Optimization of Actuated Traffic Signal Plans Using a Mesoscopic Traffic Simulation

Tomer Toledo[1]; Tamir Balasha[2]; and Mahmud Keblawi[3]

Abstract: Traffic signal plan designs become more complex with developments in the fields of sensing and communication and the introduction of features, such as transit priority or pedestrian actuation. Traffic signal optimization programs were developed mostly for the basic parameters of pretimed traffic signal plans using analytical or simple traffic models. Simulation-based optimization for other parameters related to actuated traffic signals has been very limited due to the high computational time associated with detailed traffic simulation models. This paper presents the overall structure and the various components of a simulation-based system to optimize the parameters of complex actuated traffic signal plans. The system framework incorporates a connection between a mesoscopic traffic simulation, MESCOP, traffic signal control, and a genetic algorithm as the optimization method. The integrated system is demonstrated with applications to a fully actuated signalized intersections with vehicle and pedestrian actuations and transit priority in Haifa, Israel. **DOI: 10.1061/JTEPBS.0000363.** *© 2020 American Society of Civil Engineers.*

Author keywords: Signal plans; Mesoscopic traffic simulation; Optimization; Actuated control; Transit priority.

Introduction

Traffic signals at intersections are by far the most common means to control urban traffic. Efficient design of traffic signal plans is a cost-effective method to improve accessibility and mobility (Park and Yun 2006). In recent years, increasingly complex signal plans that involve functionalities for vehicle and pedestrian actuations, transit priority, and coordination between adjacent intersections are being used. These plans involve large numbers of parameters that need to be set (e.g., minimum and maximum green times, maximum pedestrian waiting times, gap times, and detector locations). Analytic and macroscopic optimization software, such as HCM (2012), SYNCHRO (Husch and Albeck 2004), TRANSYT-7F (Hale 2005), PASSER II (Chang and Messer 1991), PASSER V (Chaudhary and Chu 2002) and MAXBAND (Little et al. 1981), are generally unable to capture the stochastic nature of arrival patterns and traffic flow and realistically represent detector states.

Stochastic optimization approaches, which rely on traffic simulation models to evaluate signal plans within an optimization framework, are alternative options. Most of these studies (e.g., Foy et al. 1992; Hadi and Wallace 1993; Park et al. 1999, 2000, 2001; Rouphail et al. 2000; Yang and Liu 2008; Hu and Chen 2011; Howell and Fu 2006; Geng and Cassandras 2012) optimized the four basic parameters of pretimed plans (cycle length, green splits, phase sequence, and offsets) to minimize delays or queue lengths. Park and Yun (2006), Branke et al. (2007), Park and Lee (2009), and Yun and Park (2012) applied the stochastic optimization approach to additional parameters of actuated plans, such as minimum and maximum green times, maximum vehicle red times, gap time, and detector locations. All showed substantial potential for improvements: Park and Yun (2006) showed a reduction of 15%–18% in average delays compared to SYNCHRO plans. Branke et al. (2007) improved stopping times by 34% and 70% for vehicles and pedestrians, respectively, compared to an engineer's design. Park and Lee (2009) reduced travel times by 2%–4% compared to the settings in the field.

Using simulation-based approaches, Park and Schneeberger (2003) and Stevanovic et al. (2007) optimized control for multiple intersections simultaneously using offsets. Park and Schneeberger (2003) reduced travel times and delays by 17% and 36%, respectively, compared to the uncoordinated plans. Stevanovic et al. (2007) optimized offsets jointly with the other plan parameters. Delays and stops were reduced by at least 5% compared to the best SYNCHRO plans. Stevanovic et al. (2008) included in the optimization transit priority parameters the maximum green time extension and the maximum early green time provision to transit phases. In a corridor of seven signalized intersections, they showed a 7% reduction in person delay compared to the initial design. In a subsequent study, Stevanovic et al. (2011) optimized a network of 70 intersections with transit priority.

A major drawback of the simulation-based optimization is the high computational effort associated with the need to run thousands of microscopic traffic simulations. Stevanovic et al. (2008, 2011, 2015) used parallel computing implementations, and they reported run times in days and even weeks. To reduce the computational effort, Wolput et al. (2015) developed a mesoscopic traffic simulation model, CAPACITEL, to be used within traffic control optimization. They optimized green splits and cycle lengths for different combinations of transit detector locations, frequency of buses, lost time, critical flow ratios, and phase sequencing. A linear regression function was then fitted to the optimization results. Adopting parameter values from the regression lines improved average delays to vehicles and buses by 29% and 22%, respectively, compared to signal plans based on Akcelik's formula. Balasha and Toledo (2015) also

[1]Associate Professor, Faculty of Civil and Environmental Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel (corresponding author). Email: toledo@technion.ac.il

[2]Research Associate, Faculty of Civil and Environmental Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel. Email: tbalasha@technion.ac.il

[3]Ph.D. Student, Faculty of Civil and Environmental Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel. Email: smahmudk@technion.ac.il

Note. This manuscript was submitted on January 22, 2019; approved on December 4, 2019; published online on March 31, 2020. Discussion period open until August 31, 2020; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering, Part A: Systems***, © ASCE, ISSN 2473-2907.

developed a mesoscopic traffic simulation model, MESCOP and used it to evaluate different plans for an intersection with transit priority and pedestrian actuation functions. Both studies showed that the mesoscopic traffic simulations are computationally very efficient compared to microscopic simulation alternatives.

In this paper, an integrated simulation-based optimization system that embeds a mesoscopic traffic simulation model is presented. The system is demonstrated with applications to fully actuated control of intersections that include transit priority and pedestrian actuations. Experiments are presented that show the potential benefits of the joint optimization compared to treating intersections separately and the effect of various characteristics of the demand.

The main contribution of the research is in the integration of a mesoscopic traffic simulation model within an optimization framework for the signal control parameters. While studies of simulation-based optimization of signal plans have been presented in the literature, they used microscopic traffic simulations that are far more complex and computationally expensive. Furthermore, previous simulation-based optimizations mostly addressed pretimed or simple actuated control plans. In contrast, the control plans optimized in this paper involve complex transit priority functions, account for pedestrian flows and their delays, and represent, in the mesoscopic model, the movement of vehicles between intersection by incorporating a platoon dispersion model.

The rest of this paper is organized as follows: The next section describes the overall structure and the components of the simulation-based optimization system. Next, two case studies and analysis of their results are presented. Lastly, a discussion and conclusions are provided.

Simulation-Based Optimization System

The simulation-based optimization framework consists of three main components that interact with each other, as shown in Fig. 1: a mesoscopic traffic simulation model, signal control logic and optimization algorithm. The mesoscopic model represents the movements of vehicles, generates detector actuations, and reports these to the control logic. The control logic generates signal indications that in turn affect the movements of vehicles in the simulation. The optimization algorithm invokes the simulation-control system with different parameter values and collects from MESCOP measures of performance that are used in the objective function.

Mesoscopic Traffic Simulation

MESCOP (Balasha and Toledo 2015) explicitly represents the movement of passenger cars, transit vehicles, and pedestrians through and between intersections. Vehicle movements are modeled as events occur at detector locations and the stop line. A vehicle enters the system when it arrives at the first decision point on the approach. This point is the farthest among a detector location, a point in which turning lanes open, or the stop line. The arrival of vehicles at this point is modeled stochastically based on a Poisson distribution of inter-arrival times. Travel times from the initial point

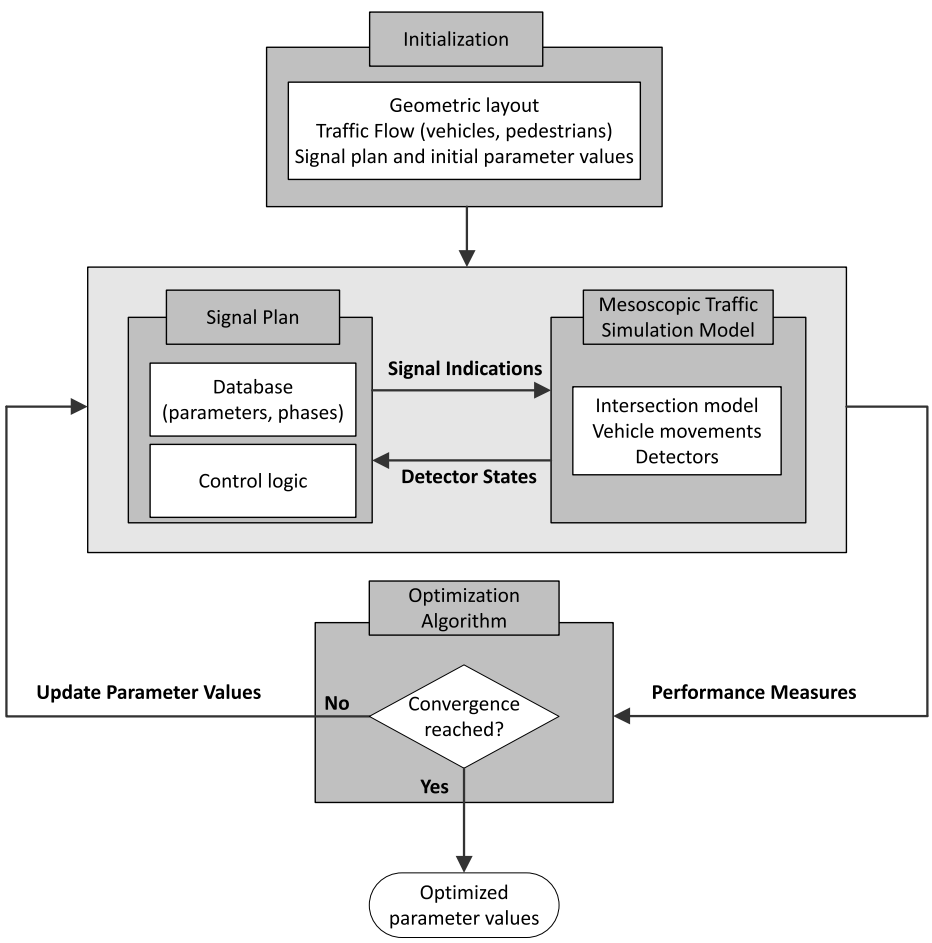

Fig. 1. Integrated simulation-based optimization overall structure.

J. Transp. Eng., Part A: Systems, 2020, 146(6): 04020041

to the next decision point are modeled as normally distributed random variables. Their mean is determined by the distance between the two points and an assumed approach speed. The standard deviation is set as a fraction of the mean. At the stop line, vehicles are allocated to a lane-specific vertical queue based on their intended turn movement. If multiple lanes are appropriate, they are allocated to the shortest queue. In addition, they are released from the queue at the saturation flow rate [the default value is 1,800 vehicles/hour (vph)] when their light is green and according to the first-in-first-out (FIFO) rule. In situations of queue spillbacks from downstream sections, overflow of traffic on turning lanes, or blockage of access to turning lanes due to queues on the through lanes, the arriving vehicles cannot move to the next road section or enter the lane they are allocated to. When a vehicle is released from the stop line, its arrival time to the next decision point (if it exists) is calculated based on the distance between the two points and Pacey's platoon dispersion model (Pacey 1956), which assumes normally distributed vehicle speeds. Pedestrians' arrival at crosswalks is assumed to follow a Poisson process with an input arrival rate. When the light for a given crosswalk is red, the first arriving pedestrian activates the relevant push button, if one exists. Once the light turns green, all waiting pedestrians cross the intersection at a constant speed (the default value is 1.2 m/s). Pedestrians that arrive when their light is green cross the intersection without any delay.

The simulation implementation is time-based with a step size of one second to fit with the resolution of the control logic. Detectors are actuated, either when a vehicle is present on them or when the length of the queue extends beyond their location. Activation states are provided to control logic functions.

Signal Control

The signal control logic is run every time step to determine the light indications. It may use the information on current and previous light indications (e.g., how long a certain light has been green) and on the activation states of vehicle detectors and pedestrian push buttons. The logic is defined using preprogrammed functions. These include detection functions that request the states of detectors and control functions to skip, extend, insert or terminate phases. Examples of transit priority functions include the following:

- Arrival expectation—a function that predicts the arrival time of a transit vehicle on a dedicated lane to the stop line. It is based on the time and location that the vehicle was detected and an assumed approach speed.
- Phase early termination—Extension of active phases may be overridden to provide priority to a transit vehicle approaching the intersection.
- Phase insertion—Activating a transit phase-out of the normal phase sequence.
- Early transit phase start—The transit phase is planned to start a few seconds before the transit vehicle is expected to arrive at the stop line so that it does not need to stop or slow down, even if it arrives earlier than expected.
- Priority cancellation—occurs when the transit vehicle is detected on a checkout detector located downstream of the stop line. Priority may also be canceled if a transit vehicle is not detected at the stop line a certain time after it was expected.
- Compensation—a function that guarantees a certain minimum green time to specific movements. It measures the cumulative green time provided to a movement within a predefined period of time (often defined in terms of a number of cycles). If needed, the green time is extended to meet the minimum threshold.
- Queue length override—a function that disables transit priority functions when a certain phase is active and the queue on the relevant (minor) approach exceeds a certain value over a predefined period of time.

Optimization

The control logic includes many parameters related to signal timings (e.g., minimum and maximum green and red durations), transit priority functions, maximum allowed pedestrians waiting times, and system layout parameters (e.g., locations of detectors). The optimization is used to set their values. It should be noted that the control logic is applied every second and results in different control actions depending on the current state of lights and detectors. However, the control parameters described above remain constant for the entire analysis period.

Various performance measures, such as delays or travel times of vehicles and pedestrians, queue lengths, throughputs, and number of stops, may be used as objective functions for the optimization. In this paper, the expected value of the average person delay is used, following several studies that suggested it in this context (Stevanovic et al. 2011; He et al. 2012; Christofa et al. 2013). Thus, the optimization problem is formulated as follows:

$$\min_{\theta} E[d(q,\theta)] = \frac{1}{R}\left[\sum_{r} \frac{\sum_i \sum_n d_{nr} N_i \delta_{nri}}{\sum_i \sum_n N_i \delta_{nri}}\right] \tag{1}$$

s.t.

$$\theta^L \leq \theta \leq \theta^U \tag{2}$$

$$\sum_{j=1}^{p} (G_{\max j} + T_j) \leq C \tag{3}$$

$$G_{\min j} \leq G_{\max j} \quad \forall\ j \tag{4}$$

where d = average person delay; q = vector of the design flows input to the model; θ = vector of decision variables (control logic parameters); θ^L and θ^U = corresponding lower and upper bounds, respectively. The expected delay is obtained by averaging R replications. d_{nr} = delay to vehicle (or pedestrian) n in replication r; N_i = number of travelers in a vehicle of type i (by definition 1 for a pedestrian); δ_{nri} = indicator variable that takes the value 1 if vehicle n is of type i (car, various bus types, pedestrian) in replication r, and 0 otherwise. $G_{\max j}$ and $G_{\min j}$ = maximum and minimum green duration of phase j, respectively; T_j = inter-green time of phase j; and C = cycle length.

The optimization is executed off-line considering all control parameters and the entire period of interest at once. The problem presented is simulation-based and consequently may have multiple local minima. A genetic algorithm (GA) (Holland 1992) was used in the case studies presented in this paper, following several studies that have shown that it is an effective method to optimize signal plans (e.g., Kovvalli and Messer 2002; Park and Yun 2006; Stevanovic et al. 2007). There were 30 generations with a population of 100 points and 20 replications used, for a total of 60,000 simulation runs. Crossover and mutation probabilities were set to 0.7 and 0.03, respectively. The rank selection method (Kovvali and Messer 2002) that assigns crossover probabilities for points based on their fitness rank was used. The combination of the search algorithm and a large number of generations increases the probability that a globally optimal solution is found.

Case Study 1: Transit Priority

Layout and Flows

The simulation-based optimization system was applied to the planned control at Palmer intersection in Haifa, Israel, which is shown schematically in Fig. 2, and it is located in the heavily congested downtown area. HaAtsmaut Avenue is the main arterial in the area. There are six vehicle movements in the intersection: Two bus rapid transit (BRT) movements (2, 6) on dedicated lanes, four nontransit vehicle movements (1, 3, 5, 7), and six signalized pedestrian crosswalks (a through g). The planned control of this intersection is fully actuated, with presence detectors on the minor approaches that are used for demand (D1, D5), extension (E1, E5) and queue detection (Q1) tasks. The demand detectors are located at the stop line. The extension detectors are located about 12 m upstream of the intersection. The queue detector Q1 serves to alleviate the concern that the queue would block a roundabout that is located upstream of it. BRT detectors (DPT21, DPT22, DPT62) identify an approaching BRT vehicle and predict its arrival time to the stop line. For eastbound vehicles, this prediction is updated when they are detected at DPT22. DPT21 is located 350 m upstream of the stop line. DPT22 and DPT62 are located 100 m upstream of the stop line. A BRT stop exists upstream of detector DPT62. This makes predictions of expected travel times to the stop line less reliable because it is more difficult to predict the dwell times at stops. For this reason, a far detector was not placed on the westbound approach in this intersection. Alternatively, an additional detector at the exit from the stop could be used, and the configuration of the detectors may affect the optimal control settings. In the simulation, the prediction error distribution can be widened to capture the larger variability in travel times to the stop line. Both BRT approaches also include priority cancellation detectors downstream of the intersection (DPT23, DPT63). The four pedestrian crosswalks on the arterial are activated by push buttons (P_d, P_e, P_f, P_g). In cases that push buttons do not exist, minimum green times must be provided to the relevant phases in every cycle.

The design traffic flows, presented in Table 1, were estimated from traffic counts for the morning peak hour (7 to 8 a.m.). The total flow is 3,295 vph, representing a saturation degree of about 0.6. Also, 30 BRT vehicles per hour arrive in each direction, and occupancies of 50 and 1.2 passengers were assumed on transit and nontransit vehicles, respectively.

Control Logic

The movements in the intersection are organized in three signal phases (A, B, and C) as shown in Fig. 3. The minor phases B and C are activated only if vehicle presence is detected on D1 and D5, respectively, or if the pedestrian push buttons (P_f or P_g for

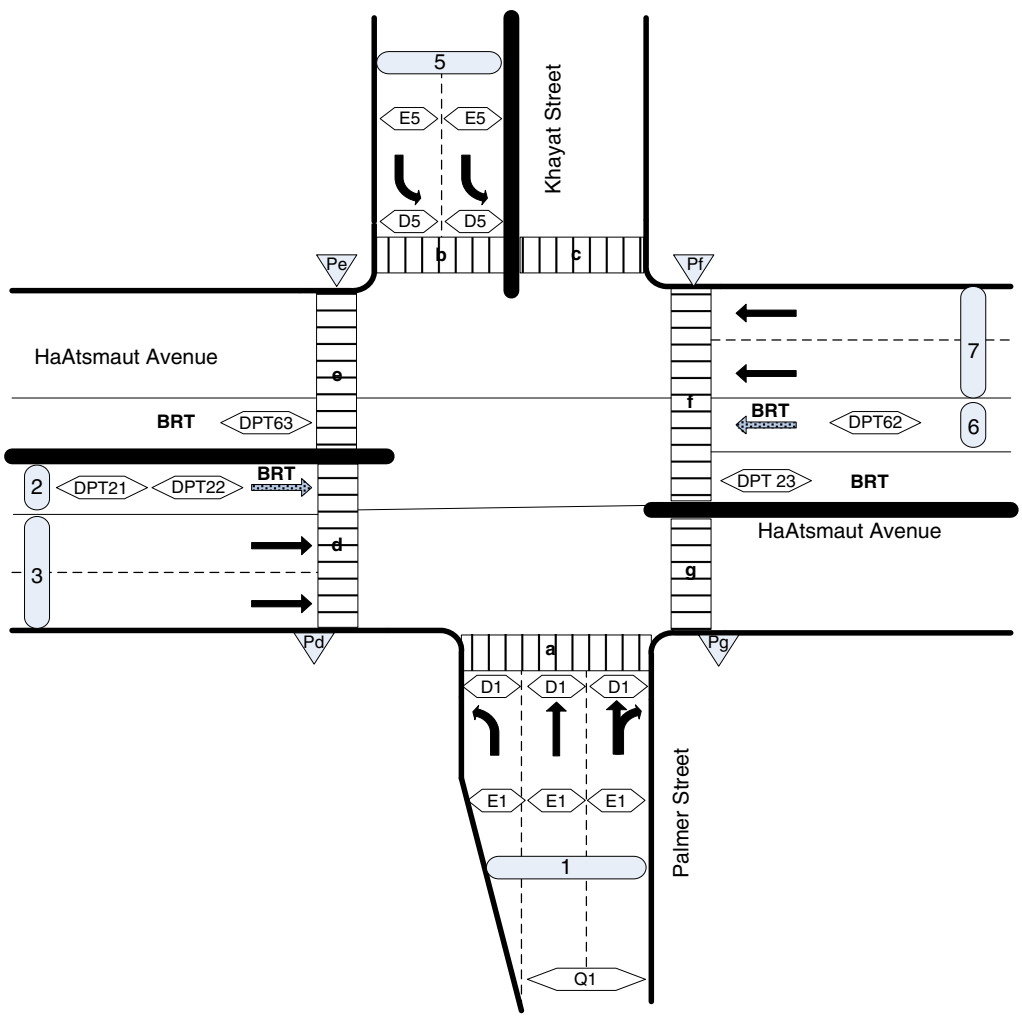

Fig. 2. Palmer intersection.

Table 1. Traffic flows at the Palmer intersection

Movement	Traffic flow (vph)
1L	115
1T	144
1R	45
2T	30
3T	1,058
5L	59
6T	30
7T	1,713

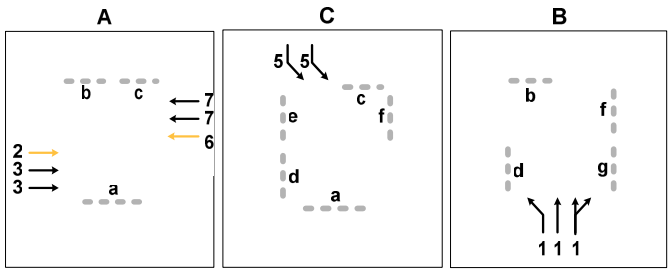

Fig. 3. Control phases at Palmer intersection.

phase B, P_d or P_e for phase C) are activated. The right turn movement in phase B must yield to pedestrians crossing crosswalk g, and the cycle length is fixed. The remaining green time, after phases B and C are served, to the end of the cycle is allocated to phase A (complementary green time).

BRT signal priority may be provided using phase extension or early phase termination (recall) depending on the time within the cycle that the transit vehicle is expected to arrive at the stop line and the currently active phase. Phase A may be extended until the BRT vehicle crosses the intersection if constraints on its maximum green time, maximum queue length at Q1, or maximum pedestrian waiting time are not violated. Alternatively, it may be terminated early so that phases B and C can be served before the BRT vehicle arrives at the stop line. The minor phases may also be terminated early if a BRT vehicle is identified, except for phase B when a queue is detected on Q1. Finally, a compensation mechanism dictates that phase B receives a certain minimum cumulative green time over several cycles to avoid overprioritizing the BRT vehicles. The engineering report presenting the complete details is available from the authors.

The control logic includes 14 parameters: The cycle length and the minimum and maximum green times for each of the three phases (7 parameters); The maximum red time for pedestrian signals (1 parameter); The early green time for transit priority, which in order to provide a smooth passage to the BRT vehicle, indicates how long before the expected arrival of the BRT vehicle to the stop line the transit phase would start (1 parameter); The compensation parameters that define the minimum cumulative green time that must be provided to phase B, and the number of cycles within which it is measured (2 parameters); The distance of the queue detector from the stop line and the time duration that a vehicle is continuously present on the detector that activates the queue condition (2 parameters). The locations of other detectors are not varied; The maximum time gap between activations of a detector that triggers the extension of the current phase (1 parameter).

The control plan was evaluated under three different sets of design parameters:

1. The initial design that was developed for the same conditions by the traffic engineers using INBAR—an Israeli commercial software (TRI 2013).
2. Full optimization of the values of all control logic parameters.
3. Partial optimization, which only included the cycle length and green splits. The other parameters are fixed at their values in the initial design. The partial optimization design is a basis to evaluate the benefits of optimizing additional parameters.

Results

Table 2 compares the average delays for the various road users with the three designs. The result shows a reduction in the average person delay of 21% and 28% from the initial design to the partial and full optimization designs, respectively. The reduction is mostly in the delays to nontransit vehicles and pedestrians. With full optimization, but not with the partial one, this reduction occurs without an increase in the delays to BRT vehicles. The results also show the benefits of optimizing additional design parameters beyond the cycle time and green splits. The average person delay is reduced by an additional 9% from the partial to the full optimization design.

Table 3 presents the parameter values in the initial, partial and full optimization designs. Parameter values in bold are those that were optimized in each case. Overall, the parameter values changed substantially in a manner that allocates more green time to pedestrians and nontransit vehicles. The cycle length decreased from 110 s in the initial design to 72 s in the optimal design. Some jurisdictions require that cycle times are constrained to increments of 5 or 10 s. This could be incorporated into the optimization that was not used in this example. The green time to phase A is divided to complementary green (the remaining time at the end of the previous cycle after phases B and C are served is allocated to phase A) and initial green time (green time to phase A in the current cycle based on its minimum and maximum green time and detector states). The shorter cycle decreases the complementary green times to phase A. With the initial design, the additional green times to phase A are underutilized, as shown in Fig. 4. The figure shows the discharge flow rates of movement 7 (phase A). In the initial design, phase A receives 73 s green time in a cycle of 110 s. Vehicles are discharged with the saturation flow rate only in the first 24 s (33% of the green time). The remaining green time is underutilized. In the full optimal design, the green time decreased to 42 s in a cycle of 72 s. Thus, the proportion of green time allocated to the phase is slightly reduced and the green time utilization is improved.

Table 2. Comparison between delays with three different designs

	Average person delay (s)			% change		
Road users	Initial design	Partial optimization design	Full optimization design	Initial to partial	Initial to full	Partial to full
BRT	1.54	1.97	1.53	28	−1	−22
Other vehicles	18.48	14.75	14.17	−20	−23	−4
Pedestrians	24.31	17.95	15.95	−26	−34	−11
All	14.41	11.43	10.40	−21	−28	−9

Table 3. Parameter values in the initial and optimized designs

Design parameter	Initial	Partial optimal	Full optimal	Change initial to full (%)
Cycle length (s)	110	**82**	**72**	−35
Minimum green phase A (s)	11	**15**	**19**	73
Maximum green phase A (s)	51	**30**	**37**	−27
Minimum green phase B (s)	2	**6**	**5**	150
Maximum green phase B (s)	17	**11**	**13**	−24
Minimum green phase C (s)	8	**5**	**5**	−38
Maximum green phase C (s)	8	**13**	**12**	50
Maximum pedestrians red (s)	138	138	**110**	−20
Early green start (s)	0	0	**5**	—
Cumulative green time (s)	27	27	**12**	−56
Compensation period (cycles)	4	4	**4**	0
Queue length (vehicles)	15	15	**27**	80
Maximum queue time (s)	5	5	**15**	200
Gap time (s)	3	3	**3**	0

The queue is discharged at the saturation flow during the first 20 s (48% of the green time).

The minimum green time to phase A increased in the full optimal plan from 11 to 19 s, ensuring that the queues of vehicles in movements 3 and 7 are completely dissipated, regardless of the complementary green time. In contrast, the maximum green time to phase A decreased from 51 to 37 s. Phase A may be extended only to provide priority to BRT vehicles. Shorter maximum green times to phase A help avoid underutilization of the intersection after the queue of nontransit vehicles has dissipated and reduced delays to vehicles in the minor approaches and to pedestrians. Furthermore, shorter maximum green times increase the flexibility of the control to transition use the recall strategy to provide BRT priority: With the full optimal design, phase A is extended in 25% of the cycles and recall is activated in 75%. With the initial plan, phase A is extended in 88% of the cycles and recall is applied only in 12% of the cycles.

The full optimal minimum green times for phases B and C were both set to 5 s, which is the lowest value that allows pedestrians with an assumed speed of 1.2 m/s. to safely cross a single crosswalk in this intersection. It is also sufficient for vehicles that are queued between the stop line and the extension detector to release from the intersection. Longer minimum green times then reduce the flexibility to apply recall when BRT vehicles are identified. The full optimal values of maximum green time to phases B and C are 13 and 12 s, respectively. These values are sufficient to accommodate longer queues that may result from the early termination of these phases to provide transit priority to phase A.

The optimal compensation settings for phase B require at least 12 s of green time within four consecutive cycles. This constraint may not be satisfied only if phase B is skipped in at least two of the cycles. The compensation mechanism was activated in 31% of the cycles with the initial design, but in less than 1% of the cycles with the optimal design. Nevertheless, the green time to phase B in the optimal design decreased by a total of only 38 s per hour. Thus, the results suggest that the compensation mechanism could be eliminated without negatively affecting performance.

The maximum pedestrians' red time parameter decreased to 110 s in the optimal design, contributing to the reduction in pedestrian delays presented in Table 2. This constraint may, in some cases, limit the extension of phase A. The early green parameter, which was not implemented in the initial design, was set to 5 s in the optimal design. This reduces BRT delays by allowing them

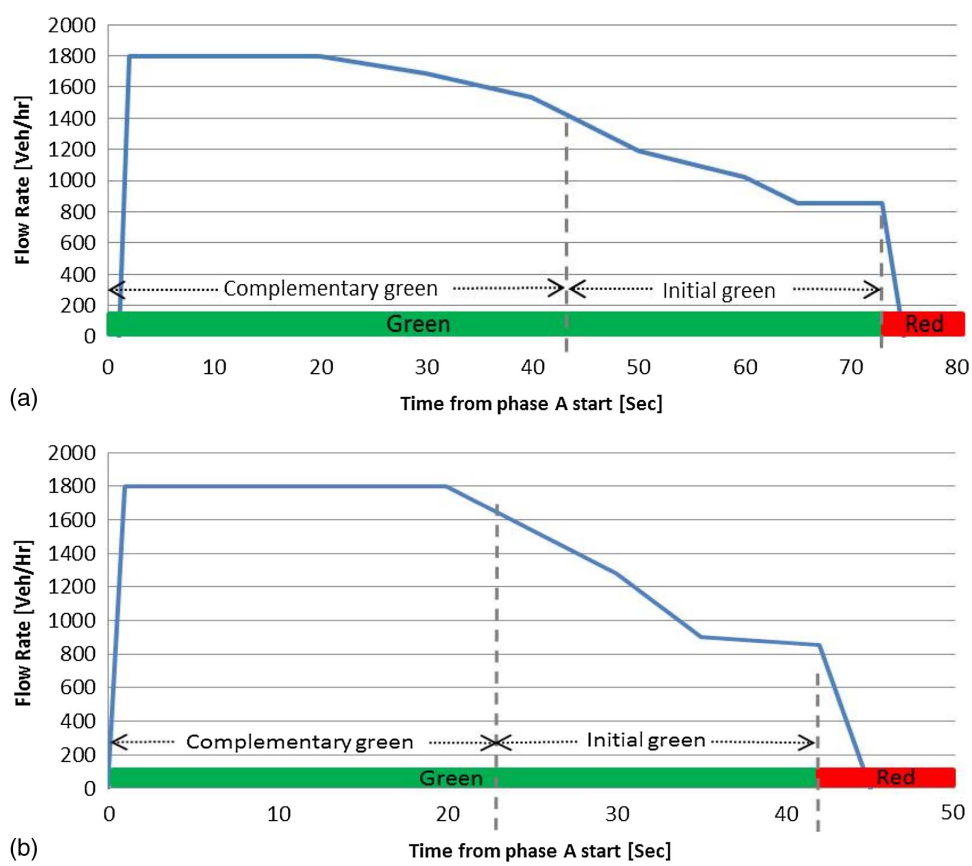

Fig. 4. Flow discharge rate in phase A with the (a) initial; and (b) optimal designs.

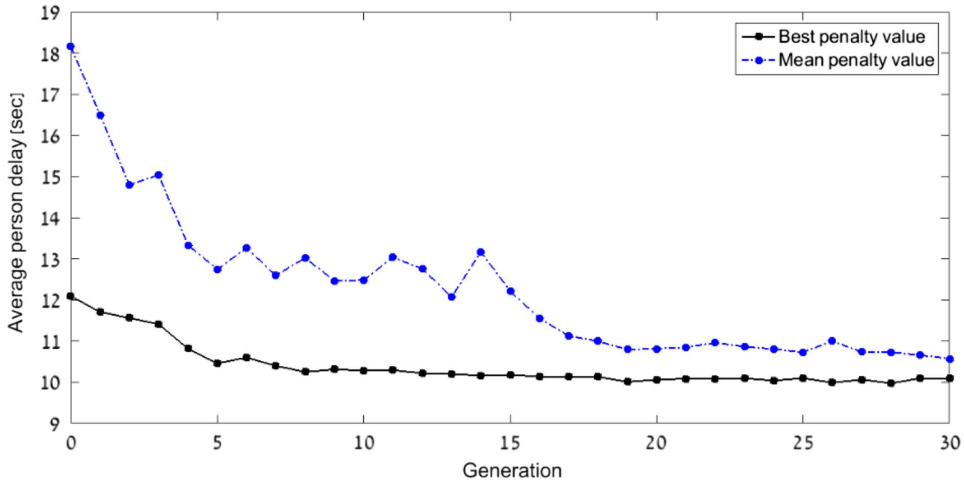

Fig. 5. Convergence properties of the genetic algorithm.

smoother movement if they arrive at the stop line earlier than expected. This does not have any noticeable effect on the delays of nontransit vehicles or pedestrians because of the low traffic demand levels on the minor approaches. Finally, there was no change in the gap time values between the initial and full optimal designs.

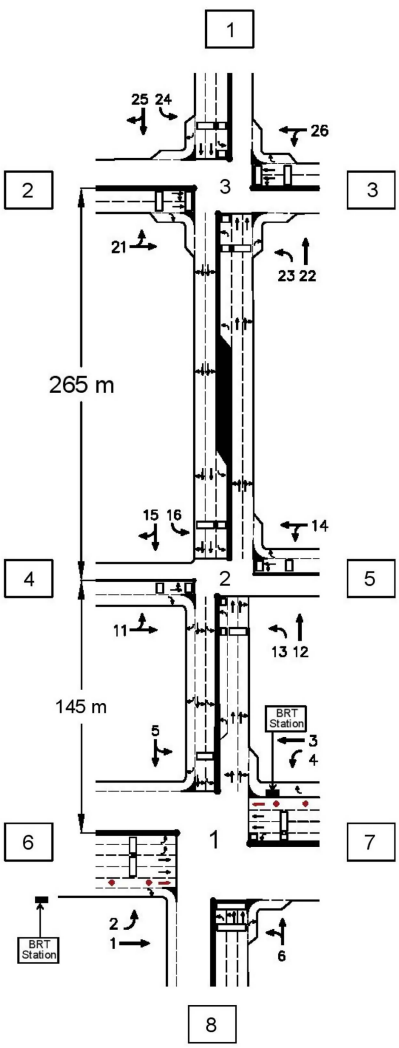

Fig. 6. Schematic structure of the network of three intersections.

This result is not surprising because of the strong relations between the gap time and the location of the extension detectors.

The progress of the GA of the full optimization is shown in Fig. 5. The value of the best solution in each generation stabilized after about 10 generations, and convergence has been achieved approximately at generation 20 when the difference between the mean and the best person delay has stabilized. The optimization took about 5 h, and comparable optimization based on microscopic traffic simulation would not be computationally feasible, with an estimated running time of about a week.

Case Study 2: Coordinated Intersections

Layout, Flows, and Control Logic

This case study concerns the simultaneous control of three adjacent intersections along a large collector road (HaEtzel St.) in Haifa. The three intersections and traffic channelings are shown schematically in Fig. 6. Intersection 1 connects the collector to a major arterial (6–7 in the drawing) that has two BRT movements with dedicated lanes. There are three BRT detectors in each direction. BRT stops are located immediately upstream of the northbound stop line and 150 m upstream of the southbound stop lime. Extension detectors are located on the major approaches to the three intersections (movements 1–4, 12–13, 15–16, 22–25). Demand detectors are located on the minor approaches to all intersections (movements 5, 6, 11, 14, 21, 26). Pedestrian push buttons are placed on the crosswalks crossing the arterial in intersection 1, the collector in intersection 3, and all crosswalks in intersection 2.

The peak hour traffic flows in the system are presented in Table 4. These values were estimated from traffic count measurements and

Table 4. Origin-destination flows in the network (vph)

O/D	1	2	3	4	5	6	7	8
1	—	272	120	8	161	235	472	64
2	112	—	300	1	28	41	83	11
3	143	292	—	1	19	28	64	8
4	75	30	9	—	42	4	9	1
5	156	61	19	47	—	41	81	11
6	88	34	11	13	36	—	1,400	106
7	228	90	28	33	95	1,400	—	55
8	38	15	5	6	16	104	58	—

represent degrees of saturation of about 0.7 in each of the three intersections. In intersection 1, the largest demands are on the main arterial (6–7). In the other two intersections, the largest demands are those in the direction of intersection 1. In addition, 15 BRT vehicles per hour cross intersection 1 in each direction. In the simulation, vehicles are generated with negative exponential inter-arrival times and the origin-destination matrix rates. Their movements to and between intersections are determined by the models within the simulation.

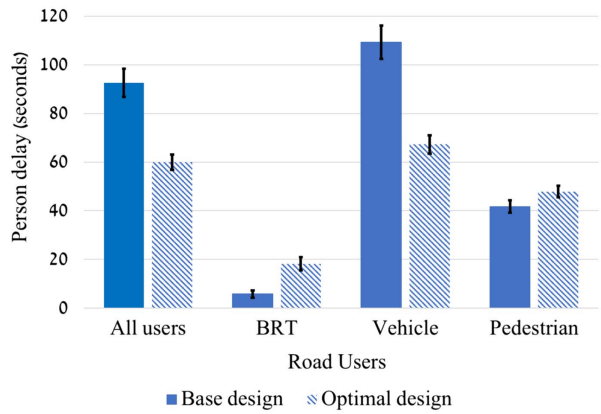

Fig. 7. Average delays and their 95% confidence intervals for the various road users in the base and optimal designs.

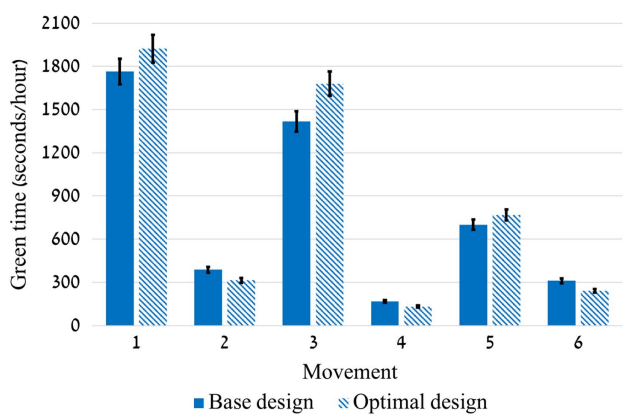

Fig. 8. Cumulative green times and their 95% confidence intervals in intersection 1 in the base and optimal designs.

The three intersections are controlled jointly and share the same cycle time with offsets to allow green light bands for through traffic along the collector in both directions. It should be noted that the optimization aims to minimize the total delays of all users and does not maximize bandwidth or other measures of coordination. Thus, coordination is possible if it supports the objective, but it is not strictly enforced. Functions of phase demand and extension, pedestrian push button requests, and transit priority and compensation (only in intersection 1) are operated in a similar manner to that of case study 1. The control logic for all three intersections includes 84 parameters, and the engineering report presenting the complete details is available from the authors.

Results

The results of optimization of all 84 parameters jointly are compared to the base design, which is implemented in the field. Fig. 7 shows the average delays to the various road users in the two designs, and the person delay decreased by 35%. This is mostly due to a decrease (38%) in the delay to vehicle passengers that constitute 77% of all users. In contrast, the delays to BRT passengers and pedestrians increased. However, their delays remain relatively low, and they are only 18% and 5% of the users, respectively.

In the control plan, the most influential differences are that the cycle time increased from 120 to 150 s and that green times for the major movements in the main intersection (movements 1 and 3) and those along the collector (movements 5, 15, 25) increased. These increases are partly a result of the lower lost times (by 7%) due to larger cycle times and partly at the expense of the minor movements. The cumulative green times for the movements in intersection 1 with the base and optimal control designs are shown in Fig. 8.

Queue lengths, in particular, along with the collector, are important because of the risk of queue spillback and blockage of upstream intersection. Fig. 9 shows the distributions and averages of queue lengths in the movements along the corridor in the high demand direction toward intersection 1 (movements 5 and 15). In movement 15, the average queue length decreased from 16.9 vehicles in the base design to 4.9 vehicles in the optimal one. The maximum queue length also decreased substantially. The differences were smaller in movement 5. The average queue length decreased from 11.3 vehicles in the base design to 7.7 vehicles

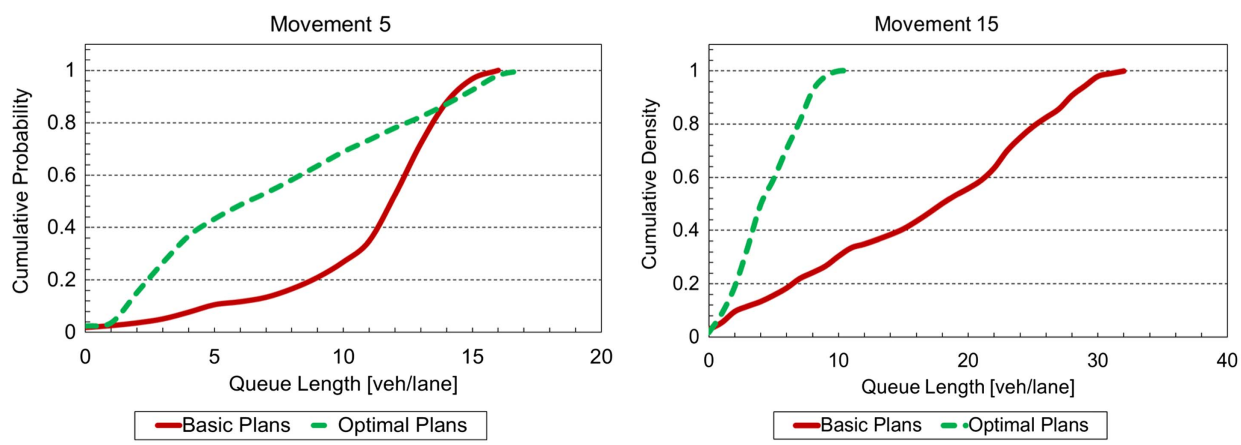

Fig. 9. Cumulative distribution and average queue lengths in movements 5 and 15.

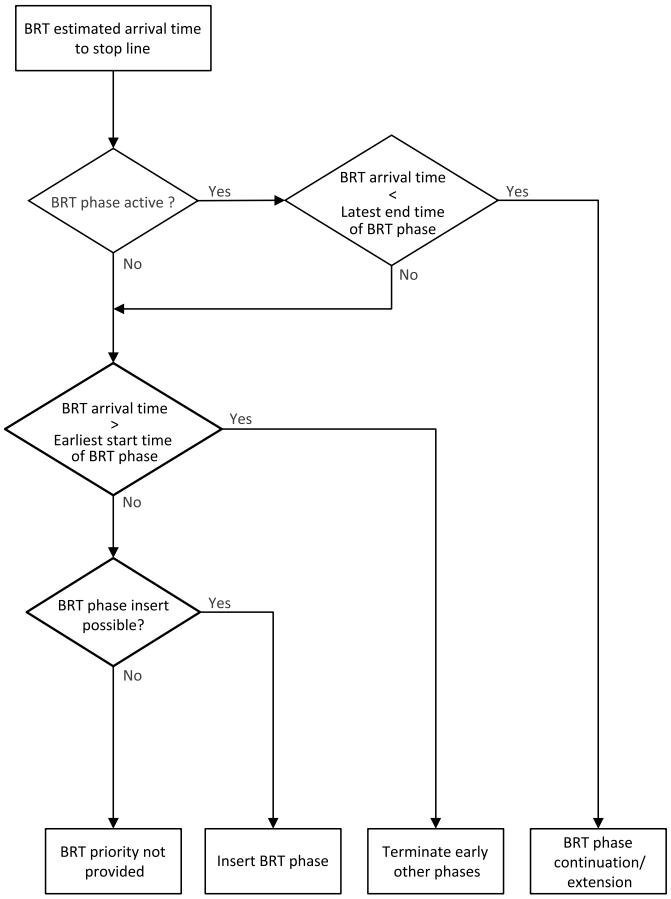

Fig. 10. BRT priority provision control logic.

in the optimal design, but the maximum queue length slightly increased.

The improvements in vehicles' delays and queues come at the expense of the BRT vehicles. As seen in Fig. 7, their delay increases by 12 s, and this is a result of a reduction in the priority coverage in the optimal design and an increase in the compensation mechanism requirements. The priority coverage is defined by the fraction of time within the cycle in which BRT priority is possible if all demand detectors and pedestrian push buttons are activated. Priority may be provided in three ways: extension of the green time to the BRT movement, early termination of the green time to other phases, or insertion of the BRT phase out of sequence. The BRT priority provision control logic showing these options is shown in Fig. 10. Fig. 11 shows the priority coverage in the base and optimal design. It decreases from 84% and 82% for movements 1 and 3, respectively, in the base design to 58% and 52% in the optimal design. The phase insert option was practically eliminated altogether, and it should be noted that the base design is aimed to maximize the coverage. With the optimal design, the relative weights assigned to the delays of different road users determine the level of BRT priority and coverage.

The compensation mechanism guarantees minimum green times to the minor movements in intersection 1 over a given number of cycles. For movement 5, this value increased from 25 s in the base design to 68 s every cycle in the optimal one. For movement 6, it slightly increased from 12 to 13 s over two cycles. As a result, BRT priority, which was never denied in the base design due to compensation constraints, was denied in 1.7% of the cycles in the optimal design.

The computation time for the optimization was about 10 h with a single PC and using ten replications for each point. This is again far shorter than would be required if a microscopic traffic simulation model would have been used.

The obtained results are optimal based on the input and modeling assumptions in the simulation model, and it is useful to evaluate their sensitivity to these assumptions. A comprehensive sensitivity analysis is beyond the scope of the current work. Nevertheless, the sensitivity to changes in the level of demand of the allocation of green times to the movements was examined. The results for intersection 1 are shown in Fig. 12. The optimal cycle time with all demand levels was 150 s, and as demand increases, the total green time to movement 5 increases that serves the collector traffic at the expense of movements 1 and 3 along the arterial. The added green time to movement 5 serves the additional demand along the

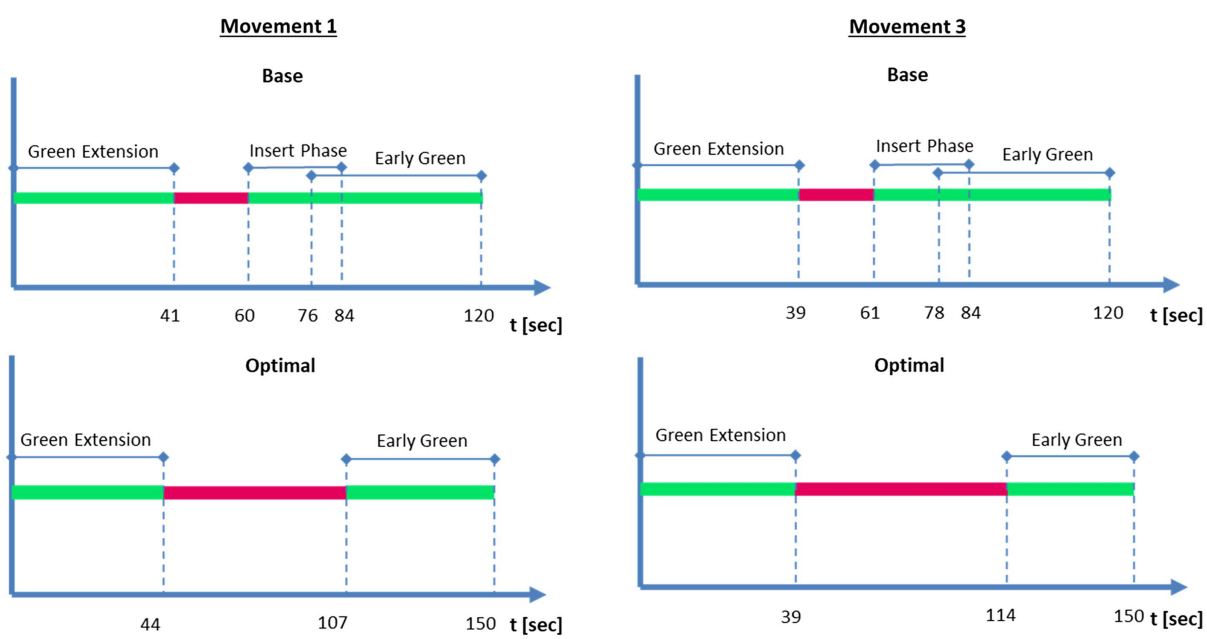

Fig. 11. BRT priority coverage in the base and optimal control plans.

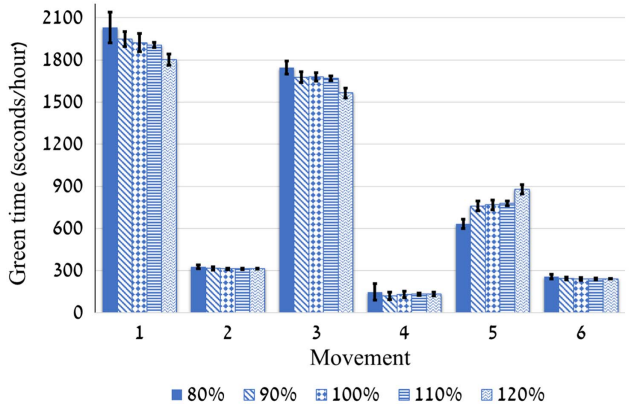

Fig. 12. Cumulative green times and their 95% confidence intervals in intersection 1 as a function of the demand level.

collector and prevents queue spillbacks and excess delays in the two minor intersections.

Summary

This paper presented the structure and application of a simulation-based optimization system for actuated traffic signal plans. This system incorporates a mesoscopic traffic simulation model that represents vehicle and pedestrian movements and an implementation of the detailed control logic. The simulation model is embedded within an optimization function that uses a GA to optimize the parameter values of actuated traffic signal plan based on prespecified performance measures. The system was applied in two case studies to optimize actuated signal plans with transit priority. The results showed potential for substantial improvement in the average person delay. They also demonstrated the usefulness of optimizing additional parameters of the signal plan beyond the cycle length and splits. The run-time advantages of the mesoscopic simulation model would be further magnified if optimizations are repeated under different conditions of the demand and other characteristics of the intersection.

The detailed results of the case studies are relevant to their specific characteristics such as movements, demands, and detector locations. However, they cannot be easily generalized. Following the approach of Wolput et al. (2015), a potential future research direction is the use of the simulation-based optimization system to investigate different control plans, specifically, those that involve transit priority functionalities to produce more general guidelines for strategy selection and to perform sensitivity analysis to different assumptions made in the models. This would involve constructing an experimental design to analyze the impact of different control strategies, travel demand, intersection layouts, and transit vehicle travel directions on the performance. More general experiment settings would also require integration of a traffic model for evaluation of the control plan performance that would be independent of the one used within the optimization. This has been done in Yun and Park (2012).

One specific detail of the case studies that may affect the implementation is that the BRT movements do not conflict. In cases that they conflict, the control logic would need to be modified, and the transit movement would be allocated to different phases. Furthermore, evaluation of the feasibility of using the transit priority functions (e.g., phase insertion, phase early termination, early transit start) would require different timing calculations for each of the phases and consideration for the provision of green time and the conflicting transit movement. This would increase the number of parameters to be optimized. However, the optimization framework and types of functions that are used within the control logic would not change.

Other specific assumptions and details of the implementation, such as that of Poisson arrivals for both vehicles and pedestrians, can be modified to provide better calibration of the model to real-world observations. The current implementation also assumes that transit vehicles receive priority travel in dedicated lanes. Transit priority without dedicated lanes is not common, but possible. The main difficulty would be in estimating when the bus would be ready to release from the stop line without knowledge of queue lengths. Case studies to evaluate the value of such practice could be undertaken. Another enhancement could involve representation of advanced detection technologies and the information they provide, such as speed measurements, within the control logic.

Data Availability Statement

Some or all data, models, or code generated or used during the study are available from the corresponding author by request. These include the control plans, intersection designs, traffic flow data, and optimization codes.

Acknowledgments

The authors thank the anonymous referees and the editor for many useful comments and suggestions. All remaining errors are our own. This research was partially funded by a grant from the Institute for Transportation Innovation at Tel Aviv University.

References

Balasha, T., and T. Toledo. 2015. "MESCOP: A mesoscopic traffic simulation model to evaluate and optimize signal control plans." *Transp. Res. Rec.* 2488 (1): 1–9. https://doi.org/10.3141/2488-01.

Branke, J., P. Goldate, and H. Prothmann. 2007. "Actuated traffic signal optimization using evolutionary algorithms." In *Proc., 6th European Congress and Exhibition on Intelligent Transportation Systems and Services*. Brussels, Belgium: ERTICO—ITS Europe.

Chang, E. C., and C. J. Messer. 1991. *Arterial signal timing optimization using PASSER II-90, program user's manual*. College Station, TX: Texas Transportation Institute.

Chaudhary, N. A., and C. L. Chu. 2002. *Passer V—Software for timing signalized arterials*. College Station, TX: Texas Transportation Institute.

Christofa, E., I. Papamichail, and A. Skabardonis. 2013. "Person-based traffic responsive signal control optimization." *IEEE Trans. Intell. Transp. Syst.* 14 (3): 1278–1289. https://doi.org/10.1109/TITS.2013.2259623.

Foy, M. D., R. F. Benekohal, and D. E. Goldberg. 1992. "Signal timing determination using genetic algorithm." *Transp. Res. Rec.* 1365: 108–115.

Geng, Y., and G. Cassandras. 2012. "Traffic light control using infinitesimal perturbation analysis." In *Proc., 51st IEEE Conf. on Decision and Control*. New York: IEEE.

Hadi, M. A., and C. E. Wallace. 1993. "Hybrid genetic algorithm to optimize signal phasing and timing." *Transp. Res. Rec.* 1421: 104–112.

Hale, D. 2005. *Traffic network study tool—TRANSYT-7F, United States version*. Gainesville, FL: McTrans Center in the Univ. of Florida.

HCM (Highway Capacity Manual). 2012. *Highway capacity manual*. Washington, DC: Transportation Research Board.

He, Q., K. L. Head, and J. Ding. 2012. "PAMSCOD: Platoon-based arterial multimodal signal control with online data." *Transp. Res. Part C* 20 (1): 164–184. https://doi.org/10.1016/j.trc.2011.05.007.

Holland, J. H. 1992. *Adaptation in natural and artificial systems*. Cambridge, MA: MIT Press.

Howell, W. C., and M. C. Fu. 2006. "Simulation optimization of traffic light signal timings via perturbation analysis." Ph.D. dissertation, Applied Mathematics and Scientific Computation Program, Univ. of Maryland.

Hu, T. Y., and L. W. Chen. 2011. "Traffic signal optimization with greedy randomized tabu search algorithm." In *Proc., 90th Annual Meeting of the Transportation Research Board*. Washington, DC: National Academies of Sciences, Engineering, and Medicine.

Husch, D., and J. Albeck. 2004. *Trafficware SYNCHRO 6 user guide*. Albany, CA: Trafficware.

Kovvali, V. G., and C. J. Messer. 2002. "Sensitivity analysis of genetic algorithm parameters in traffic signal optimization." In *Proc., 81st Annual Meeting of the Transportation Research Board*. Washington, DC: National Academies of Sciences, Engineering, and Medicine.

Little, J. D. C., M. D. Kelson, and N. H. Gartner. 1981. "Maxband: A program for setting signals on arteries and triangular networks." *Transp. Res. Rec.* 795: 40–46.

Pacey, G. M. 1956. *The progress of a bunch of vehicles released from a traffic signal*. RRL Research Note RN/2665/GMP. London: Road Research Laboratory.

Park, B., C. J. Messer, and T. Urbanik. 1999. "Traffic signal optimization program for oversaturated conditions: Genetic algorithm approach." *Transp. Res. Rec.* 1683 (1): 133–142. https://doi.org/10.3141/1683-17.

Park, B. B., and J. Lee. 2009. "Stochastic optimization of coordinated actuated traffic signal system method based on shuffled frog leaping algorithm." *Transp. Res. Rec.* 2128: 76–85. https://doi.org/10.3141/2128-08.

Park, B. B., C. J. Messer, and T. Urbanik. 2000. "Enhanced genetic algorithm for signal timings optimization of oversaturated intersections." *Transp. Res. Rec.* 1727 (1): 32–41. https://doi.org/10.3141/1727-05.

Park, B. B., N. M. Rouphail, and J. Sacks. 2001. "Assessment of stochastic signal optimization method using microsimulation." *Transp. Res. Rec.* 1748 (1): 40–45. https://doi.org/10.3141/1748-05.

Park, B. B., and J. D. Schneeberger. 2003. *Evaluation of traffic signal timing optimization methods using a stochastic and microscopic simulation program*. Research Rep. No. UVACTS-5-0-4. Charlottesville, VA: Univ. of Virginia.

Park, B. B., and I. Yun. 2006. *Evaluation of stochastic optimization methods of traffic signal control settings for coordinated actuated signal systems*. Research Rep. UVACTS-15-0-102. Charlottesville, VA: Univ. of Virginia.

Rouphail, N., Park B., and Sacks J. 2000. *Direct signal timing optimization: Strategy development and results*. NISS Technical Rep. No. 109. Research Triangle Park, NC: National Institute of Statistical Sciences.

Stevanovic, A., P. T. Martin, and J. Stevanovic. 2007. "VISGAOST: VISSIM-Based genetic algorithm optimization of signal timings." *Transp. Res. Rec.* 2035: 59–68. https://doi.org/10.3141/2035-07.

Stevanovic, A., J. Stevanovic, C. Kergaye, and P. Martin. 2011. "Traffic control optimization for multi-modal operations in a large-scale urban network." In *Proc., IEEE Forum on Integrated and Sustainable Transportation Systems*. New York: IEEE.

Stevanovic, A., J. Stevanovic, J. So, and M. Ostojic. 2015. "Multi-criteria optimization of traffic signals: Mobility, safety and environment." *Transp. Res. Part C* 55 (Jun): 46–68. https://doi.org/10.1016/j.trc.2015.03.013.

Stevanovic, J., A. Stevanovic, P. T. Martin, and T. Bauer. 2008. "Stochastic optimization of traffic control and transit priority setting in VISSIM." *Transp. Res. Part C* 16 (3): 332–349. https://doi.org/10.1016/j.trc.2008.01.002.

TRI (Transportation Research Institute). 2013. *INBAR signal-program planning system—User guide version 15*. [In Hebrew.] Haifa, Israel: Transportation Research Institute, Technion—Israel Institute of Technology.

Wolput, B., E. Christofa, A. Carbonez, A. Skabardonis, and C. Tampere. 2015. "Optimal traffic signal settings with transit signal priority." In *Proc., 94th Annual Meeting of the Transportation Research Board*, 1–19. Washington, DC: National Academies of Sciences, Engineering, and Medicine.

Yang, N., and Y. Liu. 2008. "Simulation optimization of urban arterial signals via simultaneous perturbation stochastic approximation (SPSA)." In *Proc., 15th World Congress on Intelligent Transport Systems*. Washington, DC: ITS America.

Yun, I., and B. B. Park. 2012. "Stochastic optimization for coordinated actuated traffic signal systems." *J. Transp. Eng.* 138 (7): 819–829. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000384.