Concordance ανά μέθοδο και πόλη 48

1. ΕΙΣΑΓΩΓΗ

1.1 Γενικά

Η πόλη της Αθήνας είναι η πολυπληθέστερη πόλη της Ελλάδας. Σύμφωνα με την απογραφή πληθυσμού του 2021, στην περιφέρεια Αττικής κατοικούν περίπου 3,8 εκατομμύρια άνθρωποι, από τους οποίους το ένα εκατομμύριο κατοικεί στην πόλη της Αθήνας. Οι μετακινήσεις τους είναι άρρηκτα συνδεδεμένες με το βιοτικό επίπεδο της κοινωνίας τους. Συνεπώς, η συνεχής μελέτη και έρευνα στον τομέα των μεταφορών και της εξυπηρέτησης των κατοίκων κρίνεται αναγκαία.

Η εξυπηρέτηση σε ένα οδικό δίκτυο περιγράφεται από δείκτες, όπως η ταχύτητα κυκλοφορίας, ο χρόνος ταξιδιού, η άνεση κι η ελευθερία κίνησης, οι οποίοι προσδιορίζουν την Στάθμη Εξυπηρέτησης (Pandey and Biswas, 2022). Η Στάθμη Εξυπηρέτησης μίας οδού ή ενός κόμβου καθορίζεται από χαρακτηριστικά γεωμετρικά και κυκλοφοριακά και μπορεί να συνοψίσει τα προβλήματα που ενδέχεται να υπάρχουν στην κυκλοφορία. Μπορεί ωστόσο, να προκύψει και από την απώλεια χρόνου των χρηστών, δηλαδή τις καθυστερήσεις.

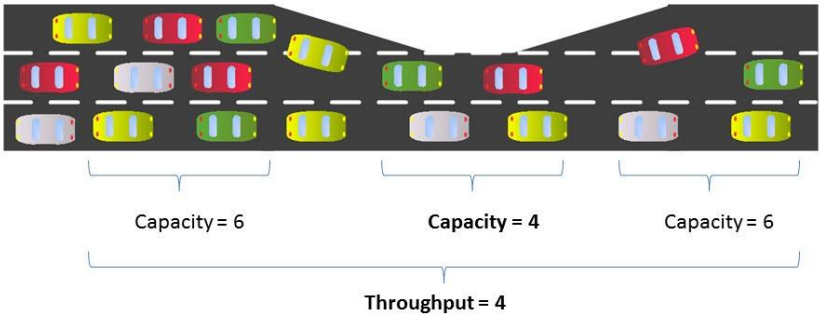
Οι καθυστερήσεις στις μετακινήσεις αποτελούν αυξημένης σημασίας παράγοντα στις αποφάσεις που λαμβάνουν οι χρήστες, πριν ή κατά τη διάρκεια του ταξιδιού τους. Η κύρια αιτία δημιουργίας καθυστερήσεων στις μετακινήσεις των πολιτών είναι η εμφάνιση κορεσμού. Ένας χρήστης επιλέγει την ώρα αναχώρησης και τη διαδρομή που θα ακολουθήσει με βάση την επιθυμητή ώρα άφιξης στον προορισμό του και με σκοπό την ελαχιστοποίηση των καθυστερήσεων και την αποφυγή της εισόδου σε κορεσμένη ροή. Τέτοιες αποφάσεις ωστόσο, λαμβάνονται με βάση την εμπειρία του χρήστη ή και την ενημέρωση σε πραγματικό χρόνο για την κατάσταση του οδικού δικτύου.

1.1.1. Η έννοια του κορεσμού και οι συνέπειες στην κοινωνία

Ως κορεσμός ορίζεται η κατάσταση μιας συγκοινωνιακής υποδομής κατά την οποία ο εισερχόμενος φόρτος οχημάτων υπερβαίνει την κυκλοφοριακή της ικανότητα (Stopher, 2004). Διακρίνεται σε δύο επιμέρους κατηγορίες: τον επαναλαμβανόμενο και το μη-επαναλαμβανόμενο κορεσμό.

<u>Επαναλαμβανόμενος κορεσμός</u>

Ο επαναλαμβανόμενος κορεσμός είναι η συμφόρηση που δημιουργείται σχεδόν καθημερινά σε συγκεκριμένα σημεία του οδικού δικτύου και σε συγκεκριμένες χρονικές περιόδους (Kumar and Raubal, 2021). Για παράδειγμα, η συμφόρηση σε στενώματα οδών (bottlenecks) (Εικόνα 1.1) ή σε βασικές αρτηρίες κατά τις ώρες αιχμής. Στην πρώτη περίπτωση προκαλείται η δημιουργία ουρών καθώς οχήματα προσπαθούν ταυτόχρονα να εισέλθουν σε τμήμα οδού στο οποίο οι λωρίδες κυκλοφορίας μειώνονται.

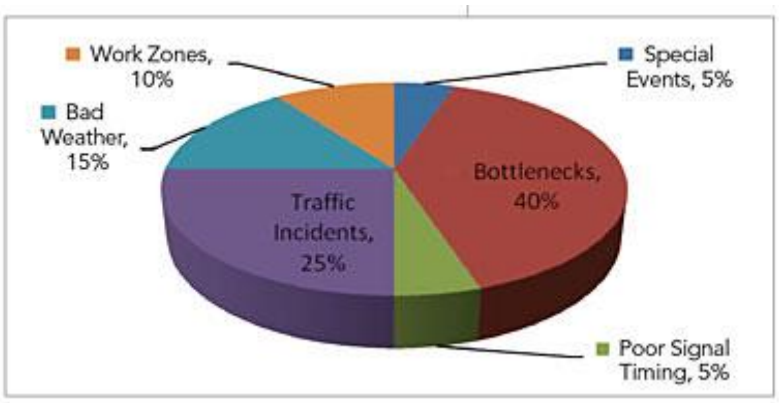


Εικόνα 1.1. Τυπική μορφή "Bottleneck"

Όσον αφορά το δεύτερο παράδειγμα, κατά τις ώρες αιχμής, οι ανάγκες των πολιτών για μετακινήσεις είναι αυξημένες, ως εκ τούτου και οι κυκλοφοριακοί φόρτοι. Ανάλογη με την κυκλοφοριακή ικανότητα κάθε οδού, είναι και η συμφόρηση που προκαλείται. Τα δύο βασικά χρονικά διαστήματα αιχμής είναι οι πρωινές ώρες κατά τις οποίες ο κόσμος πηγαίνει στην εργασία του και οι απογευματινές όταν επιστρέφει από αυτή.

Μη-επαναλαμβανόμενος κορεσμός

Σχετικά με τον μη-επαναλαμβανόμενο κορεσμό, πρόκειται για μη προγραμματισμένα συμβάντα που μπορεί να δημιουργήσουν προβλήματα στην ομαλή κυκλοφοριακή ροή (McGroarty, 2010). Μερικά παραδείγματα, όπως φαίνεται και στο γράφημα στην Εικόνα 1.2**,** είναι τα οδικά ατυχήματα, οι εργασίες επί της οδού ή τα ακραία καιρικά φαινόμενα.



Εικόνα 1.2. Αιτίες κορεσμού της κυκλοφορίας

Το φαινόμενο του κορεσμού, έχει σημαντικές συνέπειες για την κοινωνία. Οι τρεις βασικές είναι η ασταθείς συνθήκες ταξιδιού, η αύξηση εκπομπών ρύπων και ο «χαμένος χρόνος» (Stopher, 2004). Κατά τον κορεσμό, η κυκλοφορία επιβραδύνεται σημαντικά και ανά πάσα στιγμή μπορεί να ακινητοποιηθεί. Συνεπώς, οι χρόνοι ταξιδιού δεν είναι πάντα αξιόπιστοι. Ακόμα, με τη συνεχή αυξομείωση ταχυτήτων, τα οχήματα εκπέμπουν πολύ μεγαλύτερη ποσότητα ρύπων από ότι σε συνθήκες ελεύθερης ροής (Fafoutellis et al., 2020), ενώ τέλος, η εμφάνιση κορεσμού αυξάνει το «χαμένο χρόνο», δηλαδή το χρόνο κατά τον οποίο εμποδίζονται παραγωγικές διαδικασίες. Όλα τα παραπάνω συντελούν στην επιβάρυνση της οικονομίας, μιας και υπάρχει μείωση της παραγωγικότητας σε πολλούς τομείς και αύξηση των επιπτώσεων στην υγεία των πολιτών και κατ’ επέκταση αυξανόμενες δαπάνες κονδυλίων στην ιατροφαρμακευτική περίθαλψη.

1.2. Η έννοια της πρόβλεψης στις μεταφορές

Είναι γεγονός ότι οι κυκλοφοριακές συνθήκες αποτελούν σημαντική παράμετρο της ποιότητας ζωής των κατοίκων μίας πόλης (Chen et al., 2018). Κάθε πολίτης, μπορεί πολύ απλά, με τη βοήθεια ενός έξυπνου τηλεφώνου, να αναζητήσει ανά πάσα στιγμή τη βέλτιστη διαδρομή προς τον προορισμό του ή να δει απευθείας τη συμφόρηση στους δρόμους μέσω της εφαρμογής Google Maps. Μια τέτοια εφαρμογή βασίζεται σε δεδομένα τεχνολογίας GPS, αλλά και ιστορικά στοιχεία. Ουσιαστικά, πρόκειται για μία καταγραφή της υπάρχουσας κατάστασης με στοιχεία συλλεγμένα σε πραγματικό χρόνο τα οποία χρησιμοποιούνται κατάλληλα ώστε να γίνει μία πρόβλεψη του χρόνου διαδρομής.

Η μέθοδος της πρόβλεψης είναι άρρηκτα συνυφασμένη με την επιστήμη του συγκοινωνιολόγου. Η πιο συχνή πρόβλεψη, είναι αυτή της κυκλοφορίας. Υπάρχει δυνατότητα, να παραχθεί πρόβλεψη των κυκλοφοριακών μεγεθών, όπως του κυκλοφοριακού φόρτου, της ταχύτητας ή της κατάληψης, με βάση τις χρονοσειρές αυτών, στο κοντινό ή και μακρινό παρελθόν. Με ανάπτυξη μεθόδων μηχανικής μάθησης, ο ερευνητής μπορεί να διερευνήσει χαρακτηριστικά όπως η εποχικότητα, η τάση και το επίπεδο της χρονοσειράς.

Μία εξέλιξη της πρόβλεψης κυκλοφορίας είναι κι η πρόβλεψη του κορεσμού. Ειδικότερα, η πρόβλεψη κορεσμού είναι από τα πιο ενδιαφέροντα και πολύπλοκα συγκοινωνιακά ζητήματα. Όπως αναφέρθηκε παραπάνω, υπάρχουν δύο είδη κορεσμού, ο επαναλαμβανόμενος κι ο μη-επαναλαμβανόμενος. Η πρόβλεψη του επαναλαμβανόμενου κορεσμού είναι πολύ πιο εύκολη, μιας και μπορεί να βασιστεί σε ιστορικά στοιχεία, αφού συμβαίνει σε συγκεκριμένα σημεία του δικτύου (bottlenecks). Επομένως, προβλέπεται μόνο η στιγμή που θα συμβεί ο κορεσμός, δηλαδή το πότε η ζήτηση θα κορυφωθεί. Όσον αφορά τον μη-επαναλαμβανόμενο κορεσμό, η πρόβλεψη συμβάντων κορεσμού είναι δύσκολη, διότι δεν εμφανίζονται συχνά και σταθερά. Σε αυτή την περίπτωση, συχνότερη αιτία εμφάνισης κορεσμού είναι η πρόκληση τροχαίου ατυχήματος (Kumar and Raubal, 2021).

1.3. Σκοπός της εργασίας

Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη μεθοδολογίας εντοπισμού και εκτίμησης της διάρκειας συμβάντων κορεσμού σε αστικά οδικά δίκτυα με βάση ιστορικά στοιχεία κυκλοφοριακών μεγεθών.

Για τον εντοπισμό των συμβάντων κορεσμού τα οποία θα εξετασθούν για τη διάρκειά τους, εφαρμόζονται δύο μέθοδοι. Στην πρώτη χρησιμοποιείται μία εξίσωση της βιβλιογραφίας, η οποία λαμβάνει υπόψη τη σχέση κατάληψης οδοστρώματος-κυκλοφοριακού φόρτου. Στη δεύτερη μέθοδο, εντοπίζονται οι μετρήσεις ταχύτητας, οι οποίες έχουν χαμηλότερη τιμή σε σχέση με ένα όριο που τίθεται και το οποίο προκύπτει από όλες τις καταγεγραμμένες ταχύτητες στην συγκεκριμένη θέση.

Για τη μοντελοποίηση της διάρκειας κορεσμού εφαρμόζονται παραμετρικά μοντέλα επιβίωσης και αναλύεται η επιρροή παραγόντων όπως ο αριθμός των λωρίδων κυκλοφορίας ανά θέση, η ώρα αιχμής και ο κυκλοφοριακός φόρτος κατά την έναρξη του κορεσμού.

Τελικά, βασική επιδίωξη είναι η ανάλυση της επιρροής των παραπάνω μεταβλητών στην πρόβλεψη της διάρκειας κορεσμού με στόχο την πιο αποδοτική διαχείριση κυκλοφορίας σε αστικά οδικά δίκτυα.

1.4. Διάρθρωση της εργασίας

Στο 1ο Κεφάλαιο, γίνεται μία μικρή εισαγωγή στην έννοια του κορεσμού, καθορίζεται ο στόχος της διπλωματικής εργασίας και δίνεται η μορφή της δομής της.

Στο 2ο Κεφάλαιο, γίνεται αναδρομή στη βιβλιογραφική ανασκόπηση. Πρόκειται για αναφορά σε δημοσιευμένες έρευνες και μεθοδολογίες, στην ελληνική ή ξένη βιβλιογραφία, που συνδέονται με το αντικείμενο της εργασίας, ώστε να εντοπιστούν τρόποι ανάλυσης του υπό μελέτη ζητήματος.

Στο 3ο Κεφάλαιο, αναλύεται η θεωρία πίσω από τη μέθοδο Survival Analysis και τα μοντέλα πρόβλεψης που χρησιμοποιήθηκαν.

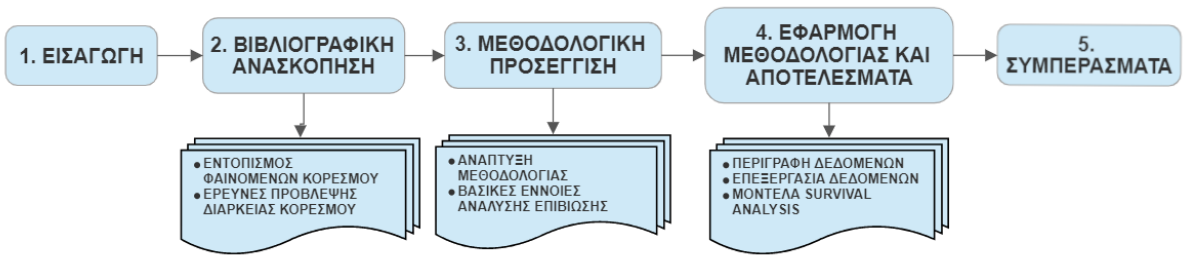
Στο 4ο Κεφάλαιο, περιγράφονται οι βάσεις δεδομένων που χρησιμοποιήθηκαν κατά την ανάλυση, καθώς επίσης και η διαδικασία που ακολουθήθηκε κατά τον εντοπισμό του κορεσμού και τον υπολογισμό της διάρκειάς του.

Στο 5^{ο} Κεφάλαιο, παρουσιάζονται τα αποτελέσματα ανά μοντέλο και μία πρώτη περιγραφή τους, ενώ γίνεται αναφορά και στην επιμέρους επιρροή των ανεξάρτητων μεταβλητών.

Στο 6^{ο} Κεφάλαιο, δίνονται τα συμπεράσματα που εξήχθησαν από την ανάλυση και μερικές προτάσεις για περαιτέρω έρευνα.

Έπειτα ακολουθεί λίστα με τη βιβλιογραφία που χρησιμοποιήθηκε και τέλος Παράρτημα με τον κώδικα που εφαρμόστηκε στη γλώσσα προγραμματισμού Python.

Η δομή της εργασίας φαίνεται στην Εικόνα 1.3.



Εικόνα 1.3. Δομή της διπλωματικής εργασίας

2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

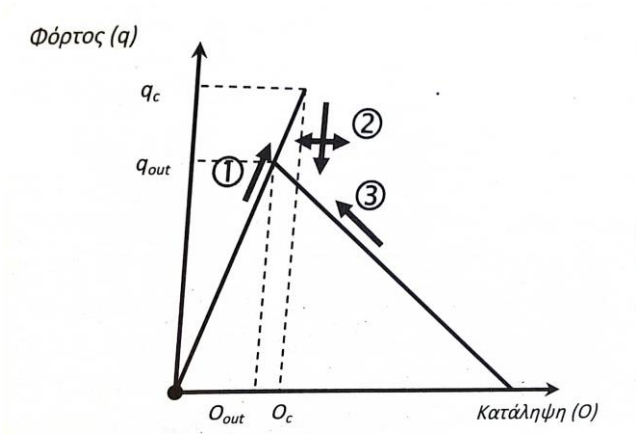
2.1. Εισαγωγή

Στο παρόν κεφάλαιο θα γίνει παρουσίαση της έρευνας που πραγματοποιήθηκε στη βιβλιογραφία. Πιο συγκεκριμένα, θα αναλυθούν μέθοδοι εντοπισμού κορεσμού της κυκλοφορίας καθώς επίσης και μοντέλα πρόβλεψης διάρκειας. Έπειτα, θα αναφερθούν τα βασικά συμπεράσματα, τα οποία θα βοηθήσουν στη βαθύτερη κατανόηση του αντικειμένου της παρούσας Διπλωματικής Εργασίας.

2.2. Εντοπισμός φαινομένων κορεσμού

Η κατάταξη μιας ροής οχημάτων ως κορεσμένη, μπορεί να φαίνεται εύκολο έργο, αλλά δεν είναι σε όλες τις περιπτώσεις. Κάθε οδός, έχει ιδιότητες που χαρακτηρίζουν τον τύπο χρήσης της και τη γεωμετρία της. Ως εκ τούτου, το πότε θεωρείται ότι λαμβάνει χώρα ένα συμβάν κορεσμού είναι ανάλογο της οδού στην οποία συμβαίνει (Stopher, 2004). Η αξιολόγηση των κυκλοφοριακών συνθηκών γίνεται μέσω της μέτρησης των βασικών κυκλοφοριακών μεγεθών (φόρτος, ταχύτητα και πυκνότητα ή κατάληψη) και συγκεκριμένα, μέσω σύγκρισης των τιμών τους με κάποιες οριακές τιμές. Για παράδειγμα, σε μία αστική αρτηρία, είναι διαφορετικό το κατώφλι για την κατάληψη, πάνω από το οποίο η ροή θεωρείται κορεσμένη, από μια τοπική οδό, όπου οι φόρτοι είναι χαμηλότεροι, η λωρίδα κυκλοφορίας είναι μία κι ενδέχεται να υπάρχει και παρόδια στάθμευση.

Οι θεμελιώδεις σχέσεις της Κυκλοφοριακής Τεχνικής μπορούν να αποτελέσουν βασικό εργαλείο εντοπισμού ενός συμβάντος κορεσμού στο οδικό δίκτυο. Η μετάβαση της κυκλοφοριακής ροής από ελεύθερη σε κορεσμένη γίνεται σε τρία στάδια (Διάγραμμα 2.1). Αρχικά, η κατάληψη και ο κυκλοφοριακός φόρτος αυξάνονται μέχρι ο φόρτος να πετύχει τη μέγιστη τιμή του (1). Έπειτα, σε περίπτωση περεταίρω αύξησης της κατάληψης, αρχίζει η συμφόρηση καθώς ο φόρτος μειώνεται απότομα (2), και τελικά σε τρίτο στάδιο συνεχίζει να μειώνεται μέχρι η κατάληψη να μεγιστοποιηθεί (3) (Ι. Μ. Φραντζεσκάκης, Ι.Κ. Γκόλιας, Μ.Χ. Πίτσιαβα-Λατινοπούλου, 2009).



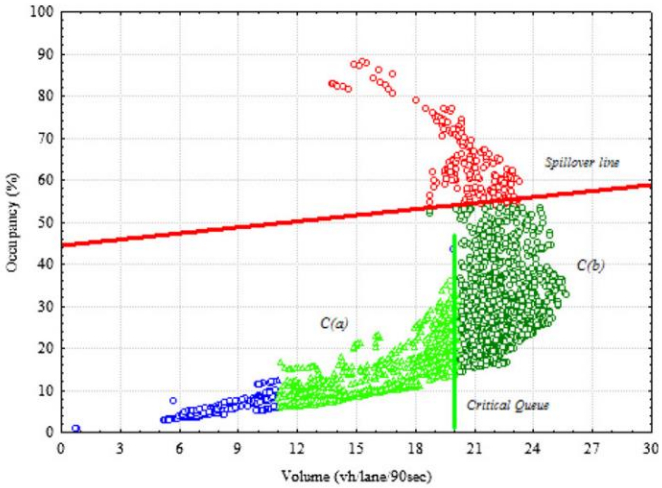
Διάγραμμα 2.1. . Διάγραμμα Κυκλοφοριακού Φόρτου-Κατάληψης - Μετάβαση ροής από ελεύθερη σε κορεσμένη (Ι. Μ. Φραντζεσκάκης, Ι.Κ. Γκόλιας, Μ.Χ. Πίτσιαβα-Λατινοπούλου, 2009)

Υπάρχει και η προσπάθεια αναλυτικού υπολογισμού του ορίου κατά το οποίο προκαλείται «διάχυση» (spillover) (Vlahogianni et al., 2007). Πρόκειται για τη στιγμή που η κατάληψη για

συγκεκριμένο φόρτο, ξεπερνά το όριο «διάχυσης», όπου και θεωρείται πως υπάρχει κυκλοφοριακή συμφόρηση. Ένα σχετικό παράδειγμα είναι η παρακάτω προσέγγιση, που λαμβάνει υπόψιν τη σχέση κατάληψης-κυκλοφοριακού φόρτου σε σηματοδοτημένες οδούς (Geroliminis and Skabardonis, 2006):

$$O_{sp} = \frac{L_{eff} * V}{u_f} + \frac{r}{C} \qquad (2.1)$$

Όπου O_{sp} η κρίσιμη τιμή για την κατάληψη, L_{eff} το τυπικό μήκος οχήματος, V ο φόρτος, u_f η ταχύτητα ελεύθερης ροής, r ο χρόνος κόκκινης ένδειξης και C ο χρόνος κύκλου του φωτεινού σηματοδότη. Η συγκεκριμένη εξίσωση παριστάνει μία ευθεία στο διάγραμμα κατάληψης-φόρτου:



Διάγραμμα 2.2. Διάγραμμα φόρτου-κατάληψης με επισήμανση του "spillover"

Όπως φαίνεται στο Διάγραμμα 2.2 (Vlahogianni et al., 2007), σε μετρήσεις στις οποίες η καταγεγραμμένη κατάληψη είναι πάνω από την ευθεία του spillover, θεωρείται πως πρόκειται για στιγμές όπου υπάρχει κορεσμός.

Παράλληλα, η ραγδαία πρόοδος της τεχνολογίας, έχει δημιουργήσει νέους τρόπους ελέγχου της κυκλοφορίας. Η ανάπτυξη νέων μεθόδων παρακολούθησης, η συνεχής καταγραφή και παροχή δεδομένων, έχει οδηγήσει τους αρμόδιους φορείς στην αναζήτηση αυτοματοποιημένων μεθόδων εντοπισμού του κορεσμού (Kumar and Raubal, 2021). Μία από αυτές τις μεθόδους είναι η χρήση καμερών (Krizhevsky et al., 2017; Redmon et al., 2016), κατά την οποία λαμβάνονται συνεχώς εικόνες και βίντεο από την κυκλοφορία των οχημάτων. Αναπτύσσοντας μοντέλα βαθιάς μάθησης, μπορούν να εξάγουν συμπεράσματα για την κατάσταση της κυκλοφορίας, την ύπαρξη ή όχι κορεσμού και τα μήκη των ουρών που δημιουργούνται (Chakraborty et al., 2018).

Σε μία άλλη έρευνα αναπτύχθηκε μία παρόμοια μέθοδος, κατά την οποία αρχικά εντοπίζονται από τις κάμερες τα οπίσθια φώτα των οχημάτων που κυκλοφορούν και αντιστοιχίζονται μεταξύ τους ώστε να πρόκειται για φώτα που ανήκουν στο ίδιο όχημα. Με αυτόν τον τρόπο, θα προκύψει ο αριθμός των οχημάτων που κινούνται της στιγμή της λήψης στο οδόστρωμα,

και με βάση το όριο που θέτει ο μελετητής, αποφαίνεται αν πρόκειται για κορεσμένη ή ελεύθερη ροή (Nidhal et al., 2014).

Μία ακόμη μέθοδος είναι η συλλογή δεδομένων GPS και η καταγραφή διαδρομών οχημάτων (D'Andrea and Marcelloni, 2017; Yong-chuan et al., 2011). Η τροχιά που θα καταγράφει το GPS ενός οχήματος δεν είναι πάντοτε η ίδια με αυτή που στην πραγματικότητα ακολουθεί. Γι'αυτό, ακολουθείται η διαδικασία Map-Matching, κατά την οποία η τροχιά του οχήματος διορθώνεται, καθώς αντιστοιχίζεται σε περιβάλλον Συστήματος Γεωγραφικών Πληροφοριών στην οδό από την οποία διέρχεται. Έπειτα, με βάση την κατανομή των ταχυτήτων κυκλοφορίας στην περιοχή μελέτης, γίνεται σύγκριση με την καταγεγραμμένη ταχύτητα του οχήματος. Σε περιπτώσεις καταγεγραμμένης ταχύτητας πολύ χαμηλότερης της μέσης, θεωρείται πως υπάρχει κορεσμός, ενώ το αντίθετο ισχύει για πολύ υψηλότερες.

Η ύπαρξη "bottleneck" είναι η κύρια αιτία εμφάνισής του επαναλαμβανόμενου κορεσμού. Όπως είναι λογικό, η αναζήτηση μεθόδων που εντοπίζουν τα "bottlenecks", τα οποία προκαλούν συμφόρηση, είναι πολύ χρήσιμη για τη διαχείριση της κυκλοφορίας. Η ανάπτυξη μοντέλων εντοπισμού των τμημάτων αυτών και η διερεύνηση της επιρροής που μπορεί να έχουν σε άλλα γειτονικά τμήματα, μπορεί να αποτελέσει χρήσιμο εργαλείο για τη μελέτη του φαινομένου του κορεσμού. Σε τέτοια μοντέλα, συνήθως γίνεται σύγκριση των τιμών κυκλοφοριακών μεγεθών ανά τμήμα, με ένα ποσοστό της κατανομής τους (Lee et al., 2011; Li et al., 2020).

Μία παρόμοια μέθοδος ακολουθήθηκε και σε μία έρευνα του μη-επαναλαμβανόμενου κορεσμού, στην οποία σκοπός ήταν η δημιουργία ενός αλγορίθμου που θα εντοπίζει ένα συμβάν, με βάση την αλλαγή που μπορεί να έχουν παρουσιάσει οι χρόνοι ταξιδιού στα γειτονικά τμήματα της οδού (Anbaroglu et al., 2014).

2.3. Έρευνες πρόβλεψης διάρκειας φαινομένων

Μπορεί η εμφάνιση κορεσμού, να αποτελεί από μόνη της σημαντική πληροφορία για την κατάσταση ενός οδικού δικτύου, αλλά η διάρκειά του είναι αυτή που διαδραματίζει πρωτεύοντα ρόλο, ιδιαίτερα στο βαθμό ικανοποίησης των χρηστών. Σε περίπτωση που οι οδηγοί εμπλακούν για μεγάλα χρονικά διαστήματα σε συμβάντα κορεσμού, δημιουργείται δυσαρέσκεια και είναι πιθανή η αλλαγή διαδρομών σε μελλοντικά τους ταξίδια ώστε να αποφύγουν παρόμοιες καταστάσεις. Χάνεται δηλαδή η αξιοπιστία του οδικού δικτύου.

Η πρόβλεψη της διάρκειας ενός φαινομένου είναι ευρέως διαδεδομένο αντικείμενο μελέτης. Τα μοντέλα διάρκειας (duration models) χρησιμοποιούνται κυρίως σε ιατρικές μελέτες, όπως ο χρόνος επιβίωσης ασθενών ή ο χρόνος μεταξύ διαδοχικών θανάτων ασθενών, αλλά και σε διάφορα άλλα ζητήματα. Η βασικότερη μέθοδος με αντικείμενο μελέτης τη διάρκεια είναι η μέθοδος Ανάλυσης Επιβίωσης (Survival Analysis).

Ένα παράδειγμα εφαρμογής της Ανάλυσης Επιβίωσης είναι η έρευνα που έγινε στην πολιτεία του Τέξας στις Η.Π.Α. για το αντίκτυπο που είχε στις επιχειρήσεις ο τυφώνας Harvey που έπληξε την περιοχή τον Αύγουστο του 2017 (Lee, 2021). Συγκεκριμένα, έγινε διερεύνηση του χρόνου μεταξύ της λήξης του φαινομένου και του μήνα επαναλειτουργίας κάθε επιχείρησης. Με βάση κάποιες ανεξάρτητες μεταβλητές, όπως την ηλικία της επιχείρησης, το μέγεθός της, το είδος της αλλά και το ποσοστό των ζημιών που προκλήθηκαν, κατασκευάστηκαν μοντέλα

τα οποία περιγράφουν την επιρροή αυτών των παραγόντων στο χρόνο που χρειάστηκε κάθε επιχείρηση μέχρι να επαναλειτουργήσει.

Τα τελευταία χρόνια, η χρήση της Ανάλυσης Επιβίωσης έχει διαδοθεί και στο συγκοινωνιακό τομέα. Η πιο συχνή χρήση των μοντέλων διάρκειας είναι στο πεδίο της οδικής ασφάλειας. Στη Μανίλα, την πρωτεύουσα των Φιλιππίνων, εφαρμόστηκε Ανάλυση Επιβίωσης για τους εμπλεκόμενους σε οδικά ατυχήματα (Ibay et al., 2019). Διερευνήθηκε η επίδραση που μπορεί να έχουν παράγοντες όπως η ηλικία, το φύλο, ο καιρός, η χρονική περίοδος και η τοποθεσία, στο χρόνο που καταφέρνει ένας ασθενής να παραμείνει στη ζωή μετά από ένα τροχαίο.

Στην πόλη Χέφει στην Κίνα διερευνήθηκε η διάρκεια των συγκρούσεων σε οδικά ατυχήματα που κατεγράφησαν από κάμερες κυκλοφορίας κατά την περίοδο 2016-2020 (Huajing et al., 2022). Η εν λόγω διάρκεια ορίζεται ως το χρονικό διάστημα μεταξύ της στιγμής που ένας οδηγός θέτει σε κίνδυνο με την οδήγησή του ένα άλλο όχημα και της στιγμής που τελικά προκαλείται η σύγκρουση. Σε αυτό το μοντέλο εξετάστηκαν βασικές συμπεριφορές οδηγών, ικανές να προκαλέσουν οδικό ατύχημα, όπως οι προσπεράσεις, η απρόσεκτη οδήγηση και η μη χρήση φώτων δείκτη κατεύθυνσης (φλας). Διερευνήθηκε επίσης και η επιρροή κάθε συμπεριφοράς ανά τύπο σύγκρουσης (μετωπική, οπίσθια, πλάγια, κλπ.).

Με τη χρονική διάρκεια συμβάντων στον αυτοκινητόδρομο M25 στο Λονδίνο ασχολήθηκαν οι Kalair and Connaughton (2021). Ουσιαστικά, διερεύνησαν όχι μόνο τη διάρκεια που το περιστατικό θα λαμβάνει χώρα, αλλά και το χρόνο που θα χρειαστεί ώστε η κυκλοφορία να επανέλθει σε κανονικές συνθήκες.

Ο χρόνος αντίδρασης σε ένα οδικό ατύχημα, διαφέρει σημαντικά ανά χρήστη. Το ζήτημα προσέγγισαν με εφαρμογή Ανάλυσης Επιβίωσης οι Parmet et al. (2014), οι οποίοι διερεύνησαν το χρόνο που χρειάζεται κάθε οδηγός από τη στιγμή που αντιλαμβάνεται τον κίνδυνο έως τη στιγμή που θα αντιδράσει.

Ακόμη μια ενδιαφέρουσα έρευνα είναι εκείνη των Oralhan and Göktolga (2018), οι οποίοι ανέπτυξαν μοντέλο πρόβλεψης του πότε ένας οδηγός θα εμπλακεί σε ένα οδικό ατύχημα στην Καισάρεια της Τουρκίας. Το υπό διερεύνηση χρονικό διάστημα, έχει ως έναρξη την ημέρα που ο οδηγός αποκτά το δίπλωμα οδήγησής του και λήξη την ημέρα που θα εμπλακεί σε κάποιο συμβάν.

Ένα ακόμα βασικό αντικείμενο έρευνας, είναι ο χρόνος αναζήτησης θέσης στάθμευσης. Το parking στην Αθήνα είναι ιδιαίτερο ζήτημα, καθώς οι χρόνοι αναζήτησης, σε συγκεκριμένες πυκνοκατοικημένες περιοχές, μπορεί να ξεπεράσουν τα 30 λεπτά. Σε αντίστοιχη έρευνα, με δεδομένα διαδρομών οχημάτων στην Αθήνα, διερευνήθηκε η επιρροή παραγόντων στη διάρκεια αναζήτησης θέσης στάθμευσης, όπως οι χρήσεις γης, οι ώρες αιχμής και η ημέρα (Mantouka et al., 2021).

Όσον αφορά τον κορεσμό, που αποτελεί και το αντικείμενο διερεύνησης της παρούσας διπλωματικής εργασίας, οι Liu et al. (2020), ανέλυσαν δεδομένα τροχιάς που συλλέχθηκαν από διαδρομές περίπου 10 χιλιάδων οχημάτων, και τα ενέταξαν σε ένα μη-παραμετρικό μοντέλο Επιβίωσης.

Οι Hu et al. (2019) διερεύνησαν τη σχέση κυκλοφοριακού φόρτου και ταχύτητας και την κατανομή της διάρκειας του κορεσμού βασισμένοι σε δεδομένα συλλεγμένα από την πόλη Σεντζέν στην Κίνα, λαμβάνοντας υπόψιν πως από τα σημαντικότερα χαρακτηριστικά της διάρκειας είναι οι ώρες αιχμής.

2.4. Σύνοψη βιβλιογραφικής ανασκόπησης

Ο κορεσμός είναι ένα περίπλοκο φαινόμενο. Υπάρχουν αρκετοί και διαφορετικοί τρόποι ώστε να εντοπιστεί και να προσδιοριστεί. Εκτός από τις αναλυτικές μεθόδους ή τις μεθόδους που βασίζονται στην κατανομή των καταγεγραμμένων ταχυτήτων, με την ραγδαία ανάπτυξη της τεχνολογίας, υπάρχουν πλέον και μέθοδοι εντοπισμού κορεσμού με τεχνικά μέσα, όπως οι κάμερες κυκλοφορίας. Η εφαρμογή των παραπάνω μεθόδων, διευκολύνει την πρόοδο ερευνών για την πρόβλεψη της διάρκειάς του. Η ανάγκη για διερεύνηση του συγκεκριμένου ζητήματος είναι επιτακτική, διότι τα μεγάλης διάρκειας συμβάντα κορεσμού δημιουργούν μεγάλα προβλήματα στη ζωή των πολιτών και την οικονομία. Η μέθοδος Ανάλυσης Επιβίωσης, είναι η πιο διαδεδομένη μέθοδος διερεύνησης της διάρκειας ενός φαινομένου, κι ενώ αρχικά εφαρμόστηκε σε έρευνες ιατρικού σκοπού, πλέον η χρήση της επεκτάθηκε και στο συγκοινωνιακό τομέα, σε έρευνες που αφορούν είτε την οδική ασφάλεια, είτε τη διαχείριση της κυκλοφορίας σε αστικά οδικά δίκτυα. Η μέθοδος Ανάλυσης Επιβίωσης θα αποτελέσει το βασικό εργαλείο της παρούσας διπλωματικής εργασίας και αναλύεται θεωρητικά στο επόμενο Κεφάλαιο.

3. ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

3.1. Ανάπτυξη μεθοδολογίας

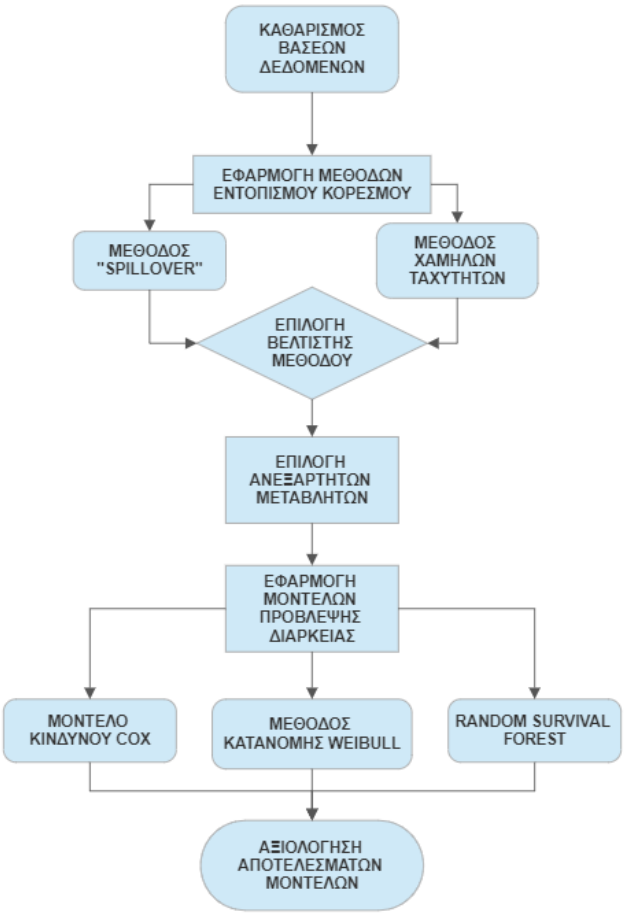
Με βάση το στόχο της διπλωματικής εργασίας, ακολουθήθηκαν τα εξής βήματα τα οποία θα καταλήξουν στην ανάπτυξη των μοντέλων πρόβλεψης διάρκειας του κορεσμού σε αστικά οδικά δίκτυα.

Αρχικά, γίνεται ο καθαρισμός των βάσεων δεδομένων, ώστε να αποφευχθεί η εισαγωγή λάθος μετρήσεων στα μοντέλα και ως εκ τούτου, η επιρροή αυτών των λαθών στα αποτελέσματα της ανάλυσης.

Έπειτα, εφαρμόζονται δύο μέθοδοι εντοπισμού κορεσμού, οι οποίες προέκυψαν μετά από εκτενή αναζήτηση στη βιβλιογραφία. Οι δύο μέθοδοι που εφαρμόζονται είναι η μέθοδος «Spillover» και η μέθοδος χαμηλών ταχυτήτων.

Για την ανάπτυξη των μοντέλων πρόβλεψης διάρκειας, επιλέγεται μία από τις δύο μεθόδους εντοπισμού και πραγματοποιείται Ανάλυση Επιβίωσης (Survival Analysis) με χρήση τριών μοντέλων και εισαγωγή των κατάλληλων ανεξάρτητων μεταβλητών.

Τέλος, αναλύονται τα αποτελέσματα ανά μοντέλο και η επιρροή των ανεξάρτητων μεταβλητών σε καθένα από αυτά.



Διάγραμμα 3.1. Διάγραμμα ροής μεθοδολογίας

3.2. Βασικές έννοιες της Ανάλυσης Επιβίωσης

Για το σκοπό της διπλωματικής εργασίας, στο μοντέλο διάρκειας που θα αναπτυχθεί, θα χρησιμοποιηθεί η μέθοδος Survival Analysis. Στο παρόν κεφάλαιο θα παρουσιαστεί η θεωρία πίσω από τη συγκεκριμένη μεθοδολογία καθώς και τα διάφορα μοντέλα που θα εφαρμοστούν (Washington, Karlaftis, Mannering, Anastasopoulos, 2004).

Η Ανάλυση Επιβίωσης (Survival Analysis) αρχικά εφαρμόστηκε σε έρευνες χρόνου επιβίωσης ασθενών (χρόνος μέχρι του θάνατο του ασθενούς), από όπου απέκτησε και το όνομά της. Στο διάγραμμα μίας καμπύλης Survival, προβάλλονται στον οριζόντιο άξονα οι διάρκειες και στον κατακόρυφο οι πιθανότητες. Στη συνάρτηση «επιβίωσης» περιγράφεται η πιθανότητα, η διάρκεια T ενός φαινομένου να είναι μεγαλύτερη ή ίση από μια χρονική στιγμή t:

$$S(t) = P(T \geq t) \quad (3.1)$$

Όπου P η πιθανότητα, T μια τυχαία μεταβλητή χρόνου και t μια συγκεκριμένη χρονική στιγμή. Παράλληλα, υπάρχουν κι οι συναρτήσεις «κινδύνου» (hazard functions) οι οποίες βασίζονται στην εξής αθροιστική συνάρτηση κατανομής:

$$F(t) = P(T < t) \quad (3.2)$$

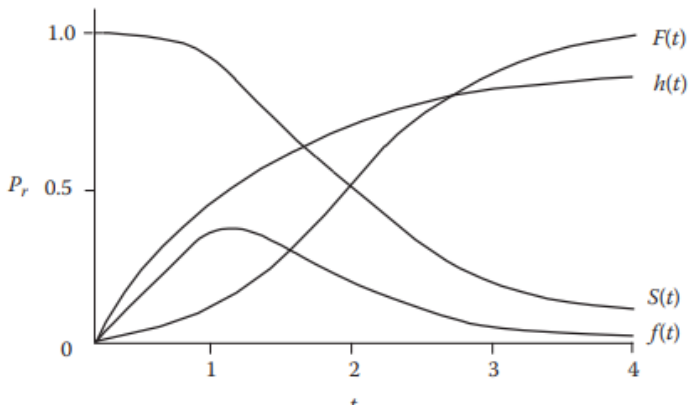
Αντίστοιχα, η συνάρτηση πυκνότητας πιθανότητας είναι:

$$f(t) = \frac{dF(t)}{dt} \quad (3.3)$$

Επομένως, προκύπτει η συνάρτηση κινδύνου:

$$h(t) = \frac{f(t)}{[1 - F(t)]} \quad (3.4)$$

Στο Διάγραμμα 3.2 φαίνεται η απεικόνιση όλων των παραπάνω συναρτήσεων:



Διάγραμμα 3.2. Απεικόνιση βασικών συναρτήσεων Survival Analysis (Washington, Karlaftis, Mannering, Anastasopoulos, 2004)

Με βάση τα παραπάνω, οι σχέσεις μεταξύ των συναρτήσεων έχουν ως εξής:

$$S(t) = 1 - F(t) = 1 - \int_0^t f(t)dt = EXP[-H(t)] \quad (3.6)$$

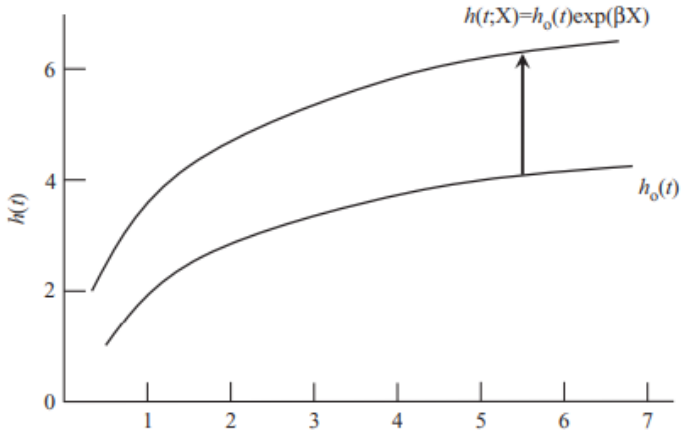
$$f(t) = \frac{d}{dt}F(t) = h(t)EXP[-H(t)] = -\frac{d}{dt}S(t) \quad (3.7)$$

$$H(t) = \int_0^t h(t)dt = -LN[S(t)] \quad (3.8)$$

$$h(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)} = \frac{d}{dt}H(t) \quad (3.9)$$

Στα μοντέλα Survival, μπορούν να ενσωματωθούν και επεξηγηματικές μεταβλητές. Ως επεξηγηματικές ορίζονται οι ανεξάρτητες μεταβλητές X οι οποίες περιγράφουν το δείγμα με βάση διάφορα χαρακτηριστικά του. Η επιρροή τους στην πιθανότητα είναι διαφορετική ανάλογα με τις τιμές που λαμβάνουν, γι'αυτό και είναι χρήσιμες στην εξαγωγή ασφαλών συμπερασμάτων. Υπάρχουν δύο μοντέλα εισαγωγής των επεξηγηματικών μεταβλητών στη συνάρτηση κινδύνου. Το πρώτο είναι το αναλογικό (proportional) μοντέλο στο οποίο θεωρείται ότι δρουν πολλαπλασιαστικά στη βασική συνάρτηση κινδύνου $h_0(t)$ και συγκεκριμένα με τη συνάρτηση $EXP(\beta X)$, όπου β ο συντελεστής της ανεξάρτητης μεταβλητής. Συνεπώς η hazard function καταλήγει ως εξής:

$$h(t; X) = h_0(t)EXP(\beta X) \quad (3.10)$$



Διάγραμμα 3.3. Εισαγωγή επεξηγηματικής μεταβλητής στη βασική συνάρτηση κινδύνου με αναλογικό μοντέλο(Washington, Karlaftis, Mannering, Anastasopoulos, 2004)

Το δεύτερο είναι το μοντέλο επιταχυνόμενης διάρκειας ζωής (accelerated lifetime) στο οποίο η επιρροή της επεξηγηματικής μεταβλητής ενσωματώνεται στη βασική συνάρτηση επιβίωσης $S_0(t)$ με τη μορφή της συνάρτησης $EXP(\beta X)$. Στο προκείμενο μοντέλο, όταν οι επεξηγηματικές μεταβλητές είναι μηδενικές, η νέα συνάρτηση επιβίωσης είναι ίση με τη βασική. Η μορφή της νέας συνάρτησης επιβίωσης είναι:

$$S(t;X) = S_0[tEXP(\beta X)] \quad (3.11)$$

Ενώ η συνάρτηση κινδύνου καταλήγει ως εξής:

$$h(t;X) = h_0[tEXP(\beta X)]EXP(\beta X) \quad (3.12)$$

Τα δεδομένα διάρκειας που χρησιμοποιούνται στην ανάπτυξη ενός μοντέλου μπορεί να είναι περιορισμένα αριστερά (left-censored), περιορισμένα δεξιά (right-censored) ή τίποτα από τα δύο. Σε μια έρευνα, που διερευνώνται οι διάρκειες ενός φαινομένου για ένα χρονικό διάστημα μεταξύ των χρονικών στιγμών t_1 και t_2, περιορισμένες αριστερά είναι οι καταγεγραμμένες διάρκειες των οποίων οι ενάρξεις έχουν συμβεί πριν τη στιγμή t_1, ενώ περιορισμένες δεξιά αυτές που λήγουν μετά τη στιγμή t_2. Σε αυτές τις περιπτώσεις, είναι δύσκολη η απόφαση για την κατανομή αυτών των διαρκειών μιας και λείπει βασικές πληροφορίες.

3.2.1. Είδη μοντέλων

Τα μοντέλα διάρκειας διακρίνονται σε τρεις κατηγορίες: τα μη-παραμετρικά, τα ημι-παραμετρικά και τα πλήρως παραμετρικά μοντέλα. Στα μη παραμετρικά μοντέλα, δε χρησιμοποιούνται επεξηγηματικές μεταβλητές και η ανάλυση γίνεται με την παραδοχή ότι δεν ακολουθείται συγκεκριμένη στατιστική κατανομή. Αυτοί είναι και οι βασικοί λόγοι για τους οποίους δε χρησιμοποιούνται αρκετά στο συγκοινωνιακό τομέα. Δύο παραδείγματα μη-παραμετρικών μοντέλων είναι η μέθοδος της Εκθετικής Κατανομής κι η μέθοδος Kaplan-Meier (Kaplan and Meier, 1958).

Στις ημι-παραμετρικές μεθόδους, ενώ ενσωματώνονται ανεξάρτητες μεταβλητές, δεν θεωρείται ότι οι διάρκειες ακολουθούν κάποια κατανομή. Η πιο γνωστή είναι η μέθοδος **Cox** (Cox, 1972), στην οποία διερευνάται η αναλογία «κινδύνου» σε κάθε παρατήρηση i του υπό μελέτη φαινομένου με τύπο:

$$\frac{EXP(\beta X_i)}{\sum_{j \in R_i} EXP(\beta X_j)} \quad (3.13)$$

Όπου R_i το σύνολο των παρατηρήσεων j.

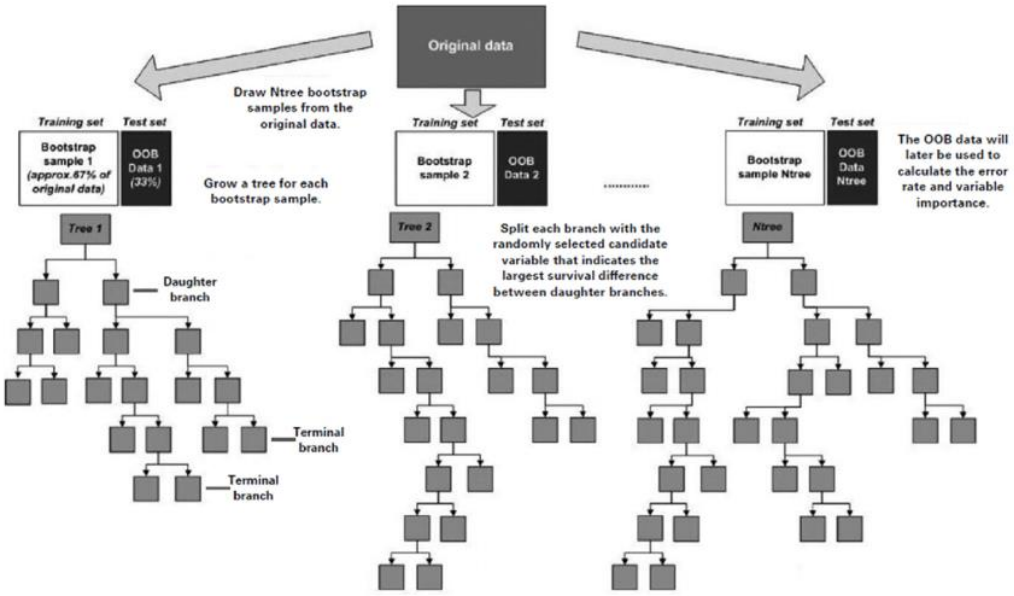
Στην περίπτωση των πλήρως παραμετρικών μοντέλων, πέρα από τις ανεξάρτητες μεταβλητές, γίνεται και η παραδοχή πως καταγεγραμμένες διάρκειες ακολουθούν μία στατιστική κατανομή, όπως η συνεχής κατανομή **Weibull**. Η συνάρτηση πυκνότητας της κατανομής Weibull είναι:

$$f(t) = (\lambda P)(\lambda t)^{P-1} EXP[(\lambda t)^{P-1}] \quad (3.14)$$

με συνάρτηση κινδύνου:

$$h(t) = (\lambda P)(\lambda t)^{P-1} \quad (3.15)$$

Τέλος, υπάρχει και η παραμετρική μέθοδος **Random Survival Forest**. Πρόκειται για ένα μοντέλο μηχανικής μάθησης, κατά το οποίο δημιουργείται ένα «δάσος» από «δέντρα» αποφάσεων τα οποία ταξινομούν τα δεδομένα διάρκειας ανάλογα με τις τιμές των μεταβλητών τους και στο τέλος γίνεται πρόβλεψη της καμπύλης επιβίωσής τους (Ishwaran et al., 2008). Το «δάσος» είναι της μορφής:



Εικόνα 3.1. Οπτικοποίηση των "δέντρων" απόφασης της Random Survival Forest (Datema et al., 2012)

3.2.2. Αξιολόγηση μοντέλων

Για να αξιολογηθεί ο συσχετισμός μεταξύ δύο δειγμάτων (διάρκειες πραγματικών συμβάντων και διάρκειες προβλεπόμενων) πρέπει να διεξαχθεί ένας στατιστικός έλεγχος. Για την αξιολόγηση των δύο κατατάξεων, συγκρίνεται η απόκλιση κάθε μεμονωμένης πρόβλεψης από την διάταξη των πραγματικών παρατηρήσεων. Ένας ενδεικτικός στατιστικός έλεγχος είναι αυτούς του συντελεστή **Concordance (C-index)**. Ο C-index είναι αυτός που δείχνει πόσο καλά προσαρμόζεται πρόβλεψη στα πραγματικά δεδομένα. Ουσιαστικά ελέγχει κατά πόσο οι τιμές που προβλέφθηκαν, έχουν την ίδια κατανομή με τις πραγματικές και αν έχουν την ίδια σχέση μεταξύ τους. Οι τιμές του κυμαίνονται από 0,50 έως 1, όπου 0,50 για τελείως τυχαία κατανομή και 1 για απόλυτη ταύτιση.

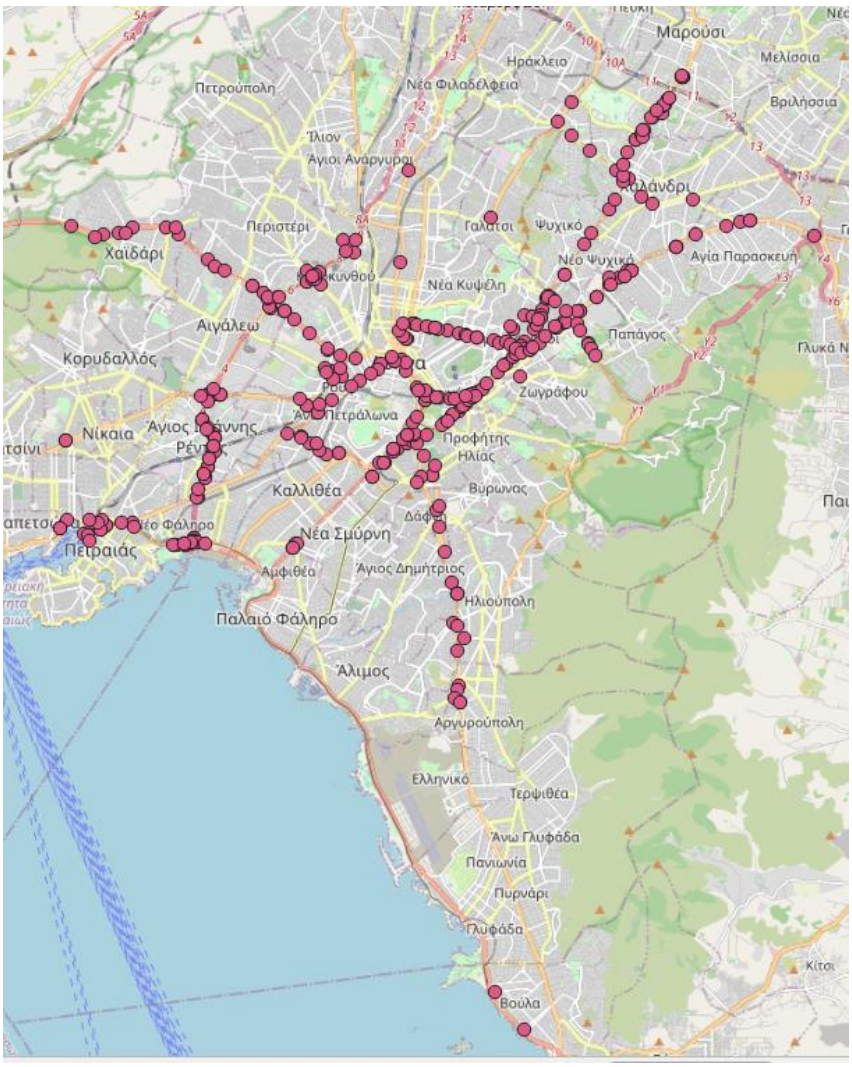
4. ΕΦΑΡΜΟΓΗ ΜΕΘΟΔΟΛΟΓΙΑΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

4.1. Περιγραφή δεδομένων

Στην παρούσα διπλωματική, λήφθηκαν υπόψη δύο βάσεις δεδομένων μετρήσεων κυκλοφοριακών μεγεθών, από δύο πόλεις της Ευρώπης, την Αθήνα και το Μάντσεστερ. Στην Αθήνα ο πληθυσμός του αστικού κέντρου ανέρχεται στο 1 εκατομμύριο κατοίκους ενώ στην ευρύτερη περιοχή αγγίζει τα 3 εκατομμύρια. Παράλληλα, στο Μάντσεστερ, ο πληθυσμός είναι 550 χιλιάδες κάτοικοι ενώ στην ευρύτερη περιοχή ξεπερνούν τα 2,8 εκατομμύρια (2η πολυπληθέστερη μητροπολιτική περιοχή στο Ηνωμένο Βασίλειο).

4.1.1. Βάση Δεδομένων Αθήνας

Τα δεδομένα τα οποία χρησιμοποιήθηκαν στην έρευνα, τα είχε στην κατοχή του το Εθνικό Μετσόβιο Πολυτεχνείο. Πρόκειται για μετρήσεις δύο ημερών (9η και 10η Νοεμβρίου 2022) από 351 φωρατές στην Αθήνα. Οι εν λόγω φωρατές είναι μαγνητικοί βρόγχοι, που βρίσκονται κάτω από το οδόστρωμα βασικών αρτηριών του οδικού δικτύου της Αθήνας και καταγράφουν, ανά 90 δευτερόλεπτα, κυκλοφοριακό φόρτο, ταχύτητα και κατάληψη οδοστρώματος.



Εικόνα 4.1. Σημεία μέτρησης κυκλοφοριακών μεγεθών στην Αθήνα

Τα δεδομένα που αναλύθηκαν είναι ομαλοποιημένα, ώστε να μην υπάρχουν ακραίες αυξομειώσεις στα μεγέθη. Η βάση δεδομένων έχει την εξής μορφή:

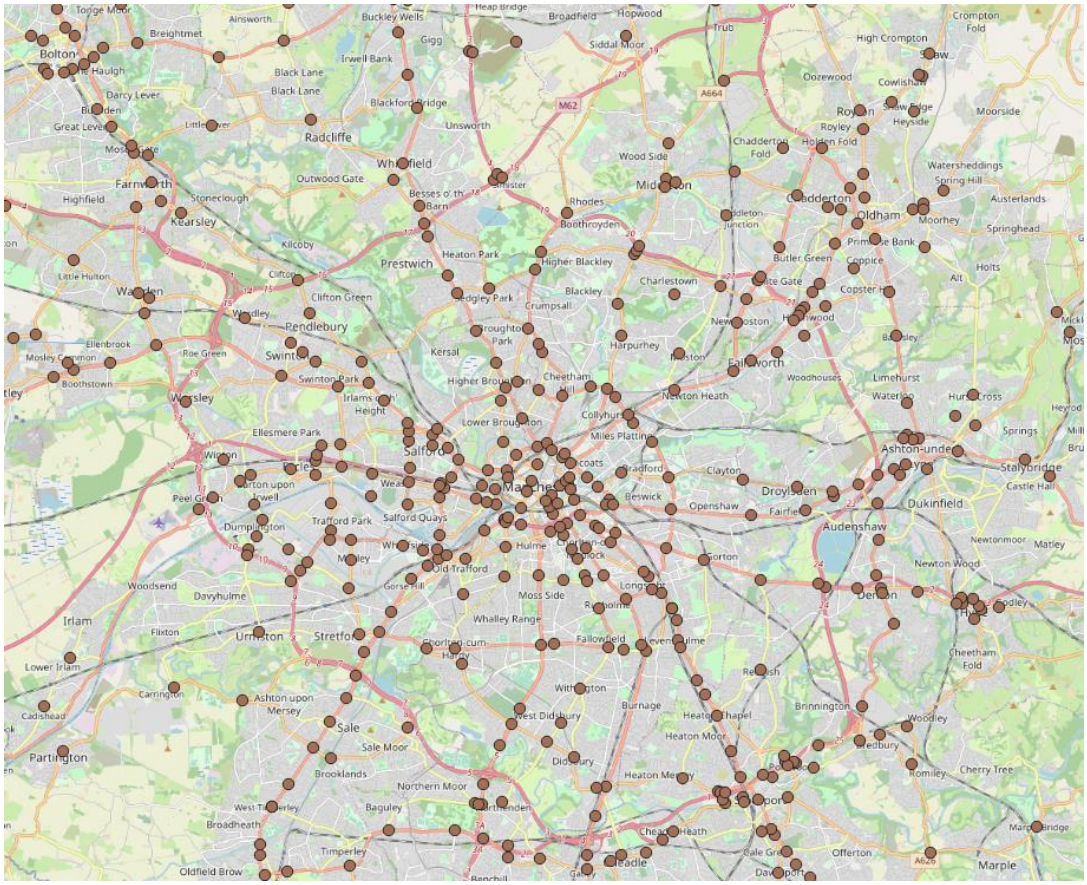
Πίνακας 4.1. Μορφή βάσης δεδομένων μετρήσεων Αθήνας

	Name	Time	smoothed_all_vol	smoothed_all_occ	smoothed_all_spd
0	MS106	09.11.2022 00:01:20	1577.9446	3.4227946	97.97528
1	MS106	09.11.2022 00:02:50	1703.4585	3.7337625	98.29494
2	MS106	09.11.2022 00:04:20	1757.5939	3.6336553	98.7096
3	MS106	09.11.2022 00:05:50	1748.1954	3.3085747	98.90688
4	MS106	09.11.2022 00:07:20	1691.1465	3.0647645	98.9486
5	MS106	09.11.2022 00:08:50	1608.3599	3.1319067	98.87189
6	MS106	09.11.2022 00:10:20	1496.2699	2.84893	98.1153
7	MS106	09.11.2022 00:11:50	1432.2024	2.7200308	98.930534
8	MS106	09.11.2022 00:13:20	1364.1519	2.6233563	98.88476
9	MS106	09.11.2022 00:14:50	1363.1139	2.5508506	98.887856

Ο φόρτος είναι ανηγμένος σε οχήματα ανά ώρα (veh/h), η ταχύτητα μετριέται σε χιλιόμετρα ανά ώρα (km/h) και η κατάληψη είναι επί τοις εκατό (%).

4.1.2. Βάση δεδομένων Μάντσεστερ

Από το Μάντσεστερ, αποκτήθηκαν μετρήσεις ταχύτητας και κυκλοφοριακού φόρτου ανά 5 λεπτά, ενός διαστήματος δεκατριών μηνών, (από 1η Ιανουαρίου 2022 έως 31η Ιανουαρίου 2023), από 137 φωρατές σε βασικούς άξονες της μητροπολιτικής περιοχής της πόλης.



Εικόνα 4.2. Σημεία μέτρησης κυκλοφοριακών μεγεθών στο Μάντσεστερ

Η βάση δεδομένων είχε τη μορφή:

Πίνακας 4.2. Μορφή βάσης δεδομένων μετρήσεων Μάντσεστερ

dyna_id	type	time	entity_id	total_count	traffic_speed
0	state	1/1/2022 00:05	000000001010	189	42.88359788
1	state	1/1/2022 00:10	000000001010	150	40.96666667
2	state	1/1/2022 00:15	000000001010	221	41.90045249
3	state	1/1/2022 00:20	000000001010	240	44.16666667
4	state	1/1/2022 00:25	000000001010	223	43.92376682
5	state	1/1/2022 00:30	000000001010	282	43.42198582
6	state	1/1/2022 00:35	000000001010	294	41.07142857
7	state	1/1/2022 00:40	000000001010	275	43
8	state	1/1/2022 00:45	000000001010	287	43.1184669
9	state	1/1/2022 00:50	000000001010	286	43.07692308

Ο φόρτος μετριέται σε οχήματα ανά 5 λεπτά και η ταχύτητα σε χιλιόμετρα ανά ώρα.

4.2. Επεξεργασία δεδομένων

Το πρώτο βήμα και για τις δύο περιπτώσεις, ήταν ο καθαρισμός των βάσεων δεδομένων. Ειδικότερα, επειδή παρατηρήθηκε μεγάλος όγκος δεδομένων όπου οι φωρατές δεν κατέγραψαν μετρήσεις, αυτές είτε αφαιρέθηκαν, είτε δε λήφθηκαν υπόψη. Επίσης, μερικοί από τους φωρατές της Αθήνας, είχαν περισσότερες καταγεγραμμένες μετρήσεις από όσες έπρεπε. Συγκεκριμένα μπορεί μεταξύ δύο μετρήσεων που απείχαν χρονικά 90 δευτερόλεπτα, να υπήρχε μία ενδιάμεση μέτρηση 3 δευτερόλεπτα μετά την πρώτη. Αυτές οι μετρήσεις αφαιρέθηκαν διότι δεν ήταν σταθερές σε όλους τους φωρατές και δεν παρείχαν κάποια άξια αναφοράς πληροφορία.

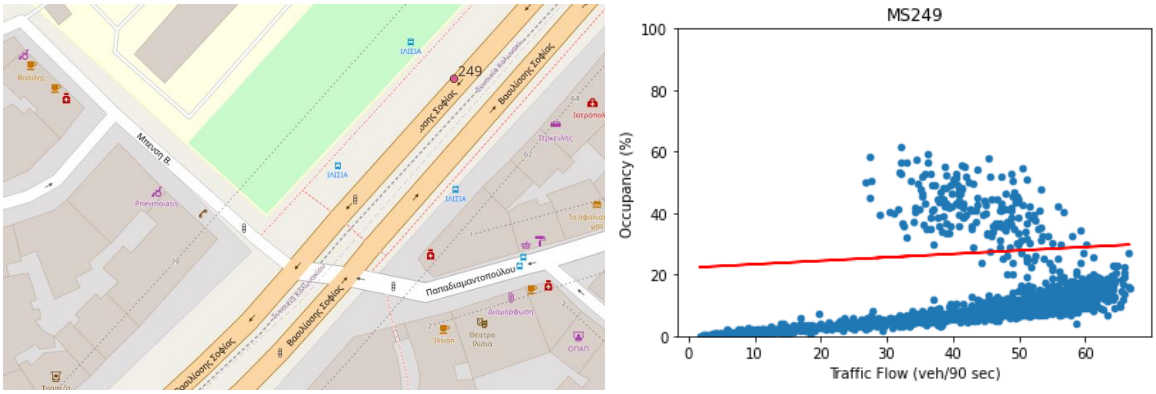
4.2.1. Εντοπισμός κορεσμού

Λόγω των διαφορετικών μεγεθών που μετρήθηκαν σε κάθε μία από τις πόλεις, δοκιμάστηκαν δύο μέθοδοι εντοπισμού κορεσμού. Μια αναλυτική, η οποία εντοπίζει το «spillover», το οποίο λαμβάνει υπόψη την κατάληψη οδοστρώματος και το φόρτο και μία που βασίζεται στα υπάρχοντα δεδομένα για τον εντοπισμό των ακραίων τιμών, στην οποία η κίνηση των οχημάτων με πολύ χαμηλές ταχύτητες για την οδό, αποτελεί κορεσμό.

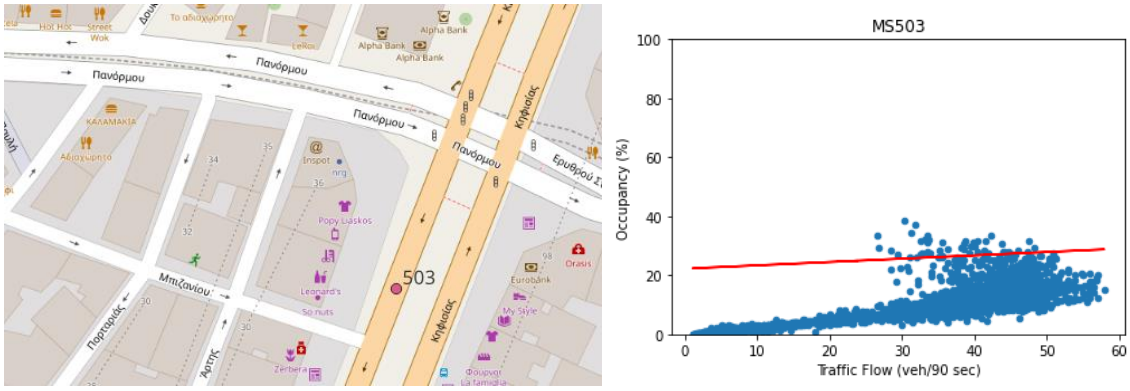
4.2.1.1. Μέθοδος «Spillover»

Η μέθοδος «spillover» εφαρμόστηκε μόνο στα δεδομένα της Αθήνας διότι μόνο σε αυτά υπάρχουν μετρήσεις κατάληψης. Στη μέθοδο αυτή, υπολογίζεται ένα ανώτατο όριο O_{sp} για την κατάληψη, πάνω από το οποίο θεωρείται πως η ροή στο οδικό δίκτυο έχει κορεστεί. Ο υπολογισμός της κατάληψης O_{sp} έγινε με την εξίσωση **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.** (Geroliminis and Skabardonis, 2006). Το μήκος οχήματος L_{eff} λήφθηκε ίσο με 5m, ως ταχύτητα ελεύθερης ροής u_f θεωρήθηκε το 90% εκατοστημόριο των καταγεγραμμένων ταχυτήτων και ως ποσοστό κόκκινης ένδειξης σε κάθε κύκλο του φωτεινού σηματοδότη, λήφθηκε μία μέση τιμή ίση με r/C=20/90.

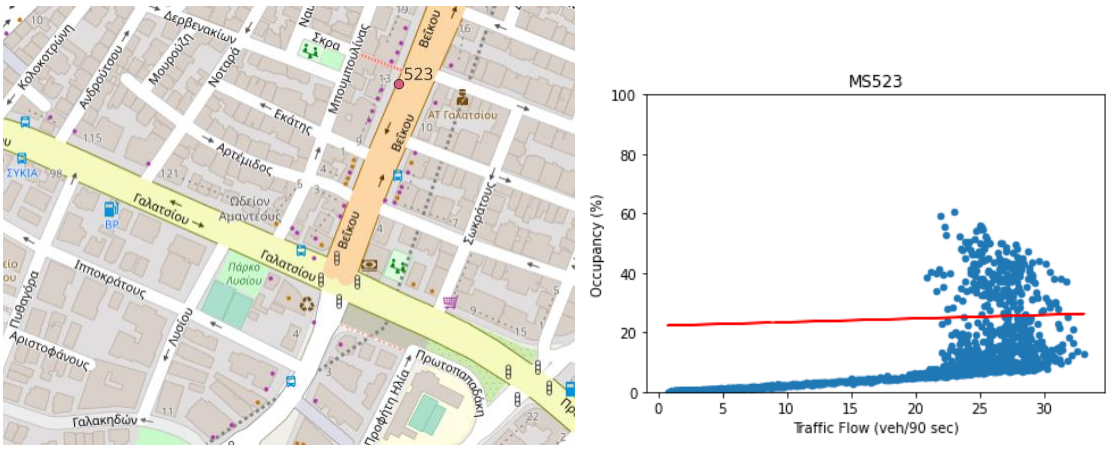
Σε γλώσσα προγραμματισμού Python, συντάχθηκε κώδικας, ο οποίος κατασκευάζει τα διαγράμματα κυκλοφοριακού φόρτου και κατάληψης οδοστρώματος σε κάθε φωρατή. Ακολουθούν μερικά παραδείγματα θέσεων και αντίστοιχων διαγραμμάτων.



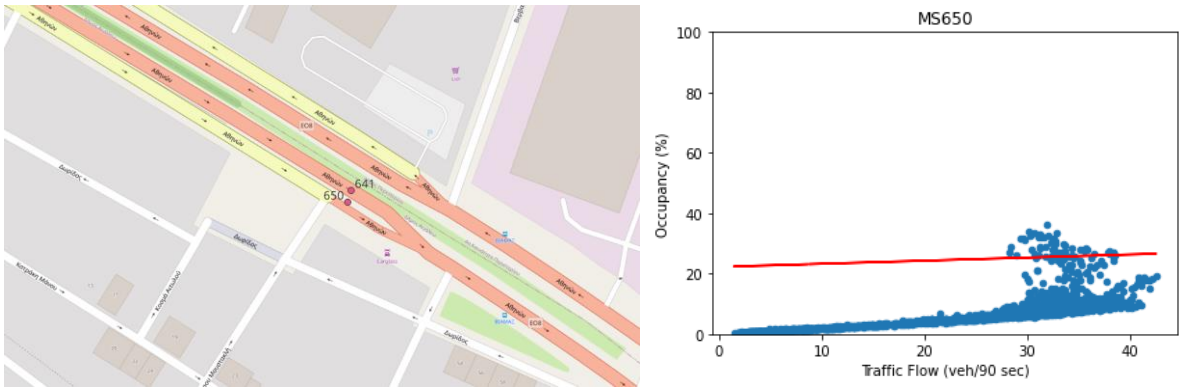
Διάγραμμα 4.1. Θέση μέτρησης MS249 και Διάγραμμα φόρτου-κατάληψης θέσης MS249 (Βασιλίσσης Σοφίας, ρεύμα προς Βασιλέως Κωνσταντίνου, πριν την Παπαδιαμαντοπούλου)



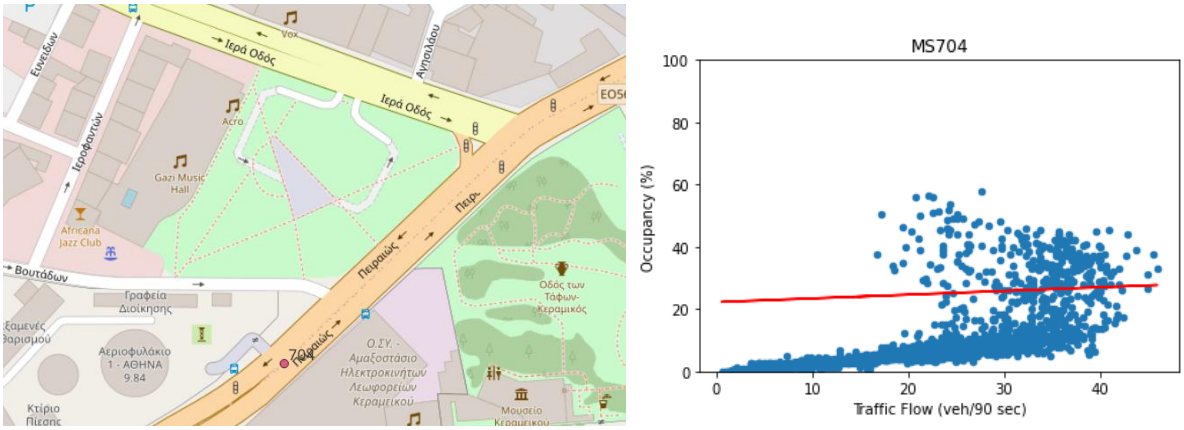
Διάγραμμα 4.2. Θέση μέτρησης MS503 και Διάγραμμα φόρτου-ταχύτητας στη θέση MS503 (Λεωφόρος Κηφισίας, ρεύμα προς Αθήνα, μετά από Πανόρμου)



Διάγραμμα 4.3. Θέση μέτρησης MS523 και Διάγραμμα φόρτου-κατάληψης θέσης MS523 (Λεωφόρος Βείκου, ρεύμα προς Αθήνα, ύψος Λεωφόρου Γαλατσίου)



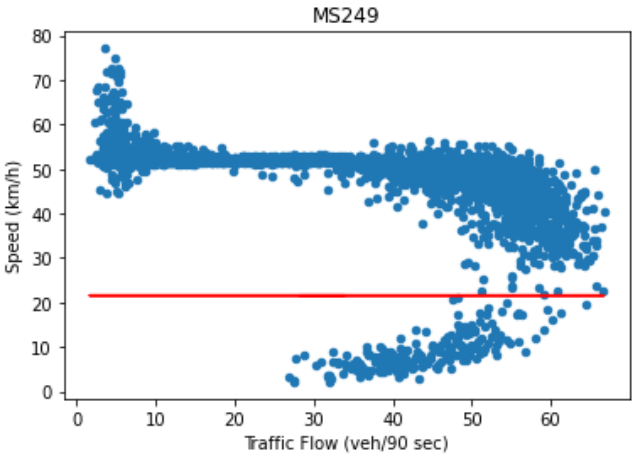
Διάγραμμα 4.4. Θέση μέτρησης MS650 και Διάγραμμα φόρτου-κατάληψης θέσης MS650 (Ράμπα εισόδου στη Λεωφόρο Αθηνών από Θηβών με κατεύθυνση προς Αθήνα)



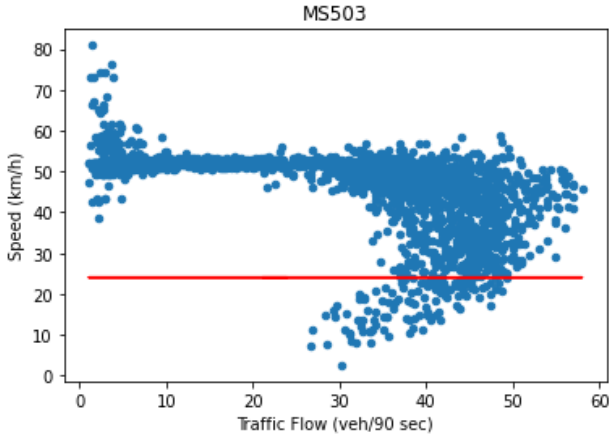
Διάγραμμα 4.5. Θέση μέτρησης MS704 και Διάγραμμα φόρτου-κατάληψης στη θέση MS704 (Λεωφόρος Πειραιώς, ρεύμα προς Αθήνα, ύψος Ιεράς Οδού)

4.2.1.2. Μέθοδος χαμηλών ταχυτήτων

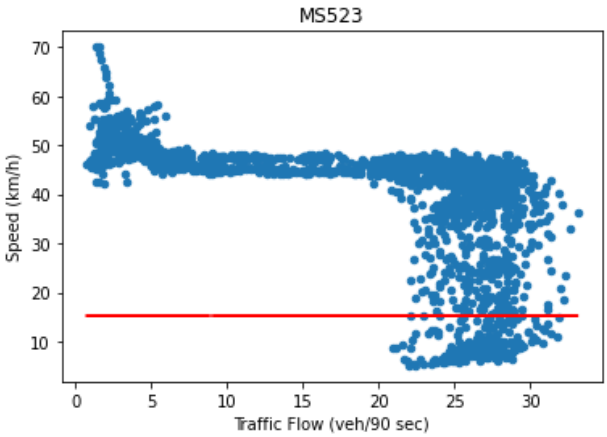
Εξαιτίας της απουσίας μετρήσεων κατάληψης από τα δεδομένα του Μάντσεστερ, αναζητήθηκε μια ακόμα μέθοδος για τον εντοπισμό του κορεσμού. Στη μέθοδο που τελικά εφαρμόστηκε, λήφθηκαν υπόψιν όλες οι καταγεγραμμένες ταχύτητες ανά φωρατή, και θεωρήθηκε πως επρόκειτο για κορεσμό, όταν η ταχύτητα ήταν μικρότερη από το 25% τεταρτημόριο των μετρήσεων μειωμένο κατά 1,5 φορές την τυπική απόκλιση, σύμφωνα με τη μέθοδο boxplot για τον εντοπισμό των ακραίων τιμών. Οι παραπάνω τιμές των παραμέτρων προέκυψαν μετά από εκτεταμένο έλεγχο για την εύρεση των βέλτιστων τιμών για τα συγκεκριμένα δεδομένα. Η μέθοδος δοκιμάστηκε και στη βάση δεδομένων της Αθήνας, και παρακάτω δίνονται τα νέα διαγράμματα φόρτου – ταχύτητας για τα ίδια σημεία μέτρησης:



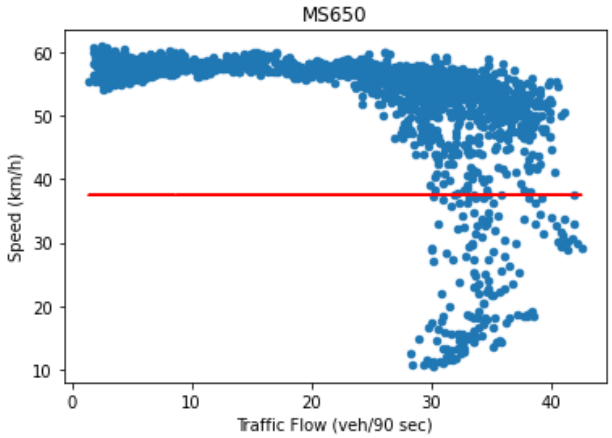
Διάγραμμα 4.6. Διάγραμμα φόρτου-ταχύτητας στη θέση MS249 (Βασιλίσσης Σοφίας, ρεύμα προς Βασιλέως Κωνσταντίνου, πριν την Παπαδιαμαντοπούλου)



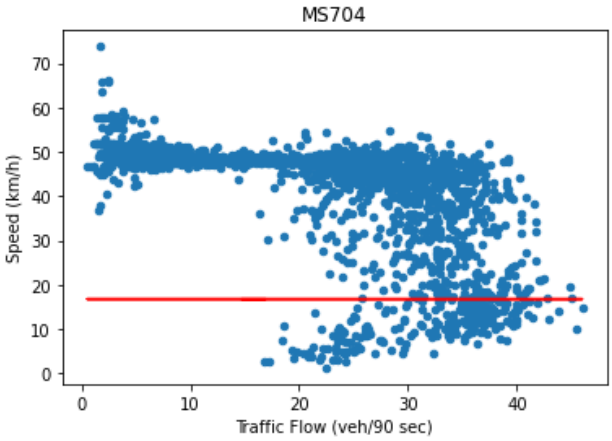
Διάγραμμα 4.7. Διάγραμμα φόρτου-ταχύτητας στη θέση MS503 (Λεωφόρος Κηφισίας, ρεύμα προς Αθήνα, μετά από Πανόρμου)



Διάγραμμα 4.8. Διάγραμμα φόρτου-ταχύτητας στη θέση MS523 (Λεωφόρος Βείκου, ρεύμα προς Αθήνα, ύψος Λεωφόρου Γαλατσίου)



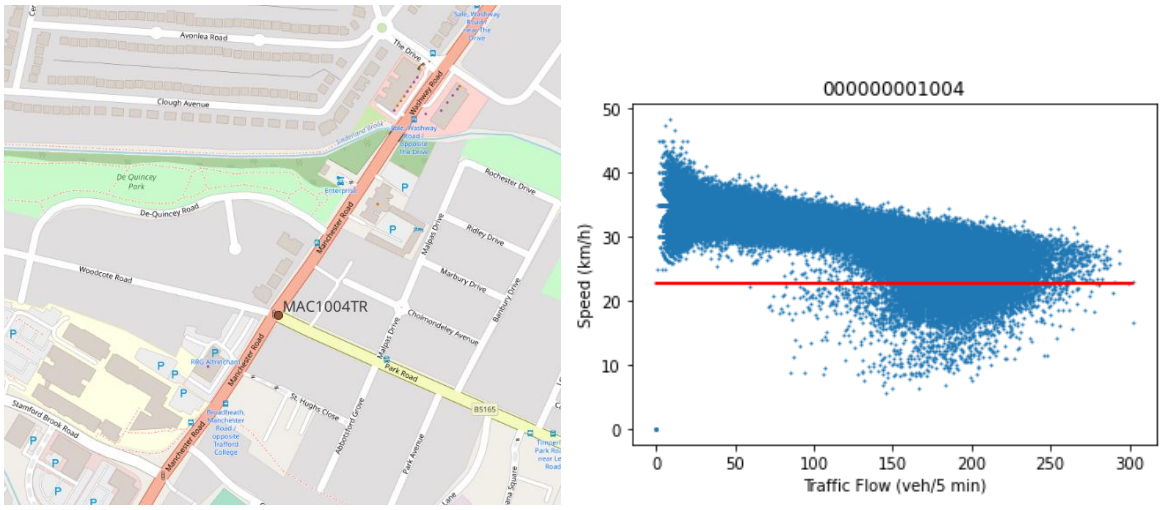
Διάγραμμα 4.9. Διάγραμμα φόρτου-ταχύτητας στη θέση MS650 (Ράμπα εισόδου στη Λεωφόρο Αθηνών από Θηβών με κατεύθυνση προς Αθήνα)



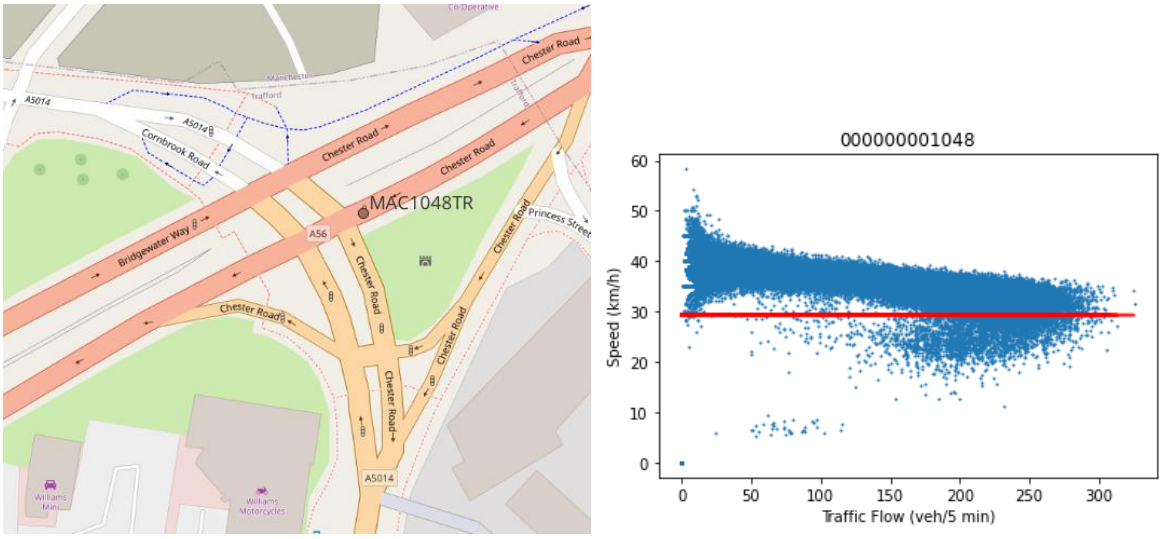
Διάγραμμα 4.10. Διάγραμμα φόρτου-ταχύτητας στη θέση MS704 (Λεωφόρος Πειραιώς, ρεύμα προς Αθήνα, ύψος Ιεράς Οδού)

Όσα σημεία βρίσκονται κάτω από την κόκκινη ευθεία, είναι μετρήσεις που έγιναν κατά τη διάρκεια του κορεσμού. Συγκρίνοντας τα διαγράμματα των σημείων της Αθήνας, και για τις δύο μεθόδους, διαπιστώθηκε διαφορά στον αριθμό των παρατηρήσεων κορεσμού. Αυτό συμβαίνει λόγω της φύσης των δύο μεθόδων, όπου η πρώτη είναι πιο αναλυτική και βασίζεται σε μια μαθηματική εξίσωση, ενώ η δεύτερη είναι πιο στατιστική καθώς λαμβάνει υπόψιν όλα τα δεδομένα κάθε σημείου.

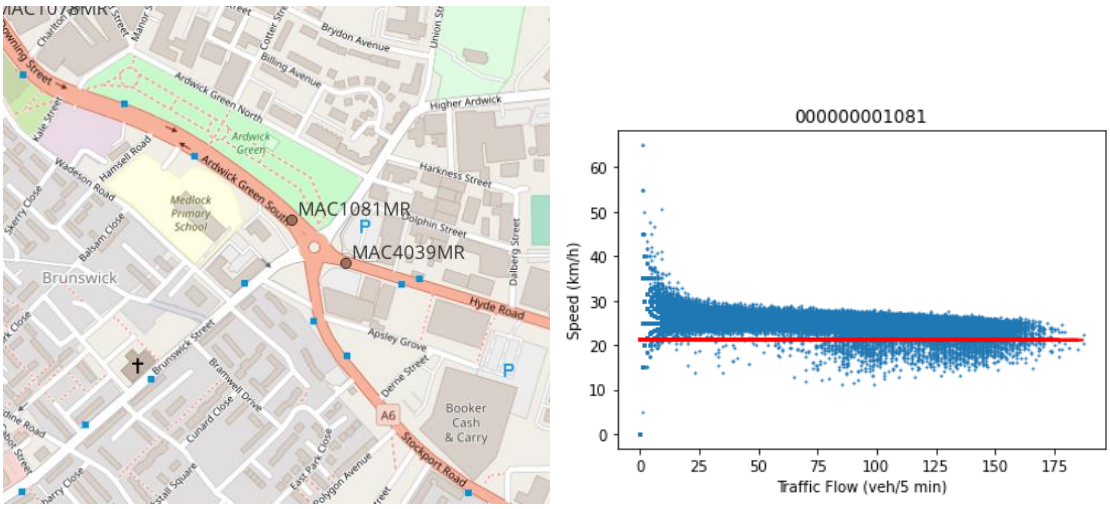
Όσον αφορά τις μετρήσεις από το Μάντσεστερ, δίνονται μερικά ενδεικτικά διαγράμματα φόρτου-ταχύτητας:



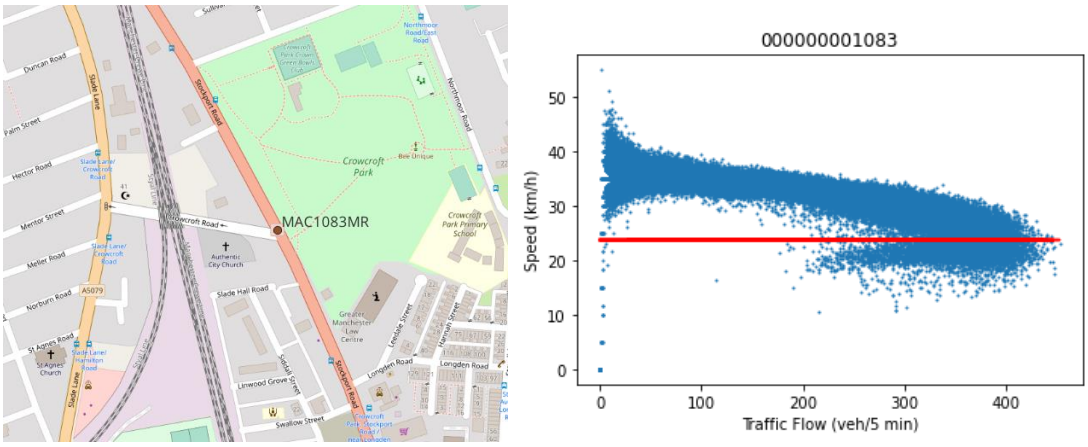
Διάγραμμα 4.11. Θέση μέτρησης 1004 και Διάγραμμα φόρτου-ταχύτητας στη θέση 1004 (Park Road, κόμβος με Manchester Road)



Διάγραμμα 4.12. Θέση μέτρησης 1048 και Διάγραμμα φόρτου-ταχύτητας στη θέση 1048 (Chester Road προς Bridgewater Way)



Διάγραμμα 4.13. Θέση μέτρησης 1081 και Διάγραμμα φόρτου-ταχύτητας στη θέση 1081 (Ardwick Green South, ρεύμα προς κόμβο Downing Street, μετά τον κυκλικό κόμβο με Hyde Road και Stockport Road)



Διάγραμμα 4.14. Θέση μέτρησης 1083 και Διάγραμμα φόρτου-ταχύτητας στη θέση 1083 (Stockport Road, κόμβος με Crowcroft Road)

Στην επόμενη ενότητα, για τον υπολογισμό της διάρκειας των συμβάντων κορεσμού που εντόπισε κάθε μέθοδος, θα χρησιμοποιηθεί η δεύτερη, με τις χαμηλές ταχύτητες. Στην Αθήνα δεν υπάρχει σε όλες τις θέσεις μέτρησης φωτεινή σηματοδότηση (παραδείγματος χάρη στον αυτοκινητόδρομό του Κηφισού), συνεπώς είναι λανθασμένη η χρήση της μεθόδου «Spillover» και της εξίσωσης **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.**, η οποία λαμβάνει υπόψιν ποσοστό κόκκινης ένδειξης στον κύκλο σηματοδότησης.

4.2.2. Υπολογισμός διάρκειας συμβάντων κορεσμού

Για τον υπολογισμό της διάρκειας των συμβάντων κορεσμού χρειάστηκε να κατασκευαστεί ο πίνακας «state». Στον πίνακα state, οι γραμμές είναι οι καταγεγραμμένες χρονικές στιγμές (90 δευτερόλεπτα για την Αθήνα και 5 λεπτά για το Μάντσεστερ) και οι στήλες, οι διάφοροι σταθμοί μέτρησης. Ο πίνακας συμπληρώνεται από τιμές 0 και 1, όπου το **0** αφορά στιγμή στην οποία **δεν υπάρχει κορεσμός** και το **1** στιγμή στην οποία **υπάρχει**. Η μορφή του πίνακα φαίνεται στην Εικόνα 4.3.

MS106	MS107	MS108	MS109	MS110	MS111	MS112	MS113	MS114
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	0	0
1	0	0	0	1	0	0	0	0
1	0	0	0	1	0	0	0	0
1	0	0	0	1	0	0	0	1
1	0	0	0	1	0	0	0	1
1	0	0	0	1	0	0	0	1
1	0	0	0	1	0	0	0	1

Εικόνα 4.3. Μορφή πίνακα state (0 ακόρεστη ροή, 1 κορεσμένη ροή)

Από τον παραπάνω πίνακα θα προκύψουν οι διάρκειες των καταγεγραμμένων συμβάντων, αθροίζοντας τις διαδοχικές μονάδες. Χρειάστηκε ωστόσο, να πραγματοποιηθούν μερικές μορφοποιήσεις στον πίνακα. Αρχικά, τέθηκε κατώτατο όριο (10 περίοδοι-15 λεπτά για την

Αθήνα και 4 περίοδοι-20 λεπτά για το Μάντσεστερ), ώστε να απαλειφθούν συμβάντα κορεσμού πολύ μικρής διάρκειας. Αυτό επετεύχθη, αντικαθιστώντας, όπου είχαμε λιγότερες από 10 ή 4 διαδοχικές μονάδες, αντίστοιχα για κάθε βάση δεδομένων, με μηδενικά. Επίσης, τα λίγα διαδοχικά μηδενικά ανάμεσα σε δύο μονάδες, δηλαδή τα μικρά χρονικά διαστήματα ανάμεσα σε δύο διαστήματα κορεσμού, αντικαταστάθηκαν με μονάδες και θεωρήθηκαν κι αυτά κορεσμός (5 περίοδοι-7,5 λεπτά για την Αθήνα και 1 περίοδος-5 λεπτά για το Μάντσεστερ), διότι είναι εξαιρετικά πιθανό ο μη εντοπισμός τους εξ αρχής να οφείλεται σε λάθος των μετρήσεων.

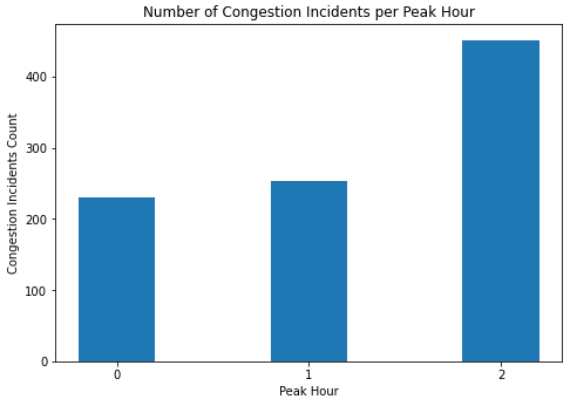
Από τις διαδοχικές μονάδες, εντοπίστηκαν οι θέσεις (indexes) των ακραίων κάθε ακολουθίας, και υπολογίζοντας τη διαφορά τους, προέκυψε η διάρκεια κάθε συμβάντος. Τελικά, κατασκευάστηκε ένας νέος πίνακας, ο οποίος περιλαμβάνει τη διάρκεια του συμβάντος και το φωρατή από τον οποίο έγινε η καταγραφή. Από τον πίνακα αυτό, απορρίφθηκαν συμβάντα μεγάλης διάρκειας (μεγαλύτερη από τρεις ώρες και για τις δύο πόλεις), ως ακραίες μη συνηθισμένες περιπτώσεις (outliers). Τελικά, από τα δεδομένα της Αθήνας προέκυψαν 926 συμβάντα κορεσμού ενώ από εκείνα του Μάντσεστερ 21370.

4.2.3. Επιλογή ανεξάρτητων επεξηγηματικών μεταβλητών

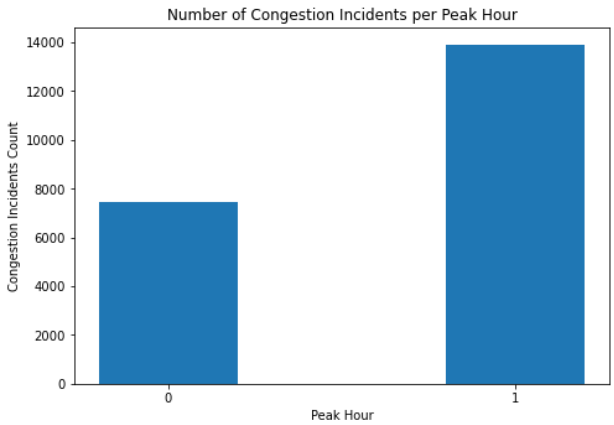
Η ανάπτυξη ενός μοντέλου διάρκειας, προϋποθέτει την ενσωμάτωση κάποιων ανεξάρτητων μεταβλητών, οι οποίες θα συνδράμουν στην καλύτερη προσαρμογή του στο δείγμα. Οι μεταβλητές που εκτιμήθηκε πως θα επηρεάσουν προς το καλύτερο την ανάλυση είναι: i) η ώρα της ημέρας που εμφανίστηκε ο κορεσμός και το αν είναι εντός ή εκτός των ωρών αιχμής, ii) ο αριθμός των λωρίδων στη συγκεκριμένη διατομή και iii) ο φόρτος κατά την έναρξη του κορεσμού:

i. <u>Ώρες αιχμής</u>

Ανάλογα με την ώρα έναρξης του εκάστοτε συμβάντος, έγινε κατηγοριοποίηση σε ώρες αιχμής. Παρατηρώντας τη χρονοσειρά του φόρτου σε Αθήνα και Μάντσεστερ, ορίστηκαν ως διαστήματα αιχμής, οι ώρες της ημέρας κατά τις οποίες ήταν εμφανής μία κορύφωση στις τιμές του. Συγκεκριμένα, ορίστηκαν δύο διαστήματα αιχμής για την Αθήνα, η πρωινή αιχμή 7:00-11:00 και η απογευματινή 14:00-19:00, ενώ στο Μάντσεστερ ένα ενιαίο διάστημα 10:00-18:00. Στο Διάγραμμα 4.15και Διάγραμμα 4.16, φαίνεται πως τα περισσότερα συμβάντα κορεσμού λαμβάνουν χώρα κατά τη διάρκεια των αιχμών.



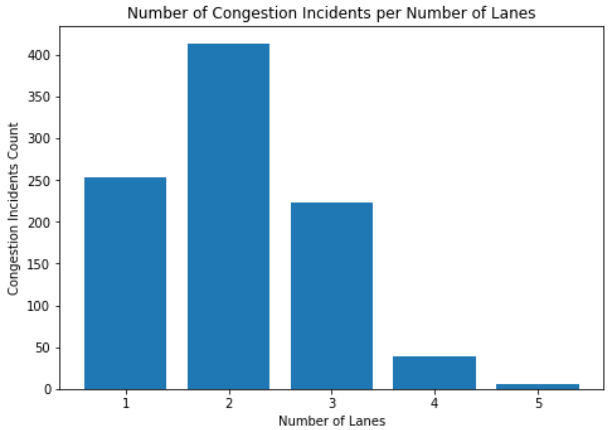
Διάγραμμα 4.15. Συμβάντα κορεσμού ανά διάστημα αιχμής στην Αθήνα (0 εκτός αιχμής, 1 πρωινή αιχμή 7:00-11:00, 2 απογευματινή αιχμή 14:00-19:00)



Διάγραμμα 4.16. Συμβάντα κορεσμό ανά διάστημα αιχμής στο Μάντσεστερ (0 εκτός αιχμής, 1 εντός αιχμής 10:00-18:00)

ii. <u>Αριθμός λωρίδων</u>

Ο αριθμός των λωρίδων θεωρήθηκε πως είναι άλλος ένας σημαντικός παράγοντας επιρροής της διάρκειας του κορεσμού. Σε μία διατομή μίας ή δύο λωρίδων κυκλοφορίας είναι πιο πιθανή η εμφάνιση κορεσμού από μια διατομή τεσσάρων λωρίδων μιας αρτηρίας, όπως άλλωστε διακρίνεται και στο Διάγραμμα 4.17. Η ανεξάρτητη μεταβλητή του αριθμού των λωρίδων, χρησιμοποιήθηκε μόνο στην ανάπτυξη του μοντέλου του οδικού δικτύου της Αθήνας, διότι δεν υπήρχαν πληροφορίες για τις διατομές των θέσεων μέτρησης στο Μάντσεστερ.



Διάγραμμα 4.17. Συμβάντα κορεσμού ανά αριθμό λωρίδων διατομής στην Αθήνα

iii. <u>Φόρτος κατά την έναρξη</u>

Ο φόρτος εισόδου της κυκλοφορίας σε κορεσμένη ροή εισήχθη στο υπό ανάπτυξη μοντέλο ως η τελευταία ανεξάρτητη μεταβλητή. Στη βάση δεδομένων της Αθήνας, ο φόρτος διαιρέθηκε με τον αριθμό των λωρίδων της αντίστοιχης διατομής.

Η τελικές μορφές των πινάκων που χρησιμοποιήθηκαν στη μέθοδο Survival Analysis είναι:

Πίνακας 4.3. Πίνακας διαρκειών συμβάντων κορεσμού Αθήνας με ανεξάρτητες μεταβλητές

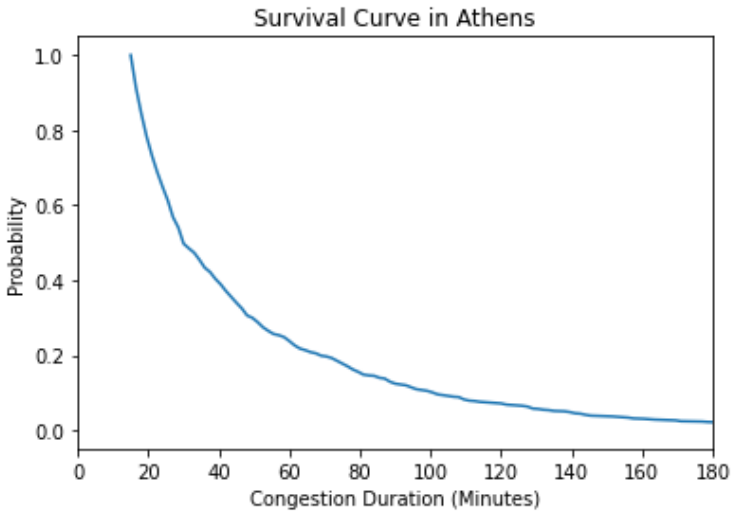
	Duration	Point	Number of Lanes	Traffic	Peak Hour
0	24	MS106	3	1471.334	2
1	34.5	MS106	3	1406.585	2
2	87	MS106	3	1487.998	2
3	145.5	MS106	3	1645.46	2
4	18	MS107	1	405.9949	1
5	72	MS107	1	549.9058	2
6	31.5	MS108	4	1140.876	1
7	16.5	MS108	4	1280.403	1
8	19.5	MS108	4	1108.76	0
9	46.5	MS108	4	1232.664	2

Πίνακας 4.4. Πίνακας διαρκειών συμβάντων κορεσμού Μάντσεστερ με ανεξάρτητες μεταβλητές

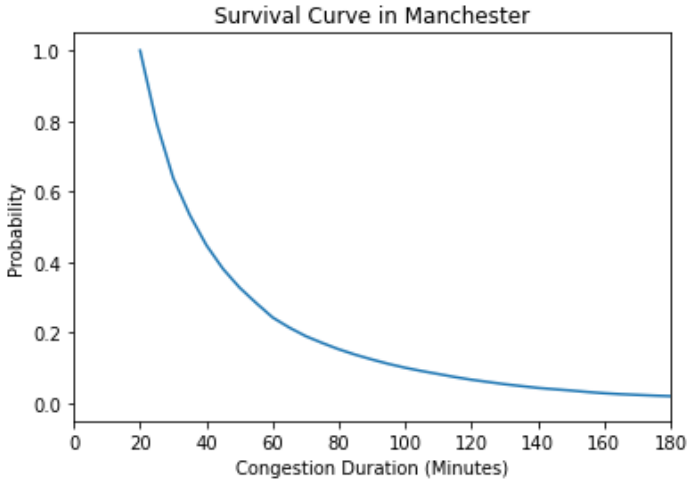
	Duration	***Point***	***Peak Hour***	***Traffic***
0	20	1304	0	93
1	20	1304	1	69
2	30	1304	0	65
3	40	1304	0	87
4	20	1304	0	94
5	20	1304	1	72
6	20	1304	0	73
7	90	1304	1	88
8	25	1304	0	93

4.3. Μοντέλα Survival Analysis

Στην παρούσα διπλωματική, εφαρμόστηκαν τρεις μέθοδοι Survival Analysis: το Μοντέλο Κινδύνου Cox, η Μέθοδος Κατανομής Weibull και η Μέθοδος Random Survival Forest, οι οποίες αναλύθηκαν θεωρητικά στην ενότητα 3.2.1. Οι καμπύλες Survival των πραγματικών μετρήσεων, με βάση την επεξεργασία που πραγματοποιήθηκε για Αθήνα και Μάντσεστερ αντίστοιχα, φαίνονται στα διαγράμματα Διάγραμμα 4.18 και Διάγραμμα 4.19. Παρατηρώντας τα, προκύπτει πως, παρά τις προφανείς ομοιότητες που παρουσιάζουν, τα συμβάντα στο Μάντσεστερ έχουν γενικά μεγαλύτερη διάρκεια. Ενδεικτικά, περίπου στο 50% των συμβάντων κορεσμού αντιστοιχεί διάρκεια τουλάχιστον 30 λεπτά στην Αθήνα και 38 λεπτά στο Μάντσεστερ.



Διάγραμμα 4.18. Καμπύλη Survival πραγματικών συμβάντων στην Αθήνα



Διάγραμμα 4.19. Καμπύλη Survival πραγματικών συμβάντων στο Μάντσεστερ

4.3.1. Μοντέλο Κινδύνου Cox

Η μέθοδος του μοντέλου κινδύνου Cox, είναι η πρώτη από τις μεθόδους που εφαρμόστηκαν. Η ανάπτυξη του μοντέλου έγινε μέσω του πακέτου “Lifelines” της Python. Τα αποτελέσματα ανά πόλη ήταν τα εξής:

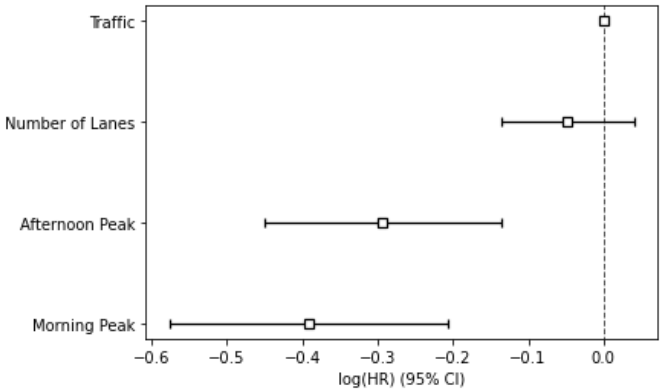
i. Αθήνα

	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	cmp to	z	p	-log2(p)
Number of Lanes	-0.05	0.95	0.05	-0.14	0.04	0.87	1.04	0.00	-1.05	0.29	1.77
Traffic	-0.00	1.00	0.00	-0.00	-0.00	1.00	1.00	0.00	-4.35	<0.005	16.17
Morning Peak	-0.39	0.68	0.09	-0.58	-0.21	0.56	0.81	0.00	-4.16	<0.005	14.94
Afternoon Peak	-0.29	0.75	0.08	-0.45	-0.14	0.64	0.87	0.00	-3.65	<0.005	11.89

Concordance	0.58
Partial AIC	10762.78
log-likelihood ratio test	52.67 on 4 df
-log2(p) of ll-ratio test	33.22

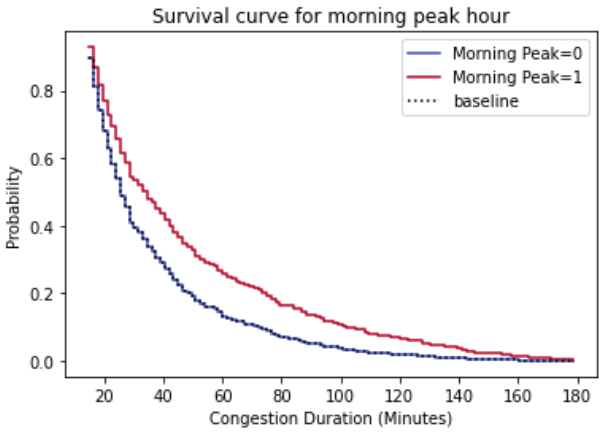
Πίνακας 4.5. Αποτελέσματα εφαρμογής μοντέλου κινδύνου Cox στην Αθήνα

Ο δείκτης Concordance είναι ίσος με 0,58.

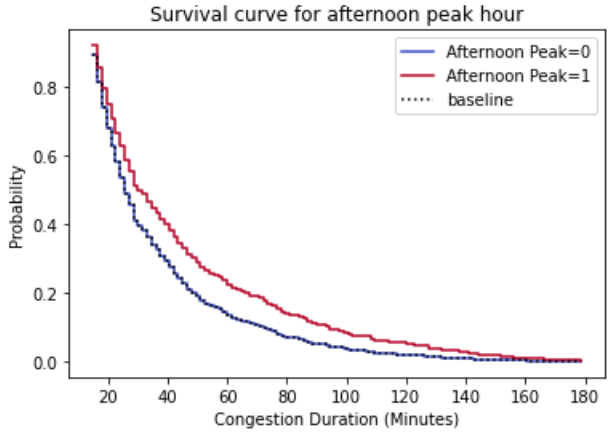


Διάγραμμα 4.20. Επιρροή ανεξάρτητων μεταβλητών στο μοντέλο κινδύνου Cox στην Αθήνα

Όπως φαίνεται στο παραπάνω διάγραμμα, μεγαλύτερη επιρροή στο μοντέλο έχουν οι ώρες αιχμής και ο αριθμός των λωρίδων της διατομής. Ωστόσο η p-value του αριθμού των λωρίδων είναι μεγαλύτερη από 0,05 που είναι το ανώτατο όριο ώστε να θεωρηθεί στατιστικά σημαντική μεταβλητή. Στα παρακάτω διαγράμματα Διάγραμμα 4.21 και Διάγραμμα 4.22 είναι ευδιάκριτη η τάση αύξησης της διάρκειας του κορεσμού, όσο τα οχήματα κινούνται κατά τις ώρες αιχμής.



Διάγραμμα 4.21. Καμπύλη Survival μοντέλου κινδύνου Cox για ώρες εντός κι εκτός πρωινής αιχμής στην Αθήνα



Διάγραμμα 4.22. Καμπύλη Survival μοντέλου κινδύνου Cox για ώρες εντός κι εκτός απογευματινής αιχμής στην Αθήνα

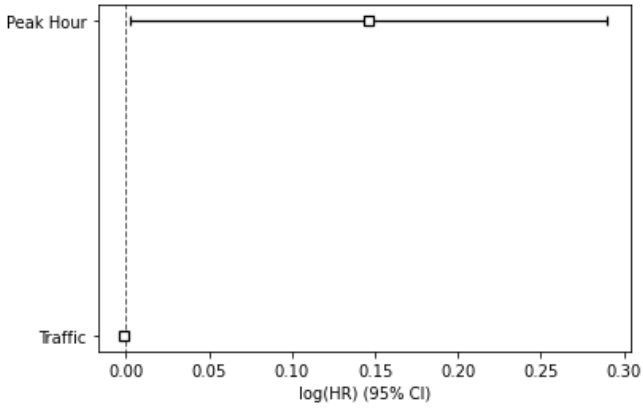
ii. Μάντσεστερ

	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	cmp to	z	p	-log2(p)
Peak Hour	0.15	1.16	0.07	0.00	0.29	1.00	1.34	0.00	2.00	0.05	4.45
Traffic	-0.00	1.00	0.00	-0.00	-0.00	1.00	1.00	0.00	-4.20	<0.005	15.20

Concordance	0.56
Partial AIC	382950.03
log-likelihood ratio test	432.60 on 2 df
-log2(p) of ll-ratio test	312.06

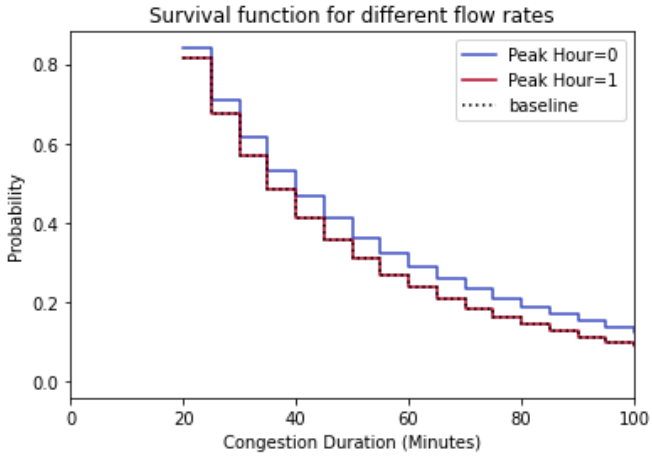
Πίνακας 4.6. Αποτελέσματα εφαρμογής μοντέλου κινδύνου Cox στο Μάντσεστερ

Ο δείκτης Concordance είναι ίσος με 0,56.



Διάγραμμα 4.23. Επιρροή ανεξάρτητων μεταβλητών στο μοντέλο κινδύνου Cox στο Μάντσεστερ

Παρατηρείται και στο Μάντσεστερ μεγάλη επιρροή της ώρας αιχμής στη διάρκεια του κορεσμού, ενώ οριακά θεωρείται στατιστικά σημαντική μεταβλητή(p-value=0,05). Ο φόρτος είναι επίσης σημαντική μεταβλητή, η οποία όμως έχει μικρότερη επιρροή.



Διάγραμμα 4.24. Καμπύλη Survival μοντέλου κινδύνου Cox για ώρες εντός κι εκτός αιχμής στο Μάντσεστερ

Στο Μάντσεστερ παρατηρείται το αντίθετο από την Αθήνα. Δηλαδή, παρότι τα συμβάντα κορεσμού είναι περισσότερα κατά το ενιαίο διάστημα αιχμής, οι διάρκειες ενδέχεται να είναι μεγαλύτερες εκτός αιχμής.

4.3.2. Μέθοδος Κατανομής Weibull

Η μέθοδος της συνεχούς κατανομής Weibull είναι πλήρως παραμετρική. Από το πακέτο "Lifelines" της Python, ανακτήθηκαν οι απαραίτητες συναρτήσεις, και τα αποτελέσματα τελικά διαμορφώθηκαν ως εξής:

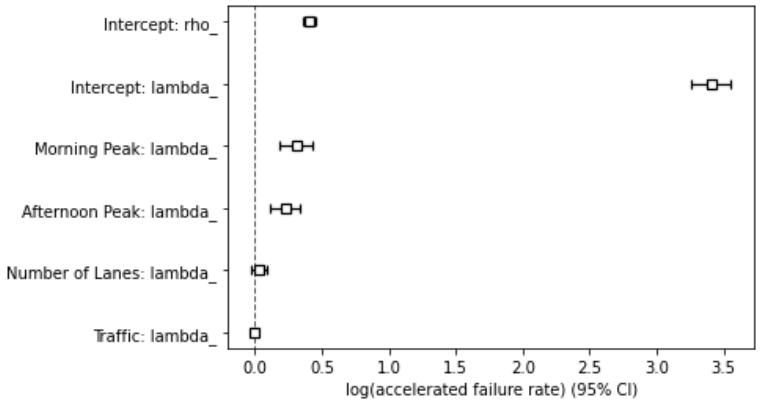
i. <u>Αθήνα</u>

		coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	cmp to	z	p	-log2(p)
lambda_	Afternoon Peak	0.22	1.25	0.06	0.12	0.33	1.12	1.40	0.00	4.03	<0.005	14.12
	Morning Peak	0.31	1.36	0.06	0.19	0.43	1.21	1.54	0.00	4.92	<0.005	20.12
	Number of Lanes	0.03	1.03	0.03	-0.02	0.09	0.98	1.09	0.00	1.13	0.26	1.96
	Traffic	0.00	1.00	0.00	0.00	0.00	1.00	1.00	0.00	4.35	<0.005	16.17
	Intercept	3.40	30.09	0.07	3.26	3.55	26.05	34.76	0.00	46.30	<0.005	inf
rho_	Intercept	0.40	1.50	0.02	0.36	0.45	1.43	1.57	0.00	17.07	<0.005	214.62

Concordance	0.58
AIC	8583.83
log-likelihood ratio test	62.49 on 4 df
-log2(p) of ll-ratio test	40.06

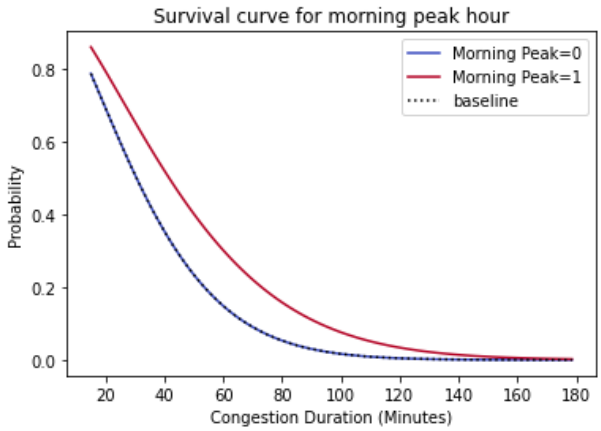
Πίνακας 4.7. Αποτελέσματα εφαρμογής μεθόδου κατανομής Weibull στην Αθήνα

Ο δείκτης Concordance είναι ίσος με 0,58.

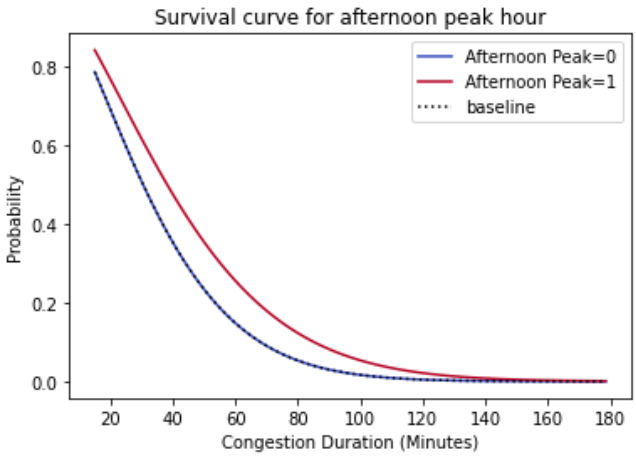


Διάγραμμα 4.25. Επιρροή ανεξάρτητων μεταβλητών στη μέθοδο κατανομής Weibull στην Αθήνα

Και στη μέθοδο αυτή, ο αριθμός των λωρίδων ανά διατομή δεν είναι στατιστικά σημαντική μεταβλητή, καθώς κι ο φόρτος δεν έχει καμία επιρροή στη διάρκεια του κορεσμού. Οι ώρες αιχμής έχουν την ίδια επιρροή με αυτή που είχαν στο μοντέλο κινδύνου Cox.



Διάγραμμα 4.26. Καμπύλη survival κατανομής Weibull για ώρες εντός κι εκτός πρωινής αιχμής στην Αθήνα



Διάγραμμα 4.27. Καμπύλη survival κατανομής Weibull για ώρες εντός κι εκτός απογευματινής αιχμής στην Αθήνα

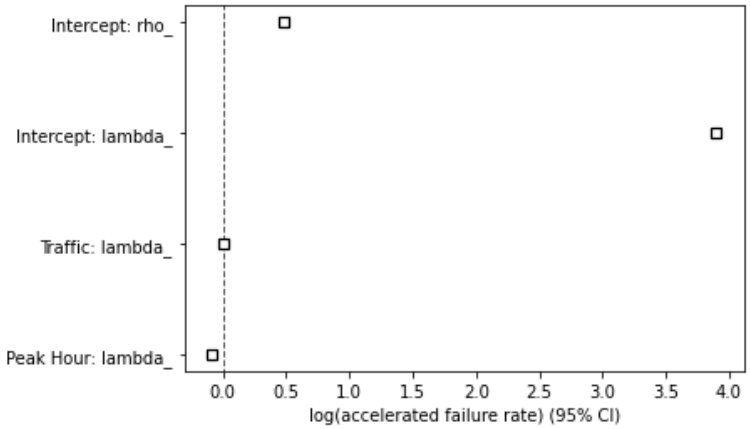
ii. Μάντσεστερ

		coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	cmp to	z	p	-log2(p)
lambda_	Peak Hour	-0.09	0.92	0.01	-0.10	-0.07	0.90	0.93	0.00	-9.59	<0.005	70.00
	Traffic	0.00	1.00	0.00	0.00	0.00	1.00	1.00	0.00	19.15	<0.005	269.18
	Intercept	3.90	49.42	0.01	3.88	3.92	48.33	50.53	0.00	344.02	<0.005	inf
rho_	Intercept	0.48	1.62	0.00	0.47	0.49	1.61	1.64	0.00	98.69	<0.005	inf

Concordance	0.56
AIC	201272.40
log-likelihood ratio test	511.63 on 2 df
-log2(p) of ll-ratio test	369.07

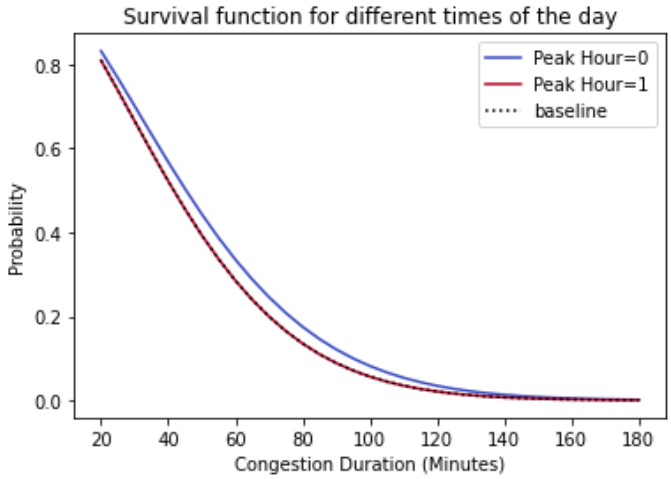
Πίνακας 4.8. Αποτελέσματα εφαρμογής μεθόδου κατανομής Weibull στο Μάντσεστερ

Ο δείκτης Concordance είναι ίσος με 0,56.



Διάγραμμα 4.28. Επιρροή ανεξάρτητων μεταβλητών στη μέθοδο κατανομής Weibull στο Μάντσεστερ

Παρά το γεγονός ότι από το Διάγραμμα 4.28 δεν διακρίνεται εύκολα η επιρροή των ανεξάρτητων μεταβλητών, από τα αποτελέσματα του μοντέλου προκύπτει πως η ώρα αιχμής είναι αυτή που το επηρεάζει περισσότερο. Η συμπεριφορά της μεταβλητής της ώρας αιχμής είναι η ίδια με αυτή που ήταν στο μοντέλο Cox (μεγαλύτερες διάρκειες εκτός αιχμής).

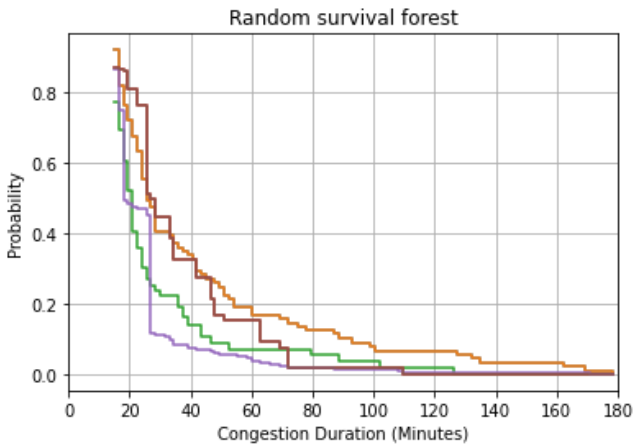


Διάγραμμα 4.29. Καμπύλη survival μεθόδου κατανομής Weibull για ώρες εντός κι εκτός αιχμής στο Μάντσεστερ

4.3.3. Μέθοδος Random Survival Forest

Για ανάπτυξη μοντέλου με τη βοήθεια της μεθόδου Random Survival Forest, χρησιμοποιήθηκαν πακέτα μηχανικής μάθησης «SciKit» στην Python. Η Random Survival Forest δε δίνει κάποια περαιτέρω πληροφορία για την επιρροή των μεταβλητών στα αποτελέσματα.

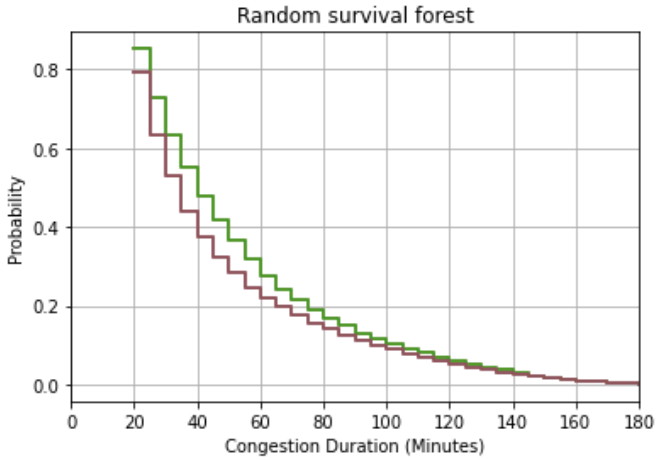
i. <u>Αθήνα</u>



Διάγραμμα 4.30. Καμπύλες Random Survival Forest για μερικά τυχαία σημεία στην Αθήνα

Ο δείκτης Concordance για τη μέθοδο Random Survival Forest στην Αθήνα, προέκυψε ίσος με 0,539. Στο Διάγραμμα 4.30 απεικονίζονται οι καμπύλες Survival για κάποια σημεία που επιλέχθηκαν τυχαία.

ii. <u>Μάντσεστερ</u>



Διάγραμμα 4.31. Καμπύλες Random Survival Forest για μερικά τυχαία σημεία στο Μάντσεστερ

Ο δείκτης Concordance για τη μέθοδο Random Survival Forest στο Μάντσεστερ, προέκυψε ίσος με 0,537. Στο Διάγραμμα 4.31 απεικονίζονται οι καμπύλες Survival για κάποια σημεία που επιλέχθηκαν τυχαία.

Τελικά τα αποτελέσματα εφαρμογής κάθε μοντέλου σε κάθε πόλη συνοψίζονται στον **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.**.

Πίνακας 4.9. Δείκτης Concordance ανά μέθοδο και πόλη

Concordance index	**Αθήνα**	**Μάντσεστερ**
Μοντέλο Κινδύνου Cox	0,58	0,56
Μέθοδος Κατανομής Weibull	0,58	0,56
Random Survival Forest	0,539	0,537

5. ΣΥΜΠΕΡΑΣΜΑΤΑ

5.1. Εισαγωγή

Ο κορεσμός της κυκλοφορίας σε ένα αστικό οδικό δίκτυο, αποτελεί ένα από τα βασικότερα προβλήματα της σύγχρονης κοινωνίας και της οικονομίας. Για τη διερεύνηση της διάρκειάς του, αρχικά εφαρμόστηκαν δύο μέθοδοι εντοπισμού κορεσμού, μία αναλυτική και μία «data-driven», οι οποίες συγκρίθηκαν ώστε να επιλεχθεί η καταλληλότερη με βάση τα χαρακτηριστικά των δεδομένων. Έπειτα, σκοπός της εργασίας ήταν η διερεύνηση της επιρροής βασικών κυκλοφοριακών χαρακτηριστικών στη διάρκεια κορεσμού με τη χρήση μοντέλων Ανάλυσης Επιβίωσης. Η παραπάνω προσέγγιση μπορεί αποτελέσει βασικό κλειδί στην διαχείριση του φαινομένου από τους ειδικούς και την Πολιτεία.

5.2. Βασικά συμπεράσματα

Από τα αποτελέσματα της εφαρμογής των μοντέλων που αναπτύχθηκαν προέκυψαν μερικά ενδιαφέροντα συμπεράσματα για τη συμπεριφορά του κορεσμού και της διάρκειάς του. Αρχικά, οι δύο μέθοδοι εντοπισμού κορεσμού που εφαρμόστηκαν, βασίζονταν σε διαφορετικά κυκλοφοριακά μεγέθη ώστε να κατατάξουν μία μέτρηση σε στιγμή που υπάρχει ή όχι κορεσμός. Πιο συγκεκριμένα, στη μέθοδο «Spillover», με βάση την εξίσωση της βιβλιογραφίας, ελέγχεται αν η καταγεγραμμένη κατάληψη υπερβαίνει το όριο που τίθεται, για το φόρτο που έχει μετρηθεί. Το γεγονός, ωστόσο, ότι η συγκεκριμένη μέθοδος λαμβάνει υπόψιν και χρόνους σηματοδότησης, περιορίζει το πεδίο εφαρμογής της. Γι' αυτό το λόγο, εφαρμόστηκε και η μέθοδος «χαμηλών ταχυτήτων», η οποία ακολουθεί μια πιο «data-driven» προσέγγιση. Ειδικότερα, ανά σημείο μέτρησης, από όλες τις καταγεγραμμένες ταχύτητες, προκύπτει ένα κατώτατο όριο ταχύτητας το οποίο αν δεν ξεπεραστεί, η χρονική στιγμή χαρακτηρίζεται ως κορεσμένη. Μετά από δοκιμές, η δεύτερη μέθοδος είχε τα καλύτερα αποτελέσματα στα μοντέλα διάρκειας που εφαρμόστηκαν.

Ο αριθμός συμβάντων που προέκυψε με την εφαρμογή της μεθόδου εντοπισμού χαμηλών ταχυτήτων στην Αθήνα ήταν πολύ μεγαλύτερος αναλογικά, από αυτόν στο Μάντσεστερ. Σε δύο ημέρες, στην Αθήνα, εντοπίστηκαν 926 συμβάντα κορεσμού ενώ στο Μάντσεστερ, για 13 μήνες μετρήσεων, εντοπίστηκαν 21370 (τα οποία αντιστοιχούν σε περίπου 108 συμβάντα το διήμερο). Μια ερμηνεία αυτής της μεγάλης απόκλισης, είναι η νοοτροπία των περισσότερων Ελλήνων να μη χρησιμοποιούν τα Μέσα Μαζικής Μεταφοράς για τις μετακινήσεις τους αλλά το προσωπικό τους όχημα, επιβαρύνοντας ακόμα περισσότερο το οδικό δίκτυο. Επίσης, όπως αναφέρθηκε νωρίτερα η Αθήνα είναι μια αρκετά μεγαλύτερη πόλη, με σημαντικά μεγαλύτερο αριθμό κατοίκων, οπότε σίγουρα το παραπάνω συμπέρασμα ήταν και αναμενόμενο ως ένα βαθμό. Να σημειωθεί επίσης, ότι οι περισσότερες θέσεις μέτρησης στο Μάντσεστερ βρίσκονται σε αυτοκινητοδρόμους, οπότε είναι λογικός ο μικρότερος αριθμός συμβάντων. Από την άλλη, στο Μάντσεστερ οι διάρκειες ήταν ελάχιστα μεγαλύτερες. Αυτό εξηγείται, μιας κι ενώ η Αθήνα είναι η μεγαλύτερη από τις δύο πόλεις και το αστικό της κέντρο έχει το διπλάσιο πληθυσμό, συνεπώς μεγαλύτερη παραγωγή μετακινήσεων, η ενιαία αιχμή που παρατηρήθηκε στο Μάντσεστερ διευκολύνει τον εντοπισμό κορεσμών μεγάλης διάρκειας. Ακόμα, στο Μάντσεστερ, δεν παρατηρούνται πολύ μικρές ταχύτητες, δηλαδή ακινητοποιημένη κυκλοφορία, άρα είναι πιθανό η ροή να κινείται για μεγάλο διάστημα στο όριο συμφόρησης που τίθεται.

Όσον αφορά τα μοντέλα διάρκειας, τα αποτελέσματα ανά μοντέλο ήταν αρκετά κοντά και για τις δύο πόλεις. Στην Αθήνα τα αποτελέσματα ήταν λίγο καλύτερα από αυτά του Μάντσεστερ. Το μοντέλο κινδύνου Cox και η μέθοδος της κατανομής Weibull είχαν ίσους δείκτες Concordance, ανά πόλη, ενώ η μέθοδος Random Survival Forest είχε σταθερά λίγο χειρότερη προσαρμογή. Σε όλα όμως, τα μοντέλα, ο συντελεστής Concordance προέκυψε σχετικά χαμηλός, αλλά αποδεκτός, αν ληφθεί υπόψη ότι οι ανεξάρτητες μεταβλητές ήταν λίγες κι ότι πρόκειται για εφαρμογή μοντέλου σε ένα μεγάλο αστικό δίκτυο, το οποίο περιλαμβάνει πολλές οδούς διαφορετικών χαρακτηριστικών η καθεμία.

Πιο σαφής έγινε κι η εικόνα της επιρροής κάθε ανεξάρτητης επεξηγηματικής μεταβλητής στη διάρκεια του κορεσμού ενός αστικού οδικού δικτύου.

i. Ώρα Αιχμής

Η μεταβλητή της Ώρας Αιχμής, ήταν εκείνη που επηρέασε σημαντικά όλα τα μοντέλα. Εκτός του ότι είχε σε κάθε μοντέλο το μεγαλύτερο συντελεστή, στα διαγράμματα της Αθήνας απεικονιζόταν αυτό που ζουν οι κάτοικοί της καθημερινά, οι οποίοι κατά τις ώρες αιχμής μπορεί «κολλήσουν» στην κυκλοφορία για μεγάλο χρονικό διάστημα. Τα διαγράμματα για τις ώρες αιχμής στο Μάντσεστερ, από την άλλη, έδειξαν ότι, παρόλο που η πλειοψηφία των συμβάντων παρατηρούνται κατά τις ώρες αιχμής, η διάρκεια των συμβάντων εκτός αιχμής είναι γενικά μεγαλύτερη.

ii. Αριθμός Λωρίδων ανά Διατομή Θέσης Μέτρησης

Η μεταβλητή του Αριθμού Λωρίδων, παρότι επηρέαζε την τιμή της διάρκειας του κορεσμού, όπως αποδείχθηκε δεν είναι στατιστικά σημαντική μεταβλητή, σε επίπεδο εμπιστοσύνης 95%, λόγω των μεγάλων τιμών που έπαιρνε η αντίστοιχη p-value. Παρόλα αυτά, όπως ήταν αναμενόμενο, αυξημένος αριθμός λωρίδων, οδηγεί σε μικρότερη διάρκεια κορεσμού.

iii. Κυκλοφοριακός Φόρτος κατά την Έναρξη του Κορεσμού

Τέλος, ο Κυκλοφοριακός Φόρτος, αν και είναι μια στατιστικά σημαντική μεταβλητή, έχει μικρότερη επιρροή στην εκτίμηση της διάρκειας. Αυτό μπορεί να εξηγηθεί από την άποψη ότι κάθε οδός χρειάζεται διαφορετικό αριθμό οχημάτων, ώστε η κυκλοφορία σε αυτή να επιβραδυνθεί και τελικά να κορεστεί (διαφορετική κυκλοφοριακή ικανότητα).

5.3. Προτάσεις για περαιτέρω έρευνα

Ο κορεσμός της κυκλοφορίας είναι ένα πολύ συχνό φαινόμενο στα αστικά οδικά δίκτυα. Η διάρκειά του αποτελεί ένα μέγεθος που διερευνάται συνεχώς από τους ερευνητές. Η προσπάθεια σύνδεσης της διάρκειας με διάφορα χαρακτηριστικά, όπως συνέβη και στην παρούσα διπλωματική εργασία, θα μπορούσε να συμβάλει στη βαθύτερη κατανόηση των αιτιών του κορεσμού αλλά και των προβλημάτων που προκαλεί στην κυκλοφορία και ακολούθως στη ζωή των οδηγών. Με βάση την έρευνα που έγινε, παρουσιάζονται μερικές προτάσεις για την περαιτέρω ανάλυση του ζητήματος.

Παρά την προσπάθεια εφαρμογής της μεθόδου «Spillover» στα δεδομένα της Αθήνας, λόγω των μόλις δύο ημερών μετρήσεων και την ύπαρξη σημείων μέτρησης σε αυτοκινητοδρόμους, δε χρησιμοποιήθηκε στην Ανάλυση Επιβίωσης. Ωστόσο, θα ήταν εξαιρετικά ενδιαφέρον το

να εφαρμοστεί σε δεδομένα ενός ή παραπάνω μηνών και μόνο σε σηματοδοτημένες οδούς, ώστε να μπορέσει να γίνει καλύτερη σύγκριση των δύο μεθόδων εντοπισμού κορεσμού.

Επιπροσθέτως, προτείνεται η εισαγωγή περισσότερων ανεξάρτητων μεταβλητών στα μοντέλα διάρκειας. Ο καιρός, όπως αναφέρθηκε παραπάνω, και ιδιαίτερα τα ακραία καιρικά φαινόμενα, αποτελούν μία από τις κύριες αιτίες εμφάνισης του μη επαναλαμβανόμενου κορεσμού. Χρήσιμη θα ήταν η εισαγωγή δεδομένων που θα περιγράφουν φαινόμενα όπως η βροχή, το χιόνι και το χαλάζι, δηλαδή όλα τα είδη των κατακρημνίσεων, ώστε να διερευνηθεί η επιρροή τους στη διάρκεια του κορεσμού. Μία ακόμα ενδιαφέρουσα προσέγγιση θα ήταν κι η μεταβλητή της ημέρας της εβδομάδας ή σε ακόμα μεγαλύτερες βάσεις δεδομένων, ο μήνας του έτους. Η κυκλοφορία διαφοροποιείται από καθημερινές σε σαββατοκύριακα (μετακινήσεις εργασίας κυρίως τις καθημερινές και αναψυχής τα σαββατοκύριακα), αλλά και από μήνα σε μήνα (τους καλοκαιρινούς μήνες η κίνηση οχημάτων είναι μειωμένη στις πόλεις λόγω διακοπών των κατοίκων).

Τέλος, η εφαρμογή ενός μοντέλου διάρκειας σε μια μικρότερη περιοχή μελέτης, όπως μία οδική αρτηρία, ή επιμέρους περιοχή μίας πόλης θα μπορούσε να έχει καλύτερα αποτελέσματα στην εκτίμηση της καμπύλης επιβίωσης.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Anbaroglu, B., Heydecker, B., Cheng, T., 2014. Spatio-temporal clustering for non-recurrent traffic congestion detection on urban road networks. Transp. Res. Part C Emerg. Technol. 48, 47–65. https://doi.org/10.1016/j.trc.2014.08.002

Chakraborty, P., Adu-Gyamfi, Y.O., Poddar, S., Ahsani, V., Sharma, A., Sarkar, S., 2018. Traffic Congestion Detection from Camera Images using Deep Convolution Neural Networks. Transp. Res. Rec. J. Transp. Res. Board 2672, 222–231. https://doi.org/10.1177/0361198118777631

Chen, M., Yu, X., Liu, Y., 2018. PCNN: Deep Convolutional Networks for Short-Term Traffic Congestion Prediction. IEEE Trans. Intell. Transp. Syst. 19, 3550–3559. https://doi.org/10.1109/TITS.2018.2835523

Cox, D.R., 1972. Regression Models and Life-Tables. J. R. Stat. Soc. Ser. B Methodol. 34, 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

D'Andrea, E., Marcelloni, F., 2017. Detection of traffic congestion and incidents from GPS trace analysis. Expert Syst. Appl. 73, 43–56. https://doi.org/10.1016/j.eswa.2016.12.018

Datema, F.R., Moya, A., Krause, P., Bäck, T., Willmes, L., Langeveld, T., Baatenburg De Jong, R.J., Blom, H.M., 2012. Novel head and neck cancer survival analysis approach: Random survival forests versus cox proportional hazards regression. Head Neck 34, 50–58. https://doi.org/10.1002/hed.21698

Fafoutellis, P., Mantouka, E.G., Vlahogianni, E.I., 2020. Eco-Driving and Its Impacts on Fuel Efficiency: An Overview of Technologies and Data-Driven Methods. Sustainability 13, 226. https://doi.org/10.3390/su13010226

Geroliminis, N., Skabardonis, A., 2006. Real Time Vehicle Reidentification and Performance Measures on Signalized Arterials, in: 2006 IEEE Intelligent Transportation Systems Conference. Presented at the 2006 IEEE Intelligent Transportation Systems Conference, IEEE, Toronto, ON, Canada, pp. 188–193. https://doi.org/10.1109/ITSC.2006.1706740

Hu, B., Lin, D., Sun, Q., Dong, X., 2019. Research on road traffic flow status based on survival analysis. J. Phys. Conf. Ser. 1187, 052056. https://doi.org/10.1088/1742-6596/1187/5/052056

Huajing, N., Yu, Y., Bai, L., 2022. Survival analysis of the unsafe behaviors leading to urban expressway crashes. PLOS ONE 17, e0267559. https://doi.org/10.1371/journal.pone.0267559

Ibay, F.J.G., Department of Mathematics, Lagro High School; ibay., De Guzman, K.J.N., Department of Mathematics and Statistics, College of Science, Polytechnic University of the Philippines - Manila;, Escobar, S.G., Department of Mathematics, Kaypian National High School, Bulacan;, Nobles, B.G., Department of Mathematics and Statistics, College of Science, Polytechnic University of the Philippines - Manila;, Oliquino, A.B., Department of Mathematics and Statistics, College of Science, Polytechnic University of the Philippines - Manila;, 2019. Survival Analysis of Persons Involved in Road Accidents in National Capital Region. Indian J. Sci. Technol. 12, 1–11. https://doi.org/10.17485/ijst/2019/v12i12/142498

Ishwaran, H., Kogalur, U.B., Blackstone, E.H., Lauer, M.S., 2008. Random survival forests. Ann. Appl. Stat. 2. https://doi.org/10.1214/08-AOAS169

Kalair, K., Connaughton, C., 2021. Dynamic and Interpretable Hazard-Based Models of Traffic Incident Durations. Front. Future Transp. 2, 669015. https://doi.org/10.3389/ffutr.2021.669015

Kaplan, E.L., Meier, P., 1958. Nonparametric Estimation from Incomplete Observations. J. Am. Stat. Assoc. 53, 457–481. https://doi.org/10.1080/01621459.1958.10501452

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM 60, 84–90. https://doi.org/10.1145/3065386

Kumar, N., Raubal, M., 2021. Applications of deep learning in congestion detection, prediction and alleviation: A survey. Transp. Res. Part C Emerg. Technol. 133, 103432. https://doi.org/10.1016/j.trc.2021.103432

Lee, J., 2021. Reopening businesses after Hurricane Harvey: evidence from a duration model with spatial effects. Disasters 45, 296–323. https://doi.org/10.1111/disa.12417

Lee, W.-H., Tseng, S.-S., Shieh, J.-L., Chen, H.-H., 2011. Discovering Traffic Bottlenecks in an Urban Network by Spatiotemporal Data Mining on Location-Based Services. IEEE Trans. Intell. Transp. Syst. 12, 1047–1056. https://doi.org/10.1109/TITS.2011.2144586

Li, C., Yue, W., Mao, G., Xu, Z., 2020. Congestion Propagation Based Bottleneck Identification in Urban Road Networks. IEEE Trans. Veh. Technol. 69, 4827–4841. https://doi.org/10.1109/TVT.2020.2973404

Liu, Y., Liu, C., Zheng, Z., 2020. Traffic Congestion and Duration Prediction Model Based on Regression Analysis and Survival Analysis. Open J. Bus. Manag. 08, 943–959. https://doi.org/10.4236/ojbm.2020.82059

Mantouka, E.G., Fafoutellis, P., Vlahogianni, E.I., 2021. Deep survival analysis of searching for on-street parking in urban areas. Transp. Res. Part C Emerg. Technol. 128, 103173. https://doi.org/10.1016/j.trc.2021.103173

McGroarty, J., 2010. Recurring and Non-Recurring Congestion.

Nidhal, A., Ngah, U.K., Ismail, W., 2014. Real time traffic congestion detection system, in: 2014 5th International Conference on Intelligent and Advanced Systems (ICIAS). Presented at the 2014 5th International Conference on Intelligent and Advanced Systems (ICIAS), IEEE, Kuala Lumpur, Malaysia, pp. 1–5. https://doi.org/10.1109/ICIAS.2014.6869538

Oralhan, B., Göktolga, Z.G., 2018. Determination of the Risk Factors That Influence Occurrence Time of Traffic Accidents with Survival Analysis. Iran J Public Health 47.

Pandey, A., Biswas, S., 2022. Assessment of Level of Service on urban roads: a revisit to past studies.

Parmet, Y., Meir, A., Borowsky, A., 2014. Survival analysis: A fresh approach for analyzing response times in driving-related hazard perception paradigms. Transp. Res. Part F Traffic Psychol. Behav. 25, 98–107. https://doi.org/10.1016/j.trf.2014.05.007

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection.

Stopher, P.R., 2004. Reducing road congestion: a reality check. Transp. Policy 11, 117–131. https://doi.org/10.1016/j.tranpol.2003.09.002

Vlahogianni, E.I., Webber, C.L., Geroliminis, N., Skabardonis, A., 2007. Statistical characteristics of transitional queue conditions in signalized arterials. Transp. Res. Part C Emerg. Technol. 15, 392–403. https://doi.org/10.1016/j.trc.2007.07.003

Washington, Karlaftis, Mannering, Anastasopoulos, 2004. Statistical and Econometric Methods for Transportation Data Analysis.

Yong-chuan, Z., Xiao-qing, Z., Li-Ting, Z., Zhen-ting, C., 2011. Traffic Congestion Detection Based On GPS Floating-Car Data. Procedia Eng. 15, 5541–5546. https://doi.org/10.1016/j.proeng.2011.08.1028

Ι. Μ. Φραντζεσκάκης, Ι.Κ. Γκόλιας, Μ.Χ. Πίτσιαβα-Λατινοπούλου, 2009. Κυκλοφοριακή Τεχνική. Εκδόσεις Παπασωτηρίου.

ΠΑΡΑΡΤΗΜΑ

Κώδικας Python

Δεδομένα Αθήνας

Μέθοδος «Spillover»

```
#%%

import pandas as pd


df = pd.read_excel(r'C:\Users\stefa\Desktop\emp\10ο
Εξάμηνο\ΔΙΠΛΩΜΑΤΙΚΗ\dokimi.xlsx')

theseis=pd.read_excel(r'C:\Users\stefa\Desktop\emp\10ο Εξάμηνο\ΔΙΠΛΩΜΑΤΙΚΗ\ΘΕΣΕΙΣ
ΜΕΤΡΗΣΗΣ_17.05.2021.xlsx')


#%%

g = df['Name'].nunique()

h = df['Name'].unique()

#fortos ana 90sec

pd.DataFrame(h).to_csv('Detectors.csv')

df['vol90'] = df['smoothed_all_vol']/40


#%%

import numpy as np

import matplotlib.pyplot as plt

leff=6

r=20

c=90

state=pd.DataFrame()

traffic=pd.DataFrame()

#kataskevi diagrammaton kai dimiourgia pinaka state (1 0...)

for ms in h:

    df2=df[df['Name']==ms]

    uf=np.percentile(df2['smoothed_all_spd'],90)

    osp=leff/uf*df2['vol90']+r/c*100

    df2.plot.scatter(x='vol90',y='smoothed_all_occ',title=ms)

    plt.plot(df2['vol90'],osp,'r')

    plt.ylim(0,100)

    plt.xlabel('Traffic Flow (veh/h)')

    plt.ylabel('Occupancy (%)')
```

```
    df2['osp']=osp
    state[ms]=np.where(df2['smoothed_all_occ']>df2['osp'],1,0)
    traffic[ms]=df2['smoothed_all_vol'].values
```

Μέθοδος χαμηλών ταχυτήτων

```
#%% diagrammata
import numpy as np
import matplotlib.pyplot as plt
leff=6
r=20
c=90
state=pd.DataFrame()
traffic=pd.DataFrame()
df2_stats=pd.DataFrame()
#kataskevi diagrammaton kai dimiourgia pinaka state (1 0...)
for ms in h:
    df2=df[df['Name']==ms]
   # uf=np.percentile(df2['smoothed_all_spd'],90)
   # osp=leff/uf*df2['vol90']+r/c*100
   # plt.plot(df2['vol90'],osp)
   # plt.ylim(0,100)
   # df2['osp']=osp
    df2_stats[ms]=df2['smoothed_all_spd'].describe()
    vel=pd.DataFrame()
    vel['vol']=df2['vol90']
    vel['vel']=df2_stats.loc['25%'][ms]-1.5*df2_stats.loc['std'][ms]
    df2.plot.scatter(x='vol90',y='smoothed_all_spd',title=ms)
    plt.plot(vel['vol'],vel['vel'],'r')
    plt.xlabel('Traffic Flow (veh/h)')
    plt.ylabel('Speed (km/h)')
    state[ms]=np.where(df2['smoothed_all_spd']<(df2_stats.loc['25%'][ms]-
1.5*df2_stats.loc['std'][ms]),1,0)
    traffic[ms]=df2['smoothed_all_vol'].values
```

Εκτίμηση διάρκειας

```
#%% endiamesa 0
for column in h:
    df3=state[column]
```

```
    df4=df3[df3==1]
    l=[]
    l2=[]
    for i in range(len(df4)-1):
        if df4.index[i+1]-df4.index[i]!=1:
            l.append((df4.index[i],df4.index[i+1]))
    for i in range(len(l)):
        if l[i][1]-l[i][0]-1<5:
            l2=l2+list(range(l[i][0]+1,l[i][1]))
    state[column][l2]=1

#%% endiamesoi 1
for ms in h:
    df3=state[ms]
    df4=df3[df3==0]
    l=[]
    l2=[]
    for i in range(len(df4)-1):
        if df4.index[i+1]-df4.index[i]!=1:
            l.append((df4.index[i],df4.index[i+1]))
    for i in range(len(l)):
        if l[i][1]-l[i][0]-1<10:
            l2=l2+list(range(l[i][0]+1,l[i][1]))
    state[ms][l2]=0

#%% ypologismos diarkeiwn kai ektypwsh survival
dur=[]
for ms in h:
    df3=state[ms]
    df4=df3[df3==0]
    for i in range(len(df4)-1):
        if df4.index[i+1]-df4.index[i]!=1:
            dur.append(((df4.index[i+1]-df4.index[i]-
1),ms,df4.index[i],df4.index[i+1]))
dur=pd.DataFrame(dur,columns=['Duration','Point','Start','End'])
durob=dur.groupby("Duration")["Point"].count()
durob=pd.DataFrame(durob)
```

```
durob=durob.sort_values(by=['Duration'], ascending=False)
durob['Possibility']=durob['Point']/sum(durob['Point'])
durob['Survival']=durob['Possibility'].cumsum()
plt.plot(durob.index,durob['Survival'])
durob=durob.sort_values(by=['Point'], ascending=False)
plt.xlabel('Congestion Duration (Steps*90sec)')
plt.ylabel('Probability')
plt.title('Survival')
plt.xlim(0,100)
durob['E']=1
durob['E']=durob['E'].astype('bool')

dur['Number of Lanes']=0
dur['Traffic']=0
dur['Hour']=0
dur['Peak Hour']=0
for i in range(len(dur)):
    stili=dur['Point'][i]
    grammi=dur['Start'][i]
    dur['Traffic'][i]=traffic.at[grammi,stili]
    if grammi<960:
        dur['Hour'][i]=grammi/40
    else:
        dur['Hour'][i]=(grammi-960)/40
    if dur['Hour'][i]<7:
        dur['Peak Hour'][i]=0
    elif dur['Hour'][i]<11:
        dur['Peak Hour'][i]=1
    elif dur['Hour'][i]<14:
        dur['Peak Hour'][i]=0
    elif dur['Hour'][i]<19:
        dur['Peak Hour'][i]=2
    else:
        dur['Peak Hour'][i]=0

dur['E']=1
```

```
dur=dur.drop(['Start','End'],axis=1)
for i in range(len(dur)):
    for j in range(len(theseis)):
        if dur['Point'][i]==theseis['MS'][j]:
            dur['Number of Lanes'][i]=theseis['ANIXNEYTEΣ'][j]
for i in range(len(dur)):
    if dur['Number of Lanes'][i]==0:
        dur['Number of Lanes'][i]=1
dur['Traffic']=dur['Traffic']/dur['Number of Lanes']
exo=[]
for i in range(len(dur)):
    if dur['Traffic'][i]<10:
        exo.append(i)
dur=dur.drop(exo,axis=0)
dur=dur.reset_index(drop=True)
ektos=[]
for i in range(len(dur)):
    if dur['Duration'][i]>150:
        ektos.append(i)
dur=dur.drop(ektos,axis=0)
dur=dur.reset_index(drop=True)
dur=dur.drop(['Hour'],axis=1)
dur['Duration']=dur['Duration']*90/60
dur.to_csv('Congestion Durations.csv')

#%%
nol=dur.sort_values(by=['Number of Lanes'], ascending=False)
lanes=dur['Number of Lanes'].unique()
numbol=nol['Number of Lanes'].value_counts()[lanes]

plt.title('Number of Congestion Incidents per Number of Lanes')
plt.bar(numbol.index,numbol)
plt.ylabel("Congestion Incidents Count")
plt.xlabel("Number of Lanes")

peak=dur.sort_values(by=['Peak Hour'], ascending=False)
ph=dur['Peak Hour'].unique()
```

```
peakcount=peak['Peak Hour'].value_counts()[ph]


plt.figure()

plt.title('Number of Congestion Incidents per Peak Hour')

plt.bar(peakcount.index,peakcount, width=0.2)

plt.ylabel("Congestion Incidents Count")

plt.xlabel("Peak Hour")


#%% Cox ola

from lifelines import CoxPHFitter

cph=CoxPHFitter()

cph.fit(dur, 'Duration','E',cluster_col='Point')

cph.print_summary()

cph.plot()


cph.plot_partial_effects_on_outcome(covariates='Peak Hour', values=[0, 1,2],
cmap='coolwarm')

plt.xlabel('Congestion Duration (Minutes)')

plt.ylabel('Probability')

plt.title('Survival function for different times of the day')


cph.plot_partial_effects_on_outcome(covariates='Traffic',
values=[0,500,1000,1500,2000], cmap='coolwarm')

plt.xlabel('Congestion Duration (Minutes)')

plt.ylabel('Probability')

plt.title('Survival function for different flow rates')


cph.plot_partial_effects_on_outcome(covariates='Number of Lanes', values=[1,2,3,4],
cmap='coolwarm')

plt.xlabel('Congestion Duration (Minutes)')

plt.ylabel('Probability')

plt.title('Survival function for different number of lanes')

plt.xlim(0,150)


#%% Weibull ola

from lifelines import WeibullAFTFitter

aft=WeibullAFTFitter()

weib=dur.drop(['Point'],axis=1)
```

```
aft.fit(weib, duration_col='Duration', event_col='E')

aft.print_summary()

aft.plot()


plt.figure()

aft.plot_partial_effects_on_outcome(covariates='Peak Hour', values=[0, 1, 2],
cmap='coolwarm')

plt.xlabel('Congestion Duration (Minutes)')

plt.ylabel('Probability')

plt.title('Survival function for different times of the day')


plt.figure()

aft.plot_partial_effects_on_outcome(covariates='Traffic',
values=[0,500,1000,1500,2000], cmap='coolwarm')

plt.xlabel('Congestion Duration (Minutes)')

plt.ylabel('Probability')

plt.title('Survival function for different flow rates')


plt.figure()

aft.plot_partial_effects_on_outcome(covariates='Number of Lanes', values=[1,2,3,4],
cmap='coolwarm')

plt.xlabel('Congestion Duration (Minutes)')

plt.ylabel('Probability')

plt.title('Survival function for different number of lanes')

plt.xlim(0,150)


#%% random survival forest

from sklearn.model_selection import train_test_split


from sksurv.ensemble import RandomSurvivalForest


x=dur.drop(['Duration','Point','E','Traffic'], axis=1 )

y=dur['Duration']

y=y.to_numpy()

E=np.array(dur['E'])

E=np.array(E,dtype=bool)

E=[str(x) for x in E]

E=np.array(E)
```

```
y=np.column_stack((E,y))

aux = [(e1,e2) for e1,e2 in y]

y = np.array(aux, dtype=[('Status', '?'), ('Survival_in_days', '<f8')])



X_train, X_test, y_train, y_test = train_test_split(x, y, test_size=0.4,
random_state=20)

rsf = RandomSurvivalForest(n_estimators=100,

                           min_samples_split=6,

                           min_samples_leaf=3,

                           n_jobs=-1,

                           random_state=20)

rsf.fit(X_train, y_train)

rsf.score(X_test, y_test)


X_test_sorted = X_test.sort_values(by=['Peak Hour'])

X_test_sel = pd.concat((X_test_sorted.head(3), X_test_sorted.tail(3)))


surv = rsf.predict_survival_function(X_test_sel, return_array=True)


for i, s in enumerate(surv):

    plt.step(rsf.event_times_, s, where="post", label=str(i))

plt.ylabel("Probability")

plt.xlabel("Congestion Duration (Minutes)")

plt.title('Random survival forest')

plt.grid(True)
```

Μάντσεστερ

Μέθοδος χαμηλών ταχυτήτων

```
#%% DIAVASMA ARXEIWN

import pandas as pd


dfvol = pd.read_csv(r'C:\Users\stefa\Desktop\emp\10ο Εξάμηνο\ΔΙΠΛΩΜΑΤΙΚΗ\man_volume.csv')

dfsp = pd.read_csv(r'C:\Users\stefa\Desktop\emp\10ο Εξάμηνο\ΔΙΠΛΩΜΑΤΙΚΗ\man_speed.csv')

dfhour = pd.read_csv(r'C:\Users\stefa\Desktop\emp\10ο Εξάμηνο\ΔΙΠΛΩΜΑΤΙΚΗ\man_hour.csv')

dfhour['peak_hour']=0

for i in range(len(dfhour)):
```

```
    if dfhour['hour'][i]>9 and dfhour['hour'][i]<19:
        dfhour['peak_hour'][i]=1
#%% diagrammata
import matplotlib.pyplot as plt
import numpy as np
state=pd.DataFrame()

forates=dfvol.columns
forates=forates.drop(forates[0])
speed_stats = dfsp.describe()
vol_stats = (dfvol.describe())

for i in forates:
    df1=dfvol[i]
    df2=dfsp[i]
    print(i, ':', len(dfvol))
    plt.figure()
    plt.scatter(df1, df2, s=1)
    vel=pd.DataFrame()
    vel['vol']=df1
    vel['vel']=speed_stats.loc['25%'][i]-1.5*speed_stats.loc['std'][i]
    plt.plot(vel['vol'],vel['vel'],'r')
    plt.title(i)
    plt.xlabel('Traffic Volume')
    plt.ylabel('Speed (km/h)')
    state[i]=np.where(dfsp[i]<(speed_stats.loc['25%'][i]-1.5*speed_stats.loc['std'][i]),1,0)
```

Εκτίμηση διάρκειας

```
#%% endiamesa 0
for column in forates:
    df3=state[column]
    df4=df3[df3==1]
    l=[]
    l2=[]
    for i in range(len(df4)-1):
        if df4.index[i+1]-df4.index[i]!=1:
            l.append((df4.index[i],df4.index[i+1]))
    for i in range(len(l)):
        if l[i][1]-l[i][0]-1<2:
```

```
            l2=l2+list(range(l[i][0]+1,l[i][1]))
    state[column][l2]=1

#%% endiamesoi 1
for ms in forates:
    df3=state[ms]
    df4=df3[df3==0]
    l=[]
    l2=[]
    for i in range(len(df4)-1):
        if df4.index[i+1]-df4.index[i]!=1:
            l.append((df4.index[i],df4.index[i+1]))
    for i in range(len(l)):
        if l[i][1]-l[i][0]-1<4:
            l2=l2+list(range(l[i][0]+1,l[i][1]))
    state[ms][l2]=0

#%% ypologismos diarkeiwn kai ektypwsh survival
dur=[]
for ms in forates:
    df3=state[ms]
    df4=df3[df3==0]
    for i in range(len(df4)-1):
        if df4.index[i+1]-df4.index[i]!=1:
            dur.append(((df4.index[i+1]-df4.index[i]-1),ms,df4.index[i],df4.index[i+1]))
dur=pd.DataFrame(dur,columns=['Duration','Point','Start','End'])
durob=dur.groupby("Duration")["Point"].count()
durob=pd.DataFrame(durob)
durob=durob.sort_values(by=['Duration'], ascending=False)
durob['Possibility']=durob['Point']/sum(durob['Point'])
durob['Survival']=durob['Possibility'].cumsum()
durob=durob.sort_values(by=['Duration'], ascending=False)
durob.index=durob.index*5
plt.plot(durob.index,durob['Survival'])
plt.xlabel('Congestion Duration (Minutes)')
plt.ylabel('Probability')
plt.title('Survival Curve in Manchester')
plt.xlim(0,180)
```

```
dur['Peak Hour']=0
durob['E']=1
durob['E']=durob['E'].astype('bool')
dur['Traffic']=0
for i in range(len(dur)):
    stili=dur['Point'][i]
    grammi=dur['Start'][i]
    dur['Traffic'][i]=dfvol.at[grammi,stili]
    dur['Peak Hour'][i]=dfhour.at[grammi,'peak_hour']
dur=dur.dropna()
dur=dur[dur['Traffic']>50]
dur=dur[dur['Duration']<=36]
dur=dur.reset_index(drop=True)
dur['E']=1
dur=dur.drop(['Start','End'],axis=1)
dur['Duration']=dur['Duration']*5
dur_stats=dur.describe()
durob_stats=durob.describe()
#%% metavlites
peak=dur.sort_values(by=['Peak Hour'], ascending=False)
ph=dur['Peak Hour'].unique()
peakcount=peak['Peak Hour'].value_counts()[ph]

fig=plt.figure()
ax = fig.add_axes([0,0,1,1])
ax.bar(peakcount.index,peakcount, width=0.4)
plt.title('Number of Congestion Incidents per Peak Hour')
ax.set_xticks(peakcount.index)
plt.ylabel("Congestion Incidents Count")
plt.xlabel("Peak Hour")

#%% Cox ola

from lifelines import CoxPHFitter
cph=CoxPHFitter()

cph.fit(dur, 'Duration','E',cluster_col='Point')
cph.print_summary()
cph.plot()
```

```
cph.plot_partial_effects_on_outcome(covariates='Traffic', values=[0,100,200,300,400,500,600],
cmap='coolwarm')

plt.xlabel('Congestion Duration (Minutes)')

plt.ylabel('Probability')

plt.title('Survival function for different flow rates')

plt.xlim(0,100)


cph.plot_partial_effects_on_outcome(covariates='Peak Hour', values=[0,1], cmap='coolwarm')

plt.xlabel('Congestion Duration (Minutes)')

plt.ylabel('Probability')

plt.title('Survival function for different flow rates')

plt.xlim(0,100)


#%% Weibull ola

from lifelines import WeibullAFTFitter

aft=WeibullAFTFitter()

weib=dur.drop(['Point'],axis=1)

aft.fit(weib, duration_col='Duration', event_col='E')

aft.print_summary()

aft.plot()


plt.figure()

aft.plot_partial_effects_on_outcome(covariates='Peak Hour', values=[0, 1], cmap='coolwarm')

plt.xlabel('Congestion Duration (Minutes)')

plt.ylabel('Probability')

plt.title('Survival function for different times of the day')


plt.figure()

aft.plot_partial_effects_on_outcome(covariates='Traffic', values=[0,100,200,300,400,500],
cmap='coolwarm')

plt.xlabel('Congestion Duration (Minutes)')

plt.ylabel('Probability')

plt.title('Survival function for different flow rates')


#%% random survival forest

from sklearn.model_selection import train_test_split


from sksurv.ensemble import RandomSurvivalForest
```

```
x=dur.drop(['Duration','Point','E','Traffic'], axis=1 )
y=dur['Duration']
y=y.to_numpy()
E=np.array(dur['E'])
E=np.array(E,dtype=bool)
E=[str(x) for x in E]
E=np.array(E)
y=np.column_stack((E,y))
aux = [(e1,e2) for e1,e2 in y]
y = np.array(aux, dtype=[('Status', '?'), ('Survival_in_days', '<f8')])


X_train, X_test, y_train, y_test = train_test_split(x, y, test_size=0.4, random_state=20)
rsf = RandomSurvivalForest(n_estimators=100,
                           min_samples_split=6,
                           min_samples_leaf=3,
                           n_jobs=-1,
                           random_state=20)
rsf.fit(X_train, y_train)
rsf.score(X_test, y_test)

X_test_sorted = X_test.sort_values(by=['Peak Hour'])
X_test_sel = pd.concat((X_test_sorted.head(3), X_test_sorted.tail(3)))

surv = rsf.predict_survival_function(X_test_sel, return_array=True)

for i, s in enumerate(surv):
    plt.step(rsf.event_times_, s, where="post", label=str(i))
plt.ylabel("Probability")
plt.xlabel("Congestion Duration (Minutes)")
plt.title('Random survival forest')
plt.grid(True)
plt.xlim(0,180)
```