



TECHNICAL UNIVERSITY OF CRETE
DEPARTMENT OF PRODUCTION ENGINEERING
& MANAGEMENT

**REAL-TIME ESTIMATION OF VEHICLE-COUNT
WITHIN SIGNALIZED LINKS**

M.Sc. Thesis

GEORGIOS VIGOS

A thesis submitted in partial fulfillment of the
requirements for the degree of

MASTER OF OPERATIONAL RESEARCH

CHANIA

2006

© Copyright by Georgios Vigos

2006

Georgios Vigos' thesis is approved by the examining committee:

Markos Papageorgiou

(Supervisor)

Elias Kosmatopoulos

Ioannis Papamichail

TABLE OF CONTENTS

LIST OF FIGURES	VII
-----------------------	-----

LIST OF TABLES.....	X
---------------------	---

ACKNOWLEDGMENTS	XI
-----------------------	----

SHORT CURRICULUM VITAE.....	XII
-----------------------------	-----

ΠΕΡΙΛΗΨΗ	XIII
----------------	------

ΕΙΣΑΓΩΓΗ.....	XIII
---------------	------

ΚΕΦΑΛΑΙΟ 1 ^ο	XIV
-------------------------------	-----

ΚΕΦΑΛΑΙΟ 2 ^ο	XVI
-------------------------------	-----

ΚΕΦΑΛΑΙΟ 3 ^ο	XVI
-------------------------------	-----

ΚΕΦΑΛΑΙΟ 4 ^ο	XIX
-------------------------------	-----

ΚΕΦΑΛΑΙΟ 5 ^ο	XX
-------------------------------	----

ΚΕΦΑΛΑΙΟ 6 ^ο	XXI
-------------------------------	-----

ΚΕΦΑΛΑΙΟ 7 ^ο	XXIII
-------------------------------	-------

ΚΕΦΑΛΑΙΟ 8 ^ο	XXIV
-------------------------------	------

ΠΑΡΑΡΤΗΜΑΤΑ	XXV
-------------------	-----

ABSTRACT	XXVI
-----------------------	-------------

CHAPTER 1 GENERAL CONCEPTS	1
-----------------------------------	----------

1.1	TRAFFIC CONGESTION	1
1.2	TRAFFIC FLOW PROCESS CONTROL.....	3
1.3	LOOP DETECTORS	4
1.4	RAMP METERING	6
1.5	KALMAN FILTER	8
1.6	MICROSCOPIC SIMULATION MODEL.....	12
1.7	THE NUMBER OF VEHICLES ESTIMATION PROBLEM	13

CHAPTER 2 TIME-OCCUPANCY AND SPACE-OCCUPANCY	15
---	-----------

2.1	INTRODUCTION	15
2.2	INDUCTIVE LOOP MEASUREMENTS	16
2.3	WARDROP'S APPROACH	18
2.4	EDIE'S DEFINITION.....	19
2.5	MEAN EFFECTIVE VEHICLE LENGTH.....	21
2.6	SPACE-OCCUPANCY	22

CHAPTER 3 SIGNALIZED LINKS	25
-----------------------------------	-----------

3.1	INTRODUCTION	25
3.2	SPACE-OCCUPANCY RELATIONSHIPS	25
3.3	BIAS-FREE ESTIMATES	27

3.4	A GENERAL EXAMPLE	29
3.5	RELAXING SOME ASSUMPTIONS.....	31
3.6	VEHICLE-COUNT IN THE LINK.....	33
 CHAPTER 4 KALMAN FILTER DEVELOPMENT		36
<hr/>		
4.1	INTRODUCTION	36
4.2	INTRODUCTION TO KALMAN FILTER.....	38
4.3	MODELLING FOR ESTIMATION.....	41
4.4	KALMAN-FILTER ESTIMATOR	45
 CHAPTER 5 EXPONENTIAL SMOOTHING		49
<hr/>		
5.1	INTRODUCTION	49
5.2	EXPONENTIAL SMOOTHED ESTIMATOR.....	50
 CHAPTER 6 MICROSCOPIC SIMULATION MODEL		52
<hr/>		
6.1	INTRODUCTION	52
6.2	SIMULATION DESCRIPTION	53
 CHAPTER 7 SIMULATION RESULTS		57
<hr/>		
7.1	INTRODUCTION	57
7.2	IMPACT OF THE SAMPLING TIME	58
7.3	IMPACT OF THE UPDATE PERIOD T	66
7.4	IMPACT OF THE NUMBER OF DETECTORS	72

7.5	NON-ZERO EFFECTIVE DETECTOR LENGTH	76
7.6	UNCERTAIN PHYSICAL VEHICLE LENGTH AVERAGE	80
7.7	LONGER LINK	84
7.8	LARGE INITIAL ESTIMATION ERROR.....	86
 CHAPTER 8 GUIDELINES, RECOMMENDATIONS AND CONCLUSIONS		89
8.1	GUIDELINES AND RECOMMENDATIONS FOR PRACTICAL APPLICATION	89
8.2	CONCLUSIONS.....	91
 APPENDIX A		93
 APPENDIX B		94
 REFERENCES		96

LIST OF FIGURES

Figure 1: The control loop.	3
Figure 2: Operation of a loop-type detector.....	5
Figure 3: (a) Local and (b) coordinated ramp metering.....	7
Figure 4: Kalman Filter application.....	11
Figure 5: Vehicle-count estimation: (a) the signal and detector configuration on the link; (b) the link vehicle-count estimation.	13
Figure 6: Time-occupancy measurements signals.	17
Figure 7: Illustration of Edie (1965; 1974) definition of traffic variables.	20
Figure 8: Placement of M internal detectors for time-occupancy.	26
Figure 9: $E\{o_s\}$ and $E\{o_s(x)\}$ for a signalized link.....	28
Figure 10: Real and estimated vehicle-counts over time for four basic scenarios	58
Figure 11: Real and KF estimated vehicle-counts over time for four basic scenarios.	59
Figure 12: Real and smoothed estimated vehicle-counts over time for four basic scenarios.	59
Figure 13: $E\{o_s\}$ and $E\{o_t(x^i)\}$ for the four basic scenarios.	62
Figure 14: Measurements and KF estimation RMSE in dependence of the KF gain K	64
Figure 15: Measurements and smoothed estimation RMSE in dependence of the smoothing parameter K_{SM}	65
Figure 16: Impact of T on RMSE for standard scenario with $M = 1, 4, 200$	66
Figure 17: Impact of T on RMSE for five scenarios ($M = 1$).....	67
Figure 18: Measurements and KF estimation RMSE in dependence of update period T for (a) the standard and (b) the stochastic scenarios.....	69
Figure 19: Measurements and smoothed estimation RMSE in dependence of update period T for (a) the standard and (b) the stochastic scenarios..	71

Figure 20: Impact of M on RMSE for the standard scenario with $T = 0.5$ s, $T = 10$ s, $T = 20$ s.	72
Figure 21: Measurements and estimation RMSE in dependence of the number of internal detectors for $T = 20$ s for (a) the standard and (b) the stochastic scenarios.....	73
Figure 22: Measurements and smoothed estimation RMSE in dependence of the number of internal detectors for $T = 20$ s for (a) the standard and (b) the stochastic scenarios.....	75
Figure 23: Average error (bias) for standard scenario with non-zero effective detector length.	77
Figure 24: KF estimation for the case of non-zero effective detector length for standard scenario.....	78
Figure 25: Smoothed estimation for the case of non-zero effective detector length for standard scenario.	79
Figure 26: KF estimation for the cases of uncertain average physical vehicle length.	81
Figure 27: KF estimation for the cases of 10% and 30% trucks with N in p.c.u. and 1 truck = 2.25 p.c.u.....	82
Figure 28: Smoothed estimation for the cases of uncertain average physical vehicle length.	83
Figure 29: Smoothed estimation for the cases of 10% and 30% trucks with N in p.c.u. and 1 truck = 2.25 p.c.u.....	83
Figure 30: KF estimation results for a longer link.....	84
Figure 31: Smoothed estimation results for a longer link.	85
Figure 32: Real and KF estimated vehicle-counts over time for longer-link scenario.....	85
Figure 33: Real and smoothed estimated vehicle-counts over time for longer-link scenario.....	86
Figure 34: KF estimation for the case of larger initial estimation error ($\hat{N}(0) = 20$ veh).	87

Figure 35: Smoothed estimation for the case of larger initial estimation error ($\hat{N}(0) = 20$ veh).	87
Figure 36: Real and KF estimated vehicle-count over time for the scenario with $\hat{N}(0) = 20$ veh.	88
Figure 37: Real and smoothed estimated vehicle-count over time for the scenario with $\hat{N}(0) = 20$ veh.	88

LIST OF TABLES

Table 1: RMSE values and $E\{N - \hat{N}\}$ for five scenarios and various cases of effective vehicle length.....	60
Table 2: KF RMSE values and $E\{N - \hat{N}_{KF}\}$ for five scenarios and various cases of effective vehicle length.	60
Table 3: SM RMSE values and $E\{N - \hat{N}_{SM}\}$ for five scenarios and various cases of effective vehicle length.	60
Table 4: Green phases of traffic signals for the standard scenario.....	94
Table 5: Green phases of traffic signals for the scenario with cycle of 40 s at the downstream traffic signal.	94
Table 6: Green phases of traffic signals for the scenario with cycle of 60 s at the downstream traffic signal.	94
Table 7: Green phases of traffic signals for the scenario with cycle of 90 s at the downstream traffic signal.	95
Table 8: Green phases of traffic signals for the stochastic scenario with cycle of x sec at the downstream traffic signal (x is an integer between 10 s and 90 s).	95

ACKNOWLEDGMENTS

After this work has finally reached to its end, I feel the obligation to express my utmost gratitude to a series of people that in their own unique way helped me with everything I have accomplished so far.

First of all I would like to thank my supervisor Prof. Markos Papageorgiou for his precious guidance and support during my thesis. The daily contact and the collaboration with this man taught me a lot and made me better at my work. He will always be a benchmark in my life.

I wish to deeply thank my parents Dimitri and Kiriaki and my brother Michali. It is very important for anyone to have the unlimited patience, support, inspiration and guidance of his family, in every step of his life. I will always be grateful to them. You did it...

I would like to thank all my non-academic friends that helped and supported me with their own way. I, also, owe many thanks to my best friends and fellow-students for their support and for being good listeners. I believe that Knowledge cannot overtake human relationships and during my stay in Chania these friendships is the most powerful qualification that I have earned. I feel privileged of knowing you guys...

Last, I have to mention that this work was supported in part by the European Commission's IST (Information Society Technologies) Programme under the project EURAMP (IST-2002-23110).

The "Journey" now begins...

SHORT CURRICULUM VITAE

Vigos Georgios (1981-)

Georgios Vigos was born in Piraeus in June 1981. Between 1999-2004 he studied Production Engineering and Management at the Technical University of Crete. Since the third year of his studies he has been working as an associated partner of the Dynamic Systems and Simulation Laboratory. In the frames of this partnership he has been involved in various research programs and has come by experience to the development and simulation of traffic control strategies. He has taught, as a teaching assistant, to undergraduates of the department.

Εισαγωγή

Ο αριθμός των οχημάτων σε μια ράμπα αυτοκινητοδρόμου ή σε έναν αστικό σηματοδοτούμενο σύνδεσμο αποτελεί πολύτιμη πληροφορία για τον έλεγχο σε πραγματικό χρόνο. Η ποσότητα αυτή είναι στενά συνδεδεμένη με τις μεταβλητές χώρου και χρόνου της ράμπας ή του συνδέσμου. Μέχρι τώρα αρκετοί ερευνητές έχουν ασχοληθεί με τις σχέσεις που συνδέουν τις στιγμιαίες χωρικές μεταβλητές και τις (εύκολα μετρήσιμες) τοπικές χρονικές μεταβλητές σε, σχετικά, ομογενή και σταθερή κυκλοφοριακή ροή, όπως αυτή συναντάται συνήθως σε συνεχείς κυκλοφοριακές συνθήκες. Όμως πώς αλλάζουν αυτές οι σχέσεις κάτω από ανομοιογενείς και ασταθείς κυκλοφοριακές συνθήκες; Για να απαντηθεί αυτό το ερώτημα πραγματοποιήθηκε μια λεπτομερής ανάλυση και διεξήχθη μια μικροσκοπική προσομοίωση για σηματοδοτούμενα δίκτυα, όπου υπάρχουν έμφυτες και ισχυρές διακυμάνσεις της κυκλοφοριακής ροής, λόγω της εναλλαγής των ενδείξεων των φωτεινών σηματοδοτών. Ένας αριθμός παραγόντων που επηρεάζουν κατά την εκτίμηση των χωρικών μεταβλητών από τις μετρήσιμες τοπικές χρονικές μεταβλητές αναλύονται και παρουσιάζονται με λεπτομέρειες. Με βάση αυτούς, εφαρμόζεται ένα φίλτρο Kalman που βασίζεται σε πραγματικού χρόνου μετρήσεις της ροής και της κατάληψης, οι οποίες παρέχονται από τρεις φωρατές κλειστού βρόχου, για την παραγωγή αξιόπιστων εκτιμήσεων του αριθμού των οχημάτων. Ο τελικός εκτιμητής του αριθμού των οχημάτων δοκιμάζεται μέσω, της ίδιας με προηγουμένως, μικροσκοπικής προσομοίωσης για διάφορα σενάρια ραμπών και κυκλοφοριακών συνθηκών. Τα αποτελέσματα της προσομοίωσης υποδηλώνουν ότι πρόκειται για μια εύρωστη

απόδοση της εκτίμησης, γεγονός που επιτρέπει την εύκολη εφαρμογή της μεθόδου.

Στη συνέχεια θα παρουσιαστεί μια εκτεταμένη περίληψη της εργασίας που πραγματοποιήθηκε δίνοντας ιδιαίτερη προσοχή στα βασικά στοιχεία που τη χαρακτηρίζουν. Στο τέλος θα αναφερθούν τα κυριότερα συμπεράσματα, τα οποία επιβεβαιώνουν τη σημασία της παρούσας και οριοθετούν την πρακτική εφαρμογή της προτεινόμενης μεθοδολογίας σε πραγματικές συνθήκες.

Κεφάλαιο 1°

Στο 1° κεφάλαιο ορίζονται και αναλύονται συνοπτικά ορισμένες βασικές έννοιες που αποτελούν τα θεμέλια ανάπτυξης ολόκληρης της εργασίας. Τέτοιες έννοιες είναι:

- Η κυκλοφοριακή συμφόρηση (traffic congestion): Πρόκειται για πρόβλημα καθημερινό στα αστικά κέντρα με ποικίλλες επιπτώσεις στην ποιότητα ζωής των κατοίκων. Ο πλέον ενδεδειγμένος και πιο οικονομικός τρόπος αντιμετώπισής του επιβάλλει την πιο αποτελεσματική εκμετάλλευση των εγκαταστάσεων που ήδη υπάρχουν.
- Ο έλεγχος της κυκλοφοριακής ροής (traffic flow control): Για τον ικανοποιητικό έλεγχο της κυκλοφοριακής ροής χρησιμοποιείται, συνήθως, η τεχνική του κλειστού βρόχου ελέγχου. Πρόκειται για μεθοδολογία ελέγχου ενός συστήματος και βελτιστοποίησης οποιουδήποτε στοιχείου του μέσω μετρήσεων που λαμβάνονται από διάφορους αισθητήρες και χρησιμοποιούνται κατάλληλα από μια στρατηγική ελέγχου.
- Φωρατές κλειστού βρόχου (loop detectors): Αποτελούν το πιο διαδεδομένο όργανο συλλογής μετρήσεων που αφορούν την κυκλοφοριακή ροή, λόγω του χαμηλού κόστους σε σχέση, για παράδειγμα, με τις κάμερες

παρακολούθησης. Η λειτουργία τους βασίζεται στην επαγωγή ηλεκτρικού ρεύματος.

- Έλεγχος ραμπών (ramp metering): Η σπουδαιότητα του ελέγχου της κυκλοφοριακής ροής σε ράμπες εισόδου/εξόδου αυτοκινητοδρόμων έχει προσελκύσει το ενδιαφέρον πολλών ερευνητών. Η ικανότητα ελέγχου της ροής των οχημάτων στα μικρά, σχετικά, αυτά τμήματα δρόμου παρέχει πολλαπλές βελτιώσεις στη ροή τόσο των αυτοκινητοδρόμων, όσο και των παρακείμενων αστικών οδικών δικτύων.
- Φίλτρο Kalman (Kalman Filter): Πρόκειται για μια μαθηματική μέθοδο που ανέπτυξε ο Rudolf E. Kalman (1960) και χρησιμοποιείται για την πρόβλεψη/εκτίμηση ενός μεγέθους. Οι εφαρμογές του επεκτείνονται από την αεροδιαστημική έως τα δημογραφικά μοντέλα και μια σειρά μηχανικών εφαρμογών (για παράδειγμα τα ραντάρ).
- Μικροσκοπικό μοντέλο προσομοίωσης (microscopic simulation model): Πρόκειται για μοντέλα που χρησιμοποιούνται ευρέως για την περιγραφή και ανάλυση της κυκλοφοριακής ροής. Η απεικόνιση μιας διαδικασίας μέσω προσομοίωσης αποτελεί βασικό εργαλείο ελέγχου αυτής πριν την εφαρμογή της σε πραγματικές συνθήκες.

Το κεφάλαιο ολοκληρώνεται με την περιγραφή του προβλήματος εκτίμησης του αριθμού των οχημάτων σε σηματοδοτούμενα οδικά δίκτυα. Σύμφωνα και με το Γράφημα 5 (Figure 5), δύο φωτεινοί σηματοδότες είναι τοποθετημένοι ανάντι και κατάντι του εξεταζόμενου συνδέσμου. Ο ανάντι φωτεινός σηματοδότης (εάν υπάρχει) καθορίζει τη ζήτηση της ροής των οχημάτων που καταφτάνει στο σύνδεσμο, ενώ ο κατάντι φωτεινός σηματοδότης καθορίζει τη ροή εξόδου των οχημάτων από το σύνδεσμο. Είναι αυτονόητο ότι όταν η ροή εισόδου των οχημάτων είναι μεγαλύτερη της ροής εξόδου, τότε μέσα στο σύνδεσμο

δημιουργείται ουρά οχημάτων. Επιπλέον, στον εξεταζόμενο σύνδεσμο έχουν εγκατασταθεί τρεις φωρατές κλειστού βρόγχου: στα δύο άκρα του συνδέσμου και στο μέσον του. Οι δύο ακραίοι φωρατές παρέχουν μετρήσεις ροής οχημάτων, ενώ ο μεσαίος φωρατής παρέχει μετρήσεις χρονικής κατάληψης του συνδέσμου. Όπως φαίνεται και στο Γράφημα 5 (Figure 5), ο εκτιμητής που αναπτύχθηκε λειτουργεί σε πραγματικό χρόνο, τόσο για να ενημερώνεται με τις μετρήσεις ροής και χρονικής κατάληψης από τους τρεις φωρατές, όσο και για να παρέχει τον εκτιμώμενο αριθμός οχημάτων μέσα στο σηματοδοτούμενο δίκτυο (ανάμεσα στους δύο ακραίους φωρατές).

Κεφάλαιο 2°

Στο 2° κεφάλαιο πραγματοποιείται μια ανασκόπηση των σχέσεων που συνδέουν τις στιγμιαίες χωρικές μεταβλητές με τις (εύκολα μετρήσιμες) τοπικές χρονικές μεταβλητές σε ομογενή και σταθερή κυκλοφοριακή ροή. Πιο συγκεκριμένα, αναλύεται ο τρόπος που λαμβάνονται οι μετρήσεις από τους φωρατές κλειστού βρόγχου και ο τρόπος που αυτές καθίστανται εκμεταλλεύσιμες, ενώ ορίζονται τα βασικά μεγέθη της κυκλοφοριακής ροής, δηλαδή η ροή οχημάτων q , η πυκνότητα ρ και η μέση ταχύτητα v , σύμφωνα με τους Wardrop (1952) και Edie (1965, 1974). Το κεφάλαιο ολοκληρώνεται με τον ορισμό της χωρικής κατάληψης o_s ενός δρόμου με βάση και το μέσο ενεργό μήκος L (mean effective length) και το φυσικό μήκος L^{Ph} (physical length) των οχημάτων.

Κεφάλαιο 3°

Στο 3° κεφάλαιο οι γνωστές σχέσεις του προηγούμενου κεφαλαίου τροποποιούνται, ώστε να ανταποκρίνονται στις ανομοιογενείς και ασταθείς κυκλοφοριακές συνθήκες που επικρατούν σε ένα σηματοδοτούμενο σύνδεσμο.

Στην προσπάθεια συσχέτισης της μετρούμενης χρονικής κατάληψης με την αντίστοιχη χωρική κατάληψη του συνδέσμου γίνονται οι εξής παραδοχές:

- Ο χρόνος ενημέρωσης T των μετρήσεων χρονικής κατάληψης είναι αρκετά μικρός, ώστε αυτές να αντιστοιχούν σε στιγμιαίες μετρήσεις και να ισχύει $o_t = 1$ εάν ένας φωρατής είναι κατειλημμένος και $o_t = 0$ διαφορετικά.
- Υπάρχουν M στον αριθμό μετρήσεις χρονικής κατάληψης κατά μήκος του συνδέσμου σύμφωνα και με το Γράφημα 8 (Figure 8) με κενό $d = \Delta/M$ μεταξύ τους, όπου Δ είναι το συνολικό μήκος του συνδέσμου.
- Ισχύει $L_j^{Ph} = L_j \ \forall j$, όπου $j = 1, \dots, N$ ο αριθμός του οχήματος στο σύνδεσμο.

Εφόσον ισχύουν αυτά, η χωρική κατάληψη $o_s(x^i)$ σε μια περιοχή μήκους d γύρω από μια θέση μέτρησης x^i , $i = 1, \dots, M$, μπορεί να μετρηθεί προσεγγιστικά μέσω της μετρήσιμης χρονικής κατάληψης $o_t(x^i)$ από τη σχέση
$$o_s \approx \sum_{i=1}^M o_s(x^i) d / \Delta = \sum_{i=1}^M o_t(x^i) / M.$$
 Όπως φαίνεται, η στιγμιαία χωρική κατάληψη του συνδέσμου μπορεί να προσεγγιστεί από μια σειρά μετρήσεων χρονικής κατάληψης όταν $L_j^{Ph} = L_j$ και για πολύ μικρό χρόνο T . Ο βαθμός ακρίβειας της προσέγγισης εξαρτάται από το διάστημα d .

Εν συνεχεία γίνεται προσπάθεια εύρεσης των καλύτερων θέσεων για τους φωρατές μέσα στο σύνδεσμο όταν αυτοί είναι λίγοι στον αριθμό ή ακόμα και όταν πρόκειται για έναν και μοναδικό φωρατή. Αποδεικνύεται ότι η τοποθέτηση των φωρατών σύμφωνα με το Γράφημα 8 (Figure 8) είναι ένας καλός τρόπος για την εκτίμηση της χωρικής κατάληψης, καθώς οδηγεί σε χαμηλό μέσο σφάλμα, παρόλο που η διακύμανσή του μπορεί να είναι υψηλή (όταν εγκαθίσταται ένας και μόνο φωρατής, η ενδεδειγμένη του θέση βρίσκεται στο μέσο, περίπου, του συνδέσμου). Μάλιστα, όσο μεγαλύτερος είναι ο αριθμός M των φωρατών στο

σύνδεσμο, τόσο καλύτερες είναι οι εκτιμήσεις χωρικής κατάληψης και σε όρους μέσου σφάλματος και σε όρους σφάλματος διακύμανσης. Προς απόδειξη αυτού, πραγματοποιείται μια λεπτομερής ανάλυση του απόλυτου και του σχετικού σφάλματος της εκτίμησης της χωρικής κατάληψης του συνδέσμου, η οποία ισχύει, τελικά, ακόμα και όταν $L_j = L_j^{Ph} + \varepsilon_j$, όπου ε_j το μη μηδενικό ενεργό μήκος των φωρατών, με $\varepsilon_j < E$ και το $E > 0$ να είναι αρκετά μικρό για να αποτρέπεται η ταυτόχρονη ενεργοποίησης ενός φωρατή από περισσότερα από ένα οχήματα. Σημαντικό είναι, επίσης, το φαινόμενο ZSZO (zero-speed zero-occupancy) που παρουσιάζεται, κατά το οποίο δύο οχήματα είναι σταματημένα (λόγω της ουράς) εκατέρωθεν ενός φωρατή με αποτέλεσμα να παρέχονται μηδενικές μετρήσεις από το φωρατή, ενώ υπάρχουν οχήματα εντός του συνδέσμου. Το συγκεκριμένο φαινόμενο μπορεί να οδηγήσει, ανάλογα με τη συχνότητα εμφάνισής του, σε, αντίστοιχα, μεγάλα σφάλματα της εκτίμησης της χωρικής κατάληψης του συνδέσμου.

Η ανάλυση του παρόντος κεφαλαίου οδηγεί σε ενδιαφέροντα αποτελέσματα. Από τη μια συμπεραίνεται ότι όταν ο χρόνος T αυξάνεται, αυξάνεται και η ποιότητα της εκτίμησης της χωρικής κατάληψης του συνδέσμου, καθώς η εκτιμώμενη χωρική κατάληψη $\hat{\delta}_s$ τείνει προς την αναμενόμενη μέση τιμή της $E\{\hat{\delta}_s\}$ (για μη μηδενική ταχύτητα οχημάτων). Από την άλλη, στην περίπτωση του σηματοδοτούμενου συνδέσμου, όπου με την πάροδο του χρόνου T υπάρχουν πολλές αλλαγές στον αριθμό των οχημάτων στο σύνδεσμο, άρα και στη χωρική κατάληψη αυτού, η αύξηση της περιόδου T , με τις αντίστοιχες σημαντικές διαφοροποιήσεις στη χωρική κατάληψη, προκαλεί πτώση της ακρίβειας της εκτιμώμενης χωρικής κατάληψης του συνδέσμου. Επομένως, υπάρχει μια βέλτιστη χρονική περίοδος T , η οποία οδηγεί στη βέλτιστη

εκτίμηση της χωρικής κατάληψης, άρα και του αριθμού των οχημάτων, του συνδέσμου.

Το κεφάλαιο ολοκληρώνεται με την εύρεση του εκτιμητή \hat{N} του αριθμού των οχημάτων μέσα στο σύνδεσμο που είναι $\hat{N} = N + \zeta$, όπου N ο πραγματικός αριθμός οχημάτων στο σύνδεσμο και $\zeta = \zeta_0 \Delta \lambda / L^{Ph}$, με Δ το μήκος του συνδέσμου και λ τον αριθμό λωρίδων στο σύνδεσμο. Ο θόρυβος ζ_0 της μέτρησης της χρονικής κατάληψης περιλαμβάνει διάφορες πηγές, όπως:

- Το θόρυβο των μετρήσεων των φωρατών.
- Το σφάλμα στη μοντελοποίηση που προέρχεται από την προσεγγιστικότητα στη σχέση $o_s \approx \sum_{i=1}^M o_s(x^i) d / \Delta = \sum_{i=1}^M o_t(x^i) / M$, κυρίως για μικρό αριθμό εγκατεστημένων φωρατών M . Εδώ εμπεριέχεται και η όποια επιρροή του χρόνου ενημέρωσης T και του φαινομένου ZSZO.
- Το σφάλμα εξαιτίας της διαφορετικότητας του ενεργού και του φυσικού μήκους των οχημάτων κατά τη μέτρηση της χρονικής κατάληψης.

Κεφάλαιο 4^ο

Το 4^ο κεφάλαιο αναφέρεται στην ανάπτυξη του εκτιμητή του αριθμού των οχημάτων N μέσα στο σηματοδοτούμενο σύνδεσμο. Καταρχήν πραγματοποιείται μια σύντομη ανασκόπηση των μαθηματικών σχέσεων του φίλτρου Kalman και γίνεται αναφορά στις προϋποθέσεις που πρέπει να ισχύουν για τα μεγέθη που εμπλέκονται σε αυτές. Στη συνέχεια γίνεται προσπάθεια εξαγωγής των απαραίτητων για την εφαρμογή του φίλτρου μοντέλων κατάστασης και μέτρησης του προβλήματος. Ο τελικός εκτιμητής που εξάγεται είναι ο εξής:

$$\hat{N}_{KF}(k) = \hat{N}_{KF}(k-1) + T[q_{in}^m(k-1) - q_{out}^m(k-1)] + K[N^m(k-1) - \hat{N}_{KF}(k-1)]$$

όπου N^m , \hat{N}_{KF} η μέτρηση και η εκτίμηση, αντίστοιχα, του αριθμού οχημάτων στο σύνδεσμο, q_{in}^m, q_{out}^m οι ροές εισόδου και εξόδου, αντίστοιχα, οχημάτων στο σύνδεσμο, T ο χρόνος ενημέρωσης και εκτίμησης του φίλτρου και K ο παράγοντας οφέλους (gain parameter) του φίλτρου. Η εκτίμηση \hat{N}_{KF} περιορίζεται στο εύρος $[0, N'_{max}]$, όπου N'_{max} ο μέγιστος αριθμός οχημάτων που μπορούν να βρεθούν μέσα στο σύνδεσμο ακίνητα, συμπεριλαμβανομένης και της συνηθισμένης απόστασης ασφαλείας μεταξύ τους D (όπου $D = 1m$). Η μέτρηση N^m δίνεται από τη σχέση $N^m(k-1) = N(k-1) + \zeta(k-1)$, όπου το ζ έχει οριστεί στο προηγούμενο κεφάλαιο.

Όσον αφορά τον παράγοντα οφέλους K , από τη θεωρία ισχύουν $K = \Pi / (\Pi + Z)$ και $\Pi = (1 - K)\Pi + T^2\Gamma$, από όπου προκύπτει, τελικά, $K = 0.5(-\alpha + \sqrt{\alpha^2 + 4\alpha})$ με $\alpha = T^2\Gamma / Z$ και Γ, Z οι διακυμάνσεις των θορύβων του συστήματος και της μέτρησης, αντίστοιχα. Αποδεικνύεται, τέλος, ότι οποιοδήποτε τάση (bias) b χαρακτηρίζει το θόρυβο της μέτρησης ζ μεταφέρεται αυτούσια στην εκτίμηση του φίλτρου, καθώς αυτό δεν μπορεί να την αποβάλλει, ισχύει, δηλαδή, $E\{\hat{N}_{KF}\} = E\{N^m\} = E\{N\} + b$.

Κεφάλαιο 5°

Στο 5° κεφάλαιο γίνεται μια απόπειρα ελέγχου της ποιότητας της εκτίμησης του αριθμού των οχημάτων στο σύνδεσμο με χρήση της μεθόδου της εκθετικής εξομάλυνσης. Χρησιμοποιώντας, πλέον, μόνο τις μετρήσεις χρονικής κατάληψης από το φωρατή που βρίσκεται στο μέσο του συνδέσμου (και όχι των δύο ακραίων φωρατών μέτρησης ροής) ο νέος εκτιμητής που αναπτύσσεται είναι $\hat{N}_{SM}(k) = K_{SM}N^m(k-1) + (1 - K_{SM})\hat{N}_{SM}(k-1)$, όπου N^m και \hat{N}_{SM} η μέτρηση και η εκθετικά εξομαλυσμένη εκτίμηση, αντίστοιχα, του αριθμού οχημάτων στο

σύνδεσμο, ενώ K_{SM} είναι η παράμετρος εξομάλυνσης. Τα αποτελέσματα των προσομοιώσεων που πραγματοποιήθηκαν με τη χρήση τόσο των μετρήσεων, όσο και του φίλτρου Kalman συγκρίνονται με αυτά που προέκυψαν από τη χρήση των εκθετικά εξομαλυμένων μετρήσεων.

Κεφάλαιο 6°

Στο 6° κεφάλαιο παρουσιάζεται το μαθηματικό μοντέλο που χρησιμοποιήθηκε στη μικροσκοπική προσομοίωση και αναφέρονται τα διάφορα σενάρια που αναπτύχθηκαν, ώστε να ανταποκρίνονται σε διάφορες κυκλοφοριακές συνθήκες.

Καταρχήν, ο εξεταζόμενος σύνδεσμος έχει μήκος 194 μέτρα και εμπεριέχει μόνο μία λωρίδα κυκλοφορίας. Το διακριτό (χρονικά) μαθηματικό μοντέλο που χρησιμοποιείται βασίζεται σε βήμα προσομοίωσης $T_{sim} = 0.25s$. Σε κάθε χρονική περίοδο $k = 1, 2, \dots$, υπολογίζεται για κάθε όχημα i η επιθυμητή του ταχύτητα $y_{d,i}$ από τη σχέση $y_{d,i} = \Lambda [\delta_i(k) - D]$, όπου $\Lambda = 0.7 s^{-1}$ και $D = 1m$, ενώ $\delta_i(k)$ ορίζεται η απόσταση (σε μέτρα) του πίσω μέρους ενός οχήματος με το μπροστινό μέρος του οχήματος που το ακολουθεί. Η μέγιστη δυνατή επιθυμητή ταχύτητα κάθε οχήματος ορίζεται ίση με $y_f = 16.5 m/s$.

Η επιτάχυνση $a_i(k)$ κάθε οχήματος i υπολογίζεται από τη σχέση $a_i(k) = \text{sat}\{g[y_{d,i}(k) - y_i(k)]\}$, όπου $g = 2 s^{-1}$ και η συνάρτηση $\text{sat}\{\cdot\}$ ορίζεται ως:

$$\text{sat}\{\eta\} = \begin{cases} a_{\max} & \text{if } \eta \geq a_{\max} \\ a_{\min} & \text{if } \eta \leq a_{\min} \\ \eta & \text{else} \end{cases}$$

με $a_{\max} = 1.5 \text{ m/s}^2$, $a_{\min} = -6 \text{ m/s}^2$ και την ταχύτητα κάθε οχήματος $y_i(k)$ να ανανεώνεται σύμφωνα με τη σχέση $y_i(k+1) = y_i(k) + a_i(k)T_{\text{sim}}$. Τελικά, η θέση κάθε οχήματος καθορίζεται από τη σχέση

$$x_i(k+1) = x_i(k) + T_{\text{sim}}y_i(k) + 0.5T_{\text{sim}}^2 a_i(k).$$

Όσον αφορά τα 5, συνολικά, διαφορετικά σενάρια που δημιουργήθηκαν, ο συνολικός χρόνος προσομοίωσης είναι ίδιος για όλα και φτάνει τα 5000 δευτερόλεπτα (περίπου 1,38 ώρες), ενώ ο χρονικός κύκλος του ανάντι φωτεινού σηματοδότη είναι πάντα 90 δευτερόλεπτα. Το μήκος των οχημάτων είναι μεταξύ $[3 \text{ m}, 5 \text{ m}]$, άρα για την εκτίμηση του αριθμού των οχημάτων χρησιμοποιείται $L^{\text{Ph}} = 4 \text{ m}$, ενώ ο χρόνος ενημέρωσης του φίλτρου είναι $T = 20 \text{ s}$. Στο βασικό σενάριο προσομοίωσης ο χρονικός κύκλος του κατάντι φωτεινού σηματοδότη είναι 20s, στα υπόλοιπα τρία είναι 40s, 60s και 90s, αντίστοιχα, ενώ σε ένα ειδικό σενάριο ο χρονικός κύκλος του κατάντι φωτεινού σηματοδότη αλλάζει στοχαστικά μεταξύ 10 και 90 δευτερολέπτων. Σε κάθε σενάριο οι διάρκειες των κόκκινων/πράσινων ενδείξεων των σηματοδοτών σε κάθε κύκλο αλλάζουν με στόχο την επίτευξη κάθε δυνατής κατανομής οχημάτων μέσα στο σύνδεσμο.

Τέλος, όσον αφορά το κριτήριο αξιολόγησης των αποτελεσμάτων της προτεινόμενης μεθοδολογίας χρησιμοποιείται το Σχετικό Μέσο Τετραγωνικό Σφάλμα (Relative Mean Square Error)

$$\text{RMSE} = 100\% \sqrt{K \sum_{k=1}^K [X - N(k)]^2} / \sum_{k=1}^K N(k)$$

όπου $N(k)$ είναι ο πραγματικός αριθμός οχημάτων στο σύνδεσμο και X είναι ο εκτιμώμενος αριθμός οχημάτων, είτε μέσω των μετρήσεων, οπότε $X = \hat{N}(k)$, είτε

μέσω του φίλτρου, οπότε $X = \hat{N}_{KF}(k)$, είτε μέσω των εκθετικά εξομαλυμένων μετρήσεων, οπότε $X = \hat{N}_{SM}(k)$.

Κεφάλαιο 7°

Στο 7° κεφάλαιο παρουσιάζονται αναλυτικά τα αποτελέσματα όλων των διερευνήσεων και γίνεται προσπάθεια να δοθούν απαντήσεις σε μια σειρά κρίσιμων ερωτημάτων:

- Πόσο βελτιώνεται η εκτίμηση των μετρήσεων (απλών ή εξομαλυμένων) με τη χρήση του φίλτρου Kalman; Πόσοι επιπλέον φωρατές απαιτούνται για να φτάσουμε την εκτίμηση του φίλτρου;
- Ποιο είναι το εύρος των κατάλληλων τιμών του παράγοντα K του φίλτρου; Ποια είναι η ευαισθησία της ποιότητας της εκτίμησης γύρω από τις τιμές αυτές;
- Πώς επηρεάζει την εκτίμηση ο χρόνος ενημέρωσης/εκτίμησης T και ποια τιμή είναι κατάλληλη για την πρακτική εφαρμογή της μεθόδου;
- Πώς αποδίδει το φίλτρο κάτω από διαφορετικές συνθήκες που αφορούν το N_{max} , τη ροή των οχημάτων και τη φωτεινή σηματοδότηση του συνδέσμου;
- Πώς επηρεάζουν την εκτίμηση οι διάφοροι θόρυβοι των μετρήσεων;

Για την απάντηση αυτών των ερωτημάτων παρουσιάζεται πληθώρα γραφημάτων και πινάκων που παρέχουν αξιόλογες πληροφορίες και οδηγούν σε σημαντικές παρατηρήσεις που αναδεικνύουν την ωφελιμότητα της χρήσης του φίλτρου Kalman.

Κεφάλαιο 8°

Στο 8° κεφάλαιο πραγματοποιείται μια σύνοψη ολόκληρης της εργασίας, τονίζονται τα σημαντικότερα στοιχεία της και προτείνονται κατάλληλες τιμές των μεγεθών που απαιτούνται για την πρακτική εφαρμογή της μεθόδου σε πραγματικό περιβάλλον. Σύμφωνα και με τα αποτελέσματα του προηγούμενου κεφαλαίου για την εφαρμογή της προτεινόμενης μεθοδολογίας σε πραγματικές συνθήκες πρέπει να γνωρίζουμε ότι:

- Το σύννηθες εύρος τιμών του χρόνου T είναι $[10s, 30s]$.
- Οι μετρήσεις χρονικής κατάληψης πρέπει να λαμβάνονται με ε ίσο με το μηδέν. Αν αυτό δεν είναι εφικτό, οι μετρήσεις πρέπει να πολλαπλασιαστούν με $L^{Ph} / (L^{Ph} + \varepsilon)$ πριν χρησιμοποιηθούν περαιτέρω, όπου το μέσο φυσικό μήκος των οχημάτων L^{Ph} είναι περίπου 4 μέτρα. Τυπικές τιμές για το ε είναι από 1 έως 2 μέτρα, ανάλογα με τους εγκατεστημένους φωρατές.
- Οι μετρήσεις χρονικής κατάληψης o_t^m πρέπει να μετατρέπονται σε αριθμό οχημάτων N^m μέσω της σχέσης $N^m = (\Delta l / L^{Ph}) o_t^m$, όπου ο παράγων $\Delta l / L^{Ph}$ ισούται με, και μπορεί να αντικατασταθεί από, το μέγιστο αριθμό οχημάτων N_{max} που μπορούν να βρεθούν σταματημένα μέσα στο σύνδεσμο. Στην περίπτωση ύπαρξης και φορτηγών μέσα στο σύνδεσμο υπάρχουν δύο δυνατότητες:
 - ο Το μήκος των φορτηγών να μη λαμβάνεται υπόψη στο χρησιμοποιούμενο μέσο μήκος οχημάτων, το οποίο παραμένει στα 4 μέτρα, οπότε το φίλτρο παρέχει εκτιμήσεις που αναφέρονται σε μονάδες επιβατικών οχημάτων (passenger car units), επιλογή η οποία και προτείνεται.

- ο Το μήκος των φορτηγών να λαμβάνεται υπόψη στον υπολογισμό του μέσου μήκους των οχημάτων, το οποίο γίνεται, φυσικά, μεγαλύτερο των 4 μέτρων, οπότε και το φίλτρο παρέχει εκτιμήσεις που αναφέρονται σε οχήματα που μπορεί να είναι είτε επιβατικά, είτε φορτηγά με μια προκαθορισμένη αναλογία.
- Οι προτεινόμενες τιμές του παράγοντα K βρίσκονται μεταξύ 0,05 και 0,3, αλλά εάν δεν είναι δυνατός ο λεπτομερής συντονισμός του (fine-tuning), τότε η τιμή $K = 0.1$ θεωρείται αρκετά ικανοποιητική και εύρωστη.
- Οι εκτιμήσεις του φίλτρου $\hat{N}_{KF}(k)$ βρίσκονται μεταξύ 0 και N'_{max} , όπου $N'_{max} = \Delta l / (L^{Ph} + D)$ είναι ο μέγιστος αριθμός ακίνητων οχημάτων που μπορεί να υπάρξουν μέσα στο σύνδεσμο, λαμβάνοντας υπόψη και την απόσταση ασφαλείας D μεταξύ δύο διαδοχικών οχημάτων.

Παραρτήματα

Τέλος, στο Παράρτημα Α παρουσιάζεται το αρχείο εισόδου δεδομένων του προγράμματος προσομοίωσης που αναπτύχθηκε, ενώ στο Παράρτημα Β δίνονται αναλυτικά οι χρονικές διάρκειες των κόκκινων/πράσινων των φωτεινών σηματοδοτών καθ' όλη τη διάρκεια των προσομοιώσεων και για κάθε σενάριο που εξετάστηκε.

ABSTRACT

The number of vehicles included in a metered motorway ramp or an urban signalized link at any time is valuable information for real-time control. This quantity is closely related to the space- and time-variables of the ramp or the link. Until now many people have been concerned with the relationships between instantaneous space-variables and (easily measurable) local time-variables in largely homogeneous and stationary traffic flow, as typically encountered in uninterrupted traffic conditions. But how do these relationships change under inhomogeneous and nonstationary traffic conditions? To answer this question a rather elaborate analysis and microscopic simulation investigation is conducted for signalized links with inherently strong traffic flow variations triggered by traffic signal switchings. A number of influencing factors when estimating space-variables from measured local time-variables is analysed and illustrated in detail. Considering these, a Kalman Filter is employed to produce reliable estimates of the number of vehicles based on real-time measurements of flow and occupancy provided by three loop detectors. The resulting vehicle-count estimator is tested via the same as before microscopic simulation for a variety of metered ramp scenarios and traffic conditions. The simulation investigations indicate a robust estimation performance with little effort required, which facilitates easy applicability of the method.

CHAPTER 1

GENERAL CONCEPTS

1.1 Traffic Congestion

Traffic congestion is a vexing problem felt by residents of most urban areas. Despite decades of effort and billions of euros worth of public spending to alleviate congestion, the problem appears to be getting worse. It seems that traffic congestion and cities go hand in hand. Everyone complains about being stuck in traffic; but no one seems to do anything about it. In particular, traffic engineers, transportation planners, and public officials responsible for metropolitan transportation systems are frequently criticized for failing to make a dent in congestion. Most importantly, while it is obvious that traffic congestion cost in time, what is less obvious, but still very real, is its cost in traffic deaths and injuries, poorer air quality, wasted fuel and lost productivity. Traffic congestion is not just a nuisance, lack of adequate capacity on the highways in these stress points actually causes traffic accidents, spews pollutants into the air in greater concentrations and requires the useless burning of fossil fuels.

Congestion occurs when traffic demand exceeds available capacity, although it is much more complex than simply stating that “too many vehicles are trying to use the road at the same time”. Congestion results from the interaction of many different factors or sources of congestion. It has several root causes that can be broken down into two main categories:

1. Too much traffic for the available physical capacity to handle.
2. Traffic influencing events like accidents, work zones, bad weather, special events and poorly timed traffic signals.

The level of congestion on a road is determined by the interaction of physical capacity with events that are taking place at a given time. For example, the effect of a traffic incident depends on how much physical capacity is present.

Nowadays, the effort in order to relieve congestion to metropolitan areas is permanent and a variety of strategies has been developed to deal with congestion. These strategies can be grouped as follows:

1. Adding road capacity where appropriate and requested. Adding more lanes to existing highways and building new ones has been the traditional response to congestion. In some metropolitan areas, however, it has become difficult to undertake major highway expansions because of funding constraints, increased construction cost, social effects and environmental constraints and opposition from local and national groups.
2. Operating existing capacity more efficiently. In recent years, new strategies that deal with the operation of existing highways have been adopted, rather than just building new infrastructure. Collectively referred to as Intelligent Transportation Systems (ITS), real-time control of transportation operations involves making changes from minute to minute, changing the operating methods or the policies that govern the use of the road and monitoring vehicles in real-time; there are numerous operations-based congestion mitigation strategies that are enhanced by the use of advanced technologies or ITS.

Although adding new physical capacity to highways is an important and effective strategy for alleviating congestion, it does not seem to ensure a long term solution and it causes manifold feedbacks. Having this in mind, the only measures that can be taken in order to control the traffic involve the application of operational improvements; a major area of great interest is the ramp metering.

1.2 Traffic Flow Process Control

Figure 1 illustrates the essential elements of a control loop which is the basis of any operational procedure in a traffic network. The traffic flow behavior in a freeway or an urban traffic network depends on some external quantities that are classified into two groups:

- Control inputs that are directly related to corresponding control devices (actuators), such as traffic signals, variable message signs, etc.; the control inputs may be

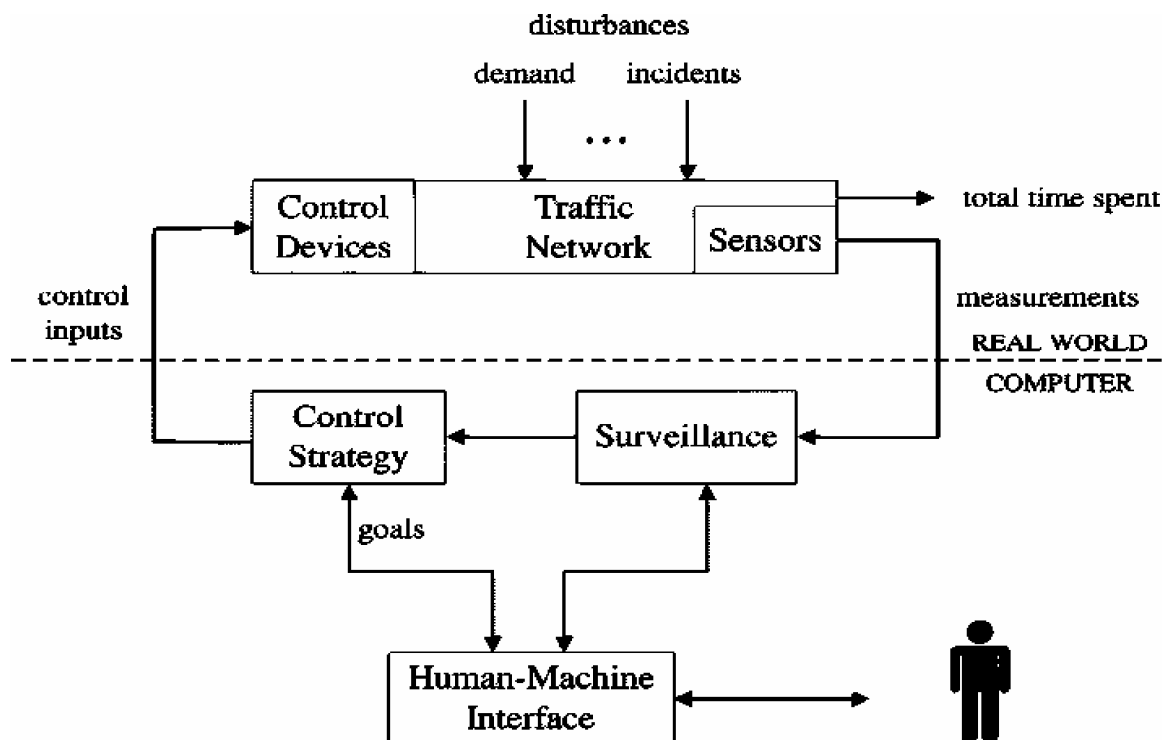


Figure 1: The control loop.

selected from an admissible control region subject to technical, physical, and operational constraints.

- Disturbances, whose values cannot be manipulated, but may, possibly, be measurable (e.g., demand) or detectable (e.g. incident) or predictable over a future time horizon.

The network's output or performance is measured via suitable indices, such as the total time spent by all vehicles in the network over a time horizon. The task of the surveillance is to enhance and to extend the information provided by suitable sensors (e.g. inductive loop detectors) as required by the subsequent control strategy and the human operators. The kernel of the control loop is the control strategy, whose task is to specify in real time the control inputs, based on available measurements/ estimations/ predictions, so as to achieve the prespecified goals (e.g. minimization of total time spent) despite the influence of various disturbances. If human operator undertakes this task, we have a manually controlled system. In an automatic control system, this task is undertaken by an algorithm (the control strategy). The relevance and efficiency of the control strategy largely determines the efficiency of the overall control system. Therefore, whenever possible, control strategies should be designed with care and via application of powerful and systematic methods of optimization and automatic control.

1.3 Loop Detectors

One of the most well used methods for measuring the flow in a link is the installation of loop detectors. A traffic loop detects metal objects such as vehicles and bicycles based on the change in inductance that they induce in the loop. The loop is an inductor in an LC circuit that is tuned to resonate at a certain frequency. A metal plate over the loop (like a vehicle) causes the magnetic flux to be shorted, reducing the inductance of the

loop. This causes a change in resonant frequency, which is detected and sent to the signal controller (see Figure 2).

Although advancements are under way that may make traffic loops obsolete some day, they still remain one of the most widespread measurement devices due to its low cost. Radar, infrared and sound detectors and systems based on video cameras are promising, but their construction and operation cost makes their implementation almost prohibitive. On the other hand, the employment of traditional loop detectors measuring time-occupancy has, also, some difficulties. In this case, a detector station (across all link lanes) positioned at a specific link location (e.g. at the signal stop line or in the middle or at the upstream end of the whole link) delivers (local) occupancy information that is not representative for the whole link. In other words, local time-occupancy measurements collected by loop detectors need to be translated into space-occupancy estimates that are directly related to the number of vehicles in the link.

Over the last years, the significant technological advances in electronic technology, coupled with excellent wire insulation for inductive loops, make both high performance

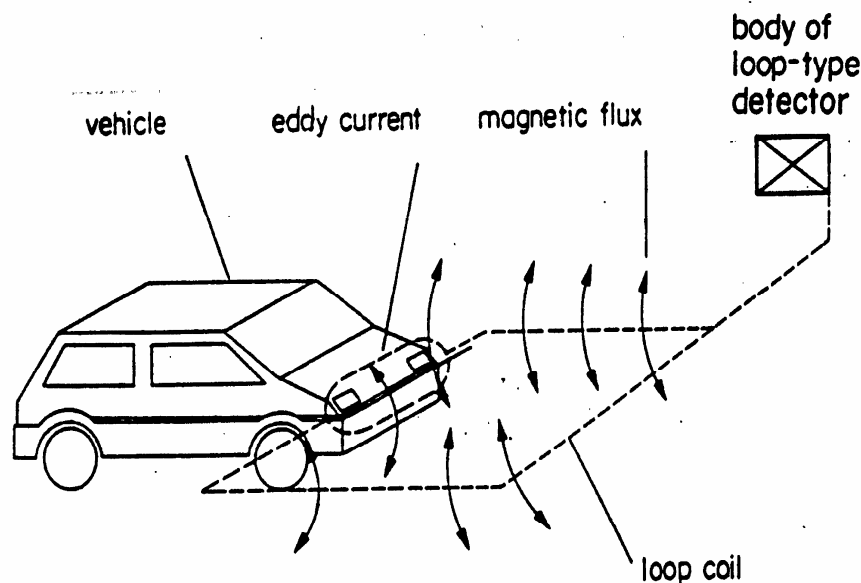


Figure 2: Operation of a loop-type detector.

and high reliability in vehicle detection possible. However, the facts of the locality of the time-occupancy measurements and their necessary modification in terms of number of vehicles may cause important inaccuracies at the final estimation of the number of vehicles in the link. That is why we first should examine more carefully some basic relationships between time-occupancy and space-occupancy.

1.4 Ramp Metering

Ramp metering is not a new freeway management technique. Various forms of ramp control were implemented during the late 1950's and through the 1960's in Chicago, Detroit and Los Angeles. By the early 1990's, ramp metering systems existed in twenty metropolitan areas within the United States, along with numerous cities around the world (Europe, New Zealand, Japan). Like in other technological areas, there is a gap between methodological advancements and practical implementations of ramp metering; this gap tends to increase as the methods become more sophisticated, but also more efficient. Nowadays Europe has a leading role in terms of methodological advancements in the area of ramp metering, but the number of operational metered ramps (less than 100 in total) is far less than in USA (about 2.500 metered ramps, thereof some 800 in Los Angeles and some 400 in Minneapolis).

Ramp metering can be defined as a method by which traffic seeking to gain access to a busy highway is controlled at the access point via traffic signals (local ramp metering-see Figure 3a). This control aims to maximize the capacity of the highway and prevent traffic flow breakdown and the onset of congestion. The operation of the traffic signals transmutes the predefined control strategy into action. Additionally, ramp metering can affect driver route choice and can be used to encourage alternative routes in corridor networks particularly where complimentary measures such as alternative rout signing

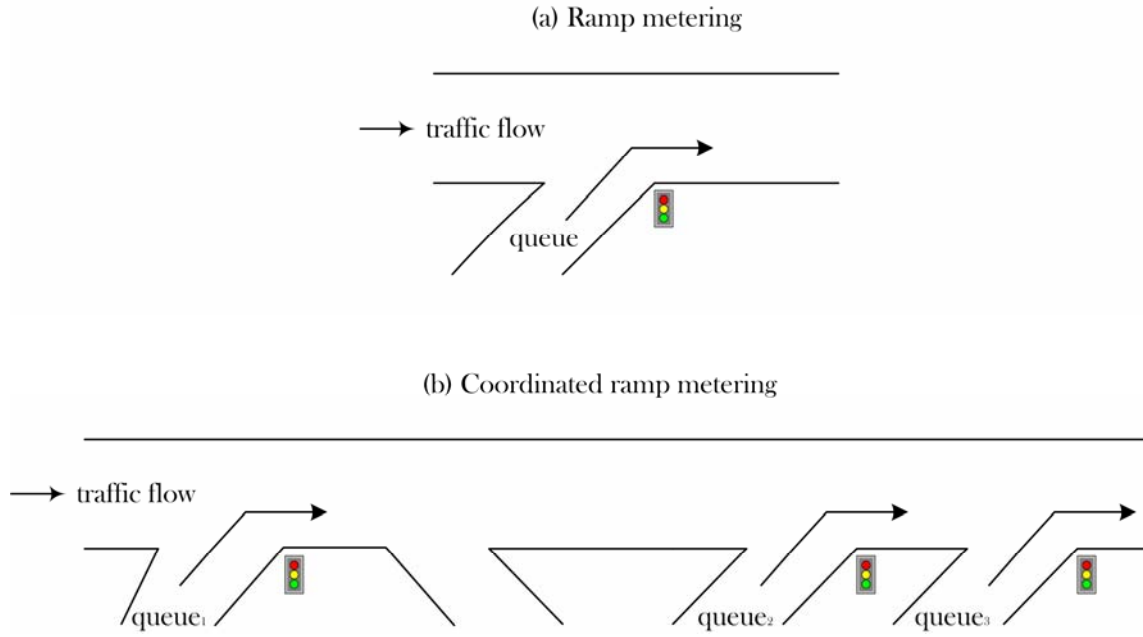


Figure 3: (a) Local and (b) coordinated ramp metering.

can be and are applied. The problem gets even more difficult when we talk about a number of on-ramps to a highway (coordinated ramp metering-see Figure 3b) and the suitable co-operation of ramp metering with signal control in order to maximize synergy and minimize mutual interference (e.g. queue spillbacks) is of major concern. Generally, the benefits of ramp meters include the reduction of congestion on the freeway at the entrance ramps, the reduction of accidents as cars merge from the entrance ramps onto the freeway, the ability to easily merge with less disruption to interstate traffic, and the reduction of vehicle emissions.

In practice, ramp metering systems have been extremely successful in reducing congestion and increasing safety. Although, ramp metering algorithms have some limitations, which researchers are working to eliminate. One problem is that existing algorithms react to rather than prevent bottlenecks. This causes oscillatory behavior, as a result of the time lag between detection and correction action. If an initial reaction to congestion leads to overly restrictive metering, excessive queue buildup may ensue. If the queue exceeds a certain length, it will interfere with the adjacent street traffic, otherwise, when a queue override is activated, freeway congestion will again increase,

and the process starts over. A proposed solution for the safeguard of the algorithmic prevention involves integrating traffic predictive capabilities into the metering logic; several such algorithms employ neural networks and Fuzzy Logic techniques, and can potentially delay or prevent bottleneck formation. The solution to the other problem, also known as the queue overflow problem, is two-fold: firstly, to estimate accurately the length of the queue on the ramp in order to accurately detect queue overflow problems and, secondly, in the case where there is a danger of queue overflow to efficiently control the ramp metering so that queue overflow is avoided.

One possible way to address the first problem (queue length estimation) is by using video sensors, which calculate quite accurately the number of the vehicles within the ramp (or the link in case of an urban network); however, video sensors may be an expensive solution especially when there is a need for placing a large number of them (coordinated ramp metering, long ramps with curvature requiring more than one video sensor, etc.). A second approach is to use the detector flow measurements in the entrance and the exit of the ramp as well as the measurements coming from detectors within the ramp in order to come up with accurate estimates of the number of vehicles within the ramp. Kalman-filter estimation techniques can be used for this purpose.

1.5 Kalman Filter

The filter is named after its inventor, Rudolf E. Kalman (1960) and its first application was by NASA to the problem of trajectory estimation for the Apollo program. Since then a wide variety of Kalman filters have been developed, from Kalman's original formulation, now called simple Kalman filter, to Schmidt's extended filter (1970), the information filter and a variety of square-root filters developed by Bierman, Thornton (1980) and many others. Kalman Filter has been the subject of extensive research and

application, particularly in the area of autonomous or assisted navigation. This is likely due in large part to advances in digital computing that made the use of the filter practical, but also to the relative simplicity and robust nature of the filter itself.

The Kalman Filter is essentially a set of mathematical equations that implement a predictor-corrector type estimator that is optimal in the sense that it minimizes the estimated error covariance when some presumed conditions are met. Rarely do the conditions necessary for optimality actually exist, and yet the filter apparently works well for many applications in spite of this situation. That is why Kalman Filter has also been applied in areas as diverse as aerospace, marine navigation, nuclear power plant instrumentation, demographic modeling, manufacturing and in a wide range of engineering applications from radar to computer vision.

Kalman Filter is optimal with respect to virtually any criterion that makes sense. According to Maybeck (1979), there are two good reasons to choose a Kalman Filter over other approaches. The first aspect of its optimality is that the Kalman Filter incorporates all information that can be provided to it. It processes all available measurements, regardless of their precision, to estimate the current value of the variables of interest, with use of:

1. knowledge of the system and measurement device dynamics,
2. the statistical description of the system noises, measurements errors, and uncertainty in the dynamic models, and
3. any available information about initial conditions of the variables of interest.

Rather than ignoring any of these outputs, a Kalman Filter could be built to combine all of this data and knowledge of the various systems' dynamics to generate an overall best estimation.

The other important reason is that Kalman Filter is a recursive filter which estimates the state of a dynamic system from a series of incomplete and noisy measurements. The term "recursive" means that only the estimated state from the previous time step and the current measurement are needed to compute the estimate for the current state. The filter does not require all previous data to be kept in storage and reprocessed every time a new measurement is taken. In contrast to batch estimation techniques, no history of observations and/or estimates is required and this is of vital importance to the practicality of filter implementation.

A Kalman Filter combines all available measurement data, plus prior knowledge about the system and measuring devices, to produce an estimate of the desired variables in such a manner that the error is minimized statistically. In other words, if we were to run a number of candidate filters many times for the same application, then the average results of the Kalman Filter would be better than the average results of any other. Conceptually, what any type of filter tries to do is obtain an "optimal", meaning that it minimizes errors in some respect.

According to Maybeck (1979) there are three basic reasons why deterministic system and control theories do not provide a totally sufficient means of performing an analysis and design of any physical system. First of all, no mathematical system is perfect, as it depicts only those characteristics of direct interest. Secondly, the dynamic systems are driven not only by our own control inputs, but also by disturbances which we can neither control nor model deterministically, and thirdly, sensors do not provide perfect

and complete data about a system. As a result, assuming perfect knowledge of all quantities necessary to describe a system completely and assuming perfect control over the system is a naive approach. This encourages the following questions:

- How do you develop system models that account for these uncertainties in a direct and proper way?
- Equipped with such models and incomplete, noise-corrupted data from available sensors, how do you optimally estimate the quantities of interest to you?
- In the face of uncertain system descriptions, incomplete and noise-corrupted data, and disturbances beyond your control, how do you optimally control a system to perform in a desirable manner?
- How do you evaluate the performance capabilities of such estimation and control systems, both before and after they are actually built?

The answer to all these is, actually, the Kalman-filtering approach. In Figure 4 is shown the general form of the filters' application. According to this, for known data $\mathbf{u}(k)$, disrupted with noise $\mathbf{g}(k)$, a process is applied and considering various possible

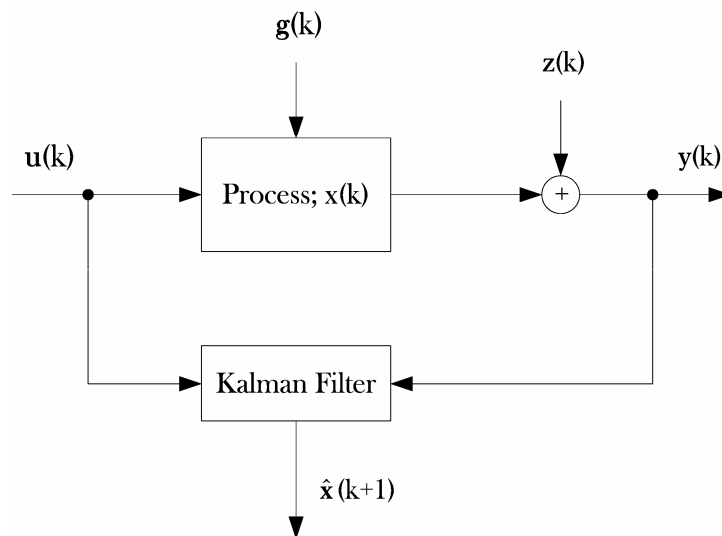


Figure 4: Kalman Filter application.

disturbances/noises $\mathbf{z}(k)$ we end to a measurement $\mathbf{y}(k)$. The use of this measurement and of the starting data give, through the application of the filter, the estimation (or prediction) $\hat{\mathbf{x}}(k+1)$ of the process state at the next time step. Notify that the filter is applied in discrete time.

1.6 Microscopic Simulation Model

Simulation is a method of imitation of operations in real-world processes and systems. It can be used to describe and analyse the behaviour of an existing and conceptual system especially in traffic and transportation. Hence this can also be a fruitful method for searching optimal solution by integrating with efficient algorithms and a part of a system supporting decision making.

One of the most common simulation methods is the microscopic simulation, a technique which provides a realistic measure of (traffic) flow on a network as well as any other parameter of interest. In the past describing the traffic was possible using macroscopic approach which perceived the traffic as a fluid flowing along the carriageway. Microscopic approach allows studying the traffic flow by modelling the motion of a particular vehicle. The mathematical models used in it are called car following models.

Every microscopic simulation model offers a set of advantages as it is possible to imitate with great accuracy the real conditions of a network. Since each vehicle can be autonomous, the realism of each vehicle's behaviour can correspond to the geometry of the road network as well as each vehicle's and its driver's behaviour can be determined by individual set of mathematical rules according to its type. Car-following, overtaking, driver's awareness and aggressiveness and every other parameter can be modelled in

this way. What must be noticed is that the more the accuracy of the model is, the higher the computational requirements and the time cost of the simulation. However, the crucial aspect of any microscopic simulation model is the calibration of parameters describing the vehicles and the environment.

1.7 The Number of Vehicles Estimation Problem

The vehicle-count estimation problem is illustrated in Figure 5. Figure 5a depicts the relevant signal and detector configuration on the signalized link. Two traffic signals are located, respectively, just upstream of the upper boundary of the link and at the downstream end of the link. The upstream signal (if it exists) determines the traffic demand approaching the link while the downstream signal controls the vehicle flow exiting the link. Obviously, whenever the link demand is larger than the link outflow, a queue is built. It is also shown in Figure 5a that three detectors are installed: at the

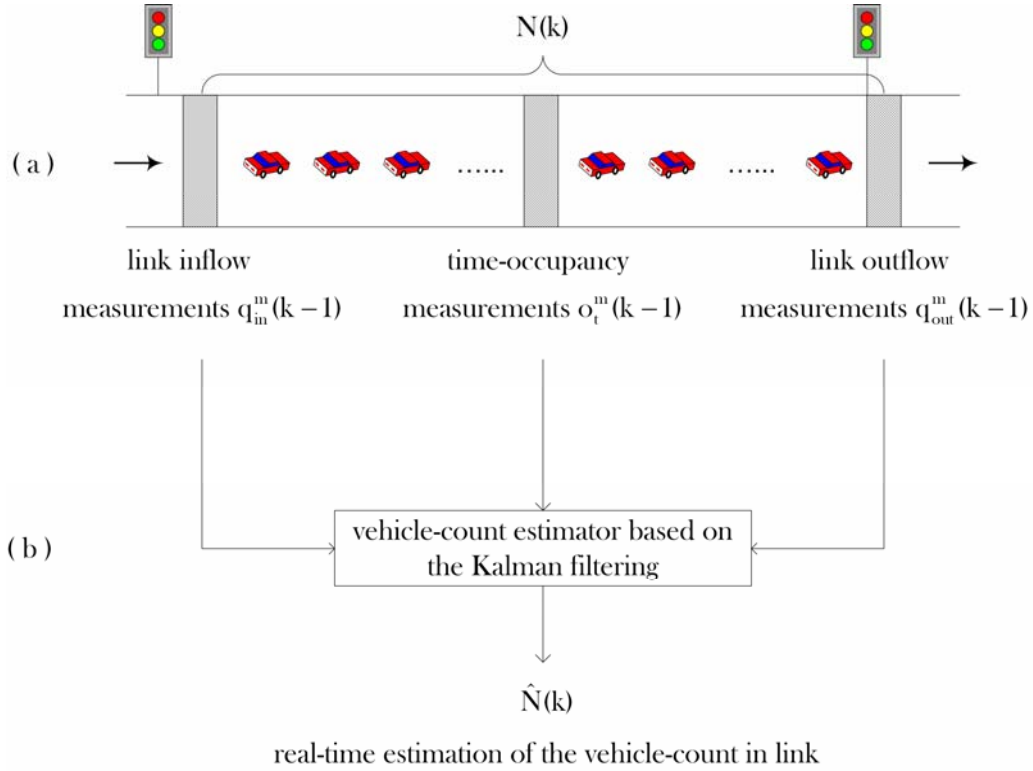


Figure 5: Vehicle-count estimation: (a) the signal and detector configuration on the link; (b) the link vehicle-count estimation.

upstream end of the link, at the downstream end of the link, and in the middle of the link. Both boundary detectors provide flow measurements, while the middle detector provides time-occupancy measurements. The basic structure of the ramp queue estimation is shown in Figure 5b:

- The estimator is fed in real time with the flow and time-occupancy measurements from the link detectors.
- The estimator delivers in real time the estimated number of vehicles in the link (between the two boundary detectors).

CHAPTER 2

TIME-OCCUPANCY AND SPACE-OCCUPANCY

2.1 Introduction

Space-occupancy (portion of highway lane length covered by vehicles), traffic density (vehicles per kilometer) and vehicle-count in (urban or freeway) links are important quantities for both traffic flow modeling and traffic control. These quantities are usually needed for highway stretches or links of few hundred meters in length, while the vast majority of available measurement devices provide only local measurements at specific highway locations. Space measurements covering a few hundred meters may in principle be collected by video sensors, but this possibility is obscured due to visibility obstacles, geometrical highway conditions, image-processing algorithm accuracy and, last but not least, cost. On the other hand, emerging technologies (see, e.g. Cheung, et. al., 2005) may soon lead to cheaper, convenient (e.g. wireless) and reliable local measurement devices which could foster a space-denser deployment of sensors along links. In any case, the investigation and understanding of the relationships between locally measurable quantities, such as time-occupancy (portion of time a local sensor is covered by vehicles), and the aforementioned space-extended quantities has a high practical significance.

Despite some related efforts in the past (Wardrop, 1952; Edie, 1974; Bhourri, et al., 1988; Banks, 1995; Cassidy and Coifman, 1997; Kim and Hall, 2004), that will be specifically addressed later, there are several issues regarding the relationships between local and space variables that are still unexplored, particularly in the case of signalized links (e.g. urban roads or freeway on-ramps) where the space-distribution of vehicles is strongly inhomogeneous due to the traffic signal switchings. After a brief review of available knowledge, the research focus is turned to signalized links to reveal some useful relationships between space-occupancy, vehicle-count and time-occupancy measurements. The derived findings are illustrated in a series of microscopic simulation experiments in CHAPTER 7. This research is a first step towards the development of a reliable low-cost estimator for the vehicle-count in signalized links by use of Kalman filtering which is reported in CHAPTER 4.

2.2 Inductive Loop Measurements

Although information on space-variables (such as density, space-occupancy, vehicle-count in links) is of major importance for various uses (including real-time control), genuine space measurements are costly or little accurate. Consequently, local measurement devices (mostly inductive loop detectors) are used to deliver traffic flow measurements in form of time-variables (e.g. time-occupancy) at specific highway locations. In fact even video sensors are frequently used as virtually local (covering few meters of pavement) or loop-emulating, rather than extended-space (few hundred meters) measurement devices.

When a vehicle j , $j = 1, \dots, N$, passes over a loop detector, an electric pulse (Figure 6a) is produced, whose shape may depend on vehicle length, height and further characteristics. The produced pulse is digitized by setting all values higher (lower) than a

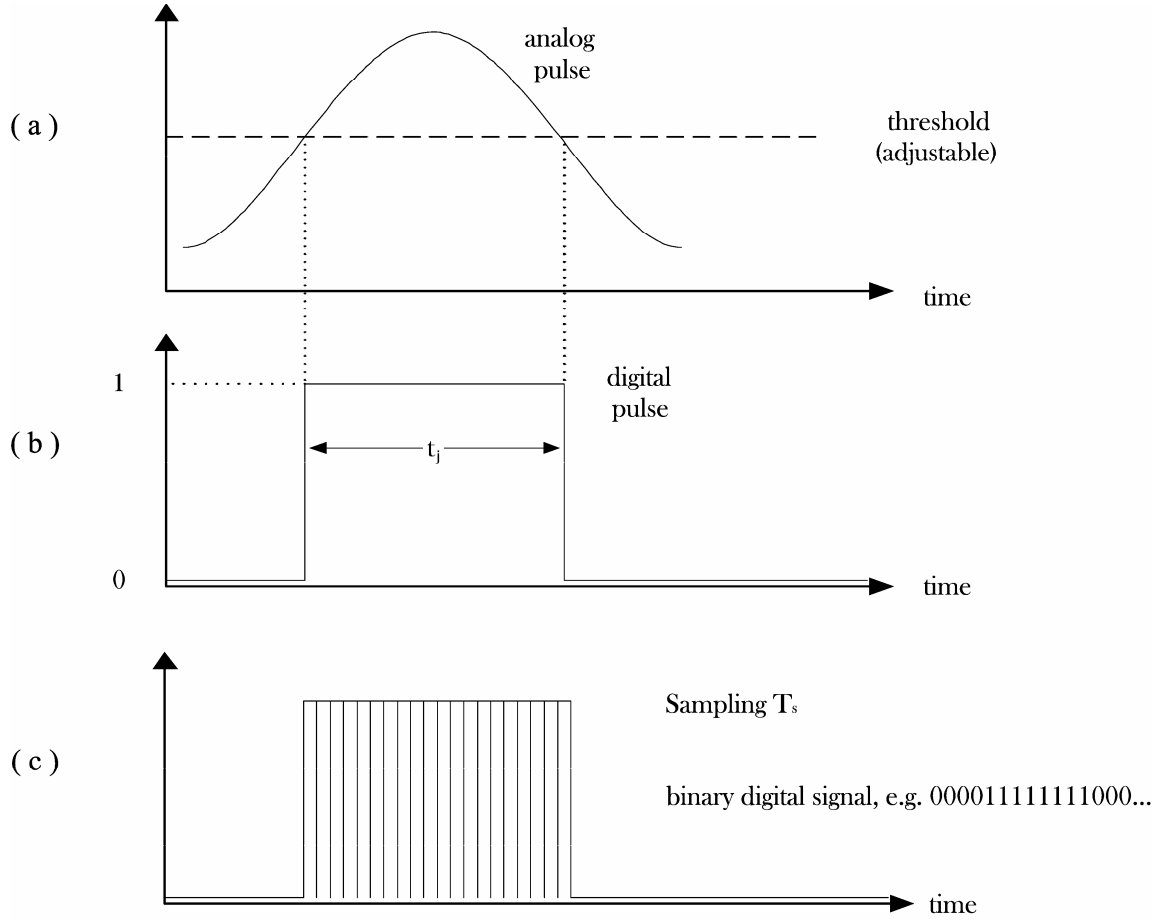


Figure 6: Time-occupancy measurements signals.

threshold equal to one (zero), see Figure 6b, whereby the threshold may be adjustable by the user. The duration t_j of the digital pulse is obviously inversely proportional to the vehicle speed y_j

$$t_j = L_j / y_j \quad (2.1)$$

where L_j is usually called the electrical or effective vehicle length which is closely related (but is not necessarily equal) to the physical vehicle length L_j^{Ph} . The digital signal is transformed to a binary series of 1's and 0's (Figure 6c) based on a sufficiently short sampling period T_s , e.g. $T_s = 10 \text{ ms}$, for further processing. The related time-measurements are produced with an update period T (e.g. $T_s = 10 \text{ s}$) and refer to the time period $[kT, (k+1)T]$ where $k = 0, 1, \dots$ is the discrete time index. Let $N(k)$ be the

number of 0-1 passages in the binary signal of Figure 6c during period k , in which case we obtain the flow measurement $q(k) = N(k)/T$; while for the time-occupancy we have

$$o_t(k) = \sum_{j=1}^{N(k)} t_j(k)/T = \sum_{j=1}^{N(k)} L_j/(y_j(k)T) \approx n(k)T_s/T \quad (2.2)$$

where $n(k)$ is the number of 1's counted during period k .

In traffic flow theory, the relevant traffic flow variables density, mean speed and flow (or traffic volume) are often defined as abstract mathematical quantities that are continuous in both space and time, i.e. $\rho(x,t)$, $v(x,t)$ and $q(x,t)$, respectively; where x and t are continuous space and time arguments, respectively. For more practical considerations, however, it is quite usual to consider density $\rho(k)$ (number of vehicles included in a section of length Δ at time instant kT , divided by T) and mean speed $v(k)$ (of vehicles in the same section at the same time instant) as instantaneous space-variables; while the flow $q(k)$ is defined for a specific location over a time period T , as mentioned earlier.

2.3 Wardrop's Approach

In an attempt to span a bridge between local time-measurements and instantaneous space-measurements, Wardrop (1952) proposed the following idealized consideration. Assume that the traffic flow consists of (space-homogeneous and stationary) sub-streams $j=1, \dots, C$, each with corresponding sub-flow q_j and equal vehicle speeds y_j ; the time-headway in each sub-stream is then $1/q_j$, hence the space-headway is y_j/q_j and the sub-stream density $\rho_j = q_j/y_j$ is the inverse of the space-headway. Taking this approach to the limit of one-vehicle sub-streams $j=1, \dots, N$ (the discrete time argument

k is suppressed for brevity), we have $q_j = 1/T$ and $\rho_j = 1/(y_j T)$ and hence the global traffic variables

$$q = \sum_{j=1}^N q_j = N/T \quad (2.3)$$

$$\rho = \sum_{j=1}^N \rho_j = \sum_{j=1}^N 1/(y_j T) \quad (2.4)$$

where y_j , $j=1, \dots, N$, are locally measured individual vehicle speeds. Thus the (space) mean speed

$$v = \sum_{j=1}^N \rho_j \Delta y_j / (\rho \Delta) = q / \rho = N / \sum_{j=1}^N 1/y_j \quad (2.5)$$

turns out to be equal to the harmonic average of passing vehicle speeds y_j while the time mean speed is equal to the arithmetic average of y_j . Note that (2.4), (2.5) actually span a bridge between locally measured quantities (passage of N vehicles and vehicle speeds y_j) on one hand and space quantities ρ , v on the other hand; albeit under the mentioned, somewhat vague stationarity and homogeneity assumptions, see also Hall (2002).

2.4 Edie's Definition

Edie (1965; 1974) proposed a generalized definition of traffic variables based on individual vehicle trajectories within a finite space-time window (Figure 7) whereby the traffic flow q equals the total vehicle mileage (in veh·km) within the window, divided by the window surface ΔT ; the density ρ equals the total time spent by all vehicles (in veh·h) within the window, divided by ΔT ; and $v = q / \rho$. Edie's definitions reduce to

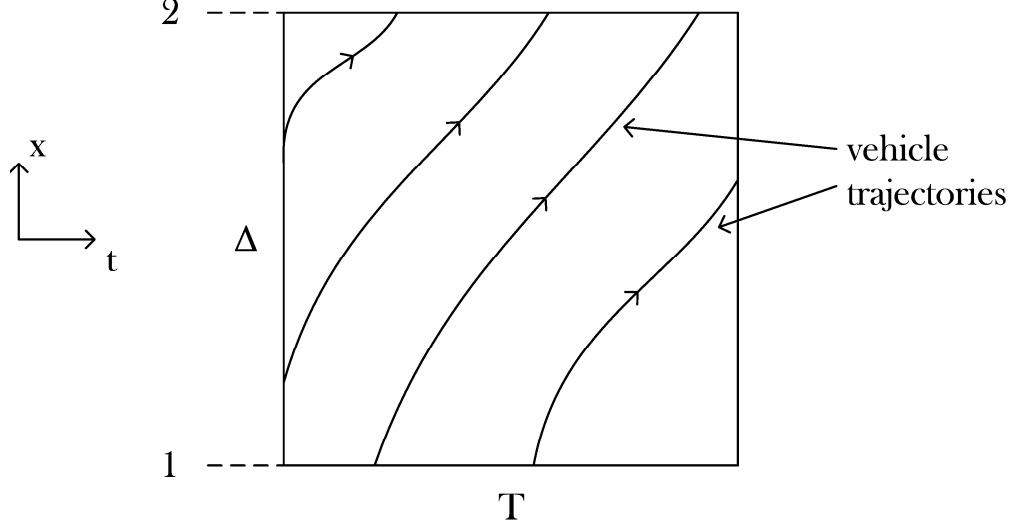


Figure 7: Illustration of Edie (1965; 1974) definition of traffic variables.

the definitions of section 2.2 for the space-variables ρ and v , for $T \rightarrow dt$; and for the time-variable q , for $\Delta \rightarrow dx$.

Assume that a local measurement device is available at location 1 (Figure 7), measuring vehicle passage and speeds over period T . We introduce the following assumptions that may actually hold if Δ is made sufficiently small:

- All vehicles crossing 1 also cross 2 within the same window and vice versa.
- Vehicles maintain a constant speed between locations 1 and 2.

Under these assumptions we obtain from Edie's definitions

$$q = N\Delta / (\Delta T) = N / T$$

$$\rho = \left(\sum_{j=1}^N \Delta / y_j \right) / (\Delta T) = \sum_{j=1}^N 1 / (y_j T)$$

$$v = q / \rho = N / \sum_{j=1}^N 1 / y_j$$

i.e. we re-gain (2.3)-(2.5), albeit under more clear (but also less general) assumptions than by use of Wardrop's approach.

Note that the above assumptions may only hold if $\Delta \rightarrow dx$, in which case ρ and v resulting from (2.4) and (2.5), respectively, become local variables, defined over a finite period T and infinitesimal space. The validity of these values over a (more or less) extended space around the measurement location depends on the level of stationarity and homogeneity of traffic flow. Thus, (2.4) and (2.5) may be useful for freeway traffic flow, where departures from stationary and homogeneous conditions are usually moderate; but in signalized links, where traffic conditions are inherently nonstationary and inhomogeneous due to the traffic light switchings, densities (2.4) and space mean speeds (2.5) calculated from local measurements are not expected to be representative for the whole link.

2.5 Mean Effective Vehicle Length

Using (2.2) and (2.4), Papageorgiou (1987) and eventually Bhourri et al. (1988) derived the following relationship between time-occupancy and density on a highway lane

$$\alpha_t = L\rho \tag{2.6}$$

where L is the mean effective vehicle length that has to be defined in a somewhat peculiar way

$$L = \sum_j (L_j / y_j) / \sum_j (1 / y_j). \tag{2.7}$$

Note that, if all vehicles have the same speed, (2.7) yields L equal to the arithmetic average of L_j ; while, if all vehicles have the same effective length \bar{L} , (2.7) yields $L = \bar{L}$

independently of the vehicle speeds y_j . The formulas (2.6), (2.7) were also included in the notes by Papageorgiou and Ioannou (1993) of a Short Course that was taught in Los Angeles and the University of California at Berkeley in 1993. The same formulas were also derived by Cassidy and Coifman (1997) on the basis of the traffic variable definitions of Edie (1965). Equations (2.6), (2.7) indicate that the measured time-occupancy is roughly proportional to traffic density, which was empirically verified by Cassidy and Coifman (1997) and, more recently, by Kim and Hall (2004) by use of freeway data. It is evident from (2.7) that, even under stationary and homogeneous conditions, departures from proportionality may be observed due to either variable traffic composition (in terms of effective vehicle lengths L_j) or inhomogeneous vehicle speeds y_j .

2.6 Space-Occupancy

Space-occupancy $o_s \in [0,1]$ may be defined as the portion of highway lane length covered by vehicles. This leads to the interesting question of possible equivalence or otherwise of space- and time-occupancies (recall that space and time mean speeds are structurally different from each other). Applying again Wardrop's approach, we assume that all vehicles belonging to sub-stream j have equal physical vehicle length L_j^{ph} and equal effective length L_j . Thus the sub-space-occupancies $o_{s,j}$ are given by

$$o_{s,j} = L_j^{ph} \rho_j. \quad (2.8)$$

On the other hand, we obtain for the sub-time-occupancies $o_{t,j}$ from (2.2)

$o_{t,j} = q_j L_j / y_j = L_j \rho_j$, and, because $o_t = \sum_j o_{t,j}$ and $o_s = \sum_j o_{s,j}$, time-occupancy o_t equals the space-occupancy o_s if the effective vehicle lengths L_j equal the

corresponding physical vehicle lengths L_j^{Ph} . Note that, if $L_j^{\text{Ph}} = L_j \forall j$, then we deduce from (2.1) that the loop detector effectively shrinks to a line of zero length and the time-occupancy signal is on for as long as that line is covered by a passing vehicle. In practical terms, if the threshold of Figure 6a is chosen such that $L_j \approx L_j^{\text{Ph}}$, then the produced time-occupancy will approximate the space-occupancy (provided the traffic conditions do not change significantly in a space-time window around the measurement location and period). Using (2.8) we have

$$o_s = \sum_j o_{s,j} = \sum_j L_j^{\text{Ph}} \rho_j = \sum_j L_j^{\text{Ph}} q_j / y_j$$

and taking the limit $q_j = 1/T$ we have

$$o_s = \frac{1}{T} \sum_j L_j^{\text{Ph}} / y_j. \quad (2.9)$$

From (2.4), (2.9) we obtain

$$o_s = \left[\sum_j (L_j^{\text{Ph}} / y_j) / \sum_j (1 / y_j) \right] \rho \quad (2.10)$$

which suggests that (2.6) applies to the space-occupancy as well, under an analogous definition of the mean physical vehicle length as in (2.7); moreover, (2.6), (2.7) and (2.10) confirm that o_s equals o_t if $L_j^{\text{Ph}} = L_j \forall j$.

Note that (2.10) may appear paradoxical because the space-occupancy on a road stretch of length Δ holding N vehicles, is by definition

$$o_s = \sum_j L_j^{\text{Ph}} / \Delta = \rho \sum_j L_j^{\text{Ph}} / N \quad (2.11)$$

which looks different than (2.10). The resolution of this paradox is that the sum in (2.10) addresses vehicles passing a detector location over time while the sum in (2.11) addresses vehicles being included in a road stretch. Under the same token, the space mean speed equals the harmonic average of the speeds of vehicles passing a detector location over time but is also equal to the arithmetic average of the speeds of vehicles included in a road stretch.

CHAPTER 3

SIGNALIZED LINKS

3.1 Introduction

All formulas derived in CHAPTER 2 to relate measurable time-variables with hardly measurable space-variables, hold approximately under reasonably homogeneous and stationary traffic conditions that are usually encountered in uninterrupted traffic flow, e.g. on freeways. These relationships, however, may break down if traffic is strongly nonstationary and inhomogeneous as, for example, in signalized links due to frequent traffic interruptions caused by the traffic signals. This section derives some relationships between space- and time-variables for this more general case.

3.2 Space-Occupancy Relationships

Consider a single-lane link of length Δ and define the instantaneous space-occupancy

$$o_s(x) = \begin{cases} 1 & \text{if } x \text{ is covered by a vehicle} \\ 0 & \text{if not} \end{cases}$$

where $0 \leq x \leq \Delta$ is the space argument. On the other hand, let o_s be the instantaneous space-occupancy for the whole link, in which case

$$o_s = \frac{1}{\Delta} \int_0^{\Delta} o_s(x) dx \quad (3.1)$$

holds. For example, if there is exactly one vehicle of length L travelling on the link, then $o_s = L/\Delta$, and, indeed, the right-hand side of (3.1) also yields

$$\frac{1}{\Delta} \int_0^{\Delta} o_s(x) dx = \frac{1}{\Delta} \int_{\chi}^{\chi+L} dx = L/\Delta$$

where χ is the location of the vehicle's rear end.

According to section 2.6, the local space-occupancies $o_s(x)$ may be approximated by time-occupancy measurements in a sufficiently small space/time-window if $L_j^{\text{Ph}} = L_j$.

More specifically, we assume that:

- The update period T of the time-occupancy measurements is sufficiently small to approximate instantaneous values, such that $o_t = 1$ if a detector is activated by a vehicle or $o_t = 0$ otherwise.
- There are M time-occupancy measurements along the link according to Figure 8 with spacing $d = \Delta/M$ among them (space sampling).
- We have $L_j^{\text{Ph}} = L_j$.

Under these assumptions, the space-occupancy $o_s(x^i)$ in an area of length d around a measurement location x^i , $i=1, \dots, M$, may be approximated by the corresponding measured time-occupancy $o_t(x^i)$ while (3.1) may be approximated by

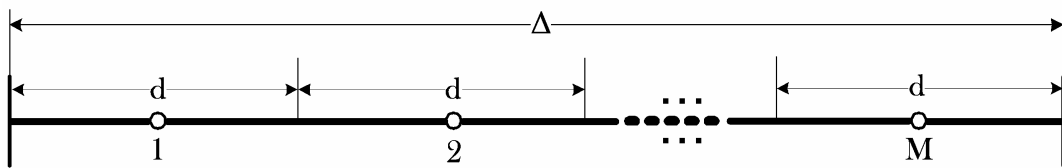


Figure 8: Placement of M internal detectors for time-occupancy.

$$o_s \approx \sum_{i=1}^M o_s(x^i) d / \Delta = \sum_{i=1}^M o_t(x^i) / M. \quad (3.2)$$

Based on (3.2), the instantaneous space-occupancy on a link may be approximated by a series of time-occupancy measurements with $L_j = L_j^{\text{Ph}}$ and very short update period T ; the level of approximation accuracy depending on the corresponding spacing d .

3.3 Bias-free Estimates

We will now turn our attention to the case where the number of utilized detectors M is small or even 1. Where should the detector(s) be best placed on the link? To elaborate on this question we denote $E\{o_s(x)\}$ to be the average over a reference period (e.g. over a signal cycle or over a peak period or over a whole day) of the space-occupancy at location x ; and we denote $E\{o_s\}$ the corresponding average of the space-occupancy for the whole link. From (3.1) we have

$$\int_0^{\Delta} E\{o_s(x)\} dx = \Delta E\{o_s\}. \quad (3.3)$$

The value of $E\{o_s(x^i)\}$ at a specific location x^i will tend to be higher (lower) if vehicles tend to spend relatively longer (shorter) times to pass location x^i as compared to other link locations. In fact, if each vehicle would travel on the link with constant speed, then all $E\{o_s(x)\}$ would be equal among them $\forall x$, and, from (3.1), we would have $E\{o_s\} = E\{o_s(x)\} \forall x$ (which verifies (3.3)).

On signalized links, the average speed of vehicles is expected to be lower for increasing x due to queue forming in the downstream part of the link during the red signal; hence the ratio $E\{o_s(x)\}/E\{o_s\}$ is expected to be monotonically increasing with x . Note,

however, that we may have different ratio profiles for different reference periods, e.g. for a.m. peak or p.m. peak or off-peak etc., whereby each profile satisfies (3.3). Figure 9 illustrates the above relationships for a certain reference period. The horizontal line reflects the average value $E\{o_s\}$ of the whole-link space-occupancy while the displayed curve corresponds to $E\{o_s(x)\}$, $0 \leq x \leq \Delta$. Because the surfaces below the horizontal line and the curve are equal due to (3.3), there is at least one location \bar{x} at which $E\{o_s\} = E\{o_s(\bar{x})\}$. Thus a bias-free estimate of the whole-link space-occupancy may be obtained from one local $o_s(\bar{x})$ or, approximately, from one local $o_t(\bar{x})$ -measurement of time-occupancy. Unfortunately the precise location \bar{x} may be unknown or may change for different reference periods.

Figure 9 also illustrates that, in view of (3.3), the difference between $E\{o_s\}$ and $E\{o_s(x)\}$ at specific locations will be smaller for flatter $E\{o_s(x)\}$ -profiles; and that one single measurement location x close to $\Delta/2$ (i.e., at the middle of the link) is likely to deliver reasonably low-biased estimates of o_s under various different practically-relevant $E\{o_s(x)\}$ -profiles. More generally, if M detectors are available (but the exact $E\{o_s(x)\}$ -profile is unknown) they should be placed symmetrically around the middle of the link

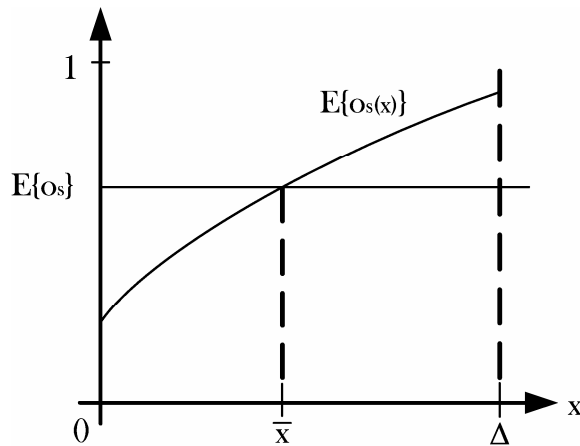


Figure 9: $E\{o_s\}$ and $E\{o_s(x)\}$ for a signalized link.

(according to Figure 8) so as to reduce the estimation bias resulting from (3.2).

In conclusion, a number M of detectors (Figure 8) measuring time-occupancy (with $L_j = L_j^{\text{Ph}}$ and T very small) may be used for estimating the instantaneous space-occupancy on a signalized link according to (3.2). The higher M in (3.2), the better the resulting estimate in terms of both average value and error variance. If M is small or even 1, then the detector placement according to Figure 8 is likely to lead to reasonably low average error (bias), but the variance of the error may be high as it will be further illustrated in the following general example.

3.4 A General Example

Consider one single vehicle travelling (with any speed) on a link of length Δ . Assume that $M = \Delta/d$ loop detectors are installed along the link with equal distances d among them according to Figure 8, and are operated with very small T . Assume also that we have $L^{\text{Ph}} = L$ for the vehicle, i.e. the effective lengths of the detectors are equal to zero. Let $L/d = n + m$ where $n = \text{int}(L/d)$.

The link's space-occupancy at any time is obviously $o_s = L/\Delta$. We wish to estimate this space-occupancy (that is difficult to measure directly) by use of the detector time-occupancies. More specifically, we produce the space-occupancy estimate \hat{o}_s as the arithmetic average of all detector time-occupancies, according to (3.2). If $m > 0$, the vehicle occupies at any time either n or $n+1$ detectors; and the respective estimation \hat{o}_s at any time is then equal to either nd/Δ (underestimation) or $(n+1)d/\Delta$ (overestimation). Assuming (for simplicity) that the vehicle has a constant speed, i.e. it may be present with the same probability at any location between two successive detectors, it is not difficult to deduce that it will activate at any time n detectors with

probability $1 - m$ and $n + 1$ detectors with probability m . Thus, while virtually all estimates \hat{o}_s are subject to (negative or positive, respectively) error at any time, the expected value of \hat{o}_s is

$$E\{\hat{o}_s\} = [(1 - m)n + m(n + 1)]d/\Delta = (n + m)d/\Delta = L/\Delta = o_s$$

i.e., the estimate \hat{o}_s is bias-free. On the other hand, the absolute error $|o_s - \hat{o}_s|$ at any time is either $|L/\Delta - (nd/\Delta)| = md/\Delta$ or $|L/\Delta - [(n + 1)d/\Delta]| = (1 - m)d/\Delta$, while the relative error $|o_s - \hat{o}_s|/o_s$ at any time is either md/L or $(1 - m)d/L$, and the variance of the absolute error is $m(1 - m)d^2/\Delta^2$. Note that $0 \leq m \leq 1$ and hence the maximum relative estimation error cannot be larger than d/L . From these we may conclude:

- The maximum absolute and relative estimation errors are proportional to the inter-detector distance d ; both errors tend to zero in the limit $d \rightarrow 0$.
- The estimation error is equal to zero at any time if $m = 0$, i.e., if the vehicle length L is an exact multiple of d ; this is because in this special case the vehicle activates exactly n detectors at any time.
- The derived results hold also for $n = 0$, i.e. $d > L$. Naturally, in this case the vehicle either activates one detector or does not activate any detector; in the first case the absolute and relative errors are $(d - L)/\Delta$ and $(d - L)/L$, respectively; while in the second case both errors are L/Δ and 1, respectively.

We will now extend the derived results for the case of N vehicles traveling on the link with respectively lengths $L_j = L_j^{\text{Ph}}$, $j = 1, \dots, N$, and corresponding n_j , m_j , $j = 1, \dots, N$, defined as before. The space-occupancy in this case is $o_s = NL/\Delta$ where L is now the arithmetic average of L_j , $j = 1, \dots, N$, and with similar reasoning as for the case $N = 1$

(and observing that no detector can be activated by more than one vehicle simultaneously) we have the following generalized results:

- The produced estimates \hat{o}_s are bias-free.
- The maximum absolute error is given by $(d/\Delta)\sum_{j=1}^N \max\{m_j, 1-m_j\}$. In the special case $L_j = L \forall j$, this error is obviously N times higher than in the one-vehicle case. The maximum relative error is given by $(d/NL)\sum_{j=1}^N \max\{m_j, 1-m_j\}$. In the special case $L_j = L \forall j$, this error is obviously the same as in the one-vehicle case. Finally, the absolute error variance is N times higher than in the single-vehicle case above if the errors relating to individual vehicles are mutually independent.

The occurrence of the maximum (absolute or relative) errors in the N -vehicle case implies that all individual vehicle errors take simultaneously their maximum value; this is an event whose likelihood decreases strongly with increasing N .

3.5 Relaxing Some Assumptions

If the utilized detectors are operated with $L_j \neq L_j^{\text{ph}}$, we may assume

$$L_j = L_j^{\text{ph}} + \varepsilon_j, \varepsilon_j < E \quad (3.4)$$

where ε_j is the (non-zero) effective detector length while the limit $E > 0$ is small enough to avoid simultaneous activation of one detector by more than one vehicles.

Assuming that the link holds N vehicles, the space-occupancy would be $o_s = NL^{\text{ph}}/\Delta$, while an, otherwise perfect, estimation (e.g. as in the example of section 3.4 based on

time-occupancies would deliver $\hat{o}_s = NL/\Delta$; where L^{ph} and L are arithmetic averages of L_j^{ph} and L_j , respectively. Thus, the additional average error (bias) due to $L_j \neq L_j^{ph}$ is readily calculated equal to $N\varepsilon/\Delta$ (where ε is the arithmetic average of ε_j).

On another issue, the analysis of sections 3.2-3.4 was conducted for instantaneous values of space- and time-occupancies. However, time-occupancy is practically measured for finite update periods $T > 0$ which gives rise to the following remarks that are explained on the basis of the example of section 3.4:

- For finite T , each vehicle may activate both n and $n+1$ detectors within a single measurement period. As a consequence, the estimation error attributed to each vehicle will be a correspondingly time-weighted average of errors resulting from activation of only n or only $n+1$ detectors. The absolute value of this time-weighted average is easily seen to be less-equal than the maximum error of the instantaneous case and, indeed, the estimated space-occupancy \hat{o}_s will tend to its expected value $E\{\hat{o}_s\}$ for sufficiently high T (and non-zero vehicle speed). This suggests that estimates tend to be more accurate with increasing period T .
- The above conclusion is only true if the number of vehicles in the link (and hence the space-occupancy) does not change during T . In the more realistic case of vehicles entries and departures, the space-occupancy o_s may be continuously changing; thus estimates \hat{o}_s calculated as averages of measured time-occupancies may become increasingly outdated with increasing T . This suggests that estimates tend to be less accurate with increasing period T if the space-occupancy changes significantly.

- Both remarks above suggest that there is an optimal value of T for which the estimates \hat{o}_s produced via averaging of time-occupancy measurements are most accurate. Clearly, this value depends on prevailing conditions (link length, signal plans, arriving demand) and can hardly be determined analytically. An important factor in this context is also the number of utilized detectors; if M is very high, we have a good averaging (over space) even with very short T ; if M is very low (or 1), the few available local measurements may not be representative for the real link occupancy, unless an averaging takes place over time due to a longer T .

A further situation that deserves some attention is when two link vehicles are standing still (due to a red traffic signal) just upstream resp. downstream of a detector, without activating the detector. In this randomly appearing case, the detector measures a zero value for time-occupancy while the real traffic situation is quite the opposite. As will be demonstrated in related simulation experiments in CHAPTER 7, this ZSZO (zero-speed zero-occupancy) situation does not distort the expected value of the estimate $E\{\hat{o}_s\}$ because the gaps between vehicles at standstill contribute to the whole-link space-occupancy o_s , hence they should also contribute to its estimate \hat{o}_s . On the other hand, this randomly appearing phenomenon may contribute to an increase of the estimation error variance due to accordingly high estimation errors at the times of its appearance.

3.6 Vehicle-count in the Link

Summarizing the findings of sections 3.2-3.5 we conclude that the estimated space-occupancy \hat{o}_s calculated as the average of M detector time-occupancies according to (3.2) may be related to the link's current space-occupancy o_s

$$\hat{o}_s = o_s + \zeta_o \quad (3.5)$$

where the random variable ζ_o incorporates several potential sources of error:

- Detector measurement noise, which may be depending on the magnitude of the measured time-occupancies.
- Modelling error due to the approximative character of (3.2), particularly for small M . Note that this error may include bias according to Figure 9. This error also includes the impact of the update period T and of the ZSZO-situation outlined in section 3.5.
- Error due to the effective vehicle lengths being different than the physical vehicle lengths while measuring time-occupancy. If the non-zero effective detector length ε is known, one may convert the collected time-occupancy measurements into non-biased representations of the space-occupancy by multiplying them with $L^{Ph}/(L^{Ph} + \varepsilon)$ according to section 3.5. This bias-rejection manipulation may have additional benefits (lower error variance in the ZSZO-case) as will be illustrated in section 7.5.

In order to span a bridge linking the time-occupancy and the vehicle-count N in the link, we observe that, by definition,

$$o_s = N L^{Ph} / (\Delta \lambda) \quad (3.6)$$

where λ is the number of lanes in the link. From (3.6) we may define the estimated vehicle count

$$\hat{N} = \frac{\Delta \lambda}{L^{Ph}} \hat{o}_s \quad (3.7)$$

and using (3.6), (3.7) in (3.5) we obtain

$$\hat{N} = N + \zeta \quad (3.8)$$

where $\zeta = \zeta_o \Delta \lambda / L^{Ph}$. Note that the transformation of the estimated space-occupancy measurement \hat{o}_s into an estimated vehicle count \hat{N} via (3.7) involves the (arithmetic) average physical vehicle length L^{Ph} which may not be accurately known; this introduces a further potential source of bias in equation (3.8). Note also that the term $N_{max} = \Delta \lambda / L^{Ph}$ appearing in (3.7) actually corresponds to the maximum number of vehicles that can be accommodated in the link in a bumper-to-bumper manner.

Equations (3.5) or (3.8) may be used for direct estimation of the link's space-occupancy o_s or vehicle-count N , respectively. The estimation accuracy depends on a number of factors, most of which have been identified in sections 3.2-3.5. As the analytical pre-calculation of all addressed error sources is rather difficult, several microscopic simulation scenarios in CHAPTER 7 will be examined in order to assess the estimation quality.

The estimation error resulting from (3.5) or (3.8) may be reduced by additional application of a Kalman Filter presented later. In particular, the Kalman-filter approach, that uses (3.5) or (3.8) as its output equation, is valuable because it leads to acceptable estimation error even by use of only $M=1$ detector measuring time-occupancy in the middle of the link. This reduces the installation and maintenance cost by replacing detector hardware by algorithmic intelligence.

CHAPTER 4

KALMAN FILTER DEVELOPMENT

4.1 Introduction

Traffic-responsive control systems require reliable real-time information on the prevailing traffic conditions to make sensible control decisions. More particularly, the number of vehicles included in signalized links (such as urban road links or metered motorway ramps) is valuable information for urban signal control and motorway ramp metering systems. Direct measurements of this quantity may in principle be collected by video sensors, but this possibility is obscured due to visibility obstacles, limited visible link length, image processing algorithm accuracy and, last not least, cost.

As an alternative possibility, traditional low-cost loop detectors measuring time-occupancy may be employed. The difficulty faced by this approach is due to the strongly inhomogeneous character of the traffic state in signalized links caused by the frequent switching of the upstream and, most importantly, downstream traffic lights. Thus, a detector station (across all link lanes) positioned at a specific link location (e.g. at the signal stop line or in the middle or at the upstream end of the link) delivers (local) occupancy information that is not representative for the whole link. In other words, local time-occupancy measurements collected by loop detectors need to be translated into space-occupancy or vehicle-count estimates that are directly related to the number of vehicles in the link.

A relatively high-cost approach, that is, e.g., practiced in some ramp metering installations in the U.K., is to install a high number of loop detectors along the link (e.g. one detector every 50 m in the U.K. ramps). Yet another (theoretical) possibility would be to install two flow-measuring detectors at the respective extreme points of the link and deduce vehicle-counts in the link from the conservation equation; however, this approach would soon lead to unacceptably high estimation errors due to time-accumulation of the inevitable measurement errors.

The estimation method presented in this section employs Kalman filtering to deliver reliable real-time estimates of vehicle-count in signalized links based on measurements of three loop detector stations located at both extreme points and at the middle of the link, respectively. This Kalman Filter turns out to be similar to the one proposed in Bhouri, et al. (1988) for the estimation of traffic density in short non-signalized motorway links.

An interesting question addressing the degree of hardware cost (for detectors, communications and maintenance) savings thanks to exploitation of low-cost algorithmic intelligence reads: How many (additional) loop detectors would be necessary to reach the estimation accuracy of the proposed method? This question as well as a thorough assessment of the proposed method under several different conditions are treated via microscopic simulation.

Although the presented method can be applied equally well to urban signalized links, the reported simulation investigations were chosen to resemble to typical metered motorway ramps. Vehicle-count estimates for motorway ramps are required within ramp metering systems for:

- Efficient ramp queue control to avoid spillback in the adjacent street network (Smaragdis and Papageorgiou, 2003; Sun and Horowitz, 2005).
- Efficient and equitable ramp metering coordination (Kotsialos and Papageorgiou, 2005).

A different approach to vehicle count estimation for motorway ramps using speed measurements and curve fitting to a high number of simulation data (that may be difficult to collect in the field) was proposed in Sun and Horowitz (2005).

4.2 Introduction to Kalman Filter

First of all it will be attempted to make an overview of the Kalman Filter's general mathematical equations and necessary assumptions for its application to any process. Then, all filters' equations will be transformed so as to develop the complete model for the estimation of the number of vehicles in a signalized link.

The Kalman Filter addresses the general problem of trying to estimate the state $\mathbf{x} \in \mathfrak{R}^n$ of a discrete-time controlled process that is governed by the linear stochastic difference equation

$$\left. \begin{array}{ll} \text{State equation:} & \mathbf{x}(k+1) = \mathbf{A}(k)\mathbf{x}(k) + \mathbf{B}(k)\mathbf{u}(k) + \mathbf{D}(k)\mathbf{g}(k) \quad (4.1) \\ \\ \text{with a measurement } \mathbf{y} \in \mathfrak{R}^n \text{ that is} & \\ \\ \text{Measurement equation:} & \mathbf{y}(k) = \mathbf{C}(k)\mathbf{x}(k) + \mathbf{z}(k) \quad (4.2) \end{array} \right\} \mathbf{\Sigma}(\mathbf{x}, \mathbf{u}, \mathbf{y})$$

The $n \times n$ matrix \mathbf{A} in the difference equation (4.1) relates the state at the current time step k to the state at the next step $k+1$, in the absence of either a driving function or process noise; in practice, \mathbf{A} might change with each time step. The $n \times \ell$ matrix \mathbf{B}

relates the optional control input $\mathbf{u} \in \mathbb{R}^\ell$ to the state \mathbf{x} and the $n \times n$ matrix \mathbf{D} enforces noise $\mathbf{g}(k)$. The $m \times n$ matrix \mathbf{C} in the measurement equation (4.2) relates the state to the measurement $\mathbf{y}(k)$; in practice \mathbf{C} might change with each time step or measurement.

The random variables $\mathbf{g}(k)$ and $\mathbf{z}(k)$ represent the process and the measurement noise respectively. These variables and the initial state $\mathbf{x}(0)$ of the process satisfy the following three conditions:

1. $\mathbf{g}(k)$ and $\mathbf{z}(k)$ are zero-mean Gaussian white random processes. For any $k \geq 0$ and $\ell \geq 0$:

$$\mathbb{E}[\mathbf{g}(k)] = \mathbf{0}$$

$$\mathbb{E}[\mathbf{z}(k)] = \mathbf{0}$$

$$\mathbb{E}[\mathbf{g}(k) \cdot \mathbf{g}^T(\ell)] = \begin{cases} \mathbf{Q}(k) & \text{if } k = \ell, \\ \mathbf{0} & \text{otherwise} \end{cases}$$

$$\mathbb{E}[\mathbf{z}(k) \cdot \mathbf{z}^T(\ell)] = \begin{cases} \mathbf{R}(k) & \text{if } k = \ell, \\ \mathbf{0} & \text{otherwise} \end{cases}$$

$$\mathbb{E}[\mathbf{g}(k) \cdot \mathbf{z}^T(\ell)] = \begin{cases} \mathbf{M}(k) & \text{if } k = \ell, \\ \mathbf{0} & \text{otherwise} \end{cases}$$

where \mathbf{Q} and \mathbf{R} are known symmetric positive semi-definite matrices, while $\mathbf{0}$ denotes zero vectors or zero matrices of appropriate dimensions. If $\mathbf{g}(k)$ and $\mathbf{z}(k)$ are correlated, then \mathbf{M} is the correlation matrix, otherwise $\mathbf{M} = \mathbf{0}$.

2. $\mathbf{x}(0)$ is a Gaussian random vector with known mean and covariance matrix \mathbf{P} :

$$\hat{\mathbf{x}}_0 = \mathbf{E}[\mathbf{x}(0)]$$

$$\mathbf{P}_0 = \mathbf{E}\left\{[\mathbf{x}(0) - \hat{\mathbf{x}}_0] \cdot [\mathbf{x}(0) - \hat{\mathbf{x}}_0]^T\right\}. \quad (4.3)$$

3. $\mathbf{x}(0)$ is uncorrelated to $\mathbf{g}(k)$ and $\mathbf{z}(k)$ at any k .

Lets consider $\Sigma(\mathbf{x}, \mathbf{u}, \mathbf{y})$ under the assumptions 1-3. At each time step k , given $\mathbf{y}(k)$ (and its available values at all previous time instants, i.e. $\mathbf{y}(k-1), \mathbf{y}(k-2), \dots$), it is the goal of the Kalman Filter to deliver state estimate $\hat{\mathbf{x}}(k+1/k)$ so as to minimize the covariance of the estimation error

$$\mathbf{E}\left\{[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1/k)]^T \cdot [\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1/k)]\right\} \rightarrow \text{Min}$$

where $\hat{\mathbf{x}}(k+1/k)$ denotes the mathematical expectation of $\mathbf{x}(k+1)$ conditional on measurements available up to the k -th time step (actually $\hat{\mathbf{x}}(k+1/k)$ is the one-step prediction of $\mathbf{x}(k+1)$).

The recursive equations of the filter are as follows:

$$\hat{\mathbf{x}}(k+1/k) = \underbrace{\mathbf{A}(k)\hat{\mathbf{x}}(k/k-1) + \mathbf{B}(k)\mathbf{u}(k)}_{\text{Model}} + \underbrace{\mathbf{K}(k) \cdot \{\mathbf{y}(k) - \mathbf{C}(k)\hat{\mathbf{x}}(k/k-1)\}}_{\text{Correction}} \quad (4.4)$$

with

$$\mathbf{K}(k) = [\mathbf{A}(k)\mathbf{P}(k/k-1)\mathbf{C}(k)^T + \mathbf{D}(k)\mathbf{M}(k)] \cdot [\mathbf{C}(k)\mathbf{P}(k/k-1)\mathbf{C}(k)^T + \mathbf{R}(k)]^{-1} \quad (4.5)$$

$$\mathbf{P}(k+1/k) = [\mathbf{A}(k) - \mathbf{K}(k)\mathbf{C}(k)]\mathbf{P}(k/k-1)\mathbf{A}^T(k) + \mathbf{D}(k)\mathbf{Q}(k)\mathbf{D}(k)^T - \mathbf{K}(k)\mathbf{M}(k)^T\mathbf{D}(k)^T \quad (4.6)$$

$$\hat{\mathbf{x}}(0/-1) \triangleq \hat{\mathbf{x}}_0$$

$$\mathbf{P}(0/-1) \triangleq \mathbf{P}_0$$

As noticed in (4.4), the Kalman Filter consists of two terms: (1) model term delivering pure model-based state estimation at each time instant k and (2) correction term based on the real-time measurements collected by each k . Both terms are essential for the satisfactory performance of the filter; the use or not of the correction term in estimation is weighted by $\mathbf{K}(k)$. The $\mathbf{K}(k)$ is the gain $n \times m$ matrix that minimizes the error covariance equation (4.3). A good way of thinking the weighting by \mathbf{K} is that as the measurement error covariance \mathbf{R} approaches zero, the actual measurement $\mathbf{y}(k)$ is “trusted” more and more (then, from equation (4.5), $\mathbf{K}(k) \rightarrow \mathbf{A}(k)\mathbf{C}(k)^{-1}$), while the predicted measurement $\mathbf{C}(k)\hat{\mathbf{x}}(k)$ is trusted less and less. On the other hand, as the estimate error covariance $\mathbf{P}(k)$ approaches zero (then, from equation (4.5), $\mathbf{K}(k) \rightarrow 0$) the actual measurement $\mathbf{y}(k)$ is trusted less and less, while the predicted measurement $\mathbf{C}(k)\hat{\mathbf{x}}(k)$ is trusted more and more.

4.3 Modelling for Estimation

An appropriate state-space model and a measurement model are needed for application of the Kalman Filter. The state-space model is the conservation-of-vehicles equation in the link while the measurement model is based on the insights gained in CHAPTER 3.

The evolution of the number of vehicles in a link obeys the following conservation equation

$$N(k) = N(k-1) + T[q_{\text{in}}(k-1) - q_{\text{out}}(k-1)] \quad (4.7)$$

Where $N(k)$ denotes the number of vehicles in the link at time kT and T is the measurement and estimation update period (or sampling time); $q_{in}(k-1)$ and $q_{out}(k-1)$ are the flows entering and exiting, respectively, the link during the period $[(k-1)T, kT]$. These flows are measured by the boundary detectors

$$q_{in}^m(k-1) = q_{in}(k-1) + \gamma_{in}^m(k-1) \quad (4.8)$$

$$q_{out}^m(k-1) = q_{out}(k-1) + \gamma_{out}^m(k-1) \quad (4.9)$$

where q_{in}^m , q_{out}^m are the related measurements while γ_{in}^m , γ_{out}^m denote the corresponding measurements noise which are assumed to be zero-mean stochastic variables. The measurement noise may realistically be modeled in proportion to the related flow value rather than independent thereof; this, however, would render the resulting state equation nonlinear and would call for application of the more complex Extended Kalman Filter. Preliminary simulation investigations indicated that the higher level of modelling realism does not lead to substantially higher estimation accuracy; hence it was decided to consider γ_{in}^m and γ_{out}^m in (4.8) and (4.9) as random variables with constant variance that is independent of the value of the measured flows. Replacing (4.8), (4.9) in (4.7) yields

$$N(k) = N(k-1) + T[q_{in}^m(k-1) - q_{out}^m(k-1)] + T\gamma(k-1) \quad (4.10)$$

where $\gamma = \gamma_{out} - \gamma_{in}$. Although, in principle, the conservation equation with inflow and outflow measurements could be directly used for estimating the vehicle-count $N(k)$, eq. (4.10) reveals that such a procedure would accumulate the unavoidable measurement noise γ leading to increasingly inaccurate estimates. Therefore, more information is

necessary to counter the accumulation of measurement noise in the state equation (4.10).

The required additional information may be provided by the middle detector in Figure 5a. According to CHAPTER 3, the time-occupancy $o_t^m(k-1)$ collected by this detector during $[(k-1)T, kT]$ may be related to the link's space-occupancy $o_s(k-1)$ at time $(k-1)T$

$$o_t^m(k-1) = o_s(k-1) + \zeta_o(k-1) \quad (4.11)$$

where the random variable ζ_o incorporates several potential sources of error:

- Detector measurement noise: As for the flow measurement noise above, this noise will be considered independent of the measured occupancy value in order to obtain a simpler estimation algorithm.
- Modelling error due to the approximative character of (3.2), particularly for small M . Note that this error may also include some (small) bias. Finally, this error also includes the impact of the update period T and of the ZSZO-phenomenon outlined in section 3.5.
- Error due to the effective vehicle lengths being different than the physical vehicle lengths while measuring time-occupancy (see section 3.5).

In order to span a bridge linking the time-occupancy and the vehicle-count N in the link, in the light of (3.6), (4.11) we may define the “measured” vehicle count

$$N^m = \frac{\Delta\lambda}{L^{Ph}} o_t^m \quad (4.12)$$

and replacing (3.6), (4.12) in (4.11) we finally obtain

$$N^m(k-1) = N(k-1) + \zeta(k-1) \quad (4.13)$$

where $\zeta = \zeta_o \Delta\lambda / L^{Ph}$. Note that the transformation of the collected time-occupancy measurement o_t^m into a “measured” vehicle-count N^m via (4.12) involves the (arithmetic) average physical vehicle length L^{Ph} which may not be accurately known; this introduces a further potential source of bias in the measurement equation (4.13). Note also that the term $N_{\max} = \Delta\lambda / L^{Ph}$ appearing in (4.12) actually corresponds to the maximum number of vehicles that can be accommodated in the link in a bumper-to-bumper manner.

The state equation (4.10) and measurement equation (4.13) have the appropriate form for Kalman-filter application. To this end we should consider the system noise γ and measurement noise ζ to be zero-mean white gaussian random variables. These assumptions may not be verified fully in practice but the Kalman Filter may nevertheless deliver practically useful, suboptimal estimates. Of particular importance is the possible appearance of biased measurements (4.13) (i.e. non-zero-mean error) because biased measurements cannot be rejected by the Kalman Filter and will lead to accordingly biased estimates.

Despite the various sources of (partly non-zero-mean) errors, it is expected that the measurement equation (4.13) contains a sufficient level of reliable information that may be exploited by the Kalman Filter in order to reduce the accumulated error that would result from the usage of the conservation equation (4.10) alone.

The quality of the measurement equation may be improved if more internal detectors are used to produce an appropriate average measurement signal according to (3.2). Based on the general scheme of Figure 8, one may employ M (instead of one) internal

detectors (separated by equal distances $d = \Delta / M$), in which case o_t^m may be calculated at each period as the arithmetic average of time-occupancies collected by the individual internal detectors according to (3.2).

4.4 Kalman-filter Estimator

Based on the state equation (4.10) and measurement equation (4.13), a Kalman-filter estimator (Jazwinsky, 1970) for the number of vehicles in the link may be immediately derived

$$\hat{N}_{KF}(k) = \hat{N}_{KF}(k-1) + T[q_{in}^m(k-1) - q_{out}^m(k-1)] + K[N^m(k-1) - \hat{N}_{KF}(k-1)] \quad (4.14)$$

where $\hat{N}_{KF}(k)$ is the delivered estimate of vehicle-count and K is the (stationary) gain parameter of the filter. The produced estimate \hat{N}_{KF} is truncated if it exceeds the range $[0, N'_{max}]$ where N'_{max} is the maximum number of vehicles that can be accommodated in the link at standstill including the usual safety distance D among vehicles (e.g. $D = 1$ m); N'_{max} may be calculated, similarly to (4.12), from $N'_{max} = \Delta\lambda / (L^{Ph} + D)$. The filter (4.14) consists of a system model (the conservation equation comprising the first two terms on the r.h.s. of (4.14)) and a correction term that attempts to reduce the estimation error resulting from the system noise γ in (4.10).

The filter equation (4.14) may be re-arranged

$$\hat{N}_{KF}(k) = K \cdot N^m(k-1) + (1-K) \cdot \hat{N}_{KF}(k-1) + T[q_{in}^m(k-1) - q_{out}^m(k-1)] \quad (4.15)$$

in which case a further interpretation may be given. The produced estimate $\hat{N}_{KF}(k)$ results from the combination of:

- An exponential smoothing (first two terms on the r.h.s. of (4.15)) based on the arriving measurements N^m .
- A prediction term that involves the most recent flow measurements.

According to the Kalman-filter theory (Jazwinsky, 1970), the value of the gain K should be selected

$$K = \Pi / (\Pi + Z)$$

where Π satisfies

$$\Pi = (1 - K)\Pi + T^2\Gamma .$$

From these equations we obtain

$$K = 0.5(-\alpha + \sqrt{\alpha^2 + 4\alpha}) \quad (4.16)$$

where $\alpha = T^2\Gamma / Z$; and Γ , Z are the variances of the system noise γ and measurement noise ζ , respectively. While the system noise variance Γ could be approximately determined based on the typical flow measurement errors, the value of the measurement noise variance Z is related to many different sub-processes and hence difficult to derive. However, (4.16) suggests that the value of K depends only on the ratio α , not on the explicit values of Γ and Z . Hence, rather than attempting to derive appropriate values for Γ , Z , one may attempt to fine-tune the ratio α , or, even more directly, the value of K to be used in the filter (4.15). Note that for $\alpha \rightarrow 0$ (i.e. zero system noise or infinite measurement noise), (4.16) yields $K = 0$ which means that the estimation (4.14) makes no use of the measurements N^m ; on the other hand, for $\alpha \rightarrow \infty$ (i.e. zero measurement noise or infinite system noise) it may be shown from

the K , Π equations above that $K = 1$, i.e. the exponential smoothing in (4.15) is based on the latest measurements N^m only. Thus, potential K values are included in the range $[0,1]$ as K is indeed monotonically increasing with α .

The aforementioned quantity Π represents the variance of the estimation error $\hat{N}_{KF} - N$ and is calculated

$$\Pi = 0.5Z(\alpha + \sqrt{\alpha^2 + 4\alpha}) \quad (4.17)$$

Thus, for $\alpha \rightarrow 0$ (and finite Z) we obtain $\Pi = 0$ (due to perfect model) while for $\alpha \rightarrow \infty$ we obtain $\Pi = T^2\Gamma$ (i.e. the estimation error in (4.15) results only from the prediction term due to perfect measurement N^m). The variance Π may be easily shown to be a monotonically increasing function of both the measurement error variance Z and the system error variance $T^2\Gamma$.

If the measurement N^m would be sufficiently accurate, it might be better to use directly $\hat{N}_{KF} = N^m$ (as in CHAPTER 3) rather than the Kalman Filter (4.14). For this to be true, we should have the variance Z of the error $\hat{N}_{KF} - N^m$ being smaller than the filter error variance Π , i.e. we should have $\Pi \geq Z$. After some calculations we derive the equivalent inequality $Z \leq 2T^2\Gamma$ which is quite unlikely to hold in practice as Z is usually much higher than Γ .

Equations (4.16), (4.17) also contain valuable information on the role of the update period T . As mentioned in section 3.5, the measurement variance Z obtains a minimum value for some value of T . Note that smaller Z can be easily shown to lead to smaller Π in (4.17). On the other hand, for higher T the system noise $T^2\Gamma$ increases,

and consequently also α and the estimation error variance Π increase. In conclusion, the optimal update period T for best Kalman-filter estimates may not coincide with the optimal update period for least measurement error in (4.13).

If the measurement noise ζ contains a bias $E\{\zeta\} = b$, i.e. from (4.13) $E\{N^m\} = E\{N\} + b$; and if $E\{q_{in}^m\} = E\{q_{out}^m\}$, i.e. there is no bias in the traffic volume measurement; then we obtain from (4.14)

$$E\{\hat{N}_{KF}\} = E\{N^m\} = E\{N\} + b \quad (4.18)$$

i.e. the measurement bias cannot be rejected by the Kalman Filter and is fully transmitted to the estimates \hat{N}_{KF} .

CHAPTER 5

EXPONENTIAL SMOOTHING

5.1 Introduction

The estimation of the number of vehicles in a link must be well managed in order to keep a balance between the high quality of the estimation and the cost of use of the available means. The Kalman Filter equation (4.14) that has been derived, demands the existence of three loop detectors in the link; one detector in the middle of the link providing time-occupancy measurements and two boundary detectors providing flow measurements. Moreover, it has been observed that the produced KF estimation, results from the combination of two terms: an exponential smoothing term and a prediction term (see section 4.4).

It has already been stated in section 4.3 that the direct estimation of the vehicle-count through the conservation equation with inflow and outflow measurements results to the accumulation of the unavoidable flow measurement noise leading to increasingly inaccurate estimates; this reveals that the exponential smoothing term is necessary for the estimation. But does the prediction term have the same importance?

In an attempt to reduce even further the cost of the vehicle-count estimation, it is quite interesting to examine the case where only the term of exponential smoothing is used for the estimation. In this way, the installation and maintenance cost of such a system is

reduced as all the necessary information for the estimation is provided by a single detector in the middle of the link measuring time-occupancy consequently, the cost of two boundary detectors is saved.

5.2 Exponential Smoothed Estimator

By the fact that the use of a single detector is economical, we are summoned to answer the question that arises: does the filter estimation become inferiorly influenced by exclusively using the exponential smoothing term? To answer the question we will use the following new exponential smoothed estimation equation:

$$\hat{N}_{SM}(k) = K_{SM}N^m(k-1) + (1 - K_{SM})\hat{N}_{SM}(k-1) \quad (5.1)$$

where $\hat{N}_{SM}(k)$ is the delivered estimate of vehicle-count and K_{SM} is the smoothing parameter. Like in the Kalman Filter, the \hat{N}_{SM} is truncated if it exceeds the range $[0, N'_{max}]$ where N'_{max} is the same as before (see section 4.3).

Now the estimation is independent of the flow measurement noise, as this measurement is not used at all, and it is affected just by the time-occupancy measurement error ζ_o that has been analyzed in section 4.3.

The estimation \hat{N}_{SM} can be improved with the suitable value of the smoothing parameter K_{SM} which, as the Kalman Filter gain K , is included in the range $[0, 1]$. Note that for $K_{SM} \rightarrow 0$, (5.1) yields $\hat{N}_{SM}(k) = \hat{N}_{SM}(k-1)$ which means that the estimation (5.1) makes no use of the measurement N^m and it depends on the unknown or biased, most of the times, $\hat{N}_{SM}(0)$; on the other hand, for $K_{SM} \rightarrow 1$ the exponential smoothing equation (5.1) is based on the latest measurements N^m only.

It can be seen from (5.1) that the update period T does not have any direct impact on the estimation but only on the measurement error, while the estimation is almost equal with the average number of vehicles that have passed the middle of the link until then. This guides the estimation on following a “predetermined” trajectory, without being affected by the creation or not of queues in the link (at least until these reach the middle of the link). On the other hand, this approach smoothes the effect of the ZSZO-phenomenon which does not influence the estimation any more.

In conclusion, it can be said that this kind of analysis/estimation would be sufficiently satisfactory in the case of a small update period T where the average vehicle-count in the link is, almost, the same with the real number of vehicles in it. Contrariwise, for longer periods the exponential smoothing analysis is inappropriate because the average number of vehicles is quite different than the real number of vehicles that remain in the link.

Some short results will be presented in CHAPTER 7 to compare the estimation by using the exponential smoothing with the simple use of measurements and the implementation of Kalman Filter. It is expected that the exponential smoothing will, of course, improve the results of the equation (3.8) and with the appropriate fine-tuned value of K_{SM} . Nevertheless, the smoothed estimation is not expected to exceed the quality of the Kalman Filter estimation (4.14) as the use of the flows measurements by the boundary detectors is important and justifies the extra cost of two additional detectors. Notice that K_{SM} is displayed in the corresponding figures as “smoothing parameter K ”.

CHAPTER 6

MICROSCOPIC SIMULATION MODEL

6.1 Introduction

In this chapter the microscopic simulation model that was developed for the testing and the confirmation of the realistic operation of the proposed estimation method will be analyzed. The model was developed in C programming language and has no special demands neither in computational time nor memory.

The microscopic simulation model that has been developed imitates the traffic flow into a link or a motorway ramp. The fact that the link (or the ramp) has only one lane and, so, overtaking is not possible, makes the model simpler; although, several other parameters that influence, more or less, the flow can be determined. The constant values that were chosen for these parameters are shown in APPENDIX A, where actually the input data file of the simulation program is shown.

As anyone may discern, the values presented there are necessary not only for the car following model, but for the general geometry of the link, the signalization, the application of the Kalman Filter etc. The simulation model gives the opportunity to examine the produced results regarding the optimization of any of its parameters. However, the purposes of the present research does not centre on the evaluation of the simulation model, even if it has been tried to resemble the real traffic phenomena. The

simulation model is just a tool to test the filter's results. This is why we will not come over in depth to minor subroutines considering, for example, the entrance of a vehicle in the link, the signalization or the way that time-occupancy measurements are collected. What is the most important for the credibility of the flow process and will be analyzed extensively below is the car following model. All the equations of the applied car following model and the constant parameter values will be presented next.

6.2 Simulation Description

A self-developed microscopic simulator was used to describe the traffic phenomena on a single-lane, 194-m long link, with both downstream and upstream traffic signals. Vehicles are generated by the simulator far upstream of the upstream signals. The vehicles are moving on one lane according to the following Bando-type (Bando, et al., 1995) discrete-time car-following equations based on a simulation time step $T_{sim} = 0.25s$. At each simulation time period $k = 1, 2, \dots$, a desired speed $y_{d,i}$ is calculated first for each vehicle i

$$y_{d,i} = \Lambda[\delta_i(k) - D] \quad (6.1)$$

with $\Lambda = 0.7 s^{-1}$ and $D = 1m$, where $\delta_i(k)$ is the distance (in m) of the front of vehicle i from the rear of the next downstream vehicle; if the desired speed calculated by (6.1) exceeds a free speed $y_f = 16.5 m/s$, then it is truncated to this value.

The acceleration $a_i(k)$ of vehicle i is calculated next via

$$a_i(k) = \text{sat}\{g[y_{d,i}(k) - y_i(k)]\} \quad (6.2)$$

where $g = 2 s^{-1}$ and the function $\text{sat}\{\cdot\}$ is defined as

$$\text{sat}\{\eta\} = \begin{cases} a_{\max} & \text{if } \eta \geq a_{\max} \\ a_{\min} & \text{if } \eta \leq a_{\min} \\ \eta & \text{else} \end{cases} \quad (6.3)$$

with $a_{\max} = 1.5 \text{ m/s}^2$, $a_{\min} = -6 \text{ m/s}^2$ and $y_i(k)$ is the speed of vehicle i which is updated via

$$y_i(k+1) = y_i(k) + a_i(k)T_{\text{sim}}. \quad (6.4)$$

Finally the updated position of the vehicle is obtained from

$$x_i(k+1) = x_i(k) + T_{\text{sim}}y_i(k) + 0.5T_{\text{sim}}^2a_i(k). \quad (6.5)$$

In some rare cases where the updated positions of a couple of subsequent vehicles indicate that their distance would be less than $D = 1 \text{ m}$, the acceleration of the following vehicle is re-calculated so that $\delta_i = D$.

Different equations than the above are applied to the first vehicle upstream of a red traffic signal, provided its distance from the signal is less than 50 m . In this case, the vehicle acceleration is calculated so as to enable the vehicle to eventually stop in front of the signal. In case the necessary acceleration is less than a_{\min} , that vehicle is allowed to pass the signal (i.e. it is updated as the other ordinary vehicles) but the special signal-treatment is applied to the next upstream vehicle.

When a vehicle passes a detector, a time-occupancy signal is produced according to section 2.2, whereby the duration of the detector occupancy depends on the vehicle physical length, the vehicle speed and the detector's effective length. The flow and

occupancy measurements for each estimation period are calculated according to section 2.2 and are eventually perturbed with random noise Δq and Δo , respectively, given by

$$\Delta q = 0.2q\psi \quad , \quad \Delta o = 0.05o\psi$$

where ψ is a white random variable with the unit normal distribution. Thus the magnitude of the measurement error depends on the current value of the measured quantity.

A standard simulation scenario with a duration of 5,000 s (≈ 1.38 h) will be defined next to be used in the simulation investigations. Modifications of this scenario will be produced later as appropriate. For the standard scenario, the simulation starts with an empty link. The upstream traffic signals are operated with a cycle of 90 s, while for the downstream signals we have a cycle of 20 s. The fixed green/red phases of both signals are chosen appropriately so as to create all possible values of vehicle-counts in the link (see Figure 10). The detector effective lengths are zero for the standard scenario while the physical vehicle lengths are uniformly distributed in the range [3 m, 5 m], hence $L^{Ph} = 4$ m is used in (3.7) to produce the estimated \hat{N} . One single internal detector is included for occupancy measurements in the middle of the link and the update period T equals 20 s.

Three further basic traffic scenarios were created by changing the traffic light settings of the standard scenario. More specifically, the three additional scenarios have different cycles times of 40 s, 60 s and 90 s, respectively, at the downstream traffic signals while the upstream signals are still operated with a cycle of 90 s. The green/red phases of both signals were again selected appropriately (see APPENDIX B) so as to create all

possible values of vehicle-counts in the link (Figure 10). A fifth stochastic scenario was created with the downstream traffic signal cycle changing stochastically between 10s and 90s during the simulation (and green/red phases changing accordingly).

The evaluation of all scenarios is based on the following Relative Mean Square Error criterion

$$\text{RMSE} = 100\% \sqrt{\frac{K \sum_{k=1}^K [\hat{N}(k) - N(k)]^2}{\sum_{k=1}^K N(k)}}$$

or

$$\text{KF RMSE} = 100\% \sqrt{\frac{K \sum_{k=1}^K [\hat{N}_{\text{KF}}(k) - N(k)]^2}{\sum_{k=1}^K N(k)}}$$

or

$$\text{SM RMSE} = 100\% \sqrt{\frac{K \sum_{k=1}^K [\hat{N}_{\text{SM}}(k) - N(k)]^2}{\sum_{k=1}^K N(k)}}$$

where $N(k)$ is the real and $\hat{N}(k)$, $\hat{N}_{\text{KF}}(k)$ and $\hat{N}_{\text{SM}}(k)$ are the estimated vehicle-counts, the Kalman Filter (KF) and the exponential smoothed estimated vehicle-counts in the link, respectively, while $k = 0, 1, 2, \dots$ is the discrete time index. All results are produced without the detector noise mentioned in section 6.2 (i.e. with $\psi = 0$) which was found to have a minor impact on the estimation accuracy compared to other sources of estimation error.

CHAPTER 7

SIMULATION RESULTS

7.1 Introduction

Although simulation of the traffic flow process may not reflect fully the real traffic phenomena, it may nevertheless provide valuable insights for a number of crucial questions that should be answered before actual field implementation:

- Does the Kalman-filter estimator improve over the mere and smoothed use of measurements? How many more internal detectors would be required to reach the estimation quality of the Kalman Filter?
- What is the range of suitable K values for the Kalman Filter? What is the sensitivity of the estimation quality around the optimal K value?
- What is the quantitative impact of the estimation period T and what value is recommended for the field implementation?
- How does the Kalman-filter estimator perform under different conditions regarding N_{\max} , traffic load, traffic light signaling?
- What is the impact of various measurement bias?

7.2 Impact of the Sampling Time

Figure 10 displays the actual and estimated vehicle-counts N and \hat{N} , respectively, while Figure 11 displays the actual and Kalman-Filter (KF) estimated vehicle-counts N and \hat{N}_{KF} , respectively, with appropriate K -values in (4.14) that will be detailed later, for the four basic scenarios described in section 6.2. Note that the N -curve is displayed with a time-resolution of 1 s while the estimates \hat{N} , \hat{N}_{KF} are updated every $T = 20$ s. In the same way, Figure 12 displays the actual and the exponential smoothed estimated vehicle-counts N and \hat{N}_{SM} , respectively, with appropriate K_{SM} -values in (5.1). It may be seen that the N -trajectory is subject to two kinds of time-variation; a high-frequency variation due to the periodical traffic signal switchings; and a low-frequency variation due to changing demand.

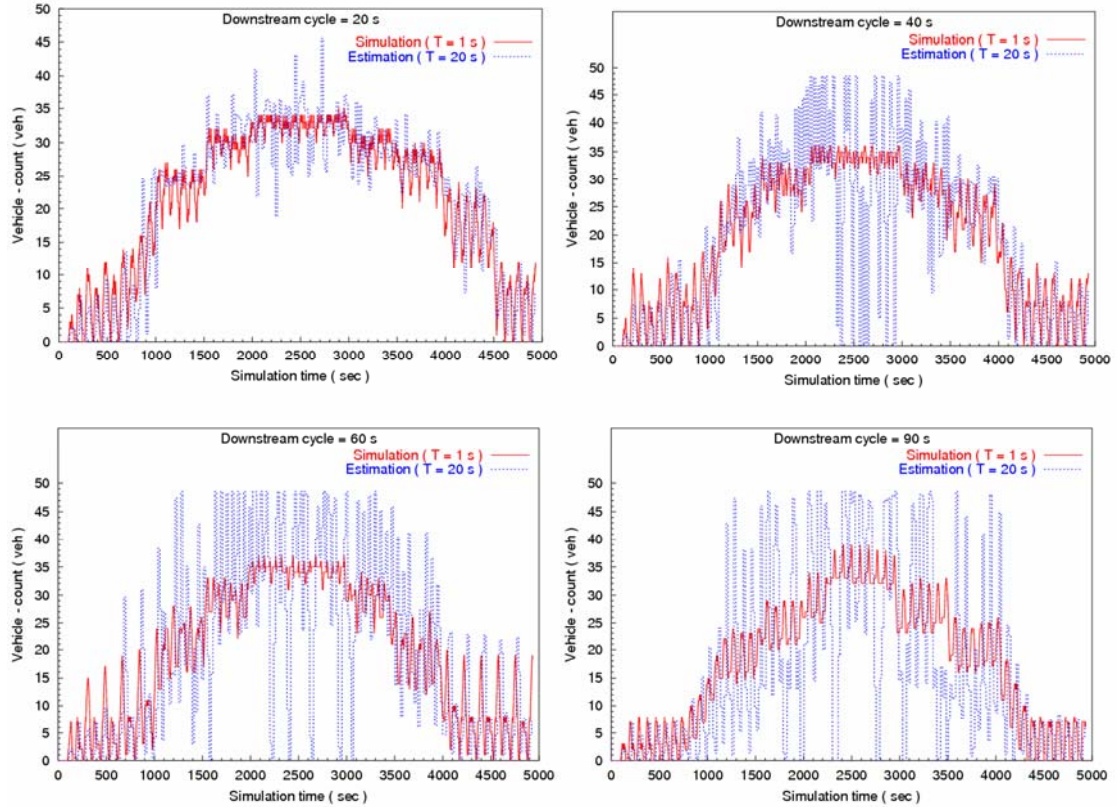


Figure 10: Real and estimated vehicle-counts over time for four basic scenarios.

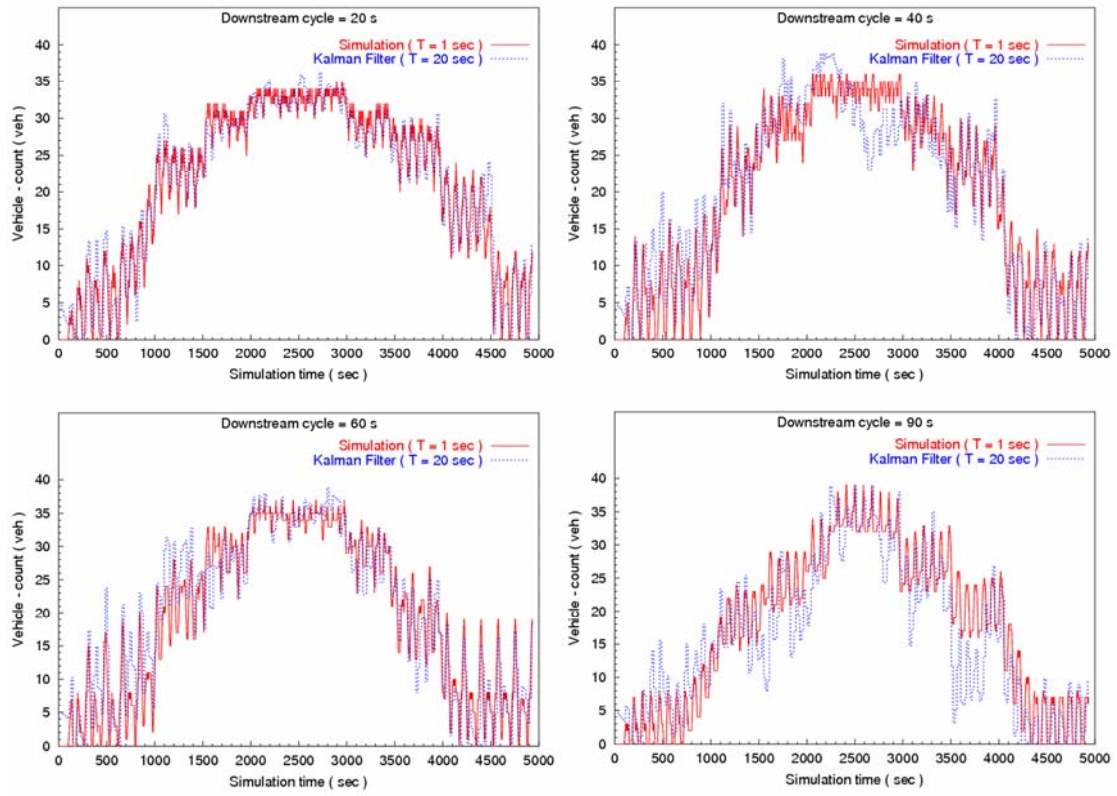


Figure 11: Real and KF estimated vehicle-counts over time for four basic scenarios.

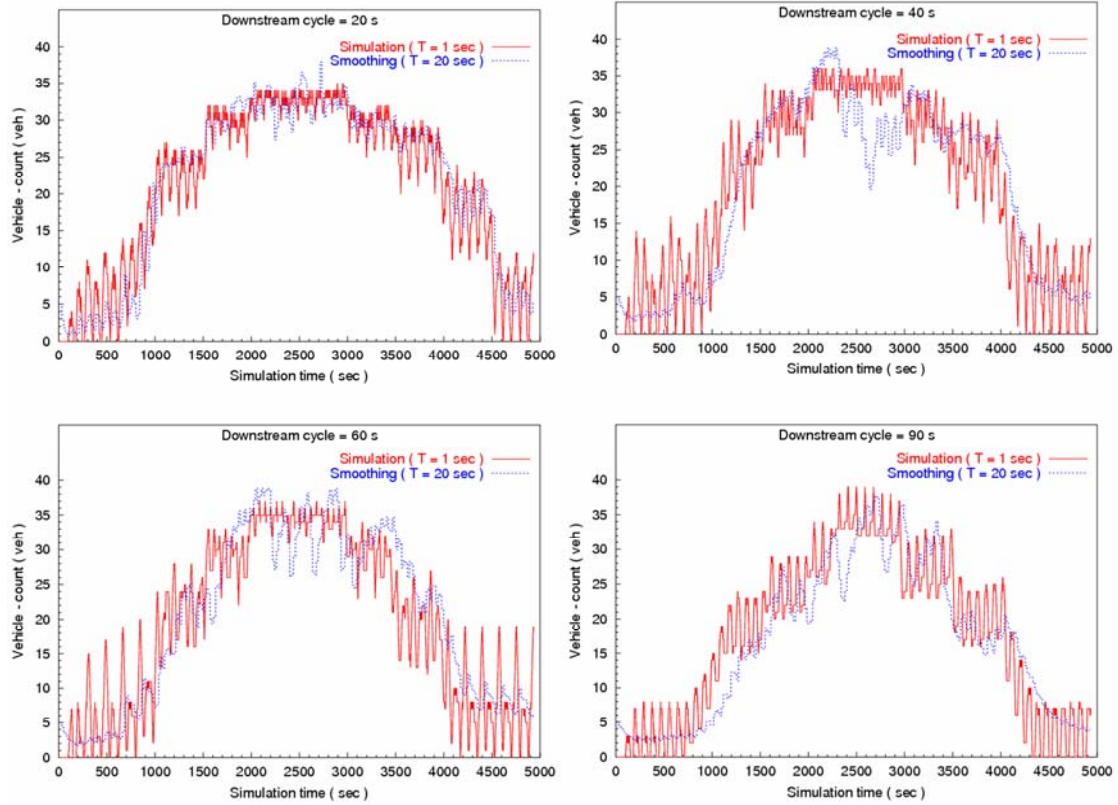


Figure 12: Real and smoothed estimated vehicle-counts over time for four basic scenarios.

Table 1, Table 2 and Table 3 ($L_j = L_j^{Ph}$) display the corresponding RMSE, KF RMSE and SM RMSE values and average errors (bias) $E\{N - \hat{N}\}$, $E\{N - \hat{N}_{KF}\}$ and $E\{N - \hat{N}_{SM}\}$ respectively. It may be seen that:

Table 1: RMSE values and $E\{N - \hat{N}\}$ for five scenarios and various cases of effective vehicle length.

Scenario cycle	$L_j = L_j^{Ph}$		$L_j = L_j^{Ph} + 1 \text{ m}$		corrected	
	RMSE (%)	$E\{N - \hat{N}\}$ (veh)	RMSE (%)	$E\{N - \hat{N}\}$ (veh)	RMSE (%)	$E\{N - \hat{N}\}$ (veh)
20 s	19.4	0.5	33.9	-5.13	16.0	0.3
40 s	48.4	0.66	44.0	-5.04	25.7	0.005
60 s	60.7	-0.49	57.1	-5.38	39.4	-0.37
90 s	69.4	0.85	60.9	-4.34	40.7	0.02

Table 2: KF RMSE values and $E\{N - \hat{N}_{KF}\}$ for five scenarios and various cases of effective vehicle length.

Scenario cycle	$L_j = L_j^{Ph}$		$L_j = L_j^{Ph} + 1 \text{ m}$		corrected	
	KF RMSE (%)	$E\{N - \hat{N}_{KF}\}$ (veh)	KF RMSE (%)	$E\{N - \hat{N}_{KF}\}$ (veh)	KF RMSE (%)	$E\{N - \hat{N}_{KF}\}$ (veh)
20 s	9.8	0.59	16.8	-2.83	9.4	0.33
40 s	17.6	-0.04	24.6	-4.07	10.7	-0.50
60 s	14.8	-0.59	22.3	-3.5	10.9	-0.64
90 s	27.5	1.25	24.4	-2.43	19.0	0.17
stochastic	22.8	0.29	23.3	-1.76	17.3	0.41

Table 3: SM RMSE values and $E\{N - \hat{N}_{SM}\}$ for five scenarios and various cases of effective vehicle length.

Scenario cycle	$L_j = L_j^{Ph}$		$L_j = L_j^{Ph} + 1 \text{ m}$		corrected	
	SM RMSE (%)	$E\{N - \hat{N}_{SM}\}$ (veh)	SM RMSE (%)	$E\{N - \hat{N}_{SM}\}$ (veh)	SM RMSE (%)	$E\{N - \hat{N}_{SM}\}$ (veh)
20 s	15.8	0.5	29.0	-4.48	14.4	0.30
40 s	27.2	0.7	34.0	-3.98	20.6	0.01
60 s	28.6	-0.41	36.2	-4.37	23.9	-0.36
90 s	28.2	0.82	35.6	-3.50	23.4	0.01
stochastic	32.7	-0.11	40.8	-3.34	29.2	0.38

- Despite the availability of only one detector, \hat{N} follows N reasonably well (Figure 10) for all scenarios and for all ranges of link occupancy.
- The strong negative estimation deviations observed occasionally in Figure 10 are due to the ZSZO-phenomenon mentioned in section 3.5. Note that this phenomenon occurs only when the link is relatively full, because only then queuing vehicles may stand still around the detector that is located in the middle of the link. Note also that the utilization of only one detector renders the estimates quite sensitive to the ZSZO-phenomenon, because, in absence of any averaging with other non-affected detector measurements, this phenomenon may strike fully on the produced estimates. Finally, the phenomenon appears more frequently for longer signal cycles because of accordingly longer red phases leading to longer queuing.
- The RMSE is seen in Table 1 to increase with increasing downstream signal cycle; this is because of the ZSZO-phenomenon being more frequent for longer signal cycles; moreover a longer cycle is connected with longer green/red phases that produce stronger stop/start-waves on the link and affect the one single measurement in an accordingly stronger way.
- The average error (bias) $E\{N - \hat{N}\}$ is seen in Table 1 to be close to zero (less than 1 veh). Figure 13 displays the average space-occupancy curves analogous to Figure 9, albeit with $E\{o_s(x)\}$ replaced here by $E\{o_t(x^i)\}$ where $o_t(x^i)$, $i = 1, \dots, 150$, are time-occupancy measurements with $T = 20$ s, for the four basic scenarios. It is seen that a measurement detector in the middle of the link is indeed a good choice in the interest of a low estimation bias. The increasingly strong oscillations observed in Figure 13 towards the downstream end of the link are again due to stochastic effects

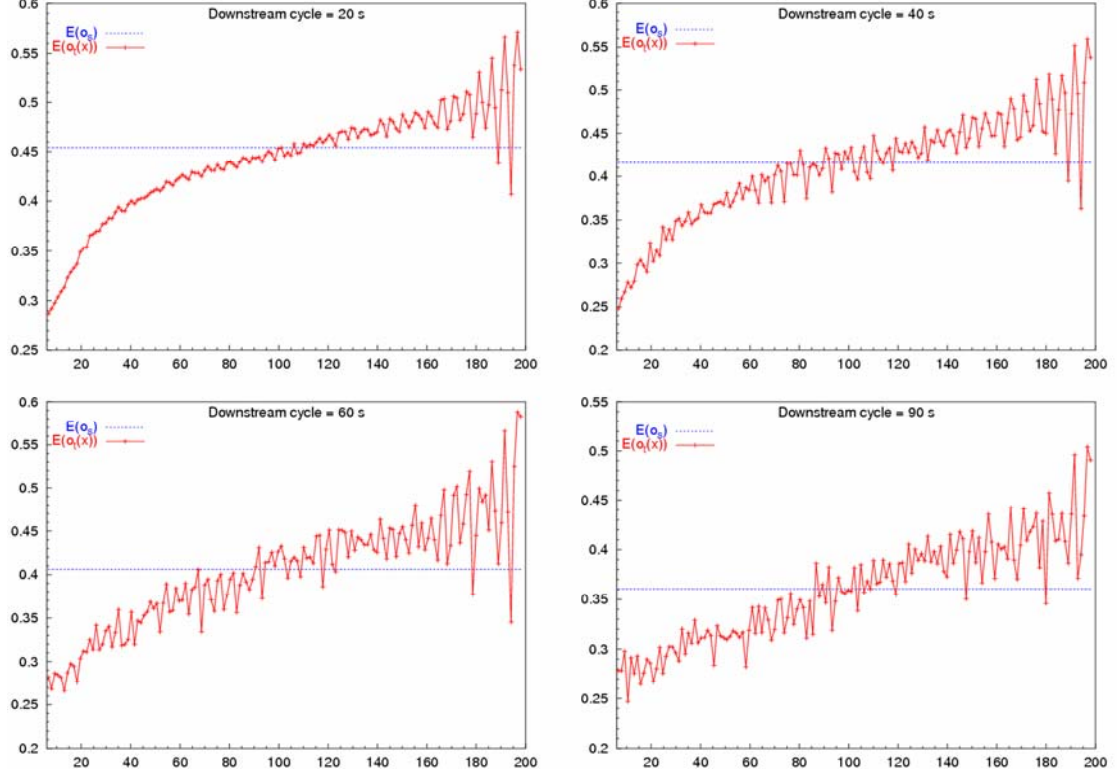


Figure 13: $E\{o_s\}$ and $E\{o_t(x)\}$ for the four basic scenarios.

related to the ZSZO-phenomenon which is more pronounced for higher x due to more frequent queuing.

As it concerns the performance of the Kalman Filter, it may be seen that:

- Despite the availability of only one link-internal detector, \hat{N}_{KF} follows N reasonably well (Figure 11) for all scenarios and for all ranges of link occupancy. The displayed results are clearly better than the corresponding measurement-only estimates reported in Table 1, which justifies the introduction of the Kalman Filter. In particular the ZSZO-phenomenon that was clearly visible in Figure 10 does not appear here thanks to the KF-imposed smoothing.
- The KF RMSE is seen in Table 2 to increase with increasing downstream signal cycle; this is due to the corresponding increase of the measurement error variance

that was attributed to the accordingly more frequent occurrence of the ZSZO-phenomenon.

- The average error (bias) $E\{N - \hat{N}_{KF}\}$ is seen in Table 2 to be close to zero (less than 1 veh) thanks to a similarly low bias of the measurement error. The initial estimation $\hat{N}_{KF}(0) = 5$ veh is rapidly reduced in all cases thanks to the correction term of the Kalman-filter equation (4.14).

The use of the exponential smoothing shows that:

- Despite the availability of only one detector, \hat{N}_{SM} follows the average N reasonably well (Figure 12) for all scenarios and for all ranges of link occupancy.
- The ZSZO-phenomenon does not appear in any scenario as much as the signal cycle is, because the estimations are not influenced by the creation of queues in the link.
- The displayed results shown in Table 3 are better than the corresponding measurement-only estimates reported in Table 1, because of the smoothing of the ZSZO-phenomenon, although the KF estimates reported in Table 2 remain better than whole.
- The SM RMSE is seen in Table 3 to increase with increasing downstream signal cycle; this is due to the corresponding increase of the measurement error variance that was attributed to the accordingly more frequent occurrence of the ZSZO-phenomenon.
- The average error (bias) $E\{N - \hat{N}_{SM}\}$ is seen in Table 3 to be close to zero (less than 1 veh) thanks to a similarly low bias of the measurement error and the initial estimation $\hat{N}_{SM}(0) = 5$ veh is rapidly reduced in all cases.

The various traffic scenarios may also be investigated in order to assess the robustness of the KF and the exponential smoothed estimators, the range of appropriate K - and K_{SM} -values in (4.14) and (5.1), respectively, as well as the sensitivity of the KF and exponential smoothed estimates to different K - and K_{SM} -values, respectively. To this end, Figure 14 and Figure 15 display the KF RMSE and the SM RMSE for each scenario in dependence of different K - and K_{SM} -values, respectively, as well as the corresponding measurement RMSE (horizontal lines) with measurement noise ψ . Note that in the case $K=0$ the Kalman Filter (4.14) exploits the boundary flow measurement (conservation equation) only, i.e. the internal occupancy measurement is not used, while in the case $K_{SM}=1$ the exponential smoothing equation (4.19) exploits the latest measurements N^m .

The results displayed in Figure 14 give rise to the following comments:

- The optimal gains K for all scenarios are in the range $[0.05, 0.25]$. In particular, the optimal K -values of the various scenarios are smaller if the measurement error

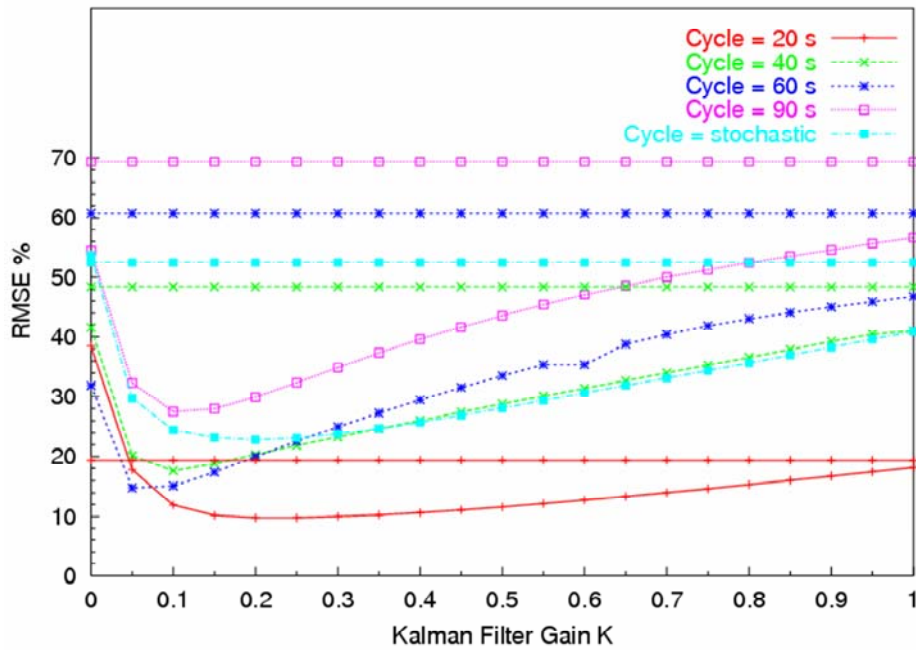


Figure 14: Measurements and KF estimation RMSE in dependence of the KF gain K .

variance is bigger, according to (4.16).

- The sensitivity of the KF RMSE within the mentioned range of K -values is rather low. This means that no elaborated fine-tuning of K is needed in practice; this is a quite significant property of the developed Kalman Filter because the exact fine-tuning of K would require exact vehicle-counts $N(k)$ that are quite cumbersome and costly to obtain in the field.
- The KF estimates (for optimal K) are much better than the conservation equation by itself ($K = 0$) or the internal measurement by itself for all scenarios, which demonstrates the utility of the KF estimator

The results displayed in Figure 15 give rise to the following comments:

- The optimal parameter K_{SM} for all scenarios are in the range $[0.1, 0.3]$. In particular, the optimal K_{SM} -values of the various scenarios are lower if the measurement error variance is bigger.

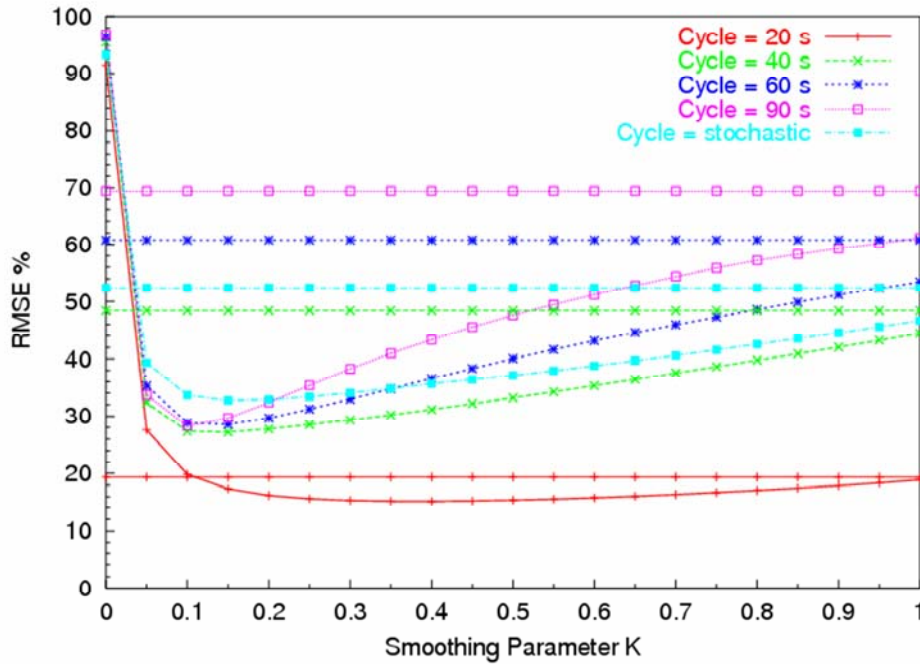


Figure 15: Measurements and smoothed estimation RMSE in dependence of the smoothing parameter K_{SM} .

- The sensitivity of the SM RMSE within the mentioned range of K_{SM} -values is insignificant. This means that no elaborated fine-tuning of K_{SM} is needed in practice.
- The exponential smoothed estimates (for optimal K_{SM}) are much better than the estimates for $K_{SM} = 0$ or the internal measurement by itself ($K_{SM} = 1$) for all scenarios, which demonstrates the utility of the exponential smoothing estimator.

7.3 Impact of the Update Period T

Figure 16 displays the RMSE in dependence of the update period T for the standard scenario for three cases of utilized detector numbers, namely $M = 1$, $M = 4$, $M = 200$.

The following observations are made:

- For $M = 1$, the RMSE is very high for very small T as expected; it reaches a plateau around $T = 20$ s due to better time-averaging; and then it increases slightly as T increases beyond 100 s (Figure 17) due to increasingly outdated estimations

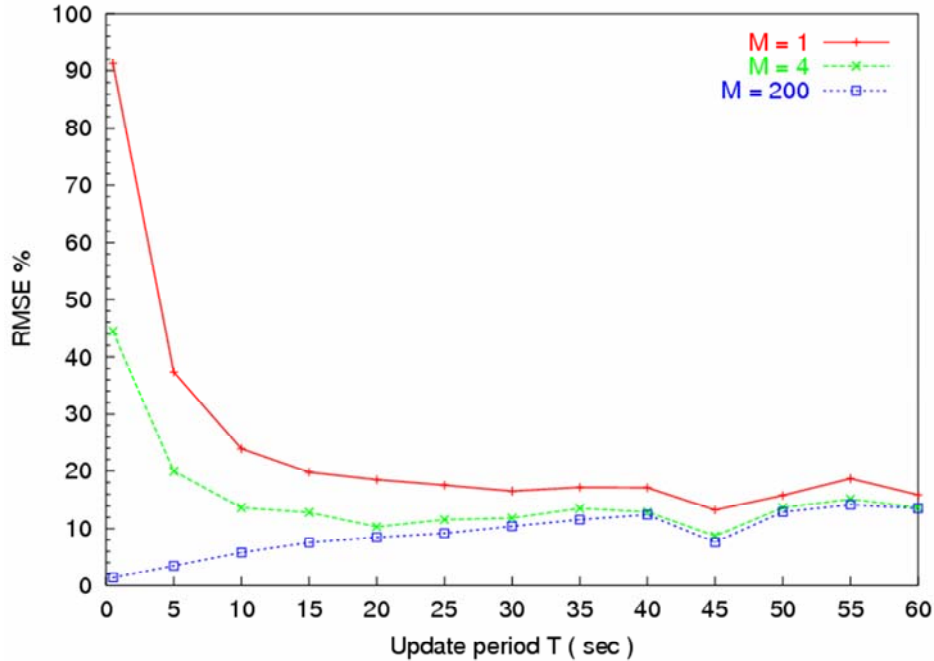


Figure 16: Impact of T on RMSE for standard scenario with $M = 1, 4, 200$.

according to section 3.5.

- For $M = 4$, the RMSE is much better than for $M = 1$ at small T , but the higher number of detectors brings only a small improvement (compared to $M = 1$) on the estimation accuracy as T increases.
- For $M = 200$, the RMSE is virtually zero at small T because the approximation (3.2) applies almost perfectly (excellent space-averaging); however, any increase of T is seen to deteriorate this excellent performance and, indeed, for high values of T , the amelioration over the cases $M = 4$ and even $M = 1$ becomes negligible.

Figure 17 displays the RMSE in dependence of T for the five scenarios ($M = 1$ for all).

The following two additional remarks are worth mentioning:

- The optimal update period T increases as the signal cycle increases. This is because the ZSZO-phenomenon and the longer-period oscillations (triggered by the green/red signal switching) that are stronger for longer signal cycles, are better smoothed for longer T .

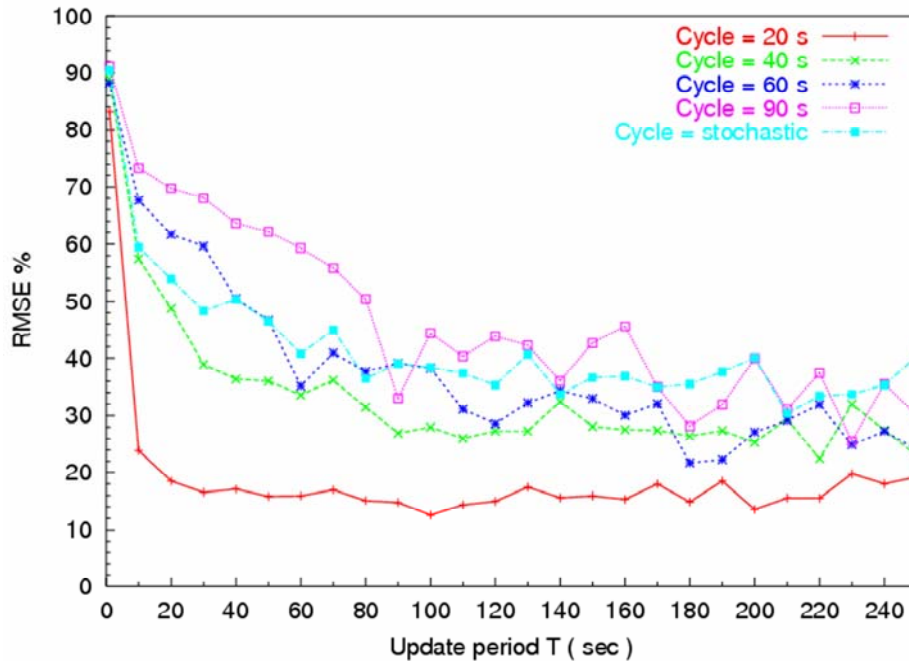


Figure 17: Impact of T on RMSE for five scenarios ($M = 1$).

- Some local minima are observed, e.g. most strikingly for the 90 s-cycle case at $T=90$ s and $T=180$ s. This is because of the synchronization of T with the signal cycles that produces estimates always at the same time within the period of the corresponding N -oscillation, which leads to a smoother estimation curve over time.

Figure 18 displays the KF RMSE of the standard and stochastic scenarios in dependence of the estimation/measurement period T for the following cases:

- $\hat{N}_{KF} = N^m$ (this means that KF RMSE equals RMSE) where N^m is produced according to (4.12) from occupancy measurements o_t^m stemming from one single internal detector without measurement noise Δo .
- \hat{N}_{KF} produced from the KF estimator (4.14) based on perturbed (i.e. $\psi \neq 0$) flow (q_{in}^m, q_{out}^m) and occupancy (o_t^m) measurements, the latter from one single internal detector. The also displayed gain parameter K was roughly fined-tuned for each individual KF case. Note that the displayed measurement RMSE reflects the corresponding values of the variance Z of the measurement error in dependence of T , while the displayed KF RMSE reflects the corresponding values of the variance Π of the KF estimation error in dependence of T . Figure 18 indicates that, as expected from section 3.5, the measurement RMSE (and hence Z) is quite high for very small T ; it reaches a plateau around $T=20$ s (Figure 18a) resp. $T=80$ s (Figure 18b); and then it increases slightly as T increases further. Clearly, the system error variance $T^2\Gamma$ increases quadratically with T since Γ is independent of T . Thus, the value of $\alpha = T^2\Gamma/Z$, appearing in (4.16), (4.17), depends on the update period T in a corresponding way. Indeed, the following observations are made from Figure 18:

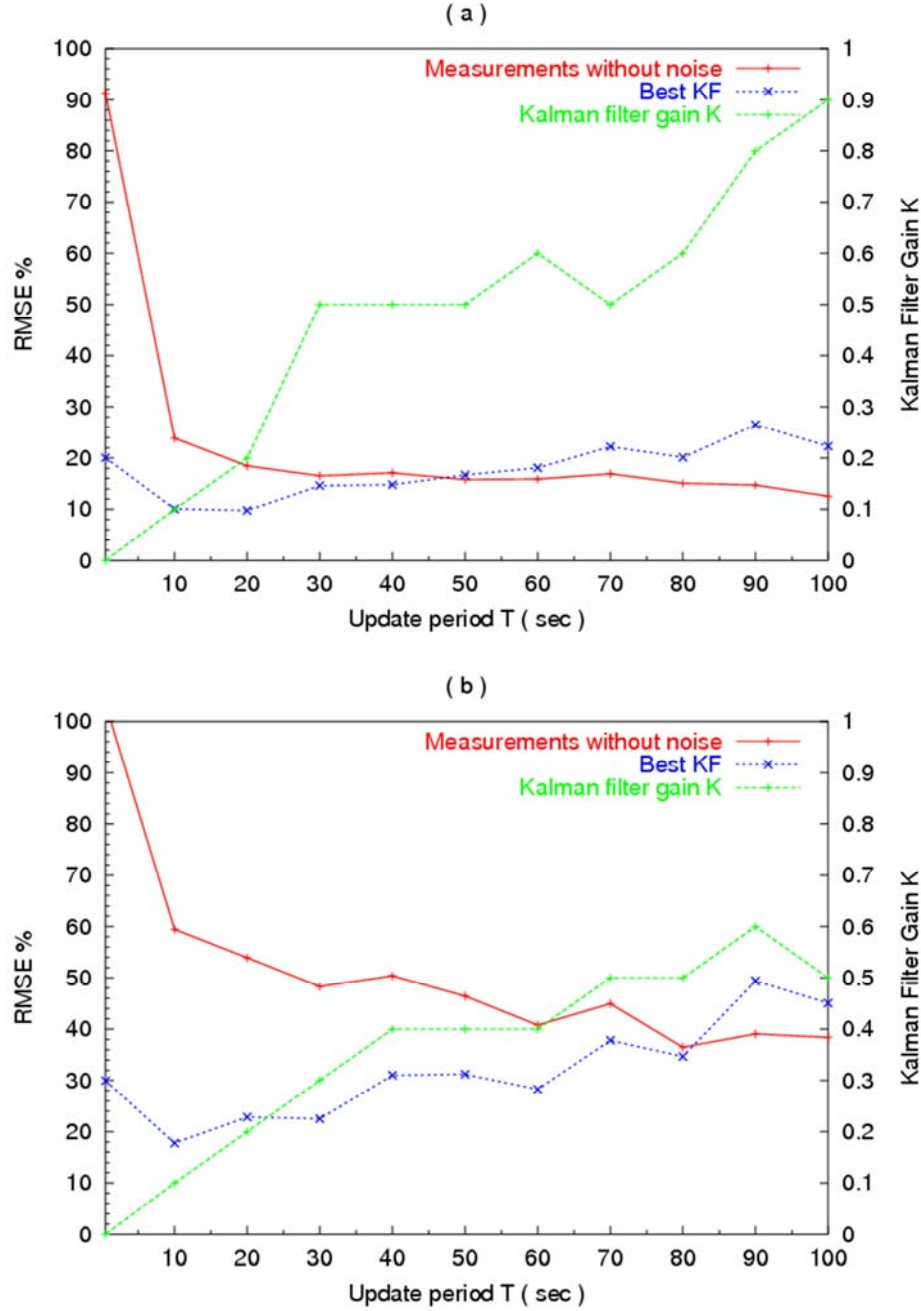


Figure 18: Measurements and KF estimation RMSE in dependence of update period T for (a) the standard and (b) the stochastic scenarios.

- In accordance with (4.17) the KF RMSE is slightly increasing with increasing T , although for $T < 20$ s (Figure 18a) resp. $T < 40$ s (Figure 18b) the increase is partly compensated by the continuously improved measurement (which leads to lower Z values).
- In accordance with (4.16), the optimal filter gain value is increasing with increasing T . More specifically, when T is small and the measurement noise variance Z is

large, the optimal gain is close to zero. As T increases and Z decreases, α and consequently the optimal gain K are increasing, the latter reaching a value around 0.5 for $T > 30$ s.

- For moderate T values, the KF produces better estimates than the measurements by themselves; but for $T > 50$ s (Figure 18a) resp. $T > 80$ s (Figure 18b) this situation is inversed due to increasing system error variance $T^2\Gamma$ while the measurement error variance Z is virtually constant.

In accordance to Figure 18, Figure 19 displays the SM RMSE of the standard and stochastic scenarios in dependence of the estimation/measurement period T . The estimations of the cases where $\hat{N}_{KF} = N^m$ and the \hat{N}_{KF} produced from the KF estimator have been hold on; there are also displayed the \hat{N}_{SM} produced from the exponential smoothed estimator and the smoothing parameter K_{SM} which was roughly fine-tuned for each individual exponential smoothed case. The following observations are made from Figure 19:

- The SM RMSE is quite low for small T and it is slightly decreasing until it reaches a plateau around $T = 20$ s, but it does not exceed the quality of the KF estimator; for the same period $T < 50$ s (Figure 19a) resp. $T < 80$ s (Figure 19b) the SM RMSE is better than the measurement RMSE; this situation is inversed for $T > 50$ s (Figure 19a) resp. $T > 60$ s (Figure 19b) as the smoothed estimations tend to follow the non-smoothed measurements which are better than the KF estimations.
- The optimal smoothing parameter value is increasing with increasing T . Even for small T , the K_{SM} -values are higher than the corresponding K -values and the more

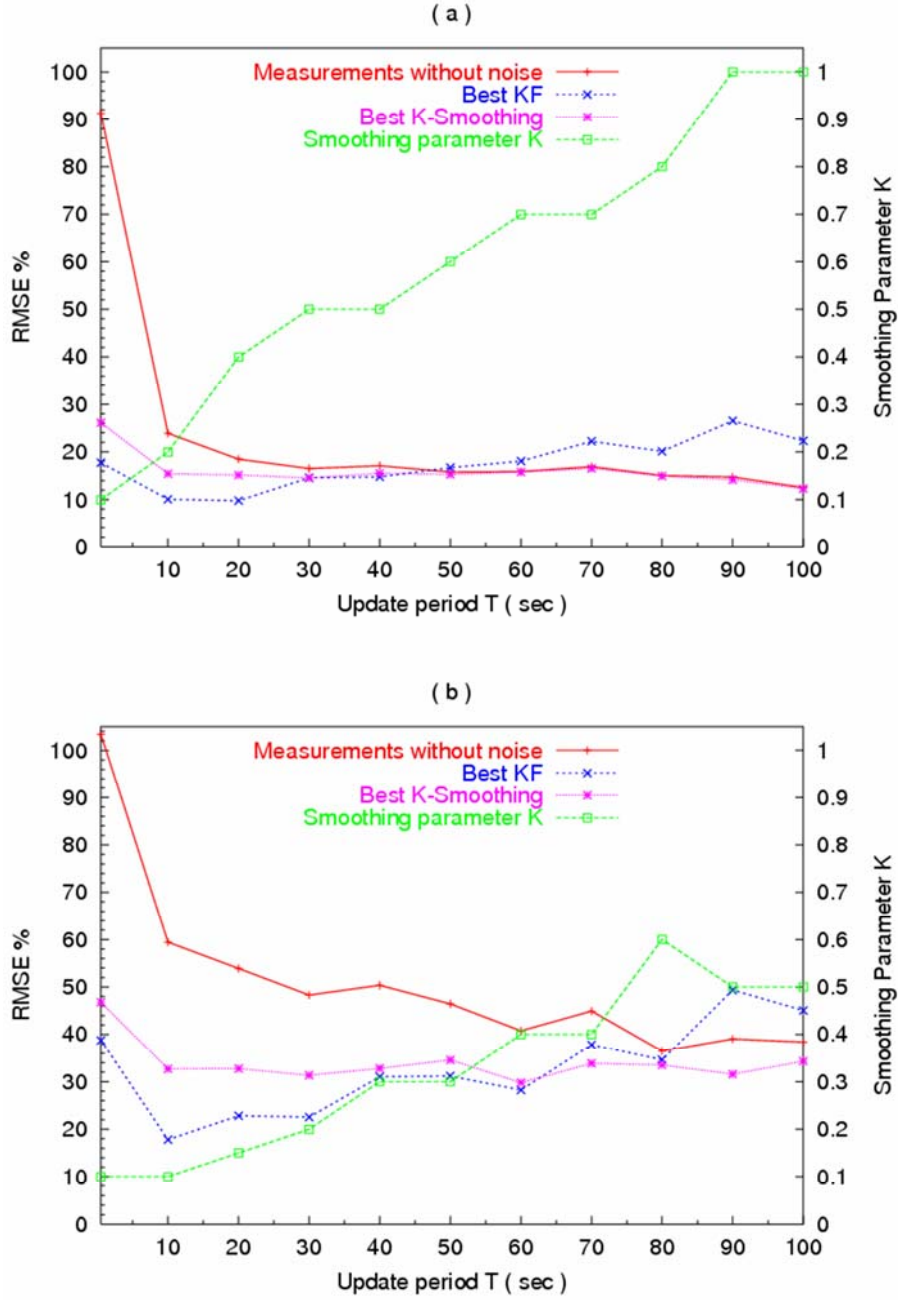


Figure 19: Measurements and smoothed estimation RMSE in dependence of update period T for (a) the standard and (b) the stochastic scenarios.

the smoothed estimation is coincided with the non-smoothed, the higher the value of K_{SM} is (the more use of the measurements).

7.4 Impact of the Number of Detectors

Figure 20 displays the RMSE of the standard scenario in dependence of the number of detectors M with $T = 20$ s, $T = 10$ s and $T = 0.5$ s. It may be seen again that, if T is very small, the RMSE increases strongly (compared to the case $T = 20$ s) if M is very small; but the RMSE decreases to become lower than in the case $T = 20$ s, as M increases; the reasons for this behaviour were already explained in section 7.3. Indeed, it is seen in Figure 20 that the curve for $T = 10$ s crosses the curve for $T = 20$ s around $M = 9$. Both curves are crossed by the curve for $T = 0.5$ s for some higher M ; e.g. at $M = 200$ (not visible in Figure 20), the RMSE for $T = 0.5$ s, 10 s, and 20 s is 1.3%, 5.7% and 8.4%, respectively.

Figure 21 displays the KF RMSE of the standard and stochastic scenarios in dependence of the number of internal detectors (measuring time-occupancy) with $T = 20$ s for the following cases:

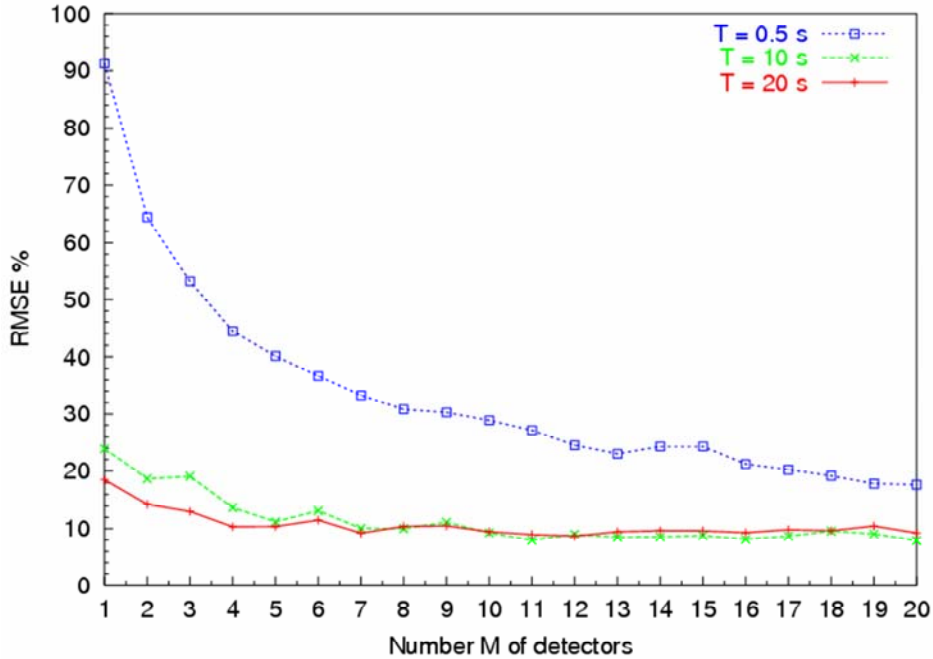


Figure 20: Impact of M on RMSE for the standard scenario with $T = 0.5$ s, $T = 10$ s, $T = 20$ s.

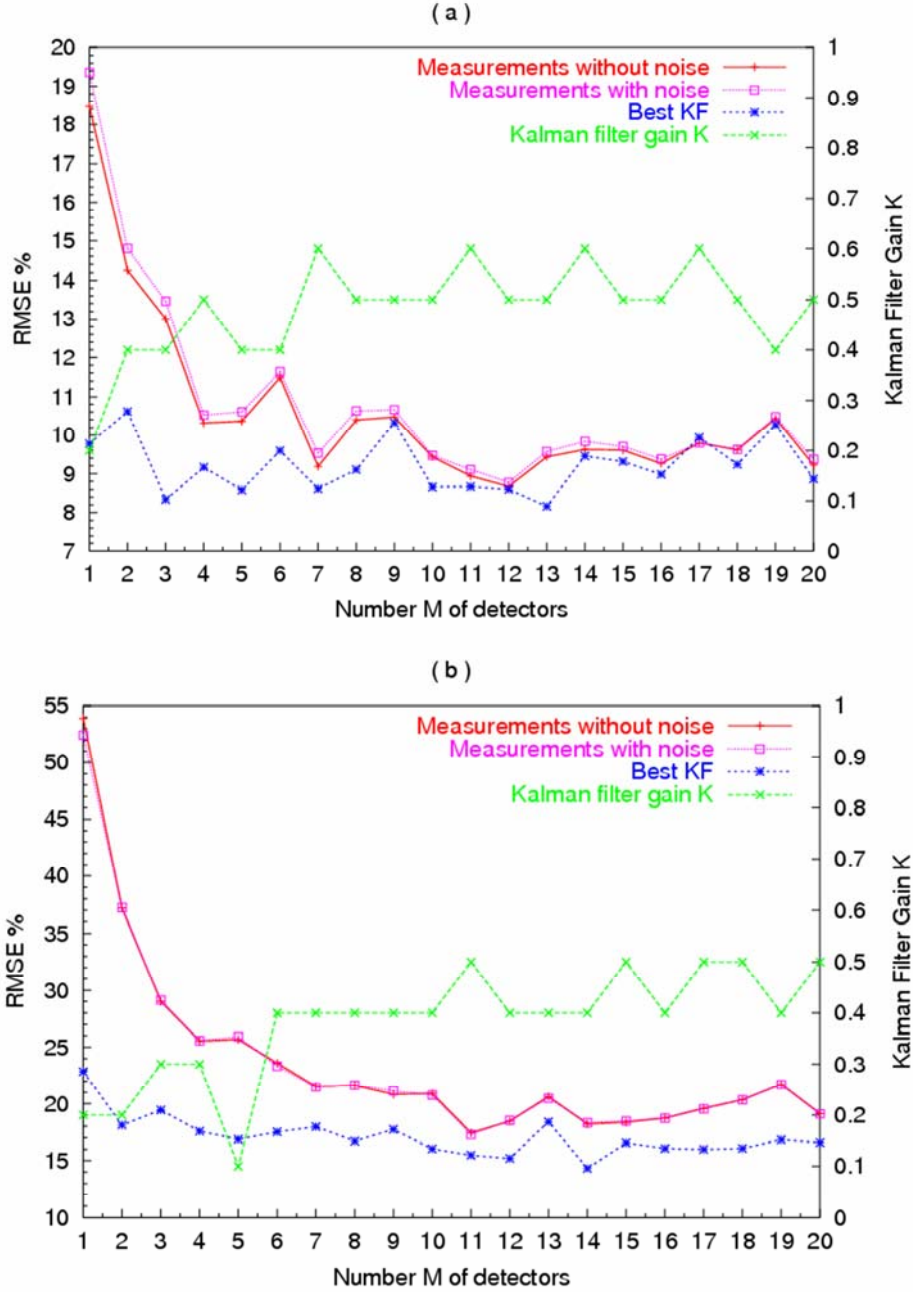


Figure 21: Measurements and estimation RMSE in dependence of the number of internal detectors for $T = 20$ s for (a) the standard and (b) the stochastic scenarios.

- $\hat{N}_{KF} = N^m$ produced from all available occupancy measurements without measurement noise Δo .
- $\hat{N}_{KF} = N^m$ as above but with occupancy measurements perturbed with measurement noise Δo .
- \hat{N}_{KF} produced from Kalman Filter (4.14) based on all available measurements (perturbed with noise) and roughly fine-tuned gains K .

The following observations are made from Figure 21:

- The impact of the occupancy measurement noise is quite negligible compared to the other error sources in the measurement equation (4.13).
- The measurement RMSE of 20% (Figure 21a) resp. 50% (Figure 21b) for one internal detector is reduced to around 10% (standard scenario) resp. 20% (stochastic scenario) for a sufficient number of internal detectors. The inclusion of more than ten internal detectors does not improve the quality of the measurements further.
- The KF RMSE is pretty stable around 9% (for the standard scenario) resp. 18% (for the stochastic scenario) for any detector number. Note that the optimal gain K takes values around 0.5 for $M > 5$ due to better measurement quality.
- Since the RMSE of 9% resp. 18% is reached with measurements only, by use of ten or more internal detectors while the KF needs only three detectors to reach this quality (two boundary and one internal detector), it may be concluded that, based on these scenarios, the application of the Kalman Filter allows for the cost of roughly seven detectors to be saved.

In accordance to Figure 21, Figure 22 displays the SM RMSE of the standard and stochastic scenarios in dependence of the number of internal detectors (measuring time-occupancy) with $T = 20$ s. The estimations of the cases where $\hat{N}_{KF} = N^m$ produced from occupancy measurements without measurement noise Δo and the \hat{N}_{KF} produced from the KF estimator have been hold; there are also displayed the \hat{N}_{SM} produced from the exponential smoothing and the roughly fine-tuned smoothing parameter K_{SM} . The following observation is made from Figure 22:

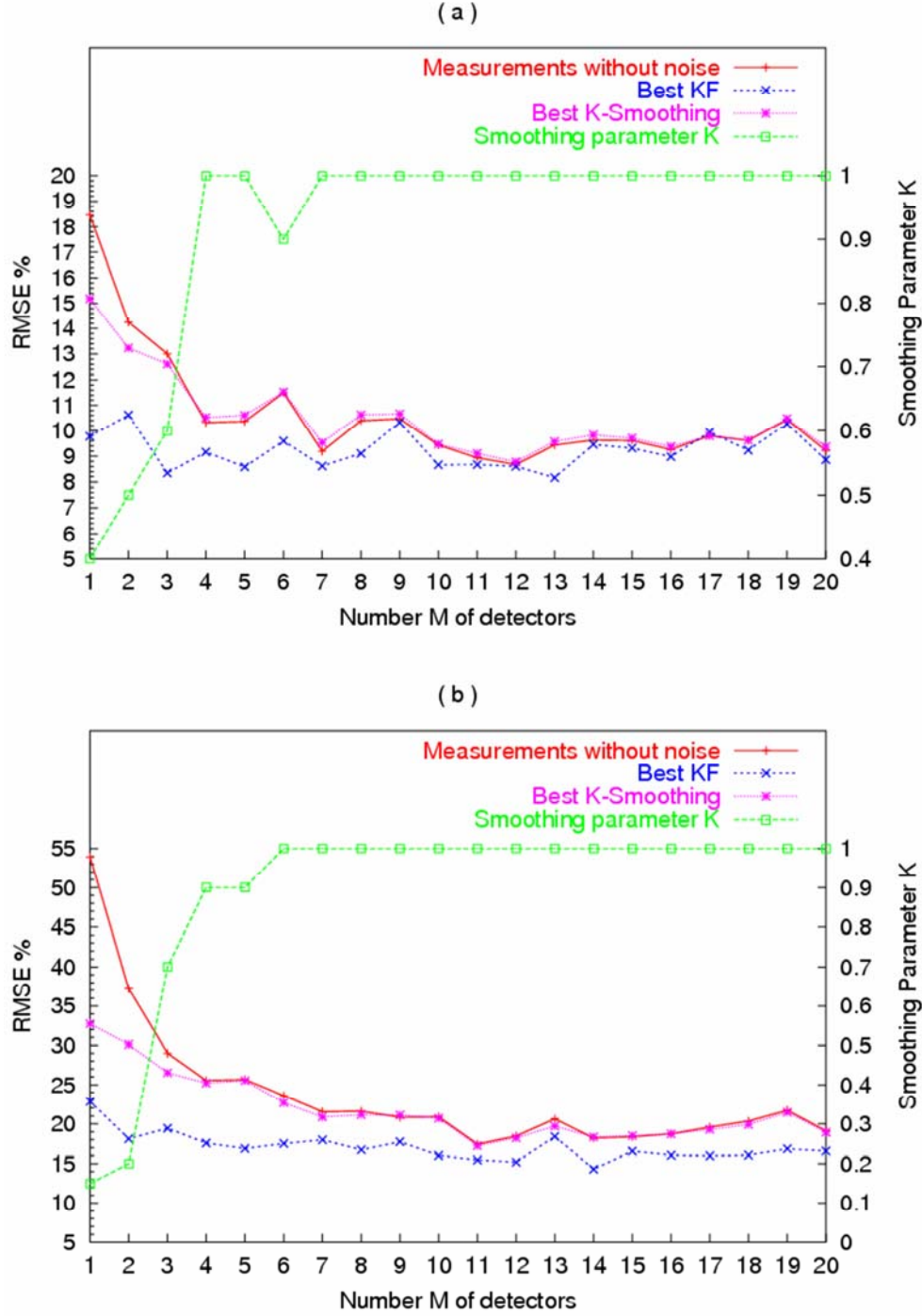


Figure 22: Measurements and smoothed estimation RMSE in dependence of the number of internal detectors for $T = 20$ s for (a) the standard and (b) the stochastic scenarios.

- The SM RMSE is very quickly coincided with the non-smoothed measurement estimation. For $M > 4$ the two curves almost become one, which means that for a high number of detectors is better just to use the measurements for the estimation of the vehicle-count in the link; the slightly differences between the estimated values are because of the measurement noise that was accounted at the exponential

smoothing estimation. Notice that the value of K_{SM} does accordingly fast reach its' maximum value.

7.5 Non-Zero Effective Detector Length

If the effective detector length is non-zero while measuring time-occupancy (or, equivalently, if the effective vehicle length is different than the physical vehicle length), then the measured time-occupancy is not a bias-free representation of the space-occupancy (section 2.3) and hence the bias of the errors ζ_o in (3.5), (4.11) and ζ in (3.8), (4.13) increases accordingly. If the non-zero effective detector length ε is known, one may convert the collected time-occupancy measurements o_t^m into virtually bias-free representations of the space-occupancy by multiplying them with $L^{Ph} / (L^{Ph} + \varepsilon)$ (section 3.5), in which case one may partly recover the results presented earlier. On the other hand, if this transformation is not performed and the biased measurements are actually used, then the produced estimates will be accordingly biased, even if the measurements are used to feed the Kalman Filter, as the filter has no means to reject the measurement bias according to (4.18).

Figure 23 displays the same information as Figure 13 for the standard scenario, now with two additional $E\{o_t(x^i)\}$ - curves; in the “biased” curve, the measurements are collected with $L_j = L_j^{Ph} + 1$ m which leads to an obvious bias in the whole curve; in the “corrected” curve, the biased measurements are multiplied with $L^{Ph} / (L^{Ph} + \varepsilon)$, in which case the original curve is recovered and the (additional) bias is virtually nullified.

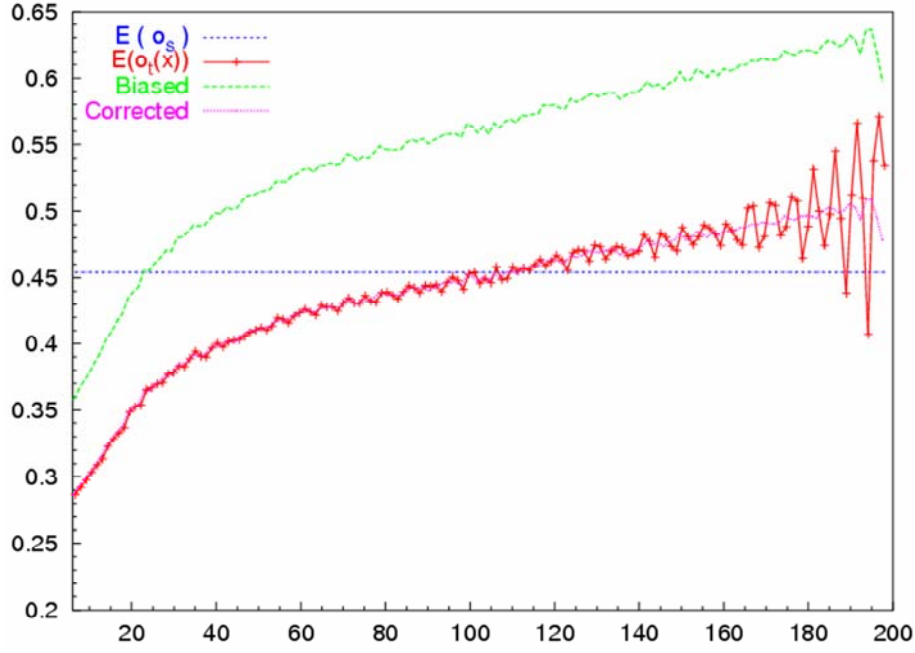


Figure 23: Average error (bias) for standard scenario with non-zero effective detector length.

Table 1 displays the RMSE and average error of the five scenarios for the cases without L_j -bias, with L_j -bias and for the corrected case. It is seen that the average error increases in the biased case but the additional bias virtually disappears in the corrected case. It is also interesting to notice that the RMSE of the corrected case (and, in some scenarios, even of the biased case) are lower than in the non-biased case; this is because simulated vehicles at standstill are separated by a gap of 1 m, and hence any effective detector length higher than 0.5 m is sufficient to suppress the occurrence of the ZSZO-phenomenon, thus reducing the estimation error variance. Clearly this RMSE improvement is more striking for the scenarios that are suffering more from the ZSZO-phenomenon, i.e. the ones with longer cycle times (see also Figure 9). These results suggest that it may be more beneficial for the estimation accuracy to employ detectors with non-zero (but approximately known) effective detector length ε and to proceed to a correction of the obtained measurements; than to adjust the detectors such that $L_j = L_j^{\text{Ph}}$.

Figure 24 presents the measurements and KF estimation results for different K values for the unbiased (standard scenario), the biased (with $\varepsilon = 1$ m) and the corrected (as mentioned above) cases. It may be seen that:

- The RMSE of the (biased) measurement of N^m is almost doubled compared to the RMSE of the measurement of the standard case as already known from section 3.5.
- Due to the lower quality of the measurement variance (higher Z), the optimal KF gain K is smaller (around 0.05) than in the non-biased case. The KF RMSE of the resulting KF estimates when using biased measurements is more than doubled compared to the unbiased case.
- In the corrected case (where measurements o_i^m are multiplied with $L^{Ph} / (L^{Ph} + \varepsilon)$), the original KF performance is virtually recovered for a gain K value similar as in the unbiased case. Indeed, Table 2 reveals (as already observed in Table 1) that, for some scenarios, the estimation RMSE of the corrected case may become lower than in the unbiased case, because a non-zero effective detector length suppresses

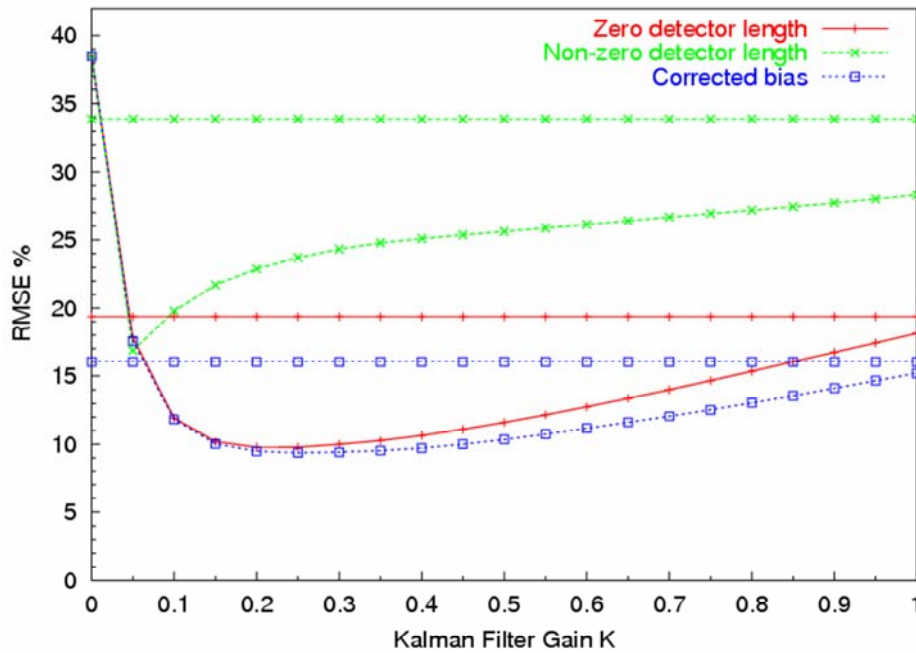


Figure 24: KF estimation for the case of non-zero effective detector length for standard scenario.

the ZSZO-phenomenon thus reducing the measurement and eventually the estimation error variance. These results amplify the suggestion that it may be more beneficial for the estimation accuracy to employ detectors with non-zero (but approximately known) effective detector length ε and to proceed to a correction of the obtained measurements; than to adjust the detectors such that $\mathbf{L}_j = \mathbf{L}_j^{\text{ph}}$.

Figure 25 presents the same as before measurements, and the exponential smoothed estimation results for different \mathbf{K}_{SM} values for the unbiased (standard scenario), the biased (with $\varepsilon = 1 \text{ m}$) and the corrected (as mentioned above) cases. It may be seen that:

- In any case, for $\mathbf{K}_{\text{SM}} = 0$ the SM RMSE has the same value (around 90%) as it depends only on the $\hat{\mathbf{N}}_{\text{SM}}(0)$ (see section 5.2).
- Due to the lower quality of the measurement variance, the optimal smoothing parameter \mathbf{K}_{SM} is bigger (around 0.5) than in the non-biased case. The SM RMSE

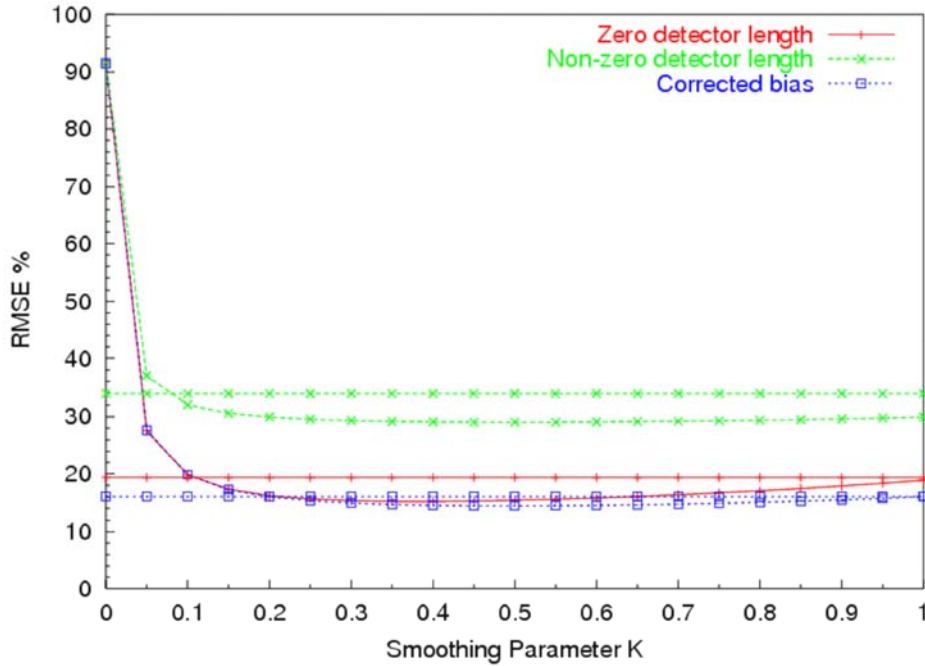


Figure 25: Smoothed estimation for the case of non-zero effective detector length for standard scenario.

of the resulting smoothed estimates when using biased measurements is almost doubled compared to the unbiased case.

- In the corrected case the original exponential smoothed performance is virtually recovered for a smoothing parameter K_{SM} value similar as in the unbiased case. Indeed, Table 3 reveals (as already observed in Table 1 and Table 2) that, for some scenarios, the estimation RMSE of the corrected case may become lower than in the unbiased case, because a non-zero effective detector length suppresses the ZSZO-phenomenon thus reducing the measurement and eventually the estimation error variance.

7.6 Uncertain Physical Vehicle Length Average

In all previous investigations we assumed that the average physical vehicle length L^{Ph} used to transform the time-occupancy measurement o_t^m into a vehicle-count N^m in (3.7), (4.12) is accurately known (equal to $L^{Ph} = 4\text{ m}$). If this value is not accurately known in practice, then an additional measurement bias will result whose effects are similar in nature as those of section 7.5. For example, the arriving traffic flow may contain varying percentages of trucks that change the average physical vehicle length accordingly. It should be emphasized, however, that the increased errors due to this bias will occur only if the results evaluation is based on real vehicle numbers N , whereby 1 truck = 1 vehicle. If the real quantity N is measured in p.c.u. (passenger-car units), then the RMSE for the estimates will be similar to the one of the standard case. Note that estimates \hat{N} , or \hat{N}_{KF} , in p.c.u. may be more useful for signal control or ramp metering applications where avoidance of queue spillback upstream of the link is a major concern.

To investigate this issue we have considered two variations of the standard scenario whereby the vehicles created in the microscopic simulation include 10% and 30% of trucks, respectively; trucks were assumed to have a length in the range [8 m,10 m] with uniform distribution. Note that in these two cases we still use $L^{\text{ph}} = 4 \text{ m}$ in (4.12) (as in the standard case) thus creating biased measurements. Figure 26 displays the results for both cases, along with the standard case for comparison. It may be seen that:

- The measurement RMSE of both cases increases with increasing percentage of trucks due to increasing bias.
- Due to the lower quality of the measurement (higher variance Z), the optimal KF gain K is accordingly smaller than in the standard case.
- The RMSE of the resulting KF estimates is increasingly higher for increasing bias, but the accuracy is still quite good for the 10% -trucks case.

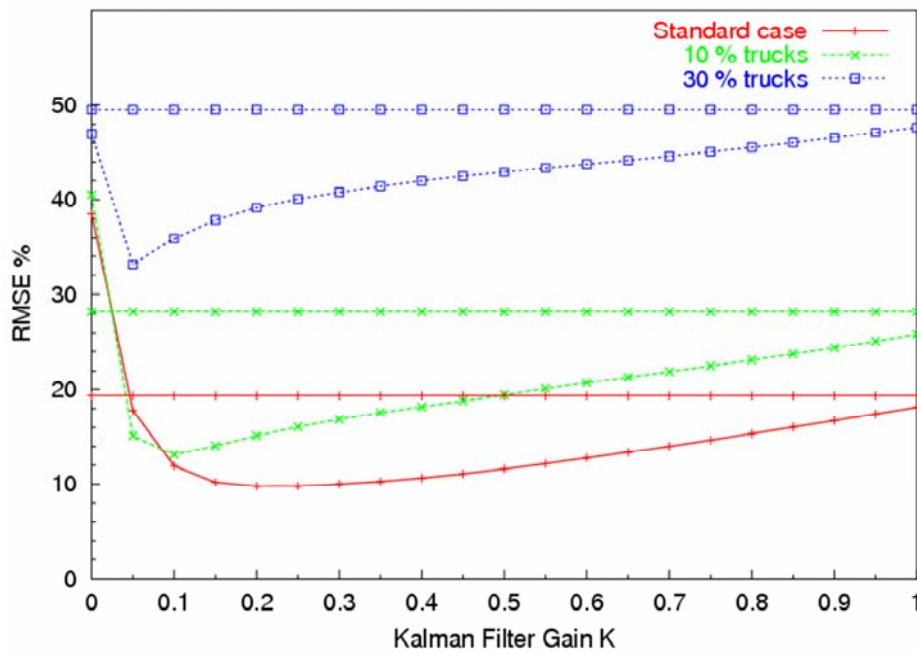


Figure 26: KF estimation for the cases of uncertain average physical vehicle length.

It should be emphasized that the above increased errors occur only if the results evaluation is based on real vehicle numbers N , whereby 1 truck = 1 vehicle. If the real quantity N is measured in p.c.u. (passenger car units) and we assume here 1 truck = 2.25 p.c.u., then the RMSE for the measurement and the estimates are similar to the standard case as Figure 27 indicates. In fact, the measurements in this case may be deemed unbiased while the variance of the state error increases slightly because the flow measurements and hence the conservation equation address vehicles, not p.c.u.; since the state error variance increases while Z is virtually the same as in the standard scenario, the resulting optimal K are slightly greater than in the standard case, but the sensitivity of the results is rather low in a broad range of K -values.

The same analysis has been done for the case of the exponential smoothing estimation. Figure 28 and Figure 29 show that:

- Due to the lower quality of the measurement, the optimal smoothing parameter K_{SM} is accordingly smaller than in the standard case.

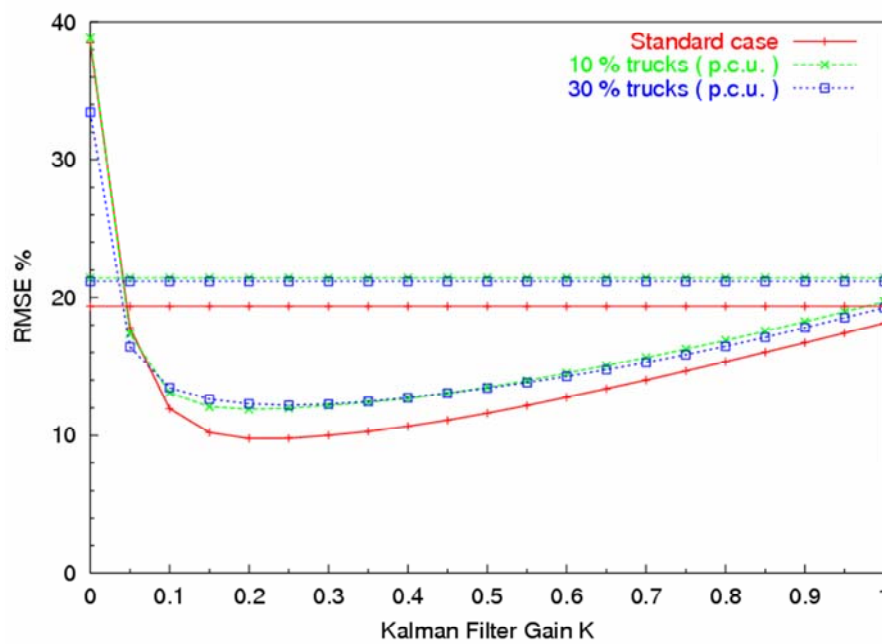


Figure 27: KF estimation for the cases of 10% and 30% trucks with N in p.c.u. and 1 truck = 2.25 p.c.u.

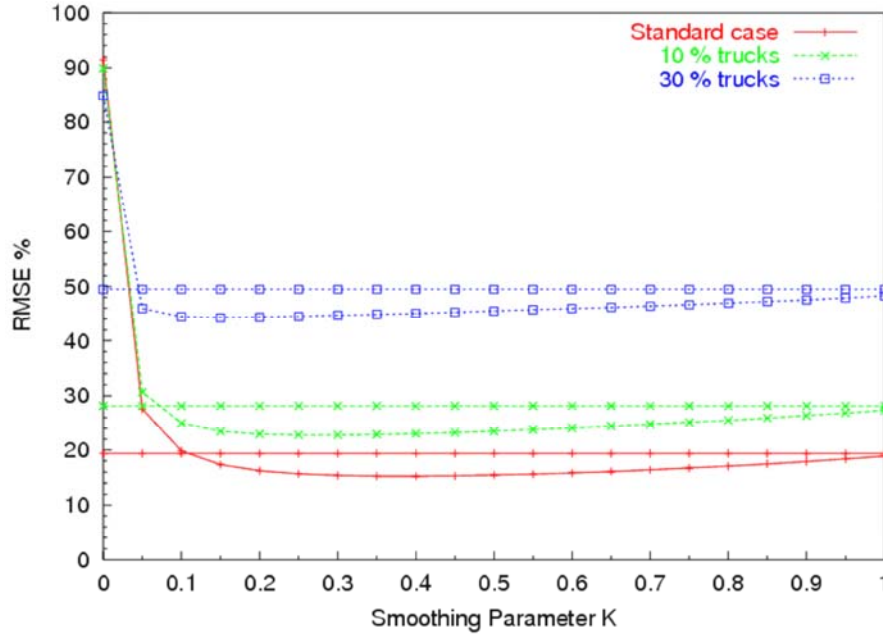


Figure 28: Smoothed estimation for the cases of uncertain average physical vehicle length.

- The SM RMSE of the resulting smoothed estimates is increasingly higher for increasing bias, but the accuracy is still quite good for the 10% -trucks case.
- If the real quantity N is measured in p.c.u. (Figure 29), again the resulting optimal K_{SM} are slightly greater than in the standard case, but the sensitivity of the results is almost zero in a broad range of K_{SM} -values.

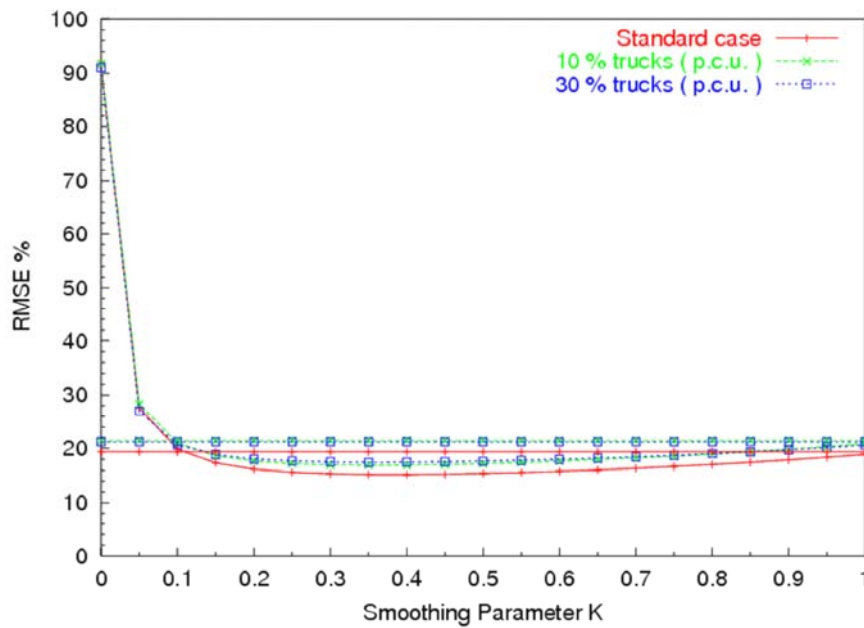


Figure 29: Smoothed estimation for the cases of 10% and 30% trucks with N in p.c.u. and 1 truck = 2.25 p.c.u.

7.7 Longer Link

To check the KF efficiency in the case of links with different geometry, the link length of the standard case was doubled to 394 m. Note that a larger link length Δ increases the variance Z of the measurement as evidenced by the definition of the measurement error ζ in (4.13) while the variance Γ of the system error remains unchanged. Hence, it is expected that the resulting optimal value of the KF gain K will be smaller than in the standard case and the results displayed in Figure 30 actually confirm this conclusion. Since RMSE is a relative error and the absolute values of the vehicle-count N are higher for a longer link, the relative estimation accuracy of both the measurement and the Kalman Filter in Figure 30 is slightly better than in the standard case. The same behaviour is appearing with the exponential smoothing analysis shown in Figure 31 where still the KF estimation remains the best. Figure 32 displays the time trajectories of the real and KF-estimated vehicle-counts N and \hat{N}_{KF} , respectively, while Figure 33 displays the time trajectories of the real and exponential smoothed estimation of the vehicle-counts N and \hat{N}_{SM} , respectively. The first confirms the excellent performance

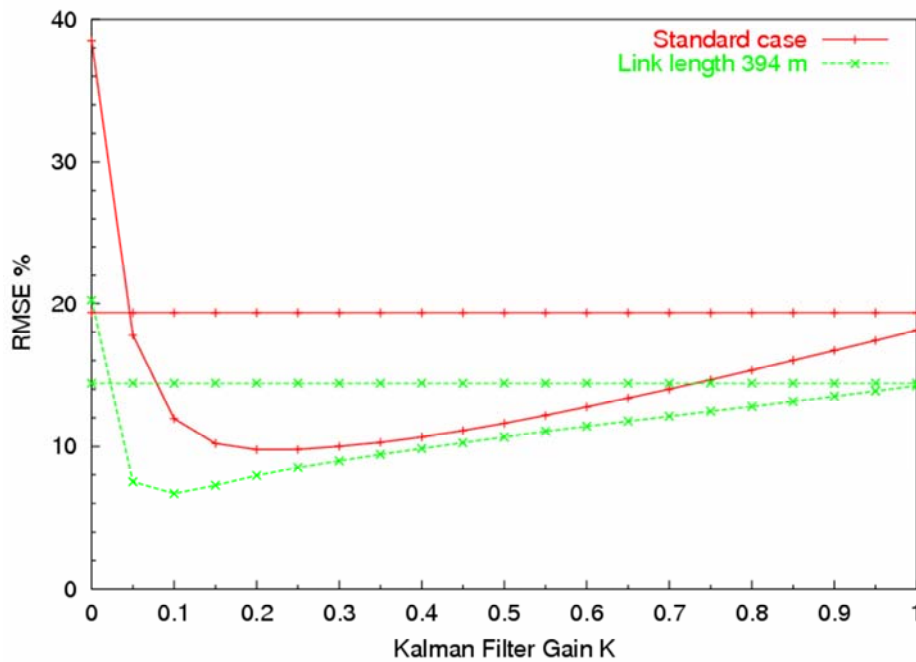


Figure 30: KF estimation results for a longer link.

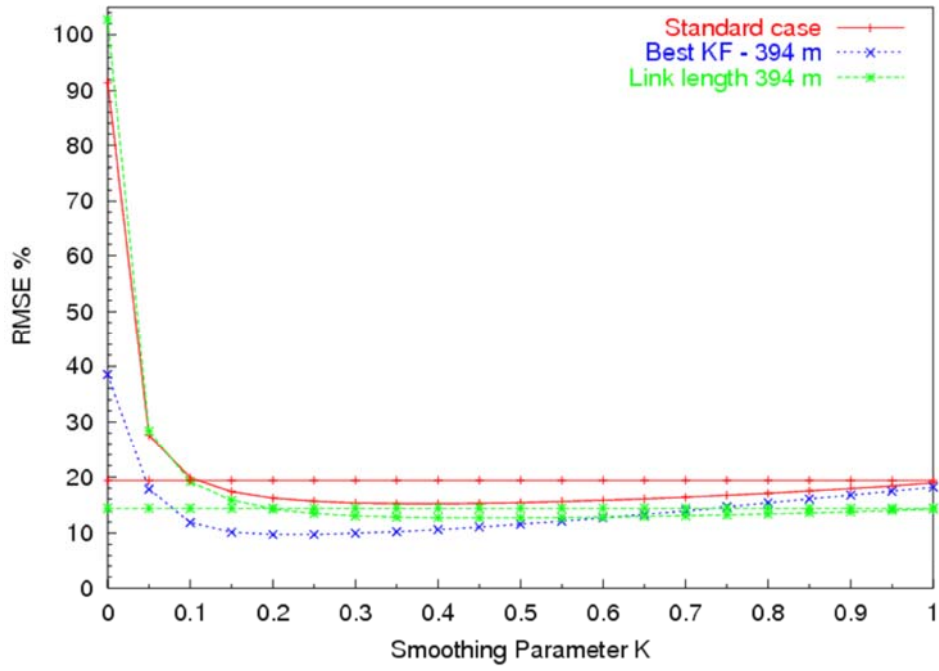


Figure 31: Smoothed estimation results for a longer link.

of the filter also for a longer link and the second shows the satisfactory performance of the exponential smoothing even for a longer link.

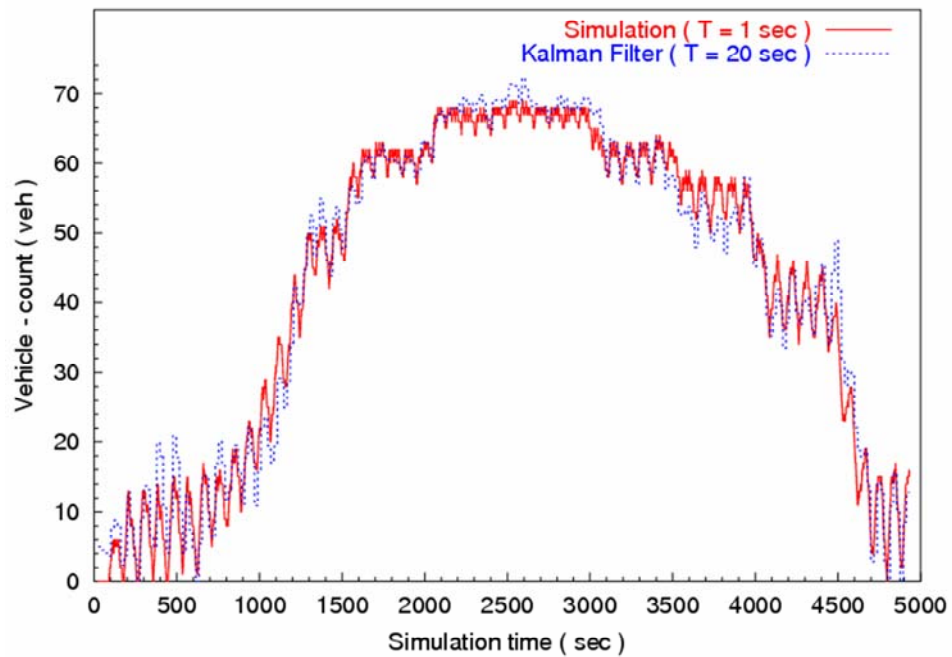


Figure 32: Real and KF estimated vehicle-counts over time for longer-link scenario.

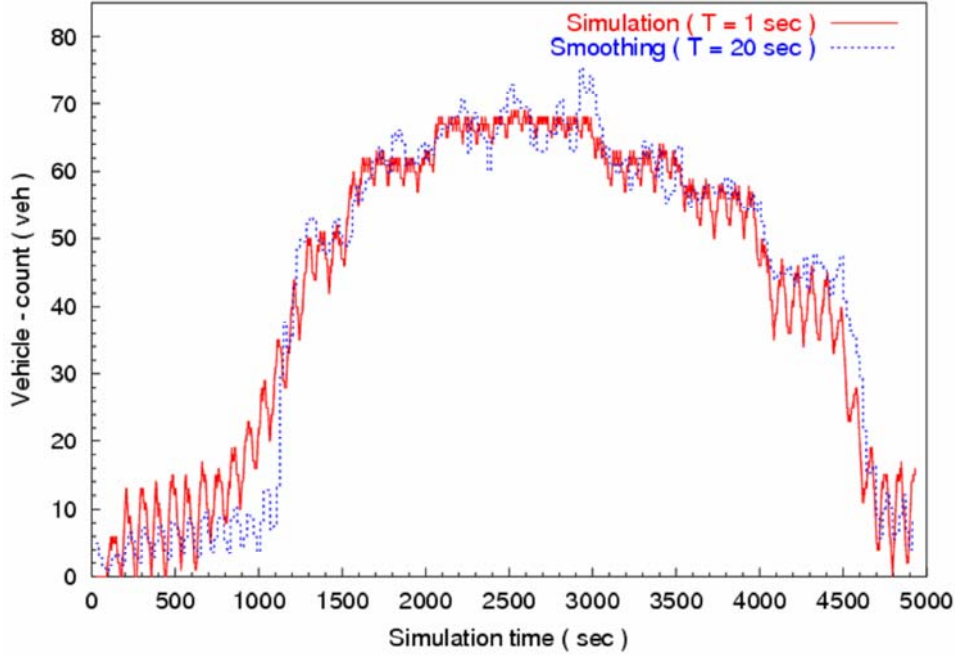


Figure 33: Real and smoothed estimated vehicle-counts over time for longer-link scenario.

7.8 Large Initial Estimation Error

In all previous investigations it was assumed that the initial estimate is $\hat{N}_{KF}(0) = 5$ veh while the real vehicle count is $N(0) = 0$ veh, i.e. an initial estimation error of 5 veh was imposed. The importance of the correction term in the filter equation (4.14) may be appreciated in a last scenario where an even larger initial estimation error is assumed with $\hat{N}_{KF}(0) = 20$ veh. Without the correction term, this initial estimation error would not be reduced and, indeed, Figure 34 indicates that the KF RMSE for $K = 0$ is much higher for $\hat{N}_{KF}(0) = 20$ veh than in the standard case. Moreover, the optimal KF gain K is seen to be slightly larger than in the standard case due to the larger initial state error. In contrast, Figure 35 shows that the SM RMSE for $K_{SM} = 0$ is much lower for $\hat{N}_{SM}(0) = 20$ veh than in the standard case probably because the initial estimation is closer to the average number of vehicles in the link. Figure 36 shows that when the filter is used, thanks to the correction term, this large initial error is rapidly reduced within a

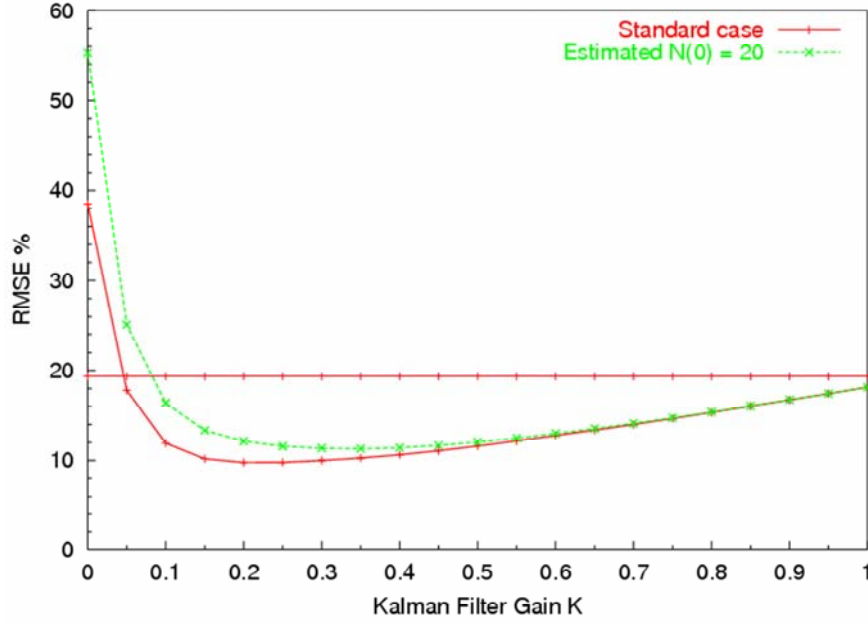


Figure 34: KF estimation for the case of larger initial estimation error ($\hat{N}(0) = 20$ veh).

few estimation time-steps, while Figure 37 shows that the exponential smoothing is not influenced by the $\hat{N}_{SM}(0)$, besides for $K_{SM} = 0$, as the estimation is similar to the one for $\hat{N}_{SM}(0) = 5$ veh and always follows the average number of vehicles that have passed the middle of the link until the time examined.

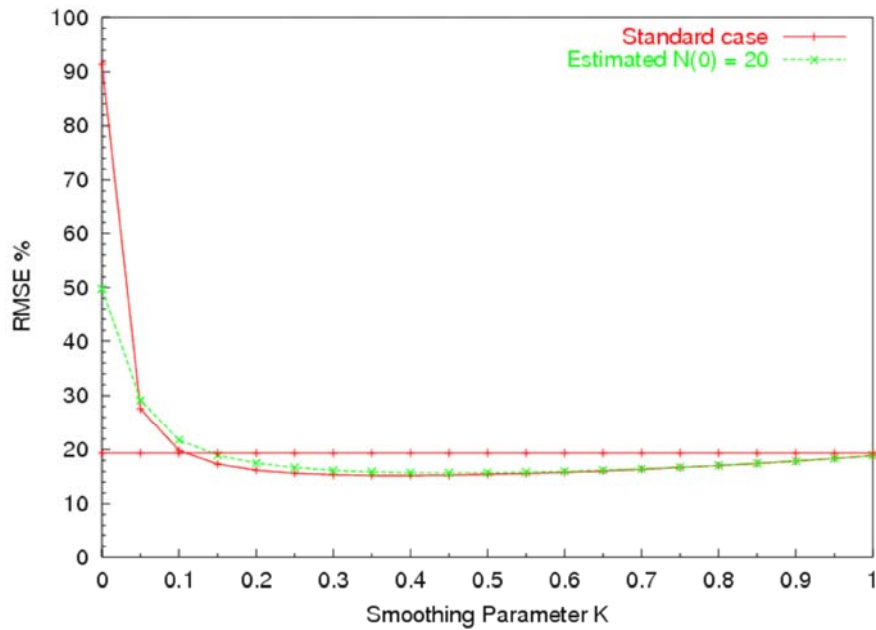


Figure 35: Smoothed estimation for the case of larger initial estimation error ($\hat{N}(0) = 20$ veh).

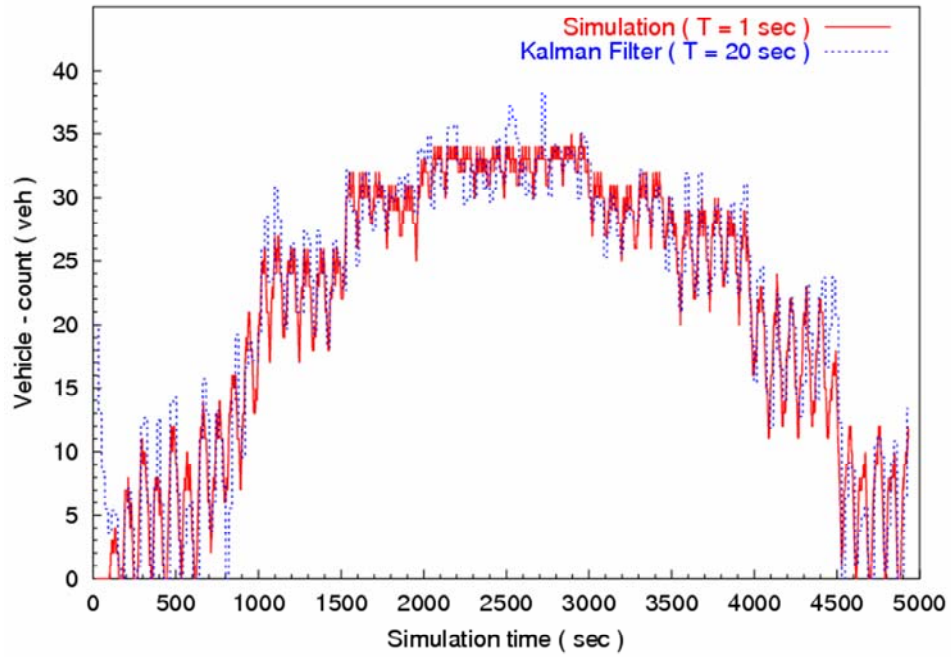


Figure 36: Real and KF estimated vehicle-count over time for the scenario with $\hat{N}(0) = 20$ veh.

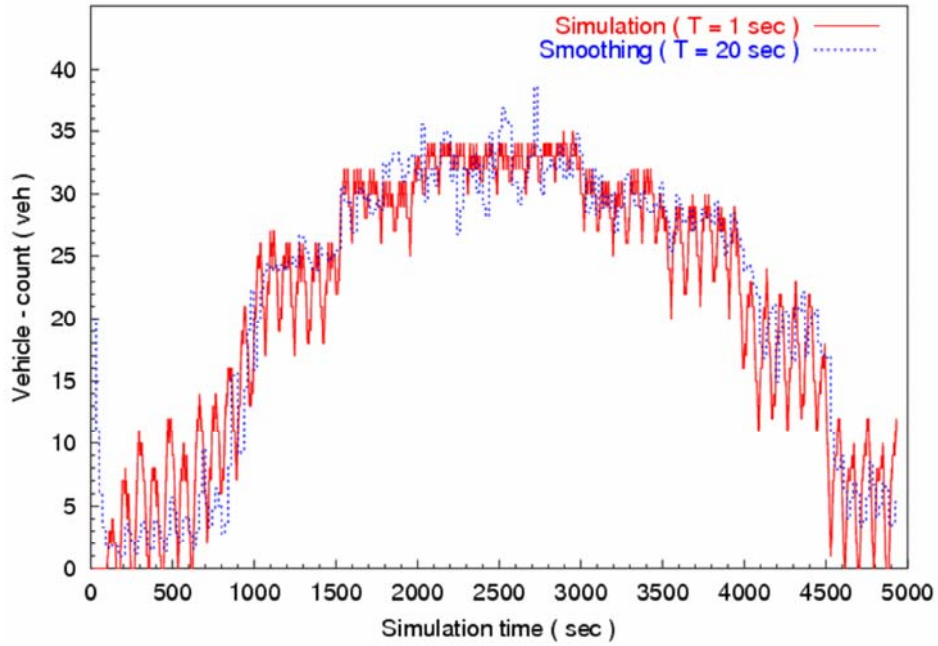


Figure 37: Real and smoothed estimated vehicle-count over time for the scenario with $\hat{N}(0) = 20$ veh.

CHAPTER 8

GUIDELINES, RECOMMENDATIONS AND CONCLUSIONS

8.1 Guidelines and Recommendations for Practical Application

We will here summarize the necessary steps, equations and recommend parameter values required if the described estimation scheme for vehicle-counts in signalized links is to be adopted and implemented.

To start with, a basic prerequisite is the availability of two boundary detectors measuring flows q_{in}^m and q_{out}^m , respectively; and one internal detector (preferably located around the middle of the link) measuring time-occupancy o_t^m , see Figure 5. In case of availability of multiple internal detectors, they should be preferably placed as Figure 8 indicates and the overall occupancy measurement o_t^m feeding the Kalman Filter should be the average of all available time-occupancies according to (3.2). All occupancies in this thesis are assumed to take values within the range $[0,1]$.

In real-time operation, the filter is fed with the latest available measurements $q_{in}^m(k-1)$, $q_{out}^m(k-1)$, $o_t^m(k-1)$, collected during the time period $[(k-1)T, kT]$, where $k=1,2,\dots$ is the discrete time index, to produce the estimated vehicle number $\hat{N}_{KF}(k)$ valid for the time instant kT . The usually required range of values for T is $[10\text{ s}, 30\text{ s}]$.

The measured occupancy o_t^m should ideally be collected with an effective detector length ε equal to zero. If the corresponding adjustment of the detectors is not possible, the collected occupancy should be multiplied with $L^{Ph}/(L^{Ph} + \varepsilon)$ before further use (see section 7.5), where L^{Ph} is the average physical vehicle length (around 4m). Typical ε -values are 1...2m depending on the employed loop detectors. Note that, according to section 7.5, the case $L_j \neq L_j^{Ph}$ with correction of the time-occupancies may lead to better results than the $L_j = L_j^{Ph}$ case.

Subsequently, the o_t^m measurement must be converted into a corresponding N^m measurement by use of (4.12). Note that the term $\Delta\lambda/L^{Ph}$ appearing in (4.12) is equal to (and could be replaced by) N_{max} , the maximum number of vehicles that could be accommodated in the link in a bumper-to-bumper manner. In case of a non-negligible percentage of trucks, there are two options:

- Truck length is not considered in the utilised average physical length L^{Ph} , which remains equal to around 4m ; in this case, the Kalman-filter estimates \hat{N}_{KF} will be automatically delivered in p.c.u., where $L^{Ph} \triangleq 1$ p.c.u. (this option is recommended).
- Truck length is considered (according to their usual percentage) when selecting the value of L^{Ph} (which would then be naturally higher than 4m), in this case, the Kalman-filter estimates \hat{N}_{KF} will be delivered in veh, that may be either passenger cars or trucks in the pre-specified proportion.

After these arrangements the Kalman-filter estimate $\hat{N}_{KF}(k)$ may be produced by use of (4.14) and subsequent possible truncation if the estimate exceeds the range $[0, N'_{\max}]$. The recommended range of values for the filter gain K is $[0.05, 0.3]$, but if no fine-tuning is effectuated, a value of $K = 0.1$ is deemed quite appropriate and robust. Finally the N'_{\max} value needed to truncate the filter results, corresponds to the maximum number of vehicles that can be accommodated in the link at standstill, including the usual safety distance D , i.e. $N'_{\max} = \Delta\lambda / (L^{Ph} + D)$.

8.2 Conclusions

The relationships between instantaneous space-variables and (easily measurable) local time-variables in largely homogeneous and stationary traffic flow, as typically encountered in uninterrupted traffic conditions, were reviewed. A quite elaborated analysis and microscopic simulation investigation was conducted for signalized links with inherently strong traffic flow variations triggered by traffic signal switchings. A number of influencing factors when estimating space-variables from measured local time-variables was analysed and illustrated in detail.

In subsequence a Kalman-filtering algorithm that may further reduce the estimation error was proposed based on the insights gained in this analysis. A rather simple Kalman-filter estimator was designed for the number of cars included in a signalized link. The estimator was found in manifold simulation investigations to be quite efficient and robust. The algorithmic intelligence of the estimator was shown to replace for several additional loop detectors that would be required in order to reach an equivalent accuracy without the estimator. Several issues and options were investigated in detail in order to come up with suitable conclusions and recommendations with respect to various aspects including:

- The appropriate measurement/estimation sampling time T .
- The impact of multiple internal detectors for occupancy.
- The proper values of the filter gain parameter K .
- The impact (and countermeasures) of non-zero effective detector length in the occupancy measurements.
- The impact of trucks.
- The impact of various traffic conditions and the ramp geometry.

Field testing is the next step in the development of this algorithm.

APPENDIX A

C	Input File				
C1	1.Duration of simulation(sec)	2.Simulation step(sec)	3.Distance of recycle(m)		
	5000	0.25	0		
C2	4.Length of road(m)	5.Number of traffic lights			
	1810	2			
C3	6.Distance observation of traffic light(m)	7.Cycle of each traffic light(sec)			
	50	90 20			
C4	8.Green period of each traffic light(sec)				
	50 13				
C5	9.Place of each traffic light(m)				
	1600 1800				
C6	10.Number of vehicles in road	11.Mean lenght of a vehicle(m)	12.Mean lenght of a truck(m)	13.Initial distance of vehicles(m)	
	0	4	9	10	
C7	14.Safety distance between vehicles,D(m)	15.Initial speed of vehicles(m/sec)	16.(Starting) Reaction time(simulation time steps - min=1)		
	1	14	1		
C8	17.Maximum acceleration(m/sec^2)	18.Maximum deceleration(m/sec^2)	19.Maximum desirable speed of vehicles(m/sec)		
	1.5	-6	16.5		
C9	20.Constant var g_acceleration(1/sec)	21.Constant var g_deceleration(1/sec)	22.Constant lambda,Λ(1/sec)		
	2	2	0.7		
C10	23.Minimum time-distance of vehicles(sec)	24.Time entrance of 1st vehicle(sec)			
	2	5			
C11	25.Number of Ot detectors	26.Length of Ot detector(m)	27.Detector upstream(1->Y/0->N)	28.Detector downstream(1->Yes/0->No)	
	1	0	0	0	
C12	29.Number of flow detectors	30.Place of each flow detector(m)	31.Stochastic vehicle length(1->Y/0->N)	32.Stochastic trucks(1->Y/0->N)	
	2	1606 1800	1	0	
C13	33.Percentage of time occupancy noise(%)	34.Percentage of flow noise(%)	35.Stochastic reaction time(1->Y/0->N)	36.Stochastic stop before traffic light(1->Y/0->N)	
	0.05	0.2	0	1	
C14	37.Period of simulation sampling(sec)	38.Update period T of KF(sec)	39.KF gain K	40.Estimated N(0)	
	1	20	0	5	
E					

APPENDIX B

In the tables below are shown the green phases (in sec) of both traffic signals during the corresponding simulation periods (also in sec):

Table 4: Green phases of traffic signals for the standard scenario.

Standard scenario									
Update period (sec)		20							
	Cycle (sec)	Simulation Periods (sec) & Corresponding Green Phases (sec)							
Upstream Signal	90	0→700	→2000		→5000				
		55	60		55				
Downstream Signal	20	0→700	→1500	→2000	→3000	→3500	→4000	→4500	→5000
		13	11	7	5	7	9	13	17

Table 5: Green phases of traffic signals for the scenario with cycle of 40 s at the downstream traffic signal.

Basic scenario 1									
Update period (sec)		20							
	Cycle (sec)	Simulation Periods (sec) & Corresponding Green Phases (sec)							
Upstream Signal	90	0→1500		→2000	→5000				
		50		55	50				
Downstream Signal	40	0→1000	→1500	→2000	→3000	→3500	→4000	→4500	→5000
		25	18	13	8	13	19	27	33

Table 6: Green phases of traffic signals for the scenario with cycle of 60 s at the downstream traffic signal.

Basic scenario 2									
Update period (sec)		20							
	Cycle (sec)	Simulation Periods (sec) & Corresponding Green Phases (sec)							
Upstream Signal	90	0→5000							
		50							
Downstream Signal	60	0→1000	→1500	→2000	→3000	→3500	→4000	→5000	
		36	28	18	10	18	30	43	

Table 7: Green phases of traffic signals for the scenario with cycle of 90 s at the downstream traffic signal.

Basic scenario 3									
Update period (sec)		20							
	Cycle (sec)	Simulation Periods (sec) & Corresponding Green Phases (sec)							
Upstream Signal	90	0→3500						→5000	
		55						50	
Downstream Signal	90	0→700	→1500	→2000	→2200	→3000	→3500	→4100	→5000
		55	45	35	25	15	30	42	55

Table 8: Green phases of traffic signals for the stochastic scenario with cycle of x sec at the downstream traffic signal (x is an integer between 10 s and 90 s).

Stochastic scenario								
Update period (sec)		20						
	Cycle (sec)	Simulation Periods (sec) & Corresponding Green Phases (sec)						
Upstream Signal	90	0→3700					→4500	→5000
		70					65	45
Downstream Signal	x (10-90)	0→14	→1900	→3000	→3500	→3700	→4500	→5000
		13	0.35x	0.05x	0.2x	0.35x	0.4x	0.55x

REFERENCES

- Bando, M., Hasebe, K., Nakayama, A., Shibata, A. and Sugiyama, Y. (1995) Dynamical model of traffic congestion and numerical simulation. *Physical Review E*, vol. 51, pp. 1035-1042.
- Banks, J.H (1995) Another look at a priori relationships among traffic flow characteristics. *Transportation Research Record* No. 1510, pp. 1-10.
- Bhouri, N., Hadj-Salem, H., Papageorgiou, M. and Blosseville, J.M. (1988) Estimation of traffic density on motorways. *Traffic Engineering and Control*, vol. 29, pp. 579-583.
- Cassidy, M.J. and Coifman, B. (1997) Relation among average speed, flow, and density and analogous relation between density and occupancy. *Transportation Research Record* No. 1591, pp. 1-6.
- Cheung, S.Y., Coleri, S., Dunder, B., Ganesh, S., Tan, C.-W., Varaiya, P. (2005) Traffic measurement and vehicle classification with a single magnetic sensor. *84th Annual Meeting of the Transportation Research Board*, Washington D.C.
- Eddie, L.C. (1965) Discussion of traffic stream measurements and definitions. *Proc. 2nd Intern. Symposium on the Theory of Traffic Flow*, Paris, France, pp. 139-154.
- Eddie L.C. (1974) Flow theories. In *"Traffic Science"*, D. Gazis, Ed., John Wiley and Sons, New York, pp. 2-108.
- Hall, F. L. (2002) Traffic stream characteristics. In *"Revised Monograph on Traffic Flow Theory"*, Federal Highway Administration, U. S. Dept. of Transportation.
- Jazwinsky, A.H. (1970) *Stochastic Processes and Filtering Theory*. Academic Press, New York.
- Kalman R.E. (1960) A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, vol. 82, pp. 35-45.

Kim, Y. and Hall, F.L. (2004) Relationships between occupancy and density reflecting average vehicles lengths. *Transportation Research Record* No. 1863, pp. 85-93.

Kotsialos, A. and Papageorgiou, M. (2005) A hierarchical ramp metering control scheme for freeway networks. *Proc. 2005 American Control Conference*, Portland, Oregon, pp. 2257-2262.

Maybeck, P.S. (1979) *Stochastic models, estimation, and control*. Volume 1, Academic Press, New York.

Papageorgiou, M. (1987) *Macroscopic Modeling of Traffic Flow on the Boulevard Périphérique de Paris*. Rapport de Contrat INRETS (Institute National de Recherche sur les Transports et leur Sécurité), Arcueil, France.

Papageorgiou, M. and Ioannou, P. (1993) *Short Course on Traffic Flow Modelling and Control and Intelligent Vehicle and Highway Systems (IVHS)*. Southern California Center for Advanced Transportation Technologies, Los Angeles, California.

Schmidt, S.F. (1970) Computational techniques in Kalman filtering. AGARDograph 139, NATO Advisory Group for Aerospace Research and Development, London.

Smaragdis, E. and Papageorgiou, M. (2003) A series of new local ramp metering strategies. *Transportation Research Record* No.1856, pp. 74-86.

Sun, X. and Horowitz, R. (2005) A localized switching ramp-metering controller with a queue length regulator for congested freeways. *Proc. 2000 American Control Conference*, Portland, Oregon, pp. 2141-2146.

Thornton, C.L. and Bierman, G.J. (1980) UDU^T covariance factorization for Kalman filtering. In C.T. Leondes (ed.), *Control and Dynamic Systems*. Academic Press, New York, pp. 177-248.

Vigos, G., Papageorgiou, M., Wang, Y. (2006) A ramp queue length estimation algorithm. *9th Intern. IEEE Conference on Intelligent Transportation Systems*, Toronto, Canada, to appear.

Wardrop, J.G. (1952) Some theoretical aspects of road traffic research. *Proc. Inst. Civil Engineers*, part II, vol. 1, pp. 325-362.