



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Πολιτικών Μηχανικών
Τομέας Μεταφορών & Συγκοινωνιακής Υποδομής

**Πολυ-επίπεδη στατιστική ανάλυση της πιθανότητας
και της σοβαρότητας ατυχήματος αξιοποιώντας
κυκλοφοριακά και μετεωρολογικά δεδομένα υψηλής
ευκρίνειας**

Διδακτορική Διατριβή

Αθανάσιος Θεοφιλάτος

Συμβουλευτική Επιτροπή:
Γ. Γιαννής, Καθηγητής ΕΜΠ
Ι. Γκόλιας, Καθηγητής ΕΜΠ
Γ. Κανελλαΐδης, Ομ. Καθηγητής ΕΜΠ

Νοέμβριος 2015

ΠΕΡΙΛΗΨΗ

Στόχος της διδακτορικής διατριβής είναι η **πολυεπίπεδη στατιστική ανάλυση της επιρροής των κυκλοφοριακών και των μετεωρολογικών παραμέτρων στην πιθανότητα και στη σοβαρότητα ατυχήματος**, με ιδιαίτερη έμφαση στους δικυκλιστές.

Αρχικά, για την ανάλυση της πιθανότητας και της σοβαρότητας του ατυχήματος πραγματοποιήθηκε μια διεξοδική βιβλιογραφική ανασκόπηση με έμφαση στην **επιρροή των κυκλοφοριακών και των μετεωρολογικών παραμέτρων**. Στη συνέχεια, πραγματοποιήθηκε εκτενής ανασκόπηση των **κρίσιμων παραγόντων επιρροής στην ασφάλεια των δικυκλιστών**. Από τα αποτελέσματα της ανασκόπησης αυτής προκύπτουν χρήσιμα συμπεράσματα αναφορικά με τις ανάγκες για περαιτέρω έρευνα και τα απαραίτητα χαρακτηριστικά σχεδιασμού της αναλυτικής μεθοδολογίας. Πιο συγκεκριμένα, διαπιστώθηκε έλλειψη στη μελέτη της σχέσης μεταξύ των κυκλοφοριακών χαρακτηριστικών και των οδικών ατυχημάτων με μοτοσυκλετιστές αν και αποτελούν μια ιδιαίτερος ευάλωτη κατηγορία χρηστών της οδού. Αξίζει να σημειωθεί ότι υπάρχει σαφής έλλειψη ερευνών οι οποίες να μελετούν τη σχέση των οδικών ατυχημάτων με τις επικρατούσες κυκλοφοριακές συνθήκες (φόρτος, ταχύτητα, σύνθεση κυκλοφορίας, κλπ) και τις καιρικές συνθήκες πριν και κατά τη στιγμή του ατυχήματος, με χρήση υψηλής ευκρίνειας δεδομένων που έχουν συλλεχθεί σε πραγματικό χρόνο σε αστικές οδούς. Επιπροσθέτως, διαπιστώθηκε η ανάγκη για εφαρμογή πολυεπίπεδης και πιο ολοκληρωμένης στατιστικής ανάλυσης για πληρέστερη κατανόηση του φαινομένου. Εναλλακτικές μεθοδολογίες πρέπει επίσης να εξεταστούν.

Λόγω του σημαντικού αριθμού δεδομένων που απαιτήθηκαν για την επίτευξη των στόχων της Διδακτορικής Διατριβής, εξετάστηκαν και αξιοποιήθηκαν διάφορες πιθανές πηγές αναλυτικών στοιχείων. Τα **δεδομένα των οδικών ατυχημάτων** με βάση τα στοιχεία της ΕΛΣΤΑΤ που συλλέγει η Τροχαία αντλήθηκαν από τη βάση δεδομένων ΣΑΝΤΡΑ του Εργαστηρίου Κυκλοφοριακής Τεχνικής του Τομέα Μεταφορών και Συγκοινωνιακής Υποδομής της Σχολής Πολιτικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου. Τα δεδομένα αυτά περιελάμβαναν γενικά χαρακτηριστικά ατυχημάτων (τύπος ατυχήματος, τύπος οχήματος, συνθήκες φωτισμού, κατάσταση οδοστρώματος, κυβισμός οχημάτων κλπ.) αλλά και χαρακτηριστικά παθόντων (φύλο, ηλικία, εθνικότητα).

Τα **κυκλοφοριακά δεδομένα** αφορούσαν δύο κεντρικές οδικές αρτηρίες της Αθήνας (Λ. Κηφισίας και Λ. Μεσογείων) και αντλήθηκαν από το Κέντρο Διαχείρισης Κυκλοφορίας της Περιφέρειας Αττικής. Επίσης συλλέχθηκαν και κυκλοφοριακά δεδομένα από το Κέντρο Διαχείρισης Κυκλοφορίας της Αττικής Οδού. Πιο συγκεκριμένα, αντλήθηκαν δεδομένα από τους κοντινότερους ανιχνευτές κυκλοφορίας και εξετάστηκαν οι εξής παράμετροι κυκλοφορίας οι οποίες καταγράφονται σε πραγματικό χρόνο: κυκλοφοριακός φόρτος (αριθμός οχημάτων ανά πέντε λεπτά), κατάληψη (σε ποσοστό %), μέση ταχύτητα χρόνου (χιλιόμετρα ανά ώρα). Στα κυκλοφοριακά δεδομένα της Αττικής Οδού συμπεριλαμβάνεται και το ποσοστό βαρέων οχημάτων στην κυκλοφορία (σε ποσοστό %).

Τα **μετεωρολογικά δεδομένα** αντλήθηκαν από το Υδρολογικό Παρατηρητήριο της Αθήνας που αναπτύχθηκε από τον Τομέα Υδρολογίας και Υδάτινων Πόρων του Εθνικού Μετσόβιου Πολυτεχνείου, και περιλαμβάνει 10 πλήρως αυτοματοποιημένους

μετεωρολογικούς σταθμούς οι οποίοι ανήκουν στο δίκτυο METEONET. Εξετάστηκαν οι εξής μετεωρολογικές παράμετροι: θερμοκρασία (σε βαθμούς κελσίου), ύψος βροχόπτωσης (σε χιλιοστά), σχετική υγρασία (σε ποσοστό %), ταχύτητα ανέμου (σε μέτρα ανά δευτερόλεπτο), διεύθυνση ανέμου (σε μοίρες) και ηλιακή ακτινοβολία (σε watt ανά τετραγωνικό μέτρο).

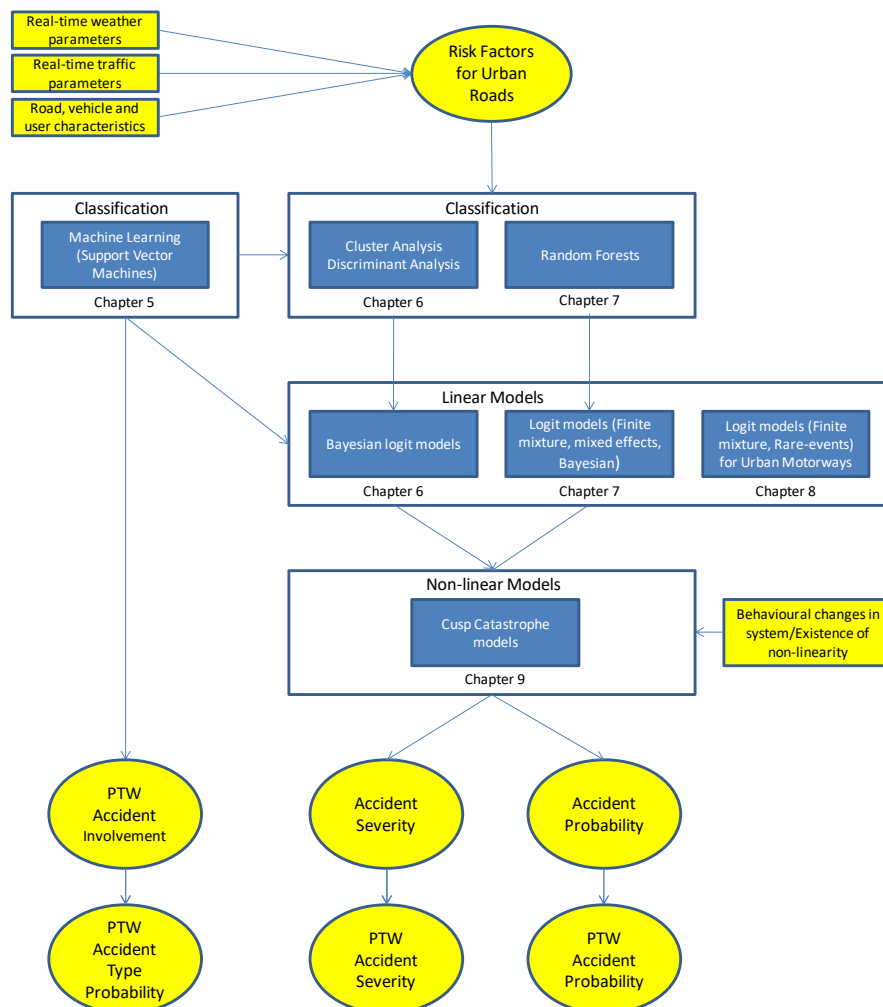
Σημειώνεται ότι με βάση τη βιβλιογραφία, έγινε συλλογή κυκλοφοριακών και μετεωρολογικών δεδομένων που αφορούσαν χρονικές περιόδους **εντός και εκτός της περιόδου των ατυχημάτων**, προκειμένου να μελετηθεί η πιθανότητα ατυχήματος (Abdel-Aty and Pande, 2005; Abdel-Aty et al., 2007; Ahmed and Abdel-Aty, 2012; Yu and Abdel-Aty, 2013a).

Μετά τη συλλογή των δεδομένων ακολούθησε η **προκαταρκτική επεξεργασία** των δεδομένων προτού δημιουργηθούν οι τελικές βάσεις δεδομένων και εφαρμοστεί η πολυεπίπεδη στατιστική ανάλυση της διδακτορικής διατριβής. Στο σημείο αυτό τυχόν λανθασμένες ή μη ρεαλιστικές μετρήσεις αφαιρούνταν (πχ. αρνητικές τιμές κατάληψης) από τη βάση δεδομένων. Στην τελική επεξεργασία των δεδομένων έλαβε χώρα μια μεσοπρόθεσμοι χαρακτήρα προσέγγιση. Πιο συγκεκριμένα, έγινε αναγωγή των αρχικών 5-λεπτων και 10-λεπτων μετρήσεων των κυκλοφοριακών και μετεωρολογικών δεδομένων αντίστοιχα, σε επίπεδο ώρας. Έτσι υπολογίστηκαν τα παρακάτω στατιστικά μεγέθη:

- μέση τιμή
- τυπική απόκλιση
- διάμεσος
- συντελεστής μεταβλητότητας.

Ο σκοπός της αναγωγής των αρχικών δεδομένων μέσα στην ώρα αλλά και ο υπολογισμός των παραπάνω στατιστικών παραμέτρων, πραγματοποιήθηκε διότι τόσο οι κυκλοφοριακές όσο και οι μετεωρολογικές παράμετροι είναι δυναμικά μεγέθη. Έτσι ως δυναμικά μεγέθη ήταν επιθυμητό να μελετηθεί η χρονική τους εξέλιξη και επιρροή και όχι η σημειακή. Επίσης, η τυπική απόκλιση και ο συντελεστής μεταβλητότητας εκφράζουν τη διακύμανση. Για παράδειγμα, η τυπική απόκλιση του κυκλοφοριακού φόρτου εκφράζει τη διακύμανση των 5-λεπτων του κυκλοφοριακού φόρτου στην ώρα. Σημειώνεται επίσης, πως η τυπική απόκλιση του ύψους βροχής προσεγγίζει την ένταση της βροχόπτωσης. Έτσι, υψηλές τιμές τυπικής απόκλισης καταδεικνύει την πιθανή ύπαρξη έντονης βροχόπτωσης.

Οι τελικές βάσεις δεδομένων αναλύθηκαν μέσα από μια **καινοτόμο μεθοδολογία**. Η επιλογή της μεθοδολογίας βασίστηκε στη διαπίστωση ότι η πιθανότητα και η σοβαρότητα ατυχήματος είναι δύο ιδιαίτερος πολύπλοκα και πολυδιάστατα φαινόμενα τα οποία δεν μπορούν να ερμηνευτούν επαρκώς από μια μόνο συγκεκριμένη μέθοδο, αλλά αντίθετα πρέπει να εξεταστούν με τη χρήση μίας πολυεπίπεδης αναλυτικής μεθόδου. Πιο συγκεκριμένα, χρησιμοποιείται μια σειρά μεθοδολογιών κάθε μια από τις οποίες έχει άλλο σκοπό, έτσι ώστε να κατανοηθούν εις βάθος τα ζητήματα της πιθανότητας και της σοβαρότητας ατυχήματος (Γράφημα 1).



Γράφημα 1: Επισκόπηση της πολυεπίπεδης στατιστικής ανάλυσης.

- Συνδυαστική εφαρμογή χρονοσειρών (time series) και μηχανών διανυσμάτων υποστήριξης (Support Vector Machines). Αυτό το κεφάλαιο είναι αφιερωμένο αποκλειστικά στους δικυκλιστές, όπου διερευνάται η πιθανότητα ατυχημάτων με δίκυκλα και ο τύπος ατυχήματος με τη χρήση μιας συνδυασμένης προσέγγισης (Zhao, 2012). Σε αυτή την προσέγγιση, οι μηχανές διανυσμάτων υποστήριξης αναπτύχθηκαν μέσα από την αξιοποίηση τόσο των πρωτότυπων όσο και μετασηματισμένων χρονοσειρών του κυκλοφοριακού φόρτου, της ταχύτητας και της κατάληψης (Κεφάλαιο 5).
- Πεπερασμένης μίξης ανάλυση σε ομάδες (finite mixture cluster analysis), ανάλυση διακριτότητας (discriminant analysis) με σκοπό να καθοριστούν κυκλοφοριακές καταστάσεις (traffic states) καθώς και λογιστικό μοντέλο παλινδρόμησης Bayes, ώστε να διερευνηθεί η επιρροή τους στην πιθανότητα και τη σοβαρότητα ατυχήματος. Πιο συγκεκριμένα, το κεφάλαιο αυτό έχει στόχο να δημιουργήσει σημαντικές ομάδες των παρατηρήσεων σύμφωνα με τα χαρακτηριστικά της κυκλοφορίας (κυκλοφοριακές καταστάσεις) με βάση τις μετρήσεις κατάληψης. Στη συνέχεια, η ανάλυση διακριτότητας κατανέμει σωστά τυχόν νέες παρατηρήσεις στις υπάρχουσες ομάδες και τρόπον τινά δρα συμπληρωματικά στην ανάλυση σε ομάδες. Τέλος, διερευνάται η επιρροή των κυκλοφοριακών καταστάσεων στην πιθανότητα ατυχήματος και σοβαρότητα ατυχήματος, με ανάπτυξη ξεχωριστών μοντέλων ειδικά για τους χρήστες των δικύκλων (Κεφάλαιο 6).

- Διερεύνηση της επιρροής των επιμέρους κυκλοφοριακών και μετεωρολογικών παραμέτρων στην πιθανότητα και τη σοβαρότητα ατυχήματος μέσω εφαρμογής μοντέλων τυχαίας διακλάδωσης (Random Forests) και πεπερασμένης μίξης λογιστικής παλινδρόμησης και λογιστικής παλινδρόμησης Bayes. Τα μοντέλα τυχαίας διακλάδωσης εφαρμόζονται αρχικώς για να καταδείξουν εν δυνάμει σημαντικές μεταβλητές, μέσα από ένα μεγάλο πλήθος ανεξαρτήτων μεταβλητών, χωρίς όμως να ποσοτικοποιήσουν την επιρροή τους. Στη συνέχεια, οι σημαντικές μεταβλητές χρησιμοποιούνται στα παραπάνω λογιστικά μοντέλα προκειμένου να ποσοτικοποιηθεί η επιρροή τους στην πιθανότητα ατυχήματος και σοβαρότητα. (Κεφάλαιο 7).
- Πεπερασμένη μίξης λογιστική παλινδρόμηση, λογιστική παλινδρόμηση Bayes και λογιστική παλινδρόμηση σπανίων γεγονότων (rare-events), με σκοπό τη διερεύνηση της πιθανότητας και τη σοβαρότητας ατυχήματος στην Αττική οδό. Συγκεκριμένα, το λογιστικό μοντέλο σπανίων γεγονότων (King and Zeng, 2001), εφαρμόστηκε ώστε να διερευνηθεί την πιθανότητα ατυχήματος όταν ατυχήματα θεωρούνται σπάνια γεγονότα (Κεφάλαιο 8).
- Εφαρμογή θεωρίας καταστροφής (catastrophe theory) στη διερεύνηση της πιθανότητας και τη σοβαρότητας ατυχήματος στις αστικές οδούς. Τα μοντέλα καταστροφής υποθέτουν την ύπαρξη ενός δυναμικού συστήματος και εξετάζουν την ύπαρξη μη γραμμικών σχέσεων, και πιο συγκεκριμένα την ύπαρξη ξαφνικών μεταβάσεων από τη μια κατάσταση (ασφαλής κατάσταση) στην άλλη (μη ασφαλής κατάσταση) και το αντίστροφο. Στη διδακτορική διατριβή, ως ασφαλείς καταστάσεις ορίζονται η μη ύπαρξη ατυχήματος, η χαμηλή σοβαρότητα και η μη εμπλοκή δικυκλιστών σε ατύχημα.

Αξίζει να σημειωθεί, πως κάποιες από τις **προτεινόμενες μεθοδολογίες είτε δεν έχουν εφαρμοστεί ποτέ στην οδική ασφάλεια έως τώρα (rare-events logit) ή ελάχιστα (θεωρία καταστροφής-catastrophe theory)**. Στο πλαίσιο αυτό, πέραν των διαφορετικών στατιστικών μεθοδολογιών που εφαρμόστηκαν, πραγματοποιήθηκε ανάλυση της επιρροής κυκλοφοριακών και μετεωρολογικών δεδομένων κάτω από διαφορετικές συνθήκες (αστικές οδοί-υπεραστικές οδοί, ΙΧ-δίκυκλα, διαφορετικοί ορισμοί για πιθανότητα-σοβαρότητα κτλ.).

Σε ό,τι αφορά στην **πιθανότητα ατυχήματος**, βρέθηκε ότι η υψηλή κατάληψη είναι συσχετισμένη με αυξημένη πιθανότητα ατυχήματος. Επιπροσθέτως, οι διακυμάνσεις της κατάληψης και του φόρτου αυξάνουν τον κίνδυνο ατυχήματος, ενώ δεν βρέθηκε καμία επιρροή των μετεωρολογικών συνθηκών. Στον αυτοκινητόδρομο, η μέση ταχύτητα έχει αρνητική συσχέτιση με την πιθανότητα ατυχήματος. Το μοντέλο καταστροφής (catastrophe model), κατέδειξε την ύπαρξη έντονης μη-γραμμικότητας στο σύστημα. Βρέθηκε ότι πολύ μικρές μεταβολές στις ανεξάρτητες μεταβλητές είναι δυνατόν να προκαλέσουν απότομες αλλαγές στην πιθανότητα ατυχήματος.

Από την άλλη πλευρά, η αξιοποίηση χρονοσειρών των κυκλοφοριακών παραμέτρων δίνει ικανοποιητική πρόβλεψη στην **πιθανότητα εμπλοκής δικυκλιστών σε ατύχημα**, με τις χρονοσειρές της ταχύτητας να δίνουν την καλύτερη πρόβλεψη. Το γεγονός ότι οι δικυκλιστές είναι πιθανότερο να εμπλακούν σε ατύχημα με 2 ή περισσότερα οχήματα και υπό συνθήκες αυξημένης κυκλοφορίας, δείχνει ότι οι δικυκλιστές είναι ευάλωτοι στην αλληλεπίδραση με τους υπόλοιπους χρήστες της οδού. Το μοντέλο καταστροφής. Στη συνέχεια φαίνεται πως ο μέσος κυκλοφοριακός φόρτος, οι διακυμάνσεις στην ταχύτητα και ο τύπος ατυχήματος επηρεάζουν την πιθανότητα εμπλοκής δικύκλου σε ατύχημα μέσω μη-γραμμικών σχέσεων, σύμφωνα με το μοντέλο καταστροφής. Σε ό,τι αφορά στον αυτοκινητόδρομο, αποδεικνύεται ότι το μη-γραμμικό μοντέλο λογιστικής

παλινδρόμησης Bayes (πολυώνυμο β' βαθμού) έχει καλύτερη στατιστική προσαρμογή έναντι του αντίστοιχου απλού γραμμικού μοντέλου.

Η **σοβαρότητα ατυχήματος** βρέθηκε να έχει θετική συσχέτιση με χαμηλά επίπεδα κατάληψης. Γενικά, διαπιστώθηκε ετερογενής επιρροή των κυκλοφοριακών παραμέτρων σε αστικές οδούς. Η κυκλοφοριακή συμφόρηση οδηγεί σε λιγότερο σοβαρά ατυχήματα στον αυτοκινητόδρομο. Συνολικά βρέθηκε σημαντική επιρροή του τύπου ατυχήματος τόσο σε αστικές οδούς όσο και στον αυτοκινητόδρομο. Το μοντέλο καταστροφής βρέθηκε ότι περιγράφει πιο αξιόπιστα τη σοβαρότητα ατυχήματος. Μάλιστα, φαίνεται ότι πολύ μικρές μεταβολές στις ανεξάρτητες μεταβλητές είναι δυνατόν να προκαλέσουν απότομες αλλαγές στη σοβαρότητα ατυχήματος.

Η χαμηλή κυκλοφορία βρέθηκε συσχετισμένη με αυξημένη **σοβαρότητα ατυχήματος με δικυκλιστές**. Επίσης τονίζεται η μη γραμμική συσχέτιση των κυκλοφοριακών και μετεωρολογικών παραμέτρων καθώς και του τύπου ατυχήματος με τη σοβαρότητα ατυχήματος με δικυκλιστές. Τέλος, στον αυτοκινητόδρομο, υπάρχει ετερογενής επιρροή των κυκλοφοριακών παραμέτρων ανάλογα με τη “λανθάνουσα” ομάδα και αρνητική συσχέτιση με τον κυβισμό του δικύκλου.

Συνολικά, τα αποτελέσματα διδακτορικής διατριβής έδειξαν ότι η **πολυ-επίπεδη στατιστική ανάλυση** που εφαρμόστηκε είναι δυνατόν να οδηγήσει σε πληρέστερη κατανόηση δύο πολύπλοκων ζητημάτων όπως είναι η πιθανότητα και η σοβαρότητα των ατυχημάτων τόσο σε αστικές οδούς όσο και στον αυτοκινητόδρομο. Επιπροσθέτως, τα υψηλής ευκρίνειας κυκλοφοριακά και μετεωρολογικά δεδομένα σε συνδυασμό με τη μεσοπρόθεσμου χαρακτήρα μεθοδολογική προσέγγιση, είναι δυνατόν να δώσουν σε νέες διαστάσεις στην ανάλυση και κατανόηση των οδικών ατυχημάτων. Είναι επίσης σημαντικό να δίνεται έμφαση στη μελέτη των ευάλωτων χρηστών της οδού όπως για παράδειγμα οι δικυκλιστές.

Από μεθοδολογικής απόψεως, εφαρμόστηκαν **καινοτόμες και εναλλακτικές μέθοδοι**, παρουσιάζοντας πολύ ικανοποιητική στατιστική προσαρμογή, και παράγοντας ενθαρρυντικά αποτελέσματα. Μάλιστα ορισμένες από αυτές τις μεθόδους εφαρμόζονται για πρώτη φορά σε συνδυασμό με τέτοιου είδους δεδομένα (πχ. μηχανές διανυσμάτων υποστήριξης, θεωρία καταστροφής) ή και για πρώτη φορά στο γνωστικό αντικείμενο της οδικής ασφάλειας. Ειδικά για τα μοντέλα της θεωρίας καταστροφής, τονίζεται ότι η υπόθεση για ύπαρξη δυναμικών συστημάτων πιθανώς να μπορούσε να οδηγήσει σε πιο γενικευμένη εφαρμογή παρόμοιων μεθόδων όπως για παράδειγμα εφαρμόζεται στη μελέτη των Frazier και Kockelman (2004).