Flexible-cycle Transit Signal Priority

Effects of variable cycle lengths on Transit Signal Priority operations

Narayan Donaldson

FLEXIBLE-CYCLE TRANSIT SIGNAL PRIORITY

EFFECTS OF CYCLE LENGTH CONSTRAINTS ON TRANSIT SIGNAL PRIORITY OPERATIONS

Master Thesis
CIE5060-09

Narayan Donaldson
5082277

January 31, 2022

Thesis Information

Author	
	N. Donaldson MSc Student, Transport & Planning TU Delft, Faculty of Civil Engineering & Geosciences
Thesis Committee	
Chair	dr. V.L. Knoop Associate Professor Director, Traffic Dynamics Modelling and Control Lab TU Delft, Faculty of Civil Engineering & Geosciences
Daily Supervisor	dr. ir. A.M. Salomons Researcher Traffic Dynamics Modelling and Control Lab TU Delft, Faculty of Civil Engineering & Geosciences
External Advisor	dr. ir. N. van Oort Assistant Professor Co-Director, Smart Public Transport Lab TU Delft, Faculty of Civil Engineering & Geosciences
Company Supervisor	M. Fick Specialist Verkeersregeltechniek Movensis B.V.

Preface

While I was working with traffic signals in the City of Toronto, one of the common areas of discussion among colleagues was whether the unique flexible-cycle Transit Signal Priority system used in the City was worth the extra effort to maintain compared to the standard fixed-cycle systems found in most other North American jurisdictions. I looked for studies which had been done to answer this question, but could not find any. It is therefore a pleasure that I have been able to spend the end of my Master's studies researching the topic of flexible-cycle TSP systems.

Working on this project has allowed me to become familiar with several items I had wanted to learn. Most importantly, I have had the opportunity to become intimately familiar with a Dutch fully-actuated *and* coordinated traffic signal system. I have also been able to become far more proficient with Vissim and with MATLAB than before, both of which would have been useful in my previous work, and will surely be useful in the future.

I could not have completed this project without the help of certain key individuals and organisations.

Movensis has provided resources to the project, namely access to Vissim modelling software and relevant add-ons, access to the code of existing CCOL traffic signal controllers, funding, and considerable help with working in both CCOL and Vissim.

Specifically, I would like to thank Marcel Fick, who has been supportive of this project since before I had selected a topic, and who has provided invaluable detailed information and insight on Dutch traffic signal operations. He also helped me solve issues with my traffic signal code when I was unable to do so myself, and coded the module which translates the input from vehicle detectors into summed vehicle volumes.

I would also like to thank Sander Boerma, a colleague at Movensis, who helped me to set up the Vissim model and connect the CCOL controller to the Vissim model.

From the TU Delft, I would like to thank my supervisor Maria Salmons, who has also been consistently supportive and has helped direct my research.

I am also highly appreciative of Victor Knoop and Niels van Oort for taking the time from their busy schedules to form the thesis committee.

Finally I must appreciate the support I have received from my parents, who enabled me to go study in another continent in order to find the knowledge I was seeking.

Summary

The typical way of coordinating a group of traffic signals with one another is to operate them all at a common, fixed, cycle length. This enables the creation of relations such as green waves, but also creates limitations for the implementation of transit signal priority (TSP) - a system which adjusts signal timings in real time to reduce delays for public transport vehicles. Whereas green waves depend on signals operating consistently, TSP depends on the signals being able to vary their operations with as few constraints as possible. This results in some conflicts between TSP objectives and coordination objectives. With a fixed cycle length, some TSP actions may not be possible to implement, if there is insufficient available time within the cycle length. And requiring all TSP interventions to be completed within a single cycle reduces the ability of a signal controller to compensate movements which were negatively affected by a priority intervention once the public transport vehicle has cleared the intersection.

The purpose of this thesis project is to study the effects of cycle length constraints on the effectiveness of transit signal priority (TSP) systems, using a Vissim microsimulation model of a simple fictional network. The project was conducted in collaboration with Movensis B.V., a traffic signal consulting company based in Den Haag.

Methodology

The core of the research is based around a new variant of a Dutch CRSV halfstarre traffic signal controller which was designed and programmed exclusively for this study. It includes a new TSP system which aims to reduce the constraints on priority interventions while still maintaining most of the benefits of fixed-cycle coordination. It allows the cycle length to temporarily vary during TSP interventions, followed by up to two cycles of "offset correction", where the any net change in cycle length is neutralised by implementing compensatory adjustments to green durations. A novel feature of the system is the concept of "offset correction credits" (OC Credits), which represent the amount of extra (or reduced) green time each phase would need to receive to compensate for any disbenefits (or benefits) it received during a TSP intervention.

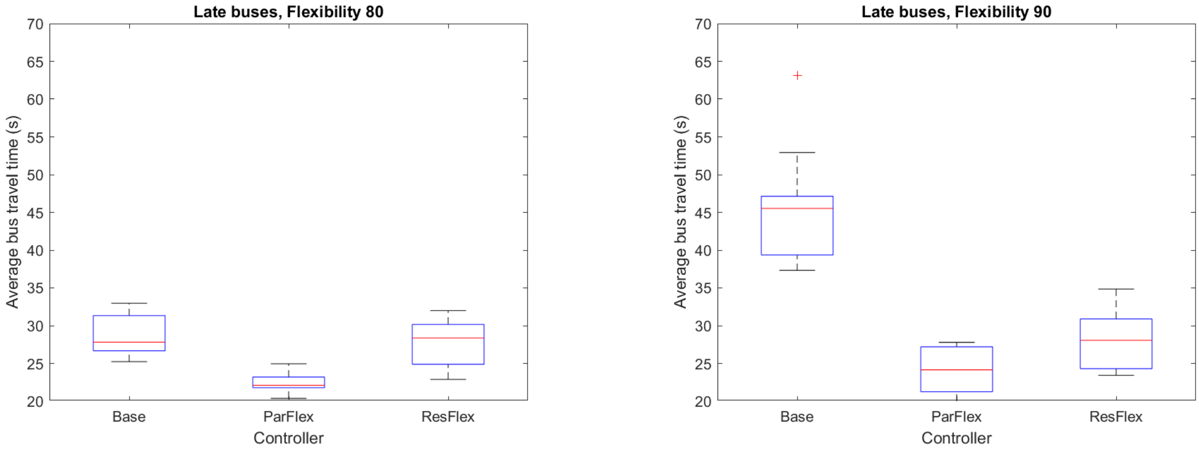
There are two versions of the new TSP system: the Parametric Flexible Cycle system (ParFlex) and the Responsive-parameter Flexible Cycle system (ResFlex). In the more basic ParFlex system, offset correction credits are distributed on a 1:1 basis, such that a one-second impact on a phase always produces one OC credit. The desired phase durations during offset correction (before redeeming OC Credits) are simply set to a fixed proportion of the normal maximum green durations, either 90% or 95% depending on the scenario. In the ResFlex system, negative offset correction credits are calculated based on the actual number of vehicles which made use of extra time during a TSP intervention. The number of OC credits represents the number of seconds of green time which would be required to clear that number of vehicles. The desired phase durations during offset correction are based on the observed traffic volumes over the previous 15 minutes, as measured by the vehicle detectors at the intersection.

The two new TSP systems are tested in a Vissim microsimulation environment, along with the existing fixed-cycle system (Base) which serves as a base case. The three TSP systems are each tested with two different sets of signal timings, for a total of six modelled scenarios. The performance of the three TSP systems is compared in terms of public transport delay, public transport reliability, ability to maintain coordination, and delay for other road users. The subject road network consists of fictional road with three coordinated traffic signals, spaced 150 metres and 400 metres apart. The difference in spacing allows the effects of intersection spacing to be observed in the results. In addition, variations in the controllers' range of flexibility to adjust green durations allow the results to capture different TSP systems' ability to cope with green time constraints.

Results

The scenarios using the new Parametric Flexible-cycle TSP system (ParFlex) and Reponsive-parameter Flexible-cycle TSP system (ResFlex) produced significantly lower bus delays than the fixed-cycle Base TSP system, and the relative difference between the two types of controller becomes increasingly large at intersections with a

lower amount of flexibility in green durations (figure 2).



(a) High-flexibility scenarios

(b) Low-flexibility scenarios

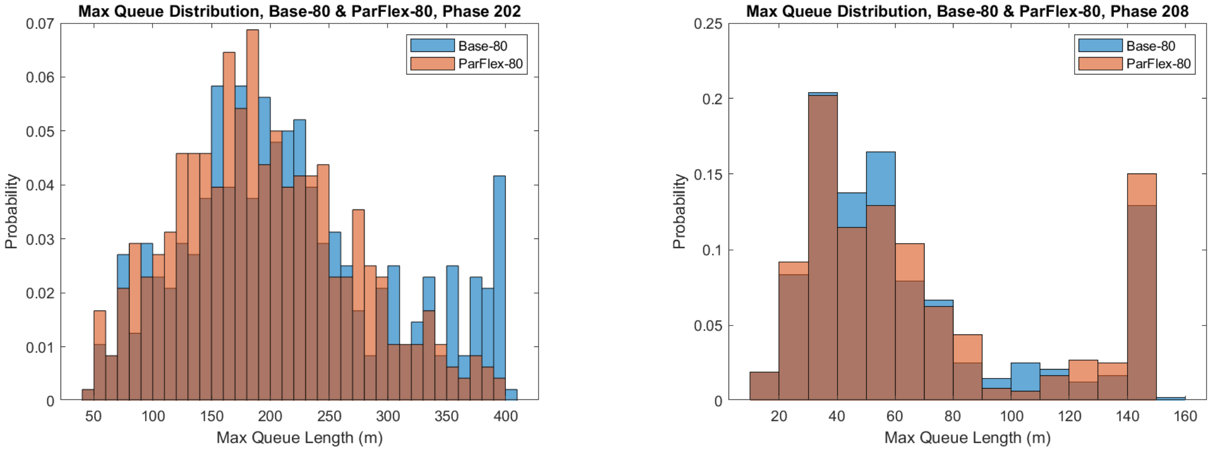
Figure 2: Bus travel times per scenario - late buses

In the scenario with high flexibility, the average delay for late buses dropped by 59% from 10.7 seconds to 4.4 seconds for the ParFlex system compared to the Base system. With low flexibility, the average delay for late buses dropped by 78% from 28.0 seconds to 6.2 seconds. The large improvements in performance for buses are due to the flexible-cycle TSP systems being able to execute more TSP actions such as phase insertions which may not fit within a fixed cycle length.

However, the controller's ability to remain in sync was negatively impacted, with 17% of cycles out of sync in both tested scenarios, compared to 16% and 5% for the Base system in the high and low flexibility scenarios, respectively.

The ResFlex system permitted a lower level of priority for buses than the ParFlex system, given the large volumes of traffic included in the modeled scenarios. The delay for late buses was 9.9 seconds for the scenario with high flexibility, which is not significantly lower than the Base system. With low flexibility, the average delay for late buses with the ResFlex system increased slightly to 10.5 seconds, but that is now 63% lower than the Base system achieved with the same level of flexibility. The new flexible-cycle systems are far less sensitive to changes in flexibility than the fixed-cycle Base system.

Queue lengths for motor traffic remained unchanged for most directions of motor traffic with the flexible-cycle TSP systems. The exceptions are that the coordinated directions with short intersection spacing faced as much as a 70% increase in the frequency of queues exceeding storage in the low-flexibility scenarios. This was also accompanied by a significant increase in the number of stops. Meanwhile, with high flexibility, the queue lengths on the long link dropped significantly, as intended by the new TSP system's green time compensation mechanism (figure 3).



(a) Phase 02 (westbound through)

(b) Phase 08 (eastbound through)

Figure 3: Distribution of queue lengths, Base and New1 controllers, high flexibility

As expected, the flexible-cycle systems increase the delay for general traffic parallel to the bus directions (labeled as "auto: priority phases" in table 1). This is due to the offset correction mechanism redistributing time to other movements to compensate for the impacts of the priority intervention. However, the coordinated phases, which are perpendicular to the bus phases, do not consistently see an improvement. This is because although the coordinated phases receive an improved distribution of green time, they are also negatively affected by the greater deviation from coordination which the flexible-cycle systems allow. The delay to other movements did not vary significantly between the TSP systems.

Table 1: Average vehicle-delay per phase type

	High flexibility			**Low flexibility**		
Delay	**Base**	**New1**	**New2**	**Base**	**ResFlex**	**ParFlex**
Bus: Late buses	10.7 s	4.4 s	9.9 s	28.0 s	6.2 s	10.5 s
Auto: Coordinated Phases	28.0 s	26.6 s	26.4 s	24.8 s	25.9 s	26.6 s
Auto: Priority Phases	47.4 s	50.0 s	50.4 s	47.6 s	51.7 s	50.7 s
Auto: Other Phases	40.0 s	40.0 s	40.2 s	39.1 s	40.1 s	40.0 s
Bicycle: All Bicycle Phases	35.0 s	34.7 s	34.5 s	34.8 s	35.5 s	35.3 s

The average person-delay per intersection approach was calculated for the network by weighting the vehicle delay by the number of vehicles and occupancy per vehicle. The exception is buses, which were only included in the calculation if they had a greater than average headway, with the rationale that buses with a shorter headway would need to be delayed regardless of the traffic signal in order to restore a regular service. When the assumed occupancy rate for buses with above average headways is 50 passengers (corresponding to a busy but not overcrowded standard bus), there was no significant difference in person-delay between any of the scenarios.

The results from the Responsive-parameter version of the flexible-cycle TSP system were not notably different than the Parametric version. The ResFlex controller provided a lower level of priority to buses resulting in higher delay for buses, and lower delays, shorter queues and fewer stops for other road users.

Regardless of the TSP system used, there were no significant differences in bus passenger waiting times downstream of the intersection, nor were there any significant differences in delay for cyclists.

Conclusions

By temporarily relaxing the constraint of a fixed cycle length during and immediately after TSP interventions, the delay for prioritised buses can be dramatically reduced. Although any TSP system can tradeoff bus delay against delay for other road users, the delay reduction observed with the flexible-cycle TSP systems cannot be replicated with a fixed-cycle TSP system given the constraints of the background signal timings.

In general, the effects on other road users were minor, despite the significant improvements for buses. The exception is for the coordinated directions between closely-spaced signals, which saw significant reductions in motor vehicle performance.

The conditions in which a flexible-cycle TSP system offers the greatest benefits over a fixed-cycle TSP system are in coordinated networks with heavy traffic volumes, signal timings with relatively little ability to reduce phase durations, and relatively long distances (>400 metres) between signals. Under those conditions, the study indicates that bus delay can be significantly reduced without increasing network person delay when the bus line is perpendicular to the coordinated directions with a frequency of up to 12 buses per hour per direction and conditional priority granting a high level of priority to approximately half of the buses.

Table 2: Criteria in which flexible-cycle TSP systems are most beneficial

Criterion	**Ideal condition**
Network control type	Fixed-cycle coordinated
Public transport direction	Perpendicular to coordination
Public transport frequency	Up to 12 buses per hour per direction
Priority requests	50% of buses granted high priority
Intersection spacing	400+ metres
Minimum green times	constrained
Traffic volume	v/c ratio up to 0.9

Next Steps

Due to the need to manually set up each scenario in the microsimulation environment, the number of trials in the study was limited, and by extension so were the number of variables which could be studied. Future research is required to see how the effects study the effects of other characteristics on the relative benefits of flexible-cycle TSP systems, such as the traffic volumes, the frequency of public transport lines, and other types of road network, such as with a bus route travelling in the coordinated direction.

Despite the relatively narrow scope of the study, there are some real-world signal networks which match the conditions which were observed to be well-suited for flexible-cycle TSP systems. For example, Movensis B.V. has identified a group of traffic signals along the Western Ring Road (Westelijke Randweg) of Haarlem which meets all of the conditions identified in the study.

Prior to implementing the new flexible-cycle TSP system on-street, some modifications would be required, including some bug fixes and a connection to the controller's user interface. Further development of the Responsive-parameter Flexible-cycle TSP system to automatically calibrate and validate its assumed parameters may be able to improve that system sufficiently for that system to be worthwhile. However in its current form, it does not provide any observable advantage over the parametric version of the flexible-cycle TSP system.

Table of Contents

List of Figures

List of Tables

1 Introduction

Reducing the travel time of public transport services is an important objective for road authorities, because attracting more riders onto public transport can bring considerable quality-of-life benefits to residents, including a reduction in energy consumption in transportation as well as an increase in the number of people who can be transported in a given amount of roadway space (NACTO, 2016). In addition to attracting more riders to the service, reducing a route's round trip travel time can also reduce operating costs by reducing the number of vehicles and operators required to operate a given frequency of service.

Numerous measures are available to improve the speed and reliability of public transport. To remove vehicles from motor vehicle traffic congestion, public transport can be provided with a dedicated right-of-way. To improve acceleration, vehicles with a higher power-to-weight ratio can be selected. And to avoid delays at intersections, a public transport right-of-way can be grade separated. The latter measure has the potential to considerably improve public transport performance because traffic signals are often the primary source of delay between stops for public transport services (CROW, 2010, NACTO, 2016). However, due the high construction and maintenance costs of grade separations, some public transport lines will inevitably need to traverse intersections at grade.

Some basic changes to the traffic signal timings can reduce delay to public transport, such as reducing the cycle length, increasing the proportion of green time dedicated to the direction of the public transport line, and adjusting the offset between coordinated signals to better suit the estimated progression of the public transport vehicles (NACTO, 2016, Wolput et al., 2016). However, the disadvantage of such passive measures is that they remain static regardless of the actual location of the public transport vehicles. This means that other users may experience a less optimal green time distribution or less optimal signal progression even when there is no public transport vehicle present. The lack of feedback mechanism from the vehicle to the intersection means that public transport vehicles are more likely to experience delays if they have a different travel time than expected, for example due to spending longer than usual serving a passengers at stop.

The measure which introduces that feedback mechanism is active Transit Signal Priority (TSP). TSP is a general term which refers to signal timing changes which are implemented in response to the detection of public transport vehicles in real time (FHWA, 2008). It is commonly implemented for tram and bus systems in cities around the world, and it can provide a dramatic reduction in delay for public transport vehicles with a relatively low implementation cost (NACTO, 2016). That said, given that TSP is such a broad term, the effectiveness of "TSP" systems varies considerably depending on the implementation. In some cases it can virtually guarantee a green light for eligible public transport vehicles, while in other implementations it provides only a slight reduction in delay.

1.1 Research Interests and Objectives

One notable situation in which TSP systems work less well is in coordinated traffic signal networks with a fixed cycle length, particularly if the intersection has a long minimum cycle length (Li et al., 2020). In most TSP systems in coordinated signal networks, the intersections' cycle length remains fixed even during TSP interventions. This may place strict limitations on the magnitude of TSP actions, such as the maximum duration of green extensions or inserted phases. It may also result in green time being redistributed at the expense of phases which conflict with the prioritised phase.

For example, the intersection of Spadina Avenue at Bremner Boulevard in Toronto Canada is shown in figure 4, where there is a streetcar (tram) line running north-south along Spadina Avenue.

Figure 4: Overhead view of Spadina Avenue at Bremner Boulevard, Toronto

The east-west pedestrian crossing distance is 40 metres, so based on the City of Toronto standard walking speed of 1.2 metres per second (City of Toronto, 2015), the pedestrian clearance time is 34 seconds. Including the other minimum intervals (listed in table 3), the westbound phase at the intersection has a minimum duration of 46 seconds. Adding in the minimum durations for the other phases at the intersection produces a minimum cycle length of 109 seconds. If Spadina Avenue operates with a fixed cycle length of 120 seconds, then the signal cycle can only be adjusted by up to 11 seconds. This means that if the cycle length remains constant during TSP interventions, it is not possible to extend a green light by more than 11 seconds, or insert a phase which has a total duration of 11 seconds.

Signal Stage	Min Walk	D.Wlk Clear	Min Green	Yellow	Red Clear	Total
North-South Left Turns	-	-	6	3	3	12
North-South Through	*7*	*23*	30	4	3	37
East-West Left Turns	-	-	6	3	3	12
East-West Through	*7*	*34*	41	4	3	48
Total			83	14	12	109

Table 3: Minimum timing intervals at Spadina & Bremner/Fort York, Toronto

Temporarily relaxing the constraint of the cycle length during TSP interventions would allow the level of priority to be significantly increased at intersections where the cycle length would otherwise limit the available TSP actions. Such a "flexible-cycle" system would also allow extra time required for the TSP phase to be partially claimed from the same phase during a later cycle, rather than from conflicting phases during the same cycle. Given the potential benefits of flexible cycle TSP systems for both public transport and general traffic, there is the potential that in some cases flexible-cycle TSP systems could reduce delays for both public transport and other traffic, which represents a more efficient TSP system.

Flexible-cycle systems have existed in practice for decades, but there is relatively little study of them in literature. Most studies of coordinated systems assume that the cycle length will always remain fixed including during TSP interventions, either because it is simpler to implement or because flexible-cycle TSP systems were not considered as an option.

What little study does exist on flexible systems is primarily focused on other TSP characteristics, and does not actively compare the performance of those systems to fixed-cycle systems. This is the research gap that this study will aim to fill.

1.2 Problem Definition

The vast majority of the literature related to TSP systems in coordinated signal systems is based on fixed-cycle systems, which constrain the cycle length to the network cycle length even during priority requests and during the recovery process after TSP actions. While that constraint does ensure that coordination is maintained at all times, it can also have negative effects on the TSP system's ability to provide a green light to the public transport vehicle which should be receiving priority, as well as the system's ability to compensate for the impacts of TSP actions.

Meanwhile, preemption systems do lift the constraint of cycle lengths during priority interventions and the subsequent offset correction process, but they too have significant impacts on general traffic due to their lack of constraints during preemption, and the lack of connection between the preemption and recovery processes.

Flexible-cycle TSP systems may offer an optimal middle ground between the highly-restricted prioritisation provided by fixed-cycle TSP systems, and the virtually unrestricted priority provided by preemption systems. There are likely situations where fixed-cycle TSP systems may preclude prioritisation actions which would have a considerable benefit to the public transport vehicle, with only minimal impacts on other road users. Conversely, preemption systems may execute priority actions which result in considerable impacts on other traffic for only a minor time savings for the prioritised vehicle.

Although there are some studies of existing flexible-cycle TSP systems, there is a lack of research which compares them to fixed-cycle systems to quantify their differences. There is also a lack of research on the effectiveness of the different offset correction strategies which become possible with flexible-cycle systems.

1.3 Research Questions

It is clear that there are some theoretical advantages to flexible-cycle TSP systems, including the potential for a higher level of priority, as well as the potential for a more effective compensation mechanism to balance green times between phases. But due to the lack of research actively comparing flexible-cycle and fixed-cycle TSP systems, it is not clear how those advantages compare to the disadvantages of temporarily disrupting signal coordination. To fill this research gap, the primary research question will be:

> *How does relaxing the constraint of cycle length during and after TSP actions affect the performance of a coordinated TSP system?*

The performance of the system will be defined using metrics such as public transport travel time and reliability, delay for other traffic, queue lengths in constrained directions, and number of stops for traffic travelling in the coordinated directions.

1.3.1 Subquestions

The research question is then divided into three primary subquestions.

- *How do changes to the cycle length during TSP interventions affect the signal's ability to promptly serve public transport vehicles?*
- *To what extent can the cycle length vary during TSP interventions while still maintaining the benefits of signal coordination, namely a reduced number of stops for vehicles in the coordinated direction, and a reduced likelihood of vehicle queues exceeding the available storage?*
- *In which situations would a flexible-cycle TSP system provide the greatest benefit compared to a fixed-cycle system?*

The first subquestion focuses on the nature of the benefits which public transport service experiences as a result of eliminating the strict constraint of cycle length. The second subquestion focuses on the disbenefits which are experienced by other road users, with a particular focus on the coordinated directions. And the third subquestion aims to determine the practical application of flexible-cycle TSP systems.

2 Background and Literature

Transit Signal Priority (TSP) is a general term used to describe changes to signal timings which reduce traffic signal delays for public transport vehicles. This can reduce travel times along the public transport line as well as increasing service reliability. This in turn will improve the attractiveness of the service, attracting new riders and potentially reducing the demand for other competing transport modes such as private automobiles. The reduced travel times and travel time variability also allows operators to schedule a shorter round-trip time for a given service frequency, thereby reducing the number of vehicles required and by extension the operating costs (NACTO, 2016).

In order to understand how TSP can modify signal timings, it is first necessary to understand how traffic signals normally work. The following section describes some key concepts which are relevant to TSP.

There is considerable variation in traffic signal terminology between different regions, and some terms have different meanings. For example, the term ’traffic-responsive’ in North America refers to a system which selects its timing plan on the basis of current traffic volumes, but the Chinese study by Mei et al. (Mei et al., 2019) uses the term to describe the process of extending a phase based on the detector (which is known as ’vehicle actuation’ in North America). The term ’phase’ also varies in meaning between the U.S. and the U.K. (Wahlstedt, 2011). In general, this paper will use the terminology as used in North America. A list of relevant terms is included in Appendix A along with the corresponding terms used in the Netherlands.

2.1 Traffic signal operations

The traffic signal operations concepts in this section describe different ways in which the operation of traffic signals can be classified.

2.1.1 Vehicle Actuation

Actuation, or vehicle actuation, is the process of determining a signal’s behaviour based on the vehicles currently detected at the intersection. This typically consists of calling a phase to serve waiting/approaching vehicles or pedestrians, and/or extending a phase which is already green to accommodate approaching vehicles. This occurs on a moment-to-moment basis, with decisions typically reevaluated ten times per second. The vehicle detector setup for a typical motor traffic lane at a traffic signal in the Netherlands is illustrated in figure 5.

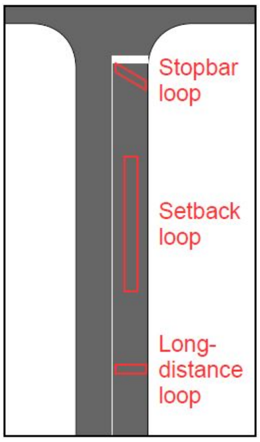


Figure 5: Typical detector setup in the Netherlands

Using vehicle actuation, the phase is called if the vehicle detector is active. If there is no call, the phase is skipped. If the calls are "locking", the call remains until the light turns green even if the detector is no longer occupied, for example if a vehicle turns right on red. If the calls are "non-locking", the call will disappear if the detector is no longer occupied (FHWA, 2008).

To principle for extending a phase through vehicle actuation is that the phase should only remain green while the approaching traffic exceeds a certain density. If the density of approaching traffic falls below the threshold, the phase should end so another conflicting direction can get a green light. If the length of the phase's vehicle detector is equal to the intended maximum distance between vehicles, the phase will continue to extend as long as no vehicles are further apart than the distance threshold. If any two vehicles are further apart than the length of the detector, the detector will momentarily unoccupied and the light will immediately turn yellow. If the detector is shorter than the intended gap between vehicles, a "passage time" is used to account for the discrepancy, by continuing to extend the phase for a certain number of seconds after the detector is no longer occupied.

The alternative to actuation is fixed-time operation, where phases are called every cycle regardless of demand, and the duration is predetermined based on the timing plan.

An intersection's level of actuation is described based on the number of phases which are actuated. At a fully-actuated intersection, every phase is actuated. At a semi-actuated intersection, all phases are actuated, except for one phase (or group of compatible phases) which are fixed-time. This operation commonly occurs along coordinated arterial roads in North America, where all phases are actuated except for the through phases for the arterial road itself, which simply receive the remainder of the time in the (fixed-length) cycle after the conflicting phases have been served. At a fixed-time intersection, at least two conflicting phases are fixed-time, causing the signal to always cycle, even in the absence of any demand (FHWA, 2008).

2.1.2 Traffic Adaptation

Adaptation, or traffic adaptation, is the process of determining the signal's behaviour based on observed traffic volumes. Whereas actuation is based on the presence or absence of individual vehicles at a particular moment, adaptation is based on traffic volumes observed over a period of time such as 10 minutes. Given the observed traffic volumes, a model of current traffic conditions is created, which is then used to determine the optimal distribution of green time given an objective such as minimizing vehicle-delay.

Adaptation and actuation can be combined by using an adaptive system to determine the parameters such as maximum phase duration, while the actuation logic determines how much of that time (if any) is actually served for each phase.

The opposite of adaptive operation is time-of-day operation, where the intersection has a set of predefined timing plans, and the plan currently in effect is determined by a predefined schedule. A different timing plan than scheduled can be manually selected by the traffic control centre based on observed incidents or planned events (Peñabaena-Niebles et al., 2014).

An intermediate between adaptive systems and time-of-day systems are traffic-responsive systems, where the intersection has a set of predefined timing plans based on hypothetical traffic volumes, and on-street detectors are used to select the most appropriate plan based on the current traffic volumes (Li et al., 2011).

2.1.3 Coordination

Coordination is the process of operating a group of signals at a common cycle length (or a multiple thereof), so that the green times at the signals can be controlled relative to each other (FHWA, 2008). The cycle length used by signals within a control group may remain constant, or it may be allowed to vary at a network level based on adaptation or actuation at the constituent intersections.

The timing of one intersection relative to another is based on the intersections' "offsets", which can be defined for each intersection relative to a central clock, or relative to one of the intersections in the control group. The offset itself represents the difference in time between a reference point within the intersection's cycle (commonly the start of a coordinated phase), and a common reference point for the entire control group (either the start of the cycle for a central clock, or a particular point in the cycle at one of the constituent intersections). With fully- or semi-actuated intersections, coordinated phases may actually start earlier than indicated in the background timing plan, but in that case the reference point for the offset is still the planned start time for the coordinated phase, not the actual start time.

Uncoordinated systems allow the cycle length at each intersection to vary from cycle to cycle based on actuation. Although the offset between adjacent signals is not fixed, there are still mechanisms available to improve traffic progression between adjacent uncoordinated intersections, such as detectors in advance of the intersection to detect approaching platoons of vehicles, or information sent directly between adjacent controllers with information such as the current signal state and/or detector occupancies (Wahlstedt, 2011).

2.2 Transit Signal Priority operations

TSP is commonly described as including both "passive" and "active" priority (Wahlstedt, 2011). Passive priority is simply when the offsets in a coordinated system's timing plans were selected based on the anticipated progression of public transport vehicles rather than the progression of cars. It is suitable in situations where the public transport frequency is extremely high, the progression of public transport vehicles is very predictable, and the network layout makes it possible to create a coordination plan which provides good coordination in both directions (if applicable) (Wahlstedt, 2011). Because passive priority does not respond to the presence of public transport vehicles, private vehicles experience suboptimal coordination regardless of whether there is a public transport vehicle approaching or not. It is also unable to vary the amount of delay based on a public transport vehicle's on-time performance in order to increase service regularity.

The more effective form of public transport prioritisation is active priority, where the position of public transport vehicles is monitored in real time in order to adjust signal timings to better suit their anticipated arrival at intersections (NACTO, 2016).

Traditionally, vehicles have been detected by physical detectors in the roadway surface, with check-in detectors in advance of the intersection, and check-out detectors beyond the stop lines. The detectors often distinguish public transport vehicles from general traffic by detecting unique radio messages transmitted from public transport vehicles. More recently, physical detectors have started to be replaced by virtual detectors, which are simply geographical coordinates where a 'check-in' or 'check-out' message is created based on the vehicle's current position determined by GPS (Ekeila et al., 2009).

Currently there is interest in migrating public transport vehicle communication to the emerging standards for connected vehicles, rather than maintaining a separate system for public transport vehicles (Seredynski et al., 2020). In contrast to the previous "check-in" and "check-out" systems, connected vehicle communication involves continuous communication between the vehicle and the wayside systems. This allows the intersection to continuously update the estimated arrival time of the vehicles at the intersection, and thereby better account for travel time variability.

The process for giving a vehicle priority using TSP follows the following steps (NACTO, 2016):

1. Vehicle sends a TSP Request to the TSP system
2. If conditions are met, system executes an a TSP intervention consisting of one or more TSP actions
3. Once the intervention is complete, the system returns to normal operation. In coordinated signal systems, this always includes an "offset correction" process.

2.2.1 TSP Actions

If a TSP system receives a priority request from a vehicle which is expected to arrive at a red light, there are five possible TSP actions the system can take to adjust the signal timings to better serve that vehicle.

All strategies for providing priority can be described as combinations of the five core TSP actions. For example, the 'green time reallocation' TSP strategy proposed by Hu et al. (2016) consists of a phase insertion of the TSP phase followed by compensatory adjustments to the TSP phase's normal green time.

Green Extension
Green extension is the action of extending the duration of a phase above its typical maximum duration. Typ-

ically this will be used for the phase used by the transit vehicle, but in some cases a preceding phase may be extended so that a short dedicated transit phase better aligns with the estimated arrival time of the vehicle.

Red Truncation
Red truncation is the action of terminating a phase earlier than under the normal signal timings. This typically occurs for phases preceding the transit phase.

Phase Insertion
Phase insertion is the action of adding an additional phase into the cycle. This may be a unique signal phase designed to serve the public transport vehicle, or an additional opportunity for a normal signal phase during a different part of the cycle than it normally occurs. Some literature distinguishes between "inserted" phases and "callable" phases(Seredynski et al., 2015). Under that distinction, a callable phase is included in the normal signal timing, but only called by the TSP system, while an inserted phase is not included in the normal signal timing at all. However, since the on-street effect of either method is virtually identical, most literature does not make this distinction. In Sweden, a variant of phase insertion known as "re-taken start" is defined as a phase being inserted into the cycle immediately after it ended. This would occur if a TSP call for that phase is received after the phase has changed to yellow, but before the following phase has turned green (Wahlstedt, 2011).

Phase Rotation
Phase rotation is the action of changing the order of phases within the cycle. It is unique among TSP actions in that it does not intrinsically have any effects on phase durations. This makes it particularly useful in systems were the cycle length is constrained. Its primary limitation is that the irregular phase order can result in very long waits between consecutive opportunities for certain phases. This in turn can have impacts on operations if the longer red light duration causes the queue of waiting traffic to exceed the available storage capacity (Zlatkovic et al., 2015).

Phase Skipping
It is theoretically possible to skip phases which precede the prioritised phase, to allow that phase to occur earlier. The impacted phases would not turn green until the following cycle. Due to the severe impacts on affected phases, this method is rarely used.

2.2.2 Offset Correction

All of the TSP actions except phase rotation would result in changes to the cycle length in the absence of some form of compensation. Green extensions and phase insertions would increase the cycle length, and red truncations would reduce the cycle length. In order to maintain coordinated operations, any time added to or subtracted from the cycle must be accompanied by a compensatory adjustment, known as offset correction (Huang et al., 2012).

In order to recover coordination following a TSP intervention which added time into the cycle, there are two possible strategies. The most intuitive option is to simply remove the same amount of time that was added to the cycle length during the intervention:

$$T_{OC} = -T_{TSP}$$

Where:

T_{OC} = *Time to be added during offset correction*
T_{TSP} = *Time that was added to cycle during TSP intervention*

It is also possible to return in sync by adding the remainder of the cycle length above the amount that was added during the TSP intervention:

$$T_{OC} = C - T_{TSP}$$

Where:

T_{OC} = *Time to be added during offset correction*
T_{TSP} = *Time that was added to cycle during TSP intervention*
C = *The cycle length during coordinated operation*

The latter option results in the total number of cycles executed to be reduced by one. Typically two cycles will occur in the time which would have accommodated three cycles under normal coordinated operation. As a result, that option tends to be undesirable unless the time added during the TSP intervention was a significant portion of the normal cycle length.

To recover coordination following a TSP truncation, the inverse of the same two options are available. But in practice, TSP truncations are nearly always recovered using OC extension, since recovering through OC reduction would require the average cycle length during the TSP intervention and associated offset correction to be half of the normal cycle length (Zlatkovic et al., 2015).

All TSP systems for coordinated networks have some form of offset correction. In some systems, such as the system proposed by Liao and Davis (2007), the added or reduced during offset correction is deliberately allocated to particular phases to compensate for the redistribution of green time during the TSP intervention. In many other systems, the time is reallocated in a trivial manner, such as by applying offset correction adjustments to phases in chronological order until the total desired offset correction amount has been reached.

The allocation of green time during offset correction is subject to two conflicting demands: coordination and compensation. From the perspective of coordination, the objective of offset correction should be to return the signal to its intended offset as quickly as possible. To this end, OC adjustments should be applied to every phase possible. But from the perspective of compensation, OC adjustments should only be applied to phases where they would help to counteract any green time redistribution which occurred during the TSP intervention. The range of options available for reduction-based offset correction following a green extension are illustrated in figure 6.

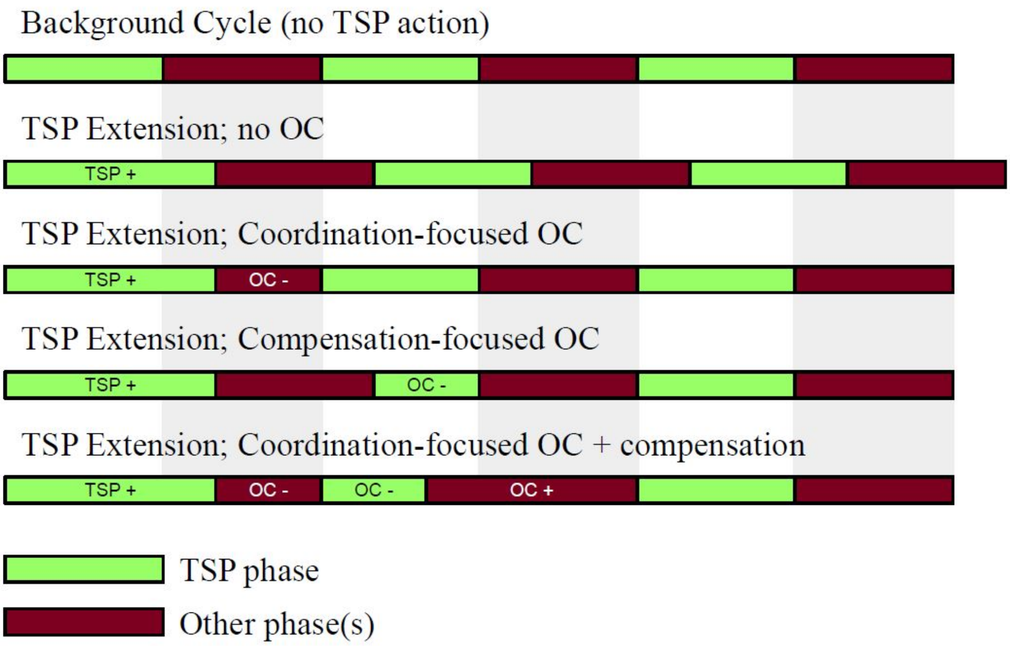


Figure 6: Offset correction strategies for green extension

The *no OC* scenario illustrates the theoretical scenario where the TSP action is not compensated for in any way. The extra time added to the TSP phase results in an increase in the cycle length, causing the signal's offset to shift. And since there is no offset correction, the signal remains out-of-sync indefinitely. This option would not be used in coordinated signal systems, because it would completely eliminate coordination.

The 'coordination-focused' offset correction scenario is the typical paradigm that is used in traffic signal controllers and in research. Time which is added through green extensions is subsequently subtracted from conflicting phases. The advantages of the coordination-focused OC approach are that it allows the signal to return to ordinary coordinated operation as quickly as possible, and that it is very simple to implement (Mei et al., 2019). In the example in figure 6, the only phase to begin later than its intended coordinated moment is the conflicting phase directly impacted by the TSP extension itself.

The coordination-focused strategy is typically implemented with the constraint that the cycle length cannot vary. As a result, it becomes impossible to have a TSP extension duration that exceeds the amount that conflicting phases can be reduced (Skabardonis and Eleni, 2011). This can place a severe limitation on the maximum possible TSP extension duration at intersections where conflicting phases include long pedestrian crossings.

The 'compensation-focused' offset correction scenario is based on the objective of avoiding net redistribution of green time. As a result, the time that is added during the green extension is compensated for by reducing the same phase the following cycle.

It is also possible to implement a compensation strategy within a fixed cycle length by executing a "compensation cycle" after the TSP call, with the inverse action from the TSP call itself (Liao and Davis, 2007). In the example above, the TSP phase is reduced by the same amount that it was extended by TSP during the previous cycle. However, this method depends on the TSP phase and conflicting phases each being able to be reduced by the duration of the TSP extension in order to avoid exceeding the cycle length during either the cycle with the TSP extension or during the compensation cycle that follows. In practice, this places a severe constraint on the maximum length of TSP extensions at intersections with long pedestrian crossings or heavy traffic volumes.

The situation following a TSP truncation is illustrated in figure 7. The same offset correction scenarios exist as with TSP extensions, but with offset correction adding time rather than reducing it.

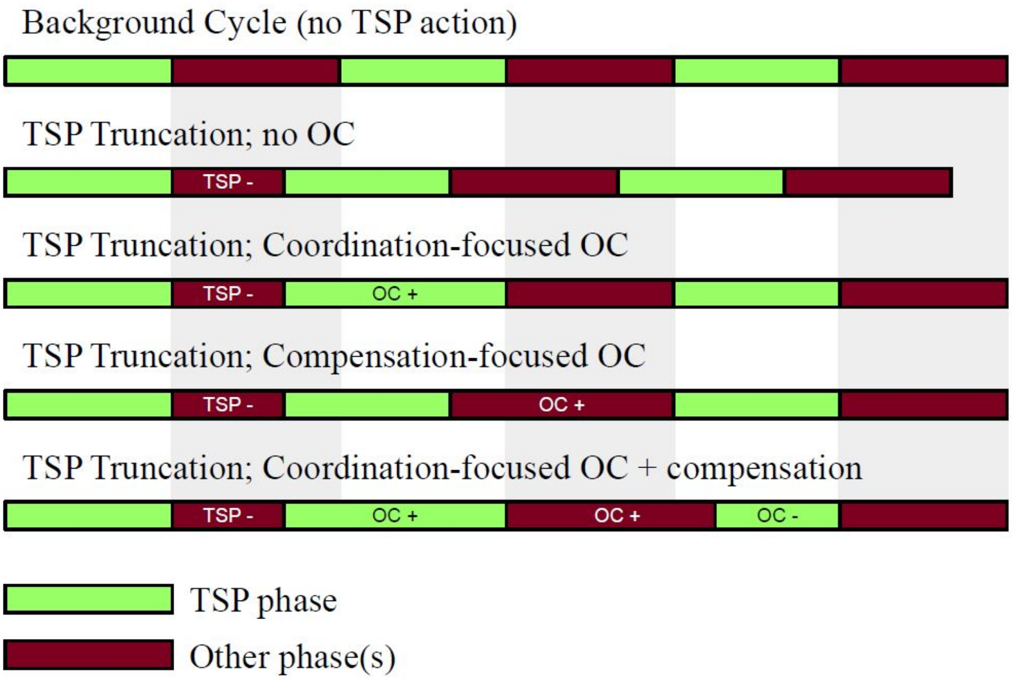


Figure 7: Offset correction strategies for red truncation

If the conflicting movements to the TSP phases are actuated or adaptive, those capabilities may be able to partially compensate for the green time reallocation associated with the coordination-focused OC paradigm

(Furth et al., 2010). For actuated signal phases, the green durations vary from cycle to cycle based on the occupancy of vehicle detectors on the approach. If a TSP action results in a longer queue for a movement, it will naturally get a longer green the following cycle, as long as there is sufficient time available within phase's maximum green duration. However, if the maximum green duration is only slightly higher than the typical green (as is usually the case), there may not be enough green time available to compensate for the TSP impact (Furth et al., 2010).

Conversely, with compensation-focused OC, a conflicting actuated phase may not necessarily use the additional time made available to it after a TSP action, if the length of the queue does not require it. In that case, the time can be removed from the cycle to allow the signal to return in sync sooner than expected.

In practice, the 'coordination-focused' and 'compensation-focused' strategies are not binary alternatives. There is also a full spectrum of possible offset correction designs ranging between those two extremes.

2.2.3 TSP Request strategies

Since TSP actions can have adverse effects on other traffic, imposing conditions for TSP requests to be granted can be beneficial for overall person-delay (Furth and Muller, 2000). For example, buses which are ahead of schedule will be expected to slow down and/or hold at bus stops in order to return to schedule. Delays that an early bus experiences at traffic signals therefore have no effect on that bus' travel time, since the bus needed to be delayed regardless. Conversely, traffic signal delays experienced by late buses can have compounding effects on that bus, since increased schedule delay results in greater passenger volumes waiting at stops, and thus longer dwell times, and a greater chance of overcrowding (Wahlstedt, 2011).

There are three primary strategies used to determine the degree of prioritisation that each public transport vehicle should receive: unconditional, conditional or adaptive.

Unconditional Requests
In an unconditional TSP system, each vehicle requests the same level of priority, regardless of its status. This strategy was historically used in systems which lacked integration between the public transport control system and the traffic signal system, and signal controllers therefore lacked information on the bus' status. TSP systems may also operate unconditionally in situations where there is no reason to differentiate between different public transport vehicles. For example at simple signalised intersection where a busway intersects with a roadway (similar to a railway crossing), all buses need to request a dedicated transit phase in order to cross the intersection, and if the intersecting street does not have any nearby traffic signals, there is not necessarily any advantage to general traffic of serving that phase at one moment rather than another.

Conditional Requests
With a conditional request system, TSP requests are categorized based on certain condition. Typically the condition will be schedule adherence, with requests categorized into either two levels (early or late), or three levels (early, on time, or late). Each intersection will have a given level of priority associated with each level of schedule adherence (Furth and Muller, 2000). For example, in a two-level system, requests from early buses may be ignored, while requests from late buses get accepted. Meanwhile in a three-level system, different punctuality levels may be associated with different levels of priority, with a greater range of permitted TSP actions and/or more generous parameters (e.g. longer maximum green extensions) available for late buses than on-time buses.

Adaptive Requests
Adaptive TSP requests determine the level of prioritisation based on the current traffic conditions. The amount of time saved by the public transport vehicle is balanced against the amount of delay incurred on other road users.

Adaptive TSP systems are not to be confused with adaptive traffic control systems. Some adaptive traffic control systems, such as SCATS, use non-adaptive TSP requests (Slavin et al., 2013).

This chapter will examine the literature to identify the current state-of-the art for TSP systems in coordinated signal networks, as well as components of TSP operation which require further research.

2.3 Types of TSP Systems in Literature

For the purposes of this research, the types of TSP systems used in literature will be classified in terms of their cycle length constraints, offset correction strategies, and use of real-time optimisation.

Papers were retrieved using a two-stage process. First a systemic keyword-based search was conducted in two online research databases: Scopus and Google Scholar. Then, using the "snowballing" method, the most relevant-sounding papers cited in those papers were also retrieved. The snowballing method refers to also seeking out the relevant papers cited in the already-retrieved papers.

The keywords were selected to identify studies related to coordinated transit signal prioritisation systems, including a variety of different terms. For example:

> ("priority" OR "preemption") AND (“traffic signal” OR "intersection" OR "junction") AND ("co-ordinated" OR "coordination")

To filter out results which are related to do with alternate definitions for the search terms (such as the circulatory "network" of the human body), results were limited to the field of Civil Engineering. The publishing date was limited to the last ten years (2011-2021), so that the retrieved research represents the current state of the practice. Important papers from more than ten years ago can be found during the ’snowballing’ step of the literature review, since they are likely to be cited within the more recent papers.

Table 4: Characteristics of TSP systems studied in selected papers

Author	**Year**	**Cycle Length**	**Offset Correction**	**Optimisation**
Chou & Nichols	2016	Preempt	Transition	no
Ekeila et al	2009	Fixed	Coordination	no
Furth & Muller	2000	Fixed	not stated	no
Furth et al	2010	Fixed	Coordination	no
Hounsell & Shrestha	2012	Fixed	not stated	no
Hu et al	2016	Fixed	Coordination	Yes
Huang et al	2012	Flexible	Transition	no
Laskaris et al	2020	Uncoordinated	N/A	Yes
Li et al	2011	Fixed	3 cycle compensation	Yes
Li et al	2020	Fixed	Coordination	Yes
Liao et al	2007	Flexible	1 cycle compensation	no
Lin et al	2013	Fixed	Coordination	Yes
Liu et al	2021	Fixed	compensation	
Mei et al	2019	Fixed	Coordination	no
Mladenoviç & Abas	2011	Preempt	Transition	no
Obenburger & Collura	2007	Preempt	Transition	no
Penabaena-Niebles et al	2014	N/A	Transition	N/A
Seredynski et al	2015	Not stated	Not stated	no
Seredynski et al	2020	Fixed	Coordination	
Shalaby et al	2006	various	various	no
Skabardonis & Christofa	2011	Fixed	Coordination	no
Slavin et al	2013	not stated	not stated	Yes
Truong et al	2019	Fixed	Coordination	Yes
Wahlstedt et al	2011	Flexible	partial compensation	no
Wolput et al	2016	Not stated	Not stated	N/A
Zeng et al	2021	Flexible	Not stated	
Zlatkovic et al	2011	Flexible	Transition	no
Zlatkovic et al	2015	Flexible	Not stated	no

2.3.1 Use of Cycle Length Constraints

The public transport prioritisation systems for coordination networks can roughly be classified into three categories in terms of the ways TSP actions are allowed to modify the normal signal cycle: fixed-cycle TSP systems, flexible-cycle TSP systems, and preemption systems. In fixed-cycle TSP systems, the cycle length is not permitted to vary during TSP requests and offset correction. As a result, the TSP actions tend to be constrained to relatively small modifications of the normal cycle. Flexible-cycle TSP systems allow the cycle length to vary during TSP requests, and then initiate an offset correction process which may last more than a cycle. As such they are able to make more significant modifications to the normal signal cycle, within certain constraints such as a maximum duration for green extensions or a certain green duration that must be served before a phase can be truncated. Pre-emption systems also allow the cycle length to vary during and after priority interventions, but there are little to no measures taken to limit priority. For example, phases are generally able to be truncated to their minimum durations regardless of traffic conditions, and green lights can be held virtually indefinitely until the vehicle passes through (Obenberger and Collura, 2007). The cycle during a TSP intervention may end up being completely different than the normal cycle, including different phases, or a different phase order. Following the intervention, the signal could be anywhere in its cycle relative to the normal in-sync cycle, and systems typically use a "transition" offset correction strategy to return in sync. This is the same strategy that is used during transitions between time-of-day plans.

Fixed-cycle TSP systems
The coordinated TSP systems studied in literature are most commonly "fixed-cycle" systems, where the cycle length is held constant at all times including during priority interventions. Some papers take a fixed cycle length as a given, stating that TSP systems in coordinated networks must always maintain a constant cycle length in order to maintain coordination (Lin et al., 2013, Skabardonis and Eleni, 2011, Truong et al., 2019). But this assumption is disproven by the flexible-cycle TSP systems in studies such as Liao and Davis (2007) which maintain good progression despite allowing the cycle length to vary during TSP interventions.

Since the constrained cycle length heavily limits the range of actions that the signal can take during TSP interventions (Mei et al., 2019, Zeng et al., 2021), there is relatively little room for variation between systems in terms of the actions taken during a TSP intervention, or in terms of offset correction. Instead, much of the research using fixed-cycle coordination systems is evaluating a different aspect of the TSP system, such as the way in which requests are generated, or the way in which conflicting requests are managed. These studies typically use a fixed-cycle system without commenting on the rationale for doing so.

One study which does make an active choice to use a fixed-cycle system is that by Mei et al. (2019), whose primary purpose was to study the effects of TSP extensions and truncations on traffic operations at a major intersection in China, as well as the effects of vehicle actuation parameters such as passage time on the effectiveness of TSP. The study selected a fixed-cycle system on the basis that it is easier to implement in the controller than a flexible-cycle system.

Most of the papers using fixed-cycle systems employ the 'coordination-focused' offset correction described in subsubsection 2.2.2 when recovering from TSP extensions or truncations. Extra time given to the TSP phase during a green extension is reduced from conflicting phases, and with truncation the time lost from the truncated phase is given to the TSP phase. The use coordination-focused offset correction is rarely described as an active choice, in fact many studies describe it as inevitable that TSP systems modify the green time distribution in favour of the TSP phase (Mei et al., 2019, Skabardonis and Eleni, 2011).

The study by Furth et al. (2010) proposes a passive form of compensation: increasing the maximum green time in the background timing plans for phases which conflict with the TSP phases, and using very tight detector settings to ensure that this time is only served when required to clear the waiting queue, and that to ensure that the green light ends as soon as possible relative to the last car in queue.

Two studies did use fixed-cycle systems which actively compensate for green time redistribution during TSP actions. The system proposed by Li et al. (2011) includes up to three cycles of compensation following a TSP intervention, where green times are adjusted in favour of the movements which were negatively impacted by the TSP action. And the system proposed by Liu et al. (2021) makes use of the coordinated phase (and TSP phase) green time which is not part of a green wave to provide additional time to conflicting movements during the cycle following a TSP intervention.

In addition to systems which compensate for redistribution, there are also systems which aim to avoid redistribution in the first place. The system proposed by Ekeila et al. (2009) includes an unusual TSP action named "cycle extension", which uniformly increases all phase lengths by 50%, such that two signal cycles are executed in the time normally occupied by three cycles. This action moves the TSP phase to a different time than when it normally occurs, which can be helpful in cases where the prioritised vehicle is expected to arrive at that time. But in order to shift the TSP phase sufficiently, the action needs to be initiated long before the vehicle arrives at the intersection. The system proposed by Ekeila et al. allows the system to start selecting and/or executing TSP actions further in advance than traditional check-in/check-out detection by using continuous communication between the public transport vehicle and the intersection, whereby the signal controller continuously updates an estimate of the vehicle's arrival time at the intersection.

Since the "cycle extension" action takes place over multiple cycles, the proposed system is technically not strictly fixed-cycle. But the remainder of the TSP actions are restricted to operating within a single cycle, and use coordination-focused offset correction, so it is still classified as a fixed-cycle system for the sake of this literature review. In fact, the potential to incorporate compensation-based offset correction is noted as a potential area for future research.

Hu et al. (2016) build upon the work by Ekeila et al. by proposing a TSP system which also continuously updates an estimate of the vehicle's arrival, while also adding a TSP action they name "green time reallocation" as a less greedy alternative to traditional phase extensions and truncations. This action, illustrated in figure 8 consists of an additional instance of the TSP phase inserted elsewhere in the cycle, for which the time is taken from the normal occurrence of that phase. The action is thus able to avoid redistribution of green time, except for the additional amber and all-red clearance time which is lost from the TSP phase's green time.

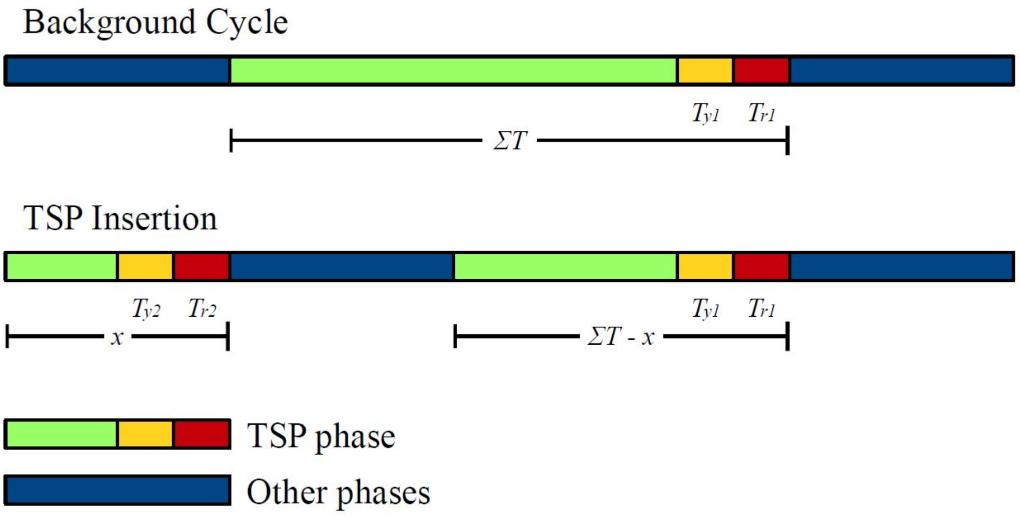


Figure 8: "Green time reallocation" as proposed by Hu et al.

The limitation of the "green time reallocation" action is that the minimum green time requirements of phases in the normal signal cycle would restrict the portions of the cycle where an additional phase could be inserted, particularly when the minimum phase times comprise a large proportion of the maximum cycle length. These constraints were avoided during the Hu et al. study by using a 160-second cycle length, and by omitting pedestrian phases from the model even though in reality the studied intersection includes long pedestrian crossings on all legs.

The remainder of the fixed-cycle TSP systems in the papers do not include any form of green time compensation. As a result, those researchers tend to put restrictions the TSP actions in order to maintain an acceptable level of service for conflicting movements. The system proposed by Lin et al. (2013) constantly monitors the side street volume to capacity ratio and adjusts the maximum allowable main-street TSP extension accordingly. Other proposed systems do not actively manage the impacts on conflicting movements, and the researchers instead suggest that the background splits be adjusted to account for the anticipated amount of green time redistribution as a result of TSP actions (Mei et al., 2019, Skabardonis and Eleni, 2011). Truong et al. (2019) found that the side-street capacity impacts of the main-street TSP calls were so severe in some scenarios that the TSP system resulted in a net increase in network person-delay. In other words, the amount of extra delay incurred to motorists and pedestrians on the side street exceeded the time saved by bus passengers.

Preemption Systems

The objective of preemption systems is to provide a green light as soon as possible to the prioritised vehicle regardless of the impacts on other traffic (e.g. preemption by emergency vehicles), or to guarantee a green light for the prioritised vehicle (e.g. preemption by trains). Whereas conventional coordinated TSP systems place considerable constraints on prioritisation in order to reduce impacts on coordination and on general traffic, preemption-based systems place minimal constraints on priority actions (Zlatkovic et al., 2011). In these systems, priority is implemented using two independent processes. First the preemption event is implemented to serve the prioritised vehicle. In many cases, the signal abandons its normal signal plan altogether and runs a separate signal plan for use during preemption, or rests indefinitely in a particular set of signal phases. Once the vehicle has cleared the intersection, the signal controller corrects its offset using the intersection's set "transition" strategy. The *transition* offset correction mechanism determines where time should add or be reduced from a signal plan solely on the basis of the signal's current offset relative to the desired offset. This is the same mechanism which is used to transition between different time-of-day plans (Obenberger and Collura, 2007).

Peñabaena-Niebles et al. (2014) conducted a literature of other studies on signal timing transition methods, to categorise them and identify common findings. They found, among other conclusions, that the bulk of research has been based on evaluating the standard signal transition strategies available in signal controllers, and relatively little research on developing new strategies which optimise phase durations to minimise delays.

Obenberger and Collura (2007) conducted a microsimulation study to evaluate the effects of the different signal transition (offset correction) strategies available in North American traffic signal controllers. The tests included preemption calls arriving randomly at different points throughout the cycle, thus evaluating the strategies' performance with different initial deviations from the desired offset. They found that several factors influenced the best strategy, namely the frequency and direction of preemption calls, the capacity of storage lanes, the level of congestion, the cycle length, the presence of pedestrian phases and the distance between signals.

The details of transition strategies vary from one type of signal controller to another, but in general they include combinations of the "dwell", "long" and "short" strategies. In the "dwell" transition strategy, the signal rests in a particular signal stage until the point in time where that stage would normally end under coordinated operation. The "long" strategy extends all phases uniformly relative to their normal phase times, and the "short" strategy reduces phases uniformly relative to their normal phase times. North American signal controllers also typically offer mechanisms which choose between two of the basic strategies based on whichever will return the signal in sync more quickly given the current offset and the transition parameters. For example, Chou and Nichols (2016) found that the quickest transition strategy available in the Econolite ASC/3 controller was the "smooth" method, which chooses between "short" and "long".

Flexible-cycle TSP systems

Of the three categories of priority systems, flexible-cycle TSP systems have the least amount of study in literature.

The TSP system proposed by Liao and Davis (2007) executes TSP actions similarly to typical fixed-cycle systems, but allows the cycle length to vary slightly during TSP actions. It then uses a very straightforward offset correction mechanism which corrects TSP extensions and truncations using the inverse action on the same phase during the following cycle. During that offset correction cycle, all TSP requests are blocked, to ensure that the signal returns in sync.

Wahlstedt (2011) conducted a microsimulation study of the PRIBUSS system - the standard TSP system in Sweden - for a small coordinated signal network in the centre of Stockholm. PRIBUSS is a flexible-cycle TSP system which includes the option for compensatory timing adjustments after the TSP intervention. It is not stated what parameters PRIBUSS has available to adjust the relative importance of returning in sync quickly versus maintaining a constant green time distribution.

The only paper which specifically focuses on coordination disruptions and offset correction from a TSP system is that by Huang et al. (2012), which evaluates the the Main Traffic Signal System (MTSS) which was used in Toronto from the 1990's until the mid 2010's. Its TSP system actually falls somewhere between a typical flexible-cycle TSP system and a "preemption" system. The actions themselves are modifications of the normal signal pattern, but after the intervention is complete, the system corrects the offset using the same transition strategies used during pattern changes, which do not consider which movements were impacted by the TSP intervention, and may even exacerbate the TSP impacts. For example, figure 9 illustrates the options MTSS would have to recover from a 20-second extension to the main street phase (shown in green), at a minor intersection with a 70-second cycle length. At minor intersections, MTSS only adds or removes time from the main street, since the side street duration is determined entirely by the pedestrian crossing time. In the OC Reduction option, the time is recovered as a 9-second, 9-second and 2-second reduction of the TSP phase, and in the OC Extension it is recovered by a 27-second and a 23-second extension of the TSP phase. In this case, the system would choose the OC Extension option, even though it exacerbates the impact on the side street and takes just as long to return in sync.

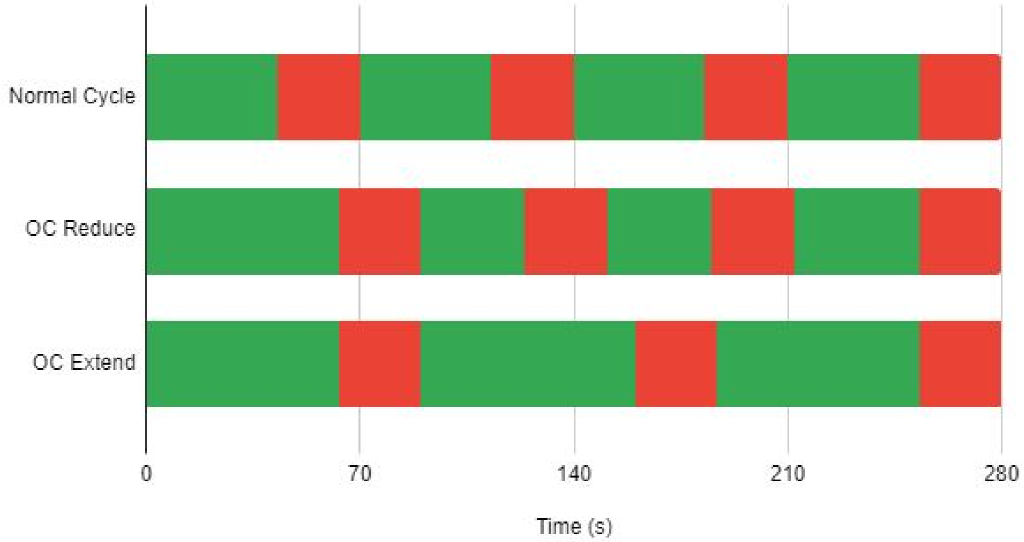


Figure 9: MTSS' options to recover from a 20-second TSP extension on the main street at a minor signal with a 70-second cycle length. In this case, MTSS would choose "OC Extend".

Like Toronto's MTSS, Salt Lake City's light rail TSP system studied by Zlatkovic et al. (2011) falls somewhere between a "TSP" and a pre-emption system. As in a pre-emption system, TSP actions are not constrained to the normal cycle length, and offset correction is done using an independent transition strategy following the end of the TSP intervention. But like a conventional TSP system, the system does aim to minimise the impacts on signal coordination by using phase rotation rather than phase insertion when possible.

Zeng et al. (2021) propose a TSP system which receives continuous position and speed updates from public transport vehicles and projects their arrival times at downstream intersections. It optimises TSP actions either locally or across the entire signal group in order to minimise public transport vehicle travel time, while penalizing deviations from coordination and green times. Although the system was initially designed as a fixed-cycle system, the cycle length constraint was removed because it prevented the system from achieving the desired public transport performance. Following a TSP intervention, the optimisation system will naturally return the signal into coordination due to the penalties in the objective function for deviations from the normal start and end of green times for the coordinated phases. It is not stated how heavily the function weighs disruptions to coordination relative to changes to the green time balance between different phases at the intersections.

Some papers do not specify what form of offset correction is used. These papers presumably use the typical coordinated offset correction method described in subsubsection 2.2.2.

2.3.2 Effects of cycle length constraints

Although there is research on a wide variety of different TSP systems, there is little research on how the different system characteristics compare to each other.

The only paper which specifically compares fixed- and flexible-cycle TSP systems is by Zeng et al. (2021), who switch their optimisation-based system TSP from fixed-cycle to flexible-cycle simply by removing the constraint in the objective function which states that the cycle length must remain constant. Whereas the basic locally-optimised system was able to reduce public transport signal delay by 23% relative to no TSP, simply removing the cycle length constraint improved the reduction to 37% relative to no TSP. For the system which simultaneously considered multiple intersections, the reduction improved from 27% to 49%.

TSP systems with a fixed cycle length inherently have a more restricted ability to implement TSP actions since not all possible actions will be able to fit within the normal cycle length. For example, a green extension cannot exceed the amount of time by which other conflicting phases can be reduced (Li et al., 2011, Skabardonis and Eleni, 2011). The same constraint also applies to inserted phases: it will only be possible to insert an additional green phase into the cycle if other phase durations can be sufficiently reduced to accommodate the phase. This limitation is evident in the difference in public transport delay reduction between fixed-cycle and flexible-cycle TSP systems. The highest-performing fixed-cycle systems in the selected papers reduced public transport signal delay by around 27% (Zeng et al., 2021) and 51% (Hu et al., 2016). In contrast, flexible-cycle systems reduced signal delay by as much as 71% (Zlatkovic et al., 2011) and 77% (Zeng et al., 2021).

The strict cycle length constraint also limits the signal's ability to compensate movements which were negatively affected by TSP truncations or extensions. Because the TSP action and offset correction must be completed within the same cycle, the time added during a green extension must be reduced from another phase rather than the same phase during a subsequent cycle (Mei et al., 2019). This results in a net redistribution of green time in favour of the TSP-prioritised phase, at the expense of conflicting phases (Skabardonis and Eleni, 2011). Nearly all papers studying fixed-cycle TSP systems found that the primary negative impact of the TSP system was the capacity reduction for phases conflicting with the TSP phase (Furth et al., 2010, Lin et al., 2013, Mei et al., 2019, Skabardonis and Eleni, 2011, Truong et al., 2019).

Preemption-based systems are not constrained by the cycle length during priority interventions, and are thus able to provide a very high level of service to the prioritised vehicle (Mladenovic and Abbas, 2011). They are also not constrained in terms of the phases which receive adjustments during offset correction and are thus not required to redistribute time in favour of the prioritised phase. However, because the priority intervention ("preemption event") and subsequent offset correction ("transition") are treated as independent processes, the priority system has little ability to use the recovery process to actively compensate for the impacts of the priority action itself. The signal timing adjustments implemented during offset correction are determined solely based on the signal's position in the cycle at the end of the preemption, relative to its intended position during normal in-sync operation (Obenberger and Collura, 2007). Using the "short" transition strategy, any extra time allocated to the cycle length during preemption will be reclaimed evenly from all phases, including the TSP phase. This results in less redistribution than the fixed-cycle coordinated-focused offset correction method, which reclaims time only from phases which conflict with the TSP phase, but does not actively compensate for it. In some cases, the transition strategy can be selected such that it partially compensates for anticipated impacts. For example, the transition process following a railway preemption event can be set up dwell in the phase(s) which were obstructed by the railway crossing until the signal returns in sync. But since the extra time allocated to those phases is dictated solely by the signal's deviation relative to the coordinated cycle, the amount of compensation provided by the transition process does not necessarily correlate with the amount of time which was lost during the preemption event.

The lack of active compensation during the offset correction process for preemption-based TSP systems is because they use the same transition process used for pattern changes and emergency vehicle preemption. In the case of a pattern change, the offset correction is starting from a steady-state situation, so there is no need for any compensation. Emergency vehicle preemption events may well redistribute time between phases, but in the case of emergency vehicle preemption the extra time received by a prioritised phase was not necessarily usable by general traffic due to vehicles pulling over to let the emergency vehicle past.

Flexible-cycle systems can have a range of different effects on other road users depending on the way in which offset correction is configured. The study of PRIBUSS by Wahlstedt (2011) found that the system increased delays for general traffic in the same direction as the prioritized bus route. This is in contrast to fixed-cycle systems where the traffic performance improves along the TSP-prioritised street, because the TSP phase receives all of its normal green time as well as extra time taken from conflicting movements. Instead, PRIBUSS maintains roughly the same amount of green time for the prioritized direction, but due to the TSP action, that green time may occur at a less optimal time for general traffic.

The Toronto and Salt Lake City systems studied by Huang et al. and Zlatkovic et al. use similar transition strategies to preemption systems. The Toronto system is found to considerably disrupt traffic signal progression, partly due to its tendency to use the "long" transition strategy which greatly increases cycle lengths and shifts the offset far from the normal offset during offset correction (Huang et al., 2012). The Salt Lake City system, on the other hand, has relatively minor impacts to coordination, which may be due to the efforts made to minimise the amount of cycle length variation during TSP actions (Zlatkovic et al., 2011).

2.3.3 Use of optimisation

In conventional non-adaptive traffic control systems, the signal's behaviour is dictated by relatively simple algorithms and decision rules, such as the typical detector-based phase extensions and predetermined maximum green times. The TSP systems for such non-adaptive control systems also tend to operate in a similar manner, with a predetermined maximum durations for TSP actions, and straightforward conditions for priority such as schedule adherence. In contrast, the behaviour of adaptive control systems is determined by optimising an objective function which includes current traffic data. Similar to adaptive control systems, many novel TSP systems proposed in literature conduct some sort of optimisation within the controller.

The TSP systems studied in the selected literature are classified in table 4 into two categories: those with optimisation as part of their operation, and those without.

Optimisation-based TSP systems are often paired with adaptive signal control systems, but this is not necessarily the case. The system proposed by Zeng et al. (2021) operates with a static background timing plan and does not require any real-time traffic data. The "cost" of TSP actions is determined in the objective function by the amount of deviation from the normal background timing plan, which is assumed to be optimal. The "benefit" of the TSP actions is reduction in estimated travel time for the bus, which takes into account arrival times at downstream intersections. This allows the system to identify situations in which a TSP action would have little overall benefit since the bus would get held up at the downstream signal regardless, as well as situations where a TSP action would have particularly high benefits due to arriving at the downstream intersection at a more ideal moment.

Some researchers even use optimisation-based systems to inform relatively straightforward TSP behaviours. The system proposed by Lin et al. (2013) has a complicated objective function simply to determine whether or not a particular bus should be granted a 10-second green extension within a fixed-time signal system.

The advantage of optimisation-based systems is that they can generally achieve better performance (i.e. lower person-delay) than systems which simply operate on the basis of set parameters (Li et al., 2011, Zeng et al., 2021). However, the disadvantage is that their increased complexity may increase maintenance costs and limit practitioners' ability to troubleshoot issues. Shalaby et al. (2006) conducted a TSP design exercise with traffic engineering practitioners in Ontario, Canada, and found that regardless of the potential efficiency advantages of more complex TSP systems, practitioners favoured simpler systems, particularly for small municipalities where maintenance costs and staff expertise are more of a limitation. Wahlstedt (2011) similarly noted that adaptive control systems are often less preferred by road authorities than systems which allow engineering practitioners to more directly influence the traffic signals' behaviour.

As part of the development of their proposed TSP system, Zeng et al. (2021) used a reactionary method whereby they adjusted the constraints, penalties and weights in the objective function in response to undesired signal behaviours. This method is similar to how traffic engineering practitioners respond to issues identified with traffic signals, but it requires access to the objective function as well as a thorough understanding of the system.

2.3.4 Traffic performance impacts

Skabardonis and Eleni (2011) propose an analytical method to quantify the capacity impacts of the green time redistribution resulting from fixed-cycle TSP systems. The method is based the Highway Capacity Manual calculations, but with average green times adjusted to reflect the redistribution of time which would occur given the intersection's TSP parameters (maximum extension and truncation durations) and the frequency of TSP requests.

Although their proposed method is easily understood and applied, it fails to account for some characteristics of a coordinated signal network. The method assumes that buses arrive randomly throughout the cycle length even though in a coordinated network, buses would tend to arrive at signals in a particular part of the cycle corresponding to the green phase at the upstream intersection.

The method also fails to account for any TSP impacts other than the green time redistribution that occurs with fixed-cycle TSP systems. Given the assumptions of the proposed model, a TSP system with a fully compensation-based TSP system would produce no impacts, since it does not result in any net reallocation of green time. But in reality, compensatory TSP systems do still have impacts on capacity and delay, as found by the study by Wahlstedt on the PRIBUSS system (Wahlstedt, 2011). That study theorises that the impacts may be due to the green times occurring at less optimal times for general traffic, but it is not clear how much of the delay is due to coordination impacts and how much is due to the reduced regularity of green and red durations.

The optimisation-based TSP system proposed by Zeng et al. (2021) assumes that a TSP action's impact on coordination is proportional to the amount that it deviates the start and end of the coordinated phase green away from the normal timing plan. But there does not appear to be other literature supporting that assumption. Most other studies investigating coordination impacts simply apply the system in a microsimulation environment and observe the changes to metrics such as the number of stops and average vehicle delay (Huang et al., 2012, Wahlstedt, 2011).

The delay experienced by pedestrians is relatively straightforward, because there are no coordination impacts between intersections, and minimal queuing effects or capacity constraints. Li et al. (2020) use a linear model to estimate pedestrian delay given the phase durations provided by the TSP system, assuming a constant arrival rate of pedestrians.

2.3.5 Public transport performance impacts

The basic effect of TSP on public transport lines is to reduce the delay that public transport vehicles face at signalised intersections, as well as the variability in that delay.

As noted in subsubsection 2.3.2, TSP systems commonly reduce signal delay by 20% to 80% for the public transport line at a given intersection, compared to the same signal operating without TSP. But the effect of TSP on public transport performance is more than just a uniform improvement in travel time.

Traffic signal delay represents a source of variation in public transport travel time, so reducing it can also improve service reliability. Reliability is also considerably improved by the use of conditional TSP requests, which vary the level of priority depending on the vehicle's punctuality (Hounsell and Shrestha, 2012).

The study by Zeng et al. (2021) measured service reliability as a metric of their proposed TSP system, but the primary focus of the study was on travel time, not reliability. As a result, the study used an extremely aggressive schedule where only 4.4% of buses were on time (which they defined as being exactly on schedule, or early) in the base case without TSP. Their proposed local intersection-based TSP system was able to improve schedule reliability to 36.8% of buses "on time", and the network-based TSP system improved reliability to 98.4%. The study also assumed a uniform random distribution for dwell times within 10 seconds of the average, which fails to account for the correlations which exist in practice between headway and dwell times for frequent services such as the one studied.

In reality, when a particular vehicle on a frequent public transport line has a longer-than-average headway, it will encounter larger than average volumes of passengers waiting at stops. The larger volumes of boarding and alighting passengers result in longer dwell times, which increases the travel time and thereby increases the headway even further (Hounsell and Shrestha, 2012). The end result of the feedback loop is that service experiences "bunching", where vehicles are grouped together rather than evenly spaced. Some researchers take this feedback loop into account in their models, such as the study by Liao and Davis (2007) which assumes a deterministic dwell time based on the bus headway (H), average passenger arrival rate at the particular stop (λ), and average boarding time per passenger ($T_{boarding}$):

$$T_{dwell} = \lambda * (H) * (T_{boarding})$$

Including the correlation between dwell time and results in a more respresentative picture of the reliability effects of TSP on the public transport line, and thus the in-vehicle travel times.

Hounsell and Shrestha (2012) proposed an analytic method for calculating public transit passenger benefits as a result of TSP which takes into account not only the in-vehicle travel time, but also the wait time. The method, designed for high-frequency services, assumes a constant arrival rate of passengers at bus stops to calculate the average passenger wait time for each bus. The arrival rate at each stop is determined based on the observed passenger volumes at that stop. When bus service is more evenly-spaced, the average wait time is lower than when the same number of buses arrive at irregular intervals.

Based on a simple numerical model which using their proposed wait time metric, they conclude that the most effective way to reduce wait times is to give priority to buses whose headway is more than the headway of the bus immediately behind them. This is in contrast to the typical practice of giving priority to buses behind schedule, or buses with an above-average headway.

2.3.6 TSP Efficiency

To assess the performance of a TSP system overall, it is necessary to weigh its benefits to public transport against the disbenefits to other road users. Any given system will typically be able improve on or the other by increasing or reducing the restrictions on TSP calls. Some studies, such as Liao and Davis (2007), qualitatively observe the changes in bus delay and private motor traffic delay in order to judge the effectiveness of the TSP system. But Zeng et al. (2021) propose a quantitative metric, "Improvement per Impact" (IPI) as a measure of efficiency for TSP systems:

$$IPI = \frac{\%change\ bus\ performance}{\%change\ passenger\ car\ delay}$$

where the *bus performance* can be any metric, such as average travel time or on-time performance, and the *passenger car delay* is an aggregated value for the entire network. The TSP systems evaluated in the study scored IPI values between 9 and 18.

Rather than explicitly comparing the public transport and passenger car performance, many studies consider all road users at once, such as by observing the network person-delay (Hu et al., 2016),

2.4 Research Gaps

The existing literature suggests that flexible-cycle cycle TSP systems can provide a better balance between public transit vehicle performance and delays to other road users than either fixed-cycle TSP systems or pre-emption systems, particularly when combined with an offset correction system that can compensate for impacts on conflicting phases. However, there is relatively little study of such flexible-cycle TSP systems in literature.

Some studies measure the impacts of flexible-cycle TSP systems relative to a baseline without any TSP (Huang et al., 2012, Liao and Davis, 2007, Slavin et al., 2013, Wahlstedt, 2011, Zlatkovic et al., 2011), but the only study which specifically compares a flexible-cycle and fixed-cycle system is that by Zeng et al. (2021). That study documented the significantly improved public transport performance as a result of removing the constraint of cycle length in an optimisation-based TSP system, but the offset correction process was not a focus the study.

There is a comprehensive body of literature examining offset correction as a standalone process (i.e. the transition following a preemption event or change in timing plan), but the lack of integration with the TSP intervention in those systems means that offset correction is unable to systematically compensate for green time redistribution.

The value of intuitive and easy-to-implement control systems should not be underestimated. In practice, road authorities do not necessarily implement the control system which has the best performance, they select a system which provides a good balance between performance, cost and ease of use. The value of simplicity is also evident in the fact that so much research uses fixed-cycle TSP systems even they tend to have inferior overall performance to flexible-cycle systems.

There is lack for research into intuitive and easy-to-implement TSP systems with a flexible cycle length and the option for compensation-based offset correction. In particular there is a need for further study on the trade-offs between side street impacts with coordination-focused offset correction and the coordination impacts of compensation-focused offset correction, particularly since there is a continuous range of possibilities between the two strategies which has not been explored in literature.

3 Study Design

The study is divided into two primary phases. During the controller development phase, a novel TSP system was designed which allows the cycle length to vary within defined parameters during and after TSP interventions. During the evaluation phase, a functional model of the novel system was tested in a microsimulation environment in order to compare it against a conventional TSP system. For each of the phases, there were multiple options available, which are evaluated in this section.

3.1 Algorithm implementation

The implementation of the algorithm consists of developing a functional TSP system using the Dutch CCOL traffic signal programming standard. An existing fixed-cycle TSP system was modified to permit the cycle length to vary according to the logic rules established in the concept development task.

Additional functionality is also added to the controller by enabling the parameters to continuously update based on the observed traffic conditions.

3.2 Modelling

To test the effectiveness of the novel TSP system relative to a conventional fixed-cycle TSP system, and to determine the effects of varied parameters, a VISSIM model was created based of an idealised hypothetical road network.

The newly developed controller and existing controller were alternately be connected to VISSIM for testing using a software-in-the-loop framework.

A set of eight test scenarios was executed, testing different TSP systems, intersection paramaeters and TSP parameters. Three different TSP systems were tested: fixed-cycle (existing), flexible-cycle, and flexible-cycle with real time parameter adaptation.

Because microsimulations include stochasticity based on pseudo-random seeds, each scenario is run ten times, each with a different random seed. The microsimulation studies included in the preliminary literature review typically used between 6 and 12 of runs per scenario.

For each scenario, the performance of the network is evaluated in terms public transport travel times, public transport headway regularity, number of stops for motor vehicles, the queue length in constrained approaches, and the network person-delay.

3.3 Controller Development

There are two primary directions the design of the novel system could take: either the system could be developed as a parametric system or as a self-optimising system. A parametric TSP system would operate based on relatively straightforward decision rules, with the system's behaviour dictated by parameters that are manually selected by the road authority as part of the signal timing plan. Alternatively, a self-optimising system would determine the most appropriate system behaviour based on minimising an objective function in real time using data collected through the signal system.

3.3.1 Parametric system

With a parametric system, the behaviour of the system would be determined primarily by adjustable thresholds, such as:

- Maximum deviation from cycle length
- Focus of offset correction, ranging from coordination-focused to compensation-focused.

- Maximum number of cycles to return in sync following a TSP action

The intent is that with a relatively small number of parameters such as those listed above, a road authority could set up the system in a way which is nearly optimal for the applicable traffic situation. An example of a parametric TSP system is the PRIBUSS system used in Sweden.

The advantage of a parametric system is that it is very intuitive for the practitioners to operate. Given the parameters selected, it would be possible to fairly accurately predict the behaviour of the system. For example, given the programmed maximum number of cycles allowed to return in sync and the permitted phase reductions, it would be possible to determine which TSP action will be executed (if any) for a given TSP request. This intuitive design makes it easy for practitioners to troubleshoot issues, or to adjust the behaviour of the system to reflect the priorities of the road authority Shalaby et al. (2006).

However, the disadvantage of a parametric system is that it does not independently find the optimal behaviour which would minimise person-delay. It is up to practitioners to select intersection parameters to best match the traffic situation, which may also require updating the parameters periodically to reflect changes in traffic patterns.

3.3.2 Self-optimising system

With a self-optimising system, decisions are evaluated based on their estimated effect on the objective of minimising network person-delay. In order to do this, the system maintains a model of the traffic situation which is used to evaluate the effects of possible actions. If a self-optimising system is selected, its primary objective would be:

Minimise network person-delay, while meeting constraints such as:

- do not exceed capacity
- do not cause spillback to adjacent intersections
- do not increase cycle length above the road authority's guidelines (e.g. 144 s)
- do not increase individual phase delay by more than a certain percentage.

An example of a self-optimising TSP system is the TSP system included in the SPOT-Utopia adaptive network control system (Wahlstedt, 2011).

3.3.3 Comparison of system concept options

The real-time responsiveness of a self-optimising system allows it to have superior performance to a parametric system whenever the actual traffic conditions deviate from the traffic conditions assumed while designing the signal plan (City of Toronto, 2015), but self-optimising systems also tend to be more technically complex and challenging for practitioners to troubleshoot. Given that road authorities tend to implement non-adaptive control systems in most cases despite the superior performance of adaptive systems (Shalaby et al., 2006), the simplicity and intuitiveness of a signal system must also be an important consideration.

To combine the benefits of a parametric system with those of a self-optimising system, this study proposes to develop a responsive-parameter system. This would be a parametric system where the parameters themselves can be set to "automatic", and the system itself would determine those parameters based on current traffic conditions. This allows some level of real-time traffic response to be implemented in the signal control while still maintaining simple and intuitive decision logic. In addition to the responsive-parameter system, a static-parameter (parametric) variant was created to maximise comparability to the static-parameter TSP system in the existing fixed-cycle controller which is serving as the base case.

Compared to a fully self-optimising system, a responsive-parameter system optimises at a higher level of abstraction. Whereas a fully self-optimising system may evaluate the person delay impacts of executing a

TSP action, a responsive-parameter system would simply act according to the decision rules established by the parameters. It is those parameters which themselves would be continuously updated to values which reflect the currently observed traffic situation.

The amount of real-time traffic information available to the traffic signal controller will vary considerably between different jurisdictions and between different intersections. Many signalised intersections in North American city centres operate completely fixed-time, and thus have no vehicle detectors at all. A responsive-parameter system can isolate actively-optimising components of the TSP system into separate modules, which allows the level of traffic responsiveness to be adjusted to reflect the amount of real-time data available. In contrast, a fully self-optimising TSP system would be unable to function effectively in a signal system with only a few vehicle detectors, because it depends on minimising an objective function for the entire intersection.

3.4 Comparison of study options

Once a novel flexible-cycle TSP system was created, the study then needed to test its performance. There are three primary options for doing so: a macroscopic model, a microscopic model, or an on-street test.

With a microsimulation model, the novel system would be tested in a computer model representing a real-world network. This method would be similar to the studies by Mei et. al. (Mei et al., 2019), Obenburger & Collura (Obenberger and Collura, 2007), Wahlstedt (Wahlstedt, 2011), and numerous others.

Using the software-in-the-loop method, the model can use the signal code of a real-world traffic controller without any simplification (Obenberger and Collura, 2007). The novel TSP system is designed to be as applicable as possible to real-world intersections, but there may still be some aspects of the program itself which would need to be tested and/or further developed in order to deploy the newly developed code at real-world intersections.

Although the microsimulation option is less flexible to test different scenarios than the macrosimulation option, it would still be possible to test several different scenarios within the given road network, such as different TSP systems, different traffic levels and different system parameters.

Given the time and budget constraints of the project, the option of an on-street test is not considered to be practical. The time required for manual correction of the data, and the high likelihood of delays caused by external sources means that there is a significant risk of the project being delayed.

The selected evaluation method is a software-in-the-loop microsimulation model, because its limited ability to study different environments is less severe than the potential inaccuracies resulting from a macrosimulation's simplification of the signal control system.

It is the researcher's opinion that it is more beneficial to determine with a high degree of confidence the efficacy of the novel TSP system in a particular environment, than to determine at a low degree of confidence the efficacy of the system in a variety of environments.

3.5 Evaluation metrics

This section outlines the metrics which are used to measure the performance of different TSP systems and setups, the TSP systems and setups which are tested, and the characteristics of the network which are modeled. The metrics which are observed are listed in table 5.

Table 5: Performance metrics

Metric type	Indicator
Bus Delay	average delay approaching central intersection
Bus Reliability	headway regularity downstream of intersection,
Signal Coordination	percentage of signal cycles in sync
Vehicle Queues	frequency of queue exceeding storage
Number of Stops	average stops per vehicle in coordinated directions
Vehicle Delay (non-bus)	average vehicle delay
TSP Efficiency	network average person-delay per approach

To determine whether differences in the metrics between scenarios are statistically significant with a 95% confidence interval, a paired T-test is used. As described in subsection 5.1, the randomness in each model run is determined by a random seed, and the same seed number is used for a given run number in all scenarios. The results for each model run in one scenario are therefore paired with the corresponding run in the other scenario being compared, with very little random variation between them.

3.5.1 Bus Delay

The public transport performance is primarily be measured in terms of the vehicles' travel time approaching the subject intersection. The number of stops the vehicles experience is not a metric because speed and dwell time advice systems such as those proposed by Seredynski et al. (2015, 2020) and Laskaris et al. (2020) can be used to avoid stopping between bus stops, regardless of the effectiveness of signal priority. However such systems do not improve the public transport travel time, so travel time savings continue to be a key metric for TSP systems.

The metric for bus delay is the average delay for buses approaching the central intersection. This consists of the difference between the free-flow travel time and the observed average travel time.

3.5.2 Bus Reliability

With improved TSP performance for late buses, the service regularity of the bus line should improve, which in turn should reduce the wait times for passengers along the route. As a metric for this improvement, the average wait time at a stop immediately downstream of the intersection is calculated, assuming a constant rate of passengers arriving at the stop. The average wait time is calculated using Equation 3.1, which is derived from figure 10 and has been previously described by Hounsell and Shrestha (2012).

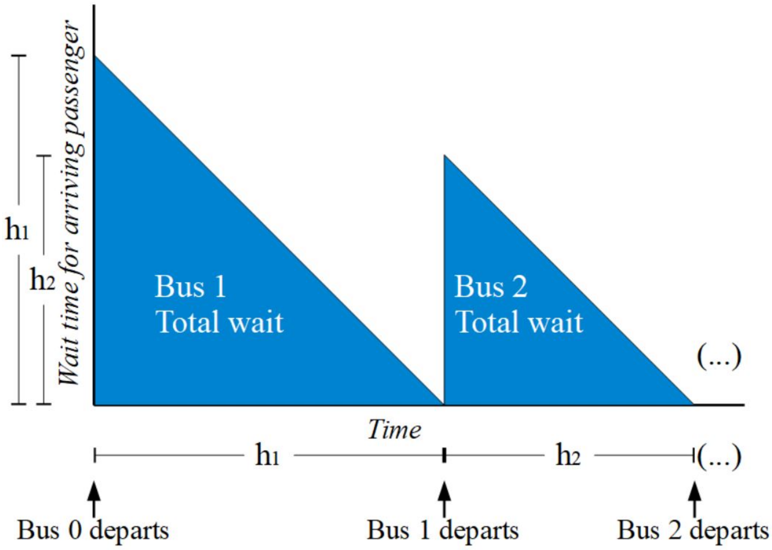


Figure 10: Passenger waiting times over time

The Y axis represents the waiting time for a passenger arriving at the bus stop at that particular time, and h_b represents the headway before each bus b. With a constant arrival rate, the average wait w_{avg} is simply the area under the curve, divided by duration of measurement (Σh).

$$w_{avg} = \frac{\sum_b \frac{1}{2} h_b^2}{\sum_b h_b} \tag{3.1}$$

where:

h_b = *headway of bus b, in seconds*

3.5.3 Signal coordination

Coordinated signal systems are generally implemented where there is a need or desire to avoid motor vehicle stops between intersections. In addition to the objective of reducing travel times for vehicles in the coordinated directions, there may also be a need to coordinate green times at closely-spaced intersections to prevent queues from exceeding the available storage space. The tested TSP systems' ability to maintain coordination is measured by their ability to maintain the programmed offset, the number of stops for vehicles in the coordinated directions, and the length of queues on intersection approaches with constrained storage space.

The signal's ability to remain in coordination is represented by the proportion of cycles which are "in sync". But the distinction between in-sync and out-of-sync operation is not black-and-white. Even during normal coordinated operations in the absence of any TSP interventions, the start and end times of the coordinated phases can vary within the cycle. The coordinated phases may start early if there happens to be less demand on preceding conflicting phases. The coordinated phases may end early if there is no longer demand in the coordinated direction, and they may end late if there is continued demand, but no demand for the subsequent conflicting phase. The one situation which does not happen in normal coordinated operation is the coordinated phase starting later than planned. In the model, the coordinated phases are on recall, which means that they are called every cycle regardless of demand. As a result, a cycle which is "not in sync" can be identified by one or more coordinated phases starting later than planned.

This metric will fail to capture cycles which are out of sync with a negative offset error, meaning that the signal is ahead of where it would normally be. This is an acceptable limitation because it is very rare that the signal would incur a negative error as a result of a TSP intervention. In nearly all cases, the TSP intervention involves adding time by inserting an additional phase into the cycle, or extending a green phase. The only case in which a TSP intervention would incur a negative offset error is when it includes green truncation and no other TSP actions. Furthermore, negative offset errors can always be recovered in a single cycle, because there is no limitation to the amount of time which can be added to a phase.

3.5.4 Vehicle queues

The primary direct concern related to vehicle queues is that the length of the queue can exceed the available storage space and spill over into other areas, obstructing an adjacent lane and/or an intersection. The key metric will therefore be the maximum queue length in metres for each measurement interval. The measurement intervals is 2.5 minutes, which is chosen to be slightly more than the cycle length (2 minutes). A measurement of 2 minutes would provide a value for each cycle under coordinated operations, but with the variable cycle lengths of the new flexible-cycle TSP systems, a 2-minute measurement period might not include a full cycle and thus might return a deceptively low queue length.

Queues are measured for all movements, but the primary focus is on the directions where the storage length is constrained: the motor vehicle approaches located on the two central links connecting the three intersections.

3.5.5 Number of stops

The average number of stops per vehicle is reported for each movement by Vissim based on the number of instances where a vehicle's speed drops below a given threshold within a given measurement area. Since the objective of measuring the number of stops is to observe how the coordinated movements are performing, stops are measured in both directions along the two roadway links connecting the three signalised intersections, for a total of 4 measurement zones. Each zone begins from the through-lanes stopbar of the upstream intersection and ends at the of the downstream intersection in the through lanes. As a result, vehicles are only measured if they are travelling straight through east-west at both the upstream and downstream intersection, which is the coordinated movement on the link. Vehicles which turn onto the link from the perpendicular street, or turn at the downstream intersection are not expected to be able to travel through non-stop, even when coordination is maintained.

3.5.6 Delay for other traffic

The primary metric for general traffic is the average delay for each movement. As with bus delay, this is simply the difference between the freeflow travel time approaching the stopbar and the observed travel time. This metric captures not only any wait times for a green light (control delay), but also the delay resulting from capacity limitations and increased queue lengths (queuing delay).

Delay is measured for all movements within a measurement zone 150 metres long for turning movements, and 300 metres long for straight-through movements. This zone is selected to be longer than the maximum likely queue. Where the previous intersection is located less than that distance away, the measurement extends from the three stopbars of the upstream intersection from which vehicles could originate. Using a measurement zone length as consistent as possible minimises the chance of capturing different amounts of background delay which simply arises from variations in cruising speed from different drivers.

3.5.7 TSP System Efficiency

The efficiency of a TSP system refers to the ratio between the improvement in performance for buses and the decrease in performance for other modes, where performance is typically defined as a reduction in delay (Zeng et al., 2021).

To evaluate the overall efficiency of the systems tested in the model, the average person-delay for the network is compared for the scenarios. This is calculated using Equation 3.2.

$$personDelay = \frac{\sum_i (avgDel_i * nVeh_i * occRate_i)}{\sum_i (nVeh_i * occRate_i)} \tag{3.2}$$

where:

$avgDel_i$ = *average delay for phase i, in seconds*
$nVeh_i$ = *number of vehicles using phase i*
$occRate_i$ = *occupancy rate: number of people per vehicle using phase i*

The assumed occupancy rates are listed in table 6.

Table 6: Assumed vehicle occupancy rates

Vehicle Type	Assumed Occupancy
Auto	1.4 people
Bicycle	1.0 people
Bus	50 people

The assumed occupancy rates for passenger cars ("auto") of 1.4 people per vehicle is the national average for the Netherlands in 2010, which was the final year the European Environmental Agency compiled passenger occupancy rate data (European Environment Agency, 2010). The bus route in the model is completely fictional, so there are no ridership statistics on which to base a bus occupancy rate. A value of 50 is chosen to represent a nearly-full standard 12-metre public transport bus. For a 2020 Mercedes-Benz Citaro NGT, this would correspond to a full seated load plus 21 passengers standing (Mercedes-Benz Buses, 2020).

The limitation of the above method is that it oversimplifies the effect of vehicle-delay on public transport passengers. For private modes such as cars, a given amount of person-delay for one vehicle is considered to be interchangeable with the same amount of delay for another vehicle of the same type. But this does not hold for public transport vehicles, because public transport lines are actively managed to produce a regular and/or punctual service. Delays to people in an early bus typically have no effect on their end-to-end travel time, because if the bus had not been delayed at the traffic signal, it would have been delayed at a subsequent bus stop in order to depart at the scheduled time or restore a regular headway. Conversely, delays experienced by late buses cause can have effects far exceeding the delay experienced by the passengers on those buses at the subject intersection, because reducing service regularity increases the average waiting time for passengers downstream, and because of the feedback loop between lengthened headways and increased dwell times at stops (Hounsell and Shrestha, 2012).

To capture these effects, the person-delay calculations will ignore delays from early buses, and assume a higher occupancy rate for late buses. It is beyond the scope of this study to determine a weighting for late buses which accurately captures the secondary effects of delays on late buses, but a sensitivity analysis is conducted to illustrate how the performance of the tested TSP systems varies depending on the weighting of delays for late buses.

The three occupancy rate values which are tested are 50, 100, and 150 passengers per bus. A value of 50 represents a full standard bus including 20 people standing, as described earlier. The values of 100 and 150 exceed the average number of passengers that would be realistically be carried in standard 12-metre buses, but increasing the occupancy rate above the number of actual passengers serves as a placeholder to represent the secondary effects which traffic delays have on the speed of late buses downstream of the signal, as well as the associated delays to waiting passengers. If those secondary effects were within the scope of this study, they would increase the relative weight of delays to late buses in the average person-delay calculation, which would produce a similar effect to artificially increasing the bus occupancy rate. Furthermore, in order to encourage people to use more space- and energy-efficient forms of transportation, road authorities may with to prioritise reducing delays to public transport users, which could involve placing a disproportionate weight on bus passenger delays (NACTO, 2016).

3.6 Variables to Test

The fundamental purpose of the study is to determine the relative impacts of cycle length constraints and coordination disruption. Furthermore, an area of interest is the potential for TSP systems to compensate for TSP actions during offset correction. Three variants of TSP systems are tested, and for each TSP system two levels of flexibility to adjust green durations are tested. The Base TSP system is the existing TSP system used by Movensis for CRSV halfstarre networks. The ParFlex and ResFlex systems use the novel system developed as part of this study. In the ParFlex system, the controller's green time compensation mechanism operates the basis of maintaining green time distribution, without taking into account green utilisation rates. The ResFlex system expands the functionality by setting parameters in real time based on the current traffic volumes and green utilisation during priority interventions. Testing two variants of the new system allows the study to differentiate the effects which are due to the flexible cycle length from the effects which are due to the responsive parameter adjustment mechanism that the ResFlex version of the controller also introduces.

The level of flexibility was selected as the parameter to study because a signal with a limited ability to reduce phase durations will also have a limited ability to fit TSP actions within a fixed cycle length. The number of varied parameters has been kept to a minimum, because it is not practical to automate the setup for microsimulation scenarios.

Table 7: Characteristics of the modeled scenarios

Scenario Number	1	2	3	4	5	6
Scenario Name	**Base-80**	**Base-90**	**ParFlex-80**	**ParFlex-90**	**ResFlex-80**	**ResFlex-90**
Controller	Base	Base	New	New	New	New
Adaptation	N/A	N/A	Off	Off	On	On
Min Max Green during OC	80%	90%	80%	90%	80%	90%
Min Green before truncate	80%	90%	80%	90%	80%	90%
Desired Green during OC	N/A	N/A	90%	95%	Adaptive	Adaptive

Note: Percentages are relative to phases' programmed maximum green times

The primary characteristics which are varied between scenarios are the use of flexible cycle lengths during and after TSP requests, the use of real-time traffic volume to determine acceptable TSP actions and green durations during offset correction, and the level of flexibility in the background signal cycle.

Scenarios 1 and 2 use the pre-existing "base" signal controller, which has a fixed cycle length including during and after TSP interventions. The remaining tests use the "new" signal controller developed as part of this study, which allows the cycle length to vary during and after TSP interventions.

Of the scenarios using the new controller, scenarios 5 to 8 use real-time traffic data to estimate the desired green times for each phase, as well as the amount of benefit or disbenefit each phase receives during a TSP intervention. Scenarios 3 and 4 depend on predetermined values for the desired green times during offset correction. These are set at 90% and 95% of the normal maximum green values, respectively.

For the purposes of this study, flexibility refers to the intersection's ability to select different maximum green values for its signal phases. For example, some intersections often have lower flexibility if they have long pedestrian crossing distances. Regardless of the optimal duration of vehicle phases, a certain amount of time needs to be reserved for pedestrians to have sufficient time to cross the street. During cycles where there is no pedestrian demand, the time can be reallocated to other phases if desired. Scenarios 1, 3, 5, and 7 represent intersections with a high degree of flexibility. In those scenarios, the maximum green duration selected by the controller for any given phase is permitted to be as low as 80% of the normal maximum green value. In the remaining "low flexibility" scenarios, the selected maximum green values must always be at least 90% of the normal maximum green value.

3.7 Bus line characteristics

The bus line in the model will operate at a frequency of 12 buses per hour (every 5 minutes) in each direction. This frequency was selected with the intent of being frequent enough that is necessary to carefully manage green time redistribution to avoid excessive traffic impacts, but not so frequent that it is impractical to provide a high level of priority to buses without causing excessive traffic impacts.

3.8 Road Network Characteristics to Model

Using the microsimulation approach, it is not practical to test a wide variety of network and intersection types, so the scope of the study is narrowed to a specific set of network characteristics, which is then be the basis for developing a single fictional road network to be modeled.

Given the time and resource limitations of the project, the study's scope is narrowed to intersections and networks with a specific set of characteristics, which are listed in table 8.

Table 8: Controller characteristics to be studied

Characteristic	**Type**
Coordination	Fixed Cycle Length
Control	Semi-Actuated (halfstarre)
Adaptivity	None
TSP Calls	Check-in / Check-out detectors
TSP Condition	Headway-based

3.8.1 Network control type

Fixed-cycle coordination systems are the focus of the study, because they are the common way of implementing signal coordination in urban networks with closely-spaced signals.

Since the study is taking place in the Netherlands, with access to Dutch control systems, the most common form of vehicle control at Dutch coordinated intersections is used: Dutch "halfstarre" control. This control method is similar in concept to semi-actuated control found in North America, but is implemented in a different way. Whereas North American "semi-actuated" intersections have a fixed end time for the coordinated phases regardless of demand for those phases(FHWA, 2008), Dutch halfstarre systems allow the end time to vary based on vehicle actuation and instead use the start of the coordinated phase as a reference point. This means that Dutch halfstarre intersections tend to include detectors for all movements, including the coordinated phases, whereas North American intersections commonly lack detection for the coordinated phases.

Dutch traffic controllers do not follow the 'ring/barrier' phase control structure which is standard in North America. Each phase's compatibility with other phases is determined solely based on the conflict matrix. During halfstarre operation, the timing and combination of phases is based on a "step raster" which defines the "primary realisation periods" of each phase. An example of a step raster from a Dutch halfstarre intersection is shown in figure 11. Each "step" represents one second of the signal cycle, and the primary realisation periods include only the phases' green times. Yellow and red clearance intervals are accommodated in the gaps between conflicting phases' primary realisation periods.

During a phase's primary realisation period, traffic demand for that phase has precedence over conflicting phases. If there is a call for a given phase, in the seconds preceding the start of that phase's primary realisation period other conflicting phases are terminated regardless of their demand, so that the phase can start at the moment specified in the step raster. If the phase is skipped or ends early due to lack of demand, the time becomes available to other phases. When a phase is served outside of its primary realisation period, it is known as an "alternative realisation".

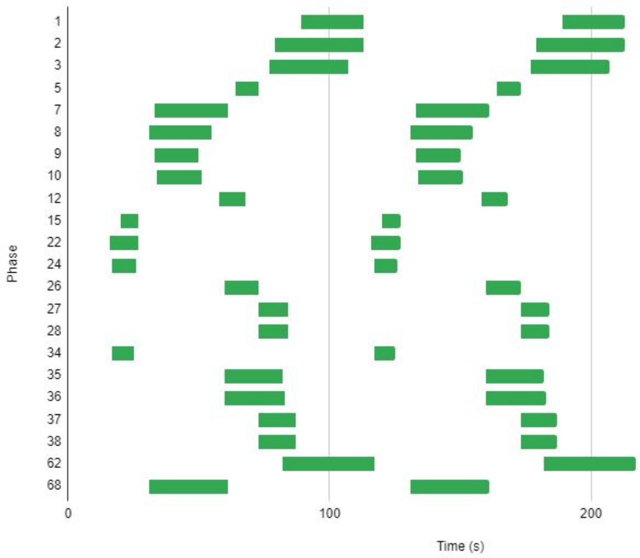


Figure 11: Example step raster for an intersection with a 100-second cycle

All-red clearance times in the Netherlands are also calculated per pair of signal phases, rather than having a fixed clearance time for each phase regardless of the phase which follows. This allows for shorter red clearances than using the North American method in situations where the time to clear the subsequent movement is shorter than the time to clear the whole intersection.

The advantage of those differences is that Dutch signal controllers have a greater ability to avoid wasted time due to artificial constraints on the allowable combinations of signal phases, or due to excessive clearance times. But the disadvantage is that it greatly increases the complexity of the controller operation. In order to determine how much time can be added to the cycle length during a TSP intervention, the novel system will need to estimate how much time can be recovered during offset correction (i.e. how short the cycle can be). Since the start and end times of phases in a block can both vary relative to each other and each has a different combination of conflicts with the phases in the subsequent block, there are many groups of conflicting phases which could potentially be the determining factor for the cycle length. It would theoretically be possible for the control program to generate the set of conflict groups automatically based on the block structure and the conflict matrix, but the halfstarre controller which is being used as the basis for this study does not currently have that feature.

To limit the complexity of the system development, this issue is avoided by manually determining the two conflict groups which are most likely to be the critical ones at the studied intersections. The system will evaluate the two groups to establish which is critical. This is the same level of controller flexibility as a standard North-American dual-ring controller, so despite the simplification, the system would have full functionality in a North-American context.

3.8.2 Traffic Adaptation

A non-adaptive (time-of-day) control system is studied, in order to maximise the applicability of the research. Despite the fact that adaptive control systems have been commercially available for decades, most traffic signals continue to use non-adaptive control systems due to their lower equipment maintenance costs and more intuitive operations (Wahlstedt, 2011). These intersections likely have a greater potential for improvements in the TSP system, since they are more vulnerable to deviations from the standard timing plan than adaptive systems which can actively respond to increased queues for particular movements. Within the timescale of an individual TSP intervention and compensation, a traffic-responsive system would behave identically to a non-adaptive system, since the traffic plans are typically only changed at most once per 10 to 15 minutes. The results of the study would therefore apply to both time-of-day and traffic-responsive systems.

It is worth noting that the although the ResFlex system includes a mechanism which adapts some TSP parameters based on the traffic volume observed over the past 15 minutes, the basic operation of the controller continues to be non-adaptive.

3.8.3 TSP specifications

Public transport vehicle detection is based on check-in and check-out detection. This is the detection method used by the existing fixed-cycle TSP system which will serve as a baseline for the novel TSP system developed in this project.

In all scenarios, TSP calls are conditional on headway with 2 levels of priority. Vehicles above the average headway will receive the full level of priority, including truncation, phase insertion and green extension. Vehicles below the average headway are only eligible for green extensions and calling phases in their normal position in the cycle. These are the two levels of priority available with the existing halfstarre control system to be used in the study, so to ensure that the results are comparable, the novel system will use the same two levels.

Conditional priority was selected because it is standard practice to use schedule-based conditional priority for conventional bus lines in the Netherlands (CROW, 2010), and because the use of headway or schedule-based conditional priority affects the distribution of TSP requests at the a signal. With conditional priority the probability of receiving multiple high-priority requests within a short period is lower because following a TSP intervention, calls will only be accepted in the opposite direction until the headway threshold is reached. Conditional priority also precludes there being more than two high-priority TSP requests in short succession, since following the second request, new TSP calls in either direction would not meet the headway threshold.

Even though priority in the Netherlands is typically conditional on schedule adherence, it would also have been possible to test the systems with non-conditional priority, but with a lower bus frequency to represent the reduced number of priority interventions. However, such an approach was not selected because it could inaccurately indicate the impacts of the proposed system on coordination, since the priority requests which are denied in a conditional system will disproportionately occur shortly following a previous intervention rather than being equivalent to a uniform reduction in the number of TSP requests.

The proposed service frequency of 12 buses per hour falls into the 'frequent service' paradigm where passengers will often head to a bus stop without checking the schedule. As a result, on routes with frequent service it becomes more important to provide regularly-spaced service than service which exactly corresponds to the schedule (Hounsell and Shrestha, 2012). For that reason, the condition for priority is based on headway regularity, not schedule adherence. The two primary methods for headway-based conditional priority are to use a set headway threshold (e.g. the average headway), or to compare the headways ahead and behind the bus in question. The advantage of having a predefined headway threshold is that it is very easy to implement, since a bus's headway can be observed at the intersection by comparing the time at which it was detected with the time at which the previous bus was detected. Comparing the headways before and after the given bus is more challenging because it requires knowledge of the headway of the following bus, which as not yet arrived at the intersection. However, the advantage of that system, according to Hounsell and Shrestha (2012) is that it is more effective at improving service regularity.

Because the focus of the study is primarily on reducing traffic signal delay rather than an optimisation of service quality along the public transport line, the simpler option is used. A fixed threshold is used, equal to the average headway between vehicles (5 minutes). Buses with a headway greater than 5 minutes receive full priority and buses with a headway less than 5 minutes receive limited priority.

4 Controller Design

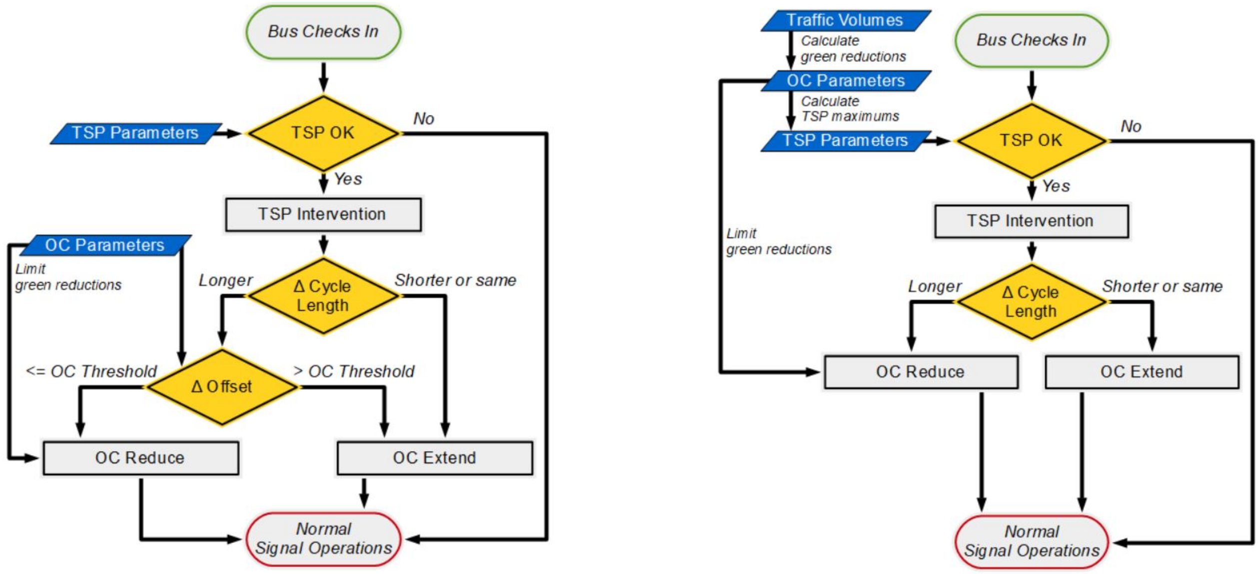
The core of the study is to evaluate comparable fixed-cycle and flexible-cycle TSP systems in a microsimulation environment, so a fully-functional signal controller is required which is able to operate either with fixed-cycle or flexible-cycle TSP. There is not currently a flexible-cycle TSP system available for the Dutch Halfstarre control system being studied, so one was developed, including active compensation for green time redistribution during priority interventions. This section describes the design decisions involved in creating the new system.

The new flexible-cycle TSP systems are based on the existing TSP system for the CRSV halfstarre controller used by Movensis, henceforth referred to as the Base system. As much as possible of the base controller's programming has been retained in the new TSP systems, to ensure that results are as comparable as possible. All of the basic functionality of the signal controller remains unchanged, such as the decisions on how to allocate unused green time and the process of executing priority actions. The primary changes required to implement a flexible-cycle TSP system in the new controller are to modify the decision logic for implementing TSP interventions, and create a new mechanism which can implement offset correction.

The objective of the new TSP system is to provide as much priority as possible to public transport vehicles, while still providing sufficient capacity to accommodate the current demand for all traffic movements, and while returning into coordination within a specified number of cycles following the TSP intervention. A novel component of the new system which was not present in the other flexible-cycle TSP systems observed in literature, is the use of "OC credits": a variable for each phase which represents the extra (or reduced) green time it should receive as compensation for an earlier TSP intervention. Compared to existing flexible-cycle systems such as that by Liao et al (Liao and Davis, 2007), which implements the inverse action from the TSP action during the cycle after the public transport vehicle has left, the use of explicit green "credits" allows gives the system a greater degree of flexibility, such as the ability to accept new TSP calls before the signal has returned in sync, and to adjust the compensation to account for the capacity impacts rather than solely the green time impacts.

4.1 System Framework

There are two possible strategies, or "system frameworks" which could be used achieve the controller's objectives: a constant-priority framework and a variable-priority framework. With a constant-priority framework, restrictions on TSP actions, such as a maximum duration for green extensions and minimum phase durations before truncation, do not vary based on traffic conditions. With a variable-priority framework, the magnitude of available TSP actions (e.g. the maximum duration for a green extension) is limited to the amount which will maintain a specified intersection performance, in terms of capacity and coordination. It is worth noting the "constant" level of priority may still vary the level of priority based on a condition such as public transport vehicle's punctuality, but the key is that it does not vary based on traffic conditions. The frameworks for the two strategies are illustrated in figure 12.



(a) Constant-priority system (b) Variable-priority system

Figure 12: Proposed framework options

4.1.1 Constant-Priority Framework

A proposed structure for a constant-priority system is illustrated in figure 12a. The defining characteristic of this structure is that the priority intervention and subsequent offset correction are treated as independent processes.

When a TSP call is received, the system takes whatever actions are available within the programmed TSP parameters to provide a green light at the estimated arrival time of the vehicle, or as soon as possible thereafter. Once the intervention is complete, the offset correction system observes the coordination impacts that have occurred, and decides the best way to return the signal into coordination given the parameters chosen by the road authority.

With this framework, the primary decision as part of the offset correction process is to determine whether to use extension-based or reduction-based offset correction following an intervention which increased the cycle length. Either option has a different set of limitations. In situations where both extension- and reduction-based offset correction are viable options, reduction-based OC tends to produce lower impacts on coordination, because with extension-based offset correction, the first cycle of offset correction moves the offset even further away from coordination than the TSP intervention. For example, in figure 9 the coordinated phase is shifted to the opposite portion of the cycle compared to the normal coordinated cycle when extension-based OC is used to recover from a TSP extension. Reduction-based OC also results in shorter waits at red lights (control delay) due to the longer cycle. However, its advantage is that the larger cycle length also produces a higher capacity, thanks to the longer phase durations and thus smaller proportion of lost time per cycle (Huang et al., 2012). The decision between implementing extension- or reduction-based offset correction is determined by a parameter known as the "OC Threshold", which is a certain amount of seconds of deviation in the offset. If the TSP intervention resulted in the offset being shifted later by more than the amount of the OC threshold, extension-based offset correction will be used, otherwise reduction-based offset correction will be used. If the TSP intervention started while the controller was in sync, the shift in offset will be equal to the cycle length increase resulting from the intervention.

If the TSP intervention reduced the cycle length (i.e. due to phase truncation), extension-based offset correction always has lower impacts on coordination and provides higher capacity. It also has a greater ability to redistribute time, since there are no limitations on the amount by which a phase's maximum green can be increased. Given the considerable disadvantages of reduction-based offset correction in this situation, extension-based OC is always used.

4.1.2 Variable-Priority Framework

The alternative option is to vary the level of restriction on TSP actions in order to prevent the TSP intervention from causing the intersection from exceeding capacity or deviating too severely from the coordinated timings. A possible structure for such a system in illustrated in figure 12b.

In this framework, the TSP parameters are determined based on the offset correction parameters. For example, the maximum amount of time which can be added to the cycle length during a TSP intervention would be limited to the amount of time which can be recovered within a specified number of cycles, given the permitted green reductions during offset correction.

This framework lacks the module for deciding between extension-based and reduction-based OC for TSP interventions which increase the cycle length, which was a key component of the constant-priority framework. It would have been possible to include that decision module, but it was deliberately omitted from this design. Instead, all TSP interventions which increase the cycle length will always use reduction-based OC. If it not possible to recover from a TSP action using reduction-based offset correction while satisfying the capacity and coordination requirements, that action will not be permitted. The option for extension-based OC in this situation was omitted since variable-priority systems will tend to be preferred due to their ability to limit the impact of TSP on other traffic. So in the situations where variable-priority is preferred, the large coordination impacts of extension-based recovery would presumably not be considered to be acceptable.

4.1.3 Comparison of Frameworks

The primary advantage of the fixed level of priority concept is that the level of service that public transport receives at the signal remains relatively constant. This is very valuable for the reliability of public transport operations because it minimises the variation in the amount of delay experienced at the intersection. With a variable level of priority, delays incurred to public transport vehicles as a result of greater than average traffic volumes will be accompanied by a reduced level of priority, which results in further additional delay relative to the typical situation.

The theoretical advantage of the variable priority system is that if the public transport line is operating in mixed traffic, maintaining sufficient capacity at the intersection can avoid queuing which would otherwise severely delay the public transport vehicles.

The framework selected for further study is a variable-priority system. Most intersections in the Netherlands are fully-actuated and uncoordinated, so if a group of intersections are coordinated with a halfstarre system, the traffic engineers have placed an unusually high level of importance on maintaining coordination, potentially due to the risk of spillback if the intersections drift out of sync with each other. A system which actively limits TSP impacts in the cases where they could result in excessive impacts to coordination and/or capacity would be more likely to be adopted than a system which could potentially exacerbate problems related to heavy traffic volumes. The intent is that even with restrictions on the TSP actions based on traffic conditions, the new controller developed as part of the study will outperform the current fixed-cycle TSP system which is being used as a base case.

4.2 Available Input Data

The base controller does not include traffic-adaptive features: the green times in the step raster are predetermined for each time-of-day plan, and the actual green times during a given cycle are determined by ordinary vehicle actuation logic. However, many signalised intersections in the Netherlands - including those without any adaptive control systems - are equipped with vehicle detectors in all directions which collect sufficient data to operate some traffic-adaptive features.

The new TSP systems have two levels of traffic adaptation, with different amounts of input data required. In the lower "parametric flexible cycle" (ParFlex) level, the priority and offset corrections operate solely with the information available at any traffic signal installation, such as actual green times and predefined parameters. It will thus not depend on any current traffic volume information.

The higher "responsive-parameter flexible cycle" (ResFlex) level would be available for intersections which vehicle detectors in all approach lanes which are able to count the number of vehicles. These traffic counts are used to continuously update the thresholds and parameters for the priority and offset recovery processes. Note that this real-time traffic volume information is only used during the priority and offset correction processes. The basic operation of the new controller (e.g. in the absence of any priority calls) is unchanged from the base controller, and thus continues to not be traffic-adaptive.

The ways in which the new TSP systems use input data in each adaptation level are detailed in subsection 3.6.

4.3 System decision structure

The new variants of the controller reuse the majority of the code of the base controller. The mechanisms which are modified or added in the new controller need to:

1. Measure traffic volumes (ResFlex system only),
2. Calculate "desired green times",
3. Estimate the impacts to each traffic movement incurred during priority interventions ("OC Credits"),
4. Determine which TSP actions are available for each TSP request,
5. Disable the cycle length constraint during priority interventions,
6. Generate timing plans in real time which implement Offset Correction while distributing OC Credits.

The process for determining OC credits is discussed before the processes to implement TSP interventions because the latter depends on an estimate of the OC credits which will be earned during the intervention.

4.3.1 Measuring traffic volumes

The ResFlex variant of the new controller uses real-time traffic volumes to determine the desired green times for each phase in a critical conflict group.

The number of vehicles counted by the stopbar detectors are summed for each minute and each phase. Every minute, the observed traffic volume is updated to include the sum of the traffic volumes from the previous 15 minutes.

Because capacity calculations use passenger car units per hour (PCU/h), the observed traffic volume in vehicles per hour then needs to be adjusted using the percentage of heavy vehicles.

The effective volume, "V_i" for a phase i is its observed traffic volume, converted to passenger car units per hour (PCU/h). Since an average heavy vehicle counts as 2 PCU (FHWA, 2008), the observed volume can be converted to from vehicles per hour to PCU/h simply by increasing the volume by the number of heavy vehicles.

$$V_i = vObs_i * (\frac{\%HGV_i}{100} + 1) \tag{4.1}$$

where:

$vObs_i$ = *observed traffic volume of phase i, in vehicles per hour*
$\%HGV$ = *percentage heavy vehicles, as an integer*

At intersections without access to real-time traffic volumes, effective traffic volumes based on past observations can be entered manually. This allows the TSP system to function correctly, but without any real-time response to traffic conditions.

4.3.2 Determining Desired Green Times

One of the key pieces of information which is used by the new priority mechanisms is the "desired green time", which represents the amount of time that a signal phase would require during a typical cycle in order to serve the traffic demand. This is not necessarily the same as the maximum green time that is specified in the timing plan, since some phases may receive longer maximum green times than necessary if they are not in a critical conflict group. The desired green time is the starting point for calculating new green durations in the offset correction process.

With the parametric ParFlex TSP system, the controller does not have access to the current traffic volumes, so the desired green time "$gReq_i$" for a phase i is simply set by the user for each phase as a percentage of the maximum green time:

$$gReq_i = gMax_i * gReqPct/100 \tag{4.2}$$

where:

$gMax$ = *programmed maximum green time for phase i, in seconds*
$gReqPct$ = *percentage of maximum green for phase i, 0-100 (user-defined)*

To enable a high level of priority for public transport, a value lower than 100% should be used, to ensure that the cycle length can be reduced during offset correction, and thus allow extended cycles during TSP interventions. If all phases were set to 100%, it would only be possible to extend the cycle by the negative of the predicted OC Credits (taking into account the net OC Credits per critical conflict group).

With the responsive-parameter ResFlex TSP system, the required green time, "$gReq_i$" for a phase i is the minimum green time that a phase requires per cycle in order to process the average traffic volume that was observed over the last 15 minutes. It is calculated as:

$$gReq_i = \frac{V_i}{S_i * satMax} * C \tag{4.3}$$

where:

V_i = *effective traffic volume of phase i, in PCU per hour*
S_i = *saturation flow rate of phase i, in PCU per hour*
$satMax$ = *maximum level of saturation, from 0 to 1 (user-defined)*
C = *The cycle length during coordinated operation, in seconds*

To ensure that the estimate is conservative, the green time is not decreased to account for the typical one-second difference between the effective green and the displayed green.

In both the parametric and responsive-parameter modes, the required green time must always be equal or greater than the minimum reduced green time. The latter is a user-defined parameter for each phase which defines the minimum green time it must be entitled to. This entire time does not necessarily need to be served if the detectors are not active, but it does always need to be reserved within the cycle length.

$$gReq_i >= gRdc_i$$

where:

$gRdc_i$ = *minimum reduced green time, in seconds (user-defined)*

4.3.3 Determining Offset Correction Credits

An unusual feature of the new TSP systems is the concept of "Offset Correction Credits", which is the controller's way of keeping track of the impact that each movement has experienced as part of the priority intervention and subsequent offset correction process. Each credit represents one second of green time that the phase is owed. Because the step raster operates in whole seconds, it is not possible to have fractions of a green credit. If a phase is negatively impacted, for example by being truncated by a priority intervention, it will earn positive OC credits to produce a compensatory positive impact during a later cycle. Conversely if a phase is positively impacted, for example due to a green extension, it will earn negative credits which reduce its entitlement to green time during one or more subsequent cycles.

There is no upper limit to the amount by which the maximum green can be increased during offset correction, so if there is a positive credit, it will all be immediately redeemed during the first cycle of offset correction. If at any time a phase ends before its maximum green due to a lack of demand, any remaining positive OC credits will be deleted. If there is a negative credit, the phase's duration is reduced as much as possible without violating the minimum green before truncation parameter (80% or 90% of the normal maximum green depending on the scenario), and the remaining negative credit is carried over to the next cycle.

With the parametric ParFlex TSP system there is little real-time information available to the signal controller, so offset correction credits are allocated at a 1:1 rate. During a TSP intervention, each second that a phase displays green outside of its primary realisation period will earn it one negative OC Credit. If a phase is truncated by a TSP intervention, the unused portion of the phase's primary realisation period is credited to the phase as positive OC credits.

With the responsive-parameter ResFlex TSP system, the signal current has access to the current and past vehicle volumes, so it is possible to take a more detailed look at the impacts of each TSP action on other phases, and create a more detailed formula for calculating OC credits. The controller only has access to 3 TSP actions: green extensions, red truncations and phase insertions, so those will each be examined.

Green Extensions or Phase Insertions

With offset correction credits allocated at a 1:1 rate and the additional green time from a green extension fully utilised, there would be no net impact on vehicle capacities, since the capacity lost during offset recovery would exactly equal the extra capacity time provided during the TSP action.

However, in practice it is most likely the extra green time will not be fully utilised, since by definition it occurs after the phase has already received the maximum green time allocated to it in the normal timing plan. As a result, the waiting queue should already have been processed, in which case the extra green time is only used by vehicles arriving intermittently.

For phase insertions, the likelihood of the green time being fully utilised is higher since it follows a red light during which a queue may have formed. But implementing a phase insertion also requires additional time in the form of yellow change and red clearance intervals.

During extra green time provided as part of a TSP intervention (i.e. a phase insertion or green extension), the number of vehicles crossing the stop line is counted for each active phase. At the end of the intervention, the OC credit is reduced by the amount of green time which would be required to clear that number of vehicles at the saturation flow rate.

If the phase received extra green time through TSP:

$$\Delta OCC_i = -3600 * \frac{nExt_i * (1 + \%HGV_i)}{S_i} \tag{4.4}$$

where:

$nExt_i$ = *number of vehicles using phase i during extra green time*
$\%HGV_i$ = *proportion of heavy vehicles for phase i, as a decimal*
S_i = *saturation flow rate of phase i, in PCU/h*

Red truncation

Assuming that the saturation flow rate remains constant throughout the duration of a phase, the capacity impact of a phase truncation can be balanced by increasing the phase duration by the same amount during the following cycle. As long as there is sufficient storage for the unserved vehicles to sit without hindering other movements, the net capacity impact over the course of two cycles should be zero.

As with the ParFlex system, the ResFlex system credits the difference between the phase's required green and the amount that was served at a 1:1 rate. If the phase had already gapped out prior to the TSP intervention, it will not receive any OC credits.

If the phase was terminated due to truncation:

$$\Delta OCC_i = gReq_i - gAct \tag{4.5}$$

where:

$gReq_i$ = *required green time for phase i, in seconds*
$gAct_i$ = *actual green time served by phase i, in seconds*

As part of the TSP action, OC credits are updated for the affected phases. Once the priority request ends, offset correction may begin. If the offset error is positive (i.e. the signal is late compared to its intended offset), the maximum green times are updated as per the allowed reductions, including redeeming OC credits as much as possible. The signal remains in offset correction until the offset error is eliminated, then the normal maximum green times are restored. If the offset error is negative (i.e. the signal is early), the maximum green times are solely adjusted according to any OC credits. The resulting green times may need to be adjusted to ensure that the signal returns in sync in one cycle.

When a phase is extended above its maximum green as part of a TSP intervention, it receives a negative OC credit (i.e a green 'penalty') which represents the number of seconds at the saturation flow rate that would be required to process the volume of motor traffic which was processed during the extra green time due to a TSP action. The rationale for deducting this time during offset correction is that that the vehicles processed during the TSP action would otherwise have needed to be processed during the following cycle (i.e. during offset correction), and the *required green time* does not take this into account.

4.3.4 Calculating allowable TSP actions

Whether or not a given TSP action will be implemented depends on whether it fits within the allowable cycle extension. For a green extension, the time required is the difference between the priority vehicle's predicted arrival time and the planned end of the phase it wishes to use. For a phase insertion, the time required is the inserted phase's minimum green time plus its yellow change interval and red clearance time.

The allowable cycle extension "*tCExt*" for a phase *i* is the maximum amount of time that a TSP action is permitted to increase the cycle length. It can restrict green extensions to a lower maximum than the *maximum green extension* parameter, if doing so is necessary in order to meet the coordination and capacity requirements selected by the road authority. The maximum allowable TSP extension is based on the amount of time which can be added to the cycle length while still returning in sync within the specified maximum cycles of offset correction, and which does not cause the offset to shift further than the *maximum offset error* value:

$$tCExt_i = min(tOCRdc_i\,,\, eCMax - eC) \tag{4.6}$$

where:

$tOCRdc_i$ = *amount of time which can be reduced during offset correction following a green extension of phase i, in seconds*
$eCMax$ = *maximum offset error, in seconds*
eC = *current offset error when the TSP call was received, in seconds*

For each conflict group *k*, the total time required during the offset correction process needs to be calculated. The actual amount of time that can be reclaimed during offset correction is then the difference between the normal time served during that number of cycles (i.e. *cycle length * number of cycles*) and the largest time required during offset correction for any of the conflict groups. In Equation 4.7, phases *j* are the phases in the conflict group other than the TSP phase, *i*. If the TSP phase is not in the conflict group, the phase which is located in the same stage as the TSP phase is used instead.

$$tOCRdc_{ik} = nOC(C) - \left[gOC_i + nOC(y_i + r_{ik}) + \sum_j [gOC_j + nOC(y_j + r_{jk})] \right] \tag{4.7}$$

where:

C = *normal cycle length, in seconds*
nOC = *maximum number of cycles of offset correction*
gOC_i = *total green served during OC for TSP phase, in seconds*
y_i = *yellow change interval for TSP phase, in seconds*
r_{ik} = *red clearance interval for TSP phase in conflict group k, in seconds*
gOC_j = *total green served during OC for non-TSP phase j, in seconds*
y_j = *yellow change interval for non-TSP phase j, in seconds*
r_{jk} = *red clearance interval for non-TSP phase j in conflict group k, in seconds*

The green time that is served by the TSP phase during the offset correction process is either the total required green time and (negative) OC credits, or the total minimum reduced green time, whichever is greater.

$$gOC_i = gReq_i * nOC + OCC_i + \Delta OCCe_i \tag{4.8}$$

subject to the constraint:

$$gOC_i >= gRdc_i * nOC$$

where:

$gReq_i$ = *required green for phase i, in seconds*
nOC = *maximum number of cycles of offset correction*
OCC_i = *OC credits before start of intervention, in seconds*
$\Delta OCCe_i$ = *estimated OC credits from intervention, in seconds*
$gRdc_i$ = *minimum reduced green time for phase i, in seconds*

Because the amount of time that can be recovered during offset correction is also related to the extension duration itself (through the *OC credit*), calculating the maximum allowable TSP extension requires an estimate of the OC credits which will be obtained as a result of that TSP extension.

The net OC credit resulting from a phase extension will be estimated using the assumption that the queue has already cleared, and the vehicles entering the intersection are randomly arriving at the intersection. The ratio of the movement's traffic volume to its saturation flow rate is used as a conversion factor from the green extension time to the equivalent duration of saturated green:

If the phase was extended above its maximum by TSP:

$$\Delta OCCe_i = -(gAct_i - gReq_i) * \frac{V_i}{S_i * satMax} \tag{4.9}$$

where:

$gReq_i$ = *required green time for phase i, in seconds*
$gAct_i$ = *actual green time served by phase i, in seconds*
V_i = *traffic volume of phase i, in PCU/h*
S_i = *saturation flow rate of phase i, in PCU/h*
$satMax$ = *maximum saturation rate for the intersection*

Since *gAct* the green extension will allow the TSP phase to receive *less* time during offset correction, ΔOCC_e is negative.

For movements with an upstream signal in the same coordinated group, this estimate will tend to overestimate the number of vehicles making use of the extended green time, because traffic will disproportionately arrive during the portion of the cycle corresponding to the normal green time.

Substituting Equation 4.9 into Equation 4.8 gives:

$$gOC_i = gReq_i * nOC + OCC_i - (gAct_i - gReq_i) * \frac{V_i}{S_i * satMax} \tag{4.10}$$

In order to be able to solve for the green extension, that extension will be assumed to occur exclusively during extra time, beyond the cycle length. This is a reasonable simplification since it will underestimate the negative OC credits received by the TSP phase when the extension occurs partly using spare time within the cycle, but if there is spare time within the normal cycle length, the cycle length is less of a constraint to begin with. Substituting Equation 4.10 into Equation 4.7 and setting the actual green extension time $(gAct - gReq)$ and total time reclaimed during offset correction $(tOCRdc_{ik})$ equal to $tCExt_i$ produces:

$$tCExt_i = nOC(C - gReq_i - y_i - r_{ik}) - OCC_i + tCExt_i * \frac{V_i}{S_i * satMax} - \sum_j [gOC_j + nOC(y_j + r_{jk})] \tag{4.11}$$

Solving for $tCExt_i$ produces:

$$tCExt_i = \frac{nOC(C - gReq_i - y_i - r_{ik}) - OCC_i - \sum_j [gOC_j + nOC(y_j + r_{jk})]}{1 - \frac{V_i}{S_i * satMax}} \tag{4.12}$$

where:

nOC = *maximum number of cycles of offset correction*
C = *normal cycle length, in seconds*
$gReq_i$ = *required green time for phase i, in seconds*
y_i = *yellow change interval for phase i, in seconds*
r_{ik} = *red clearance interval for phase i in conflict group k, in seconds*
OCC_i = *initial OC credits for phase i, in seconds*
gOC_j = *total green served during OC for non-TSP phase j, in seconds*
y_j = *yellow change interval for non-TSP phase j, in seconds*
r_{jk} = *red clearance interval for non-TSP phase j in conflict group k, in seconds*
V_i = *traffic volume of phase i, in PCU/h*
S_i = *saturation flow rate of phase i, in seconds*
$satMax$ = *maximum level of saturation, from 0 to 1*

The calculation for the green time during offset correction of other phases in the conflict group, gOC_j, is similar to the calculation for that for the TSP phase (gOC_i), but without the calculation for net OC credits, since green extensions do not result in any net change in OC credits phases in other stages than the TSP phase.

4.3.5 Varying the cycle length

In order to effectively vary the cycle length, the new systems include four modules which are each capable of generating step rasters in real time to achieve a given objective. These each operate in one of the coordination statuses, which are: In Sync, TSP Active, OC Reduce and OC Extend. In a typical TSP intervention, the status will step through the modes in that order, but it is also possible to skip steps if certain conditions are met. The logical process for determining the currently active coordination status is illustrated in figure 13.

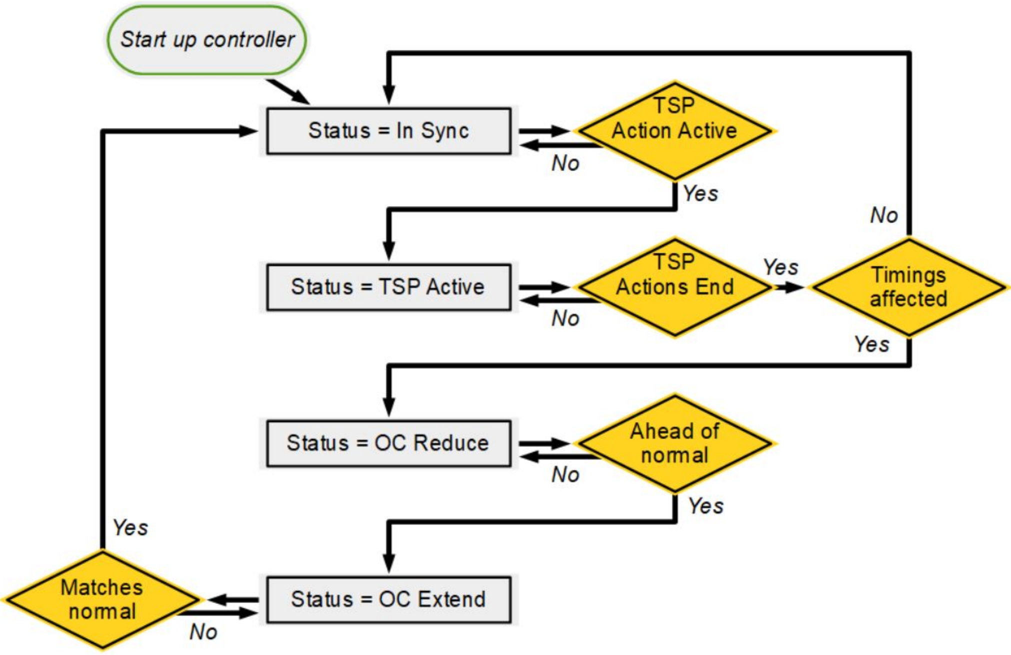


Figure 13: New controller's coordination status logic

In Sync is the default status. In this state, the step raster which was input by the user is used directly and the signal's cycle length and offset are fixed.

If a TSP request meets the requirements for one or more TSP actions, the signal status will change to *TSP Active*. In this status, the background cycle is effectively paused. In practice the clock continues operating so this is achieved by shifting the start and end time of every phase by one second each second (thus shifting the signal's offset by one second as well).

Once all approved TSP actions are complete, the signal evaluates the net impacts. OC Credits are calculated and distributed to phases based on the number of extra seconds received (ParFlex system), or the counted number of vehicles benefiting from the intervention (ResFlex system). If the same phases are still in their primary realisation periods as in the background (In Sync) timings, then there has been no deviation from the cycle length. The signal will revert directly back to In Sync and all OC credits will be discarded. Otherwise the signal will begin offset correction.

4.3.6 Implementing Offset Correction

Offset correction is implemented through two modes: OC Reduce and OC Extend. Each of these modes needs to continuously generate signal timing plans which achieve a given objective, while also accommodating the desired green times and distributing OC credits.

In order to do so comprehensively the controller would need to calculate all of the relevant combinations of conflicting phases ("conflict groups"), which could potentially determine the cycle lengths during offset correction. But the number of conflict groups increases exponentially with increased number of phases, and with 18 phases at the intersection, this would be too computationally taxing if a brute force method is used. Movensis is currently developing a module which uses a more efficient method to generate a list of relevant conflict groups for the intersection. But at the time of the study, that module was not yet ready for deployment. So for the sake of simplicity, that module is omitted in this version of the design, and instead the four conflict groups which are capacity-critical given the model's traffic volumes are hard-coded in the controller. The rest of the conflict groups are therefore not considered as part of cycle length estimates. For this study, the conflict groups considered are those including the through and left turn movements, as illustrated in figure 14 along with their phase ID numbers.

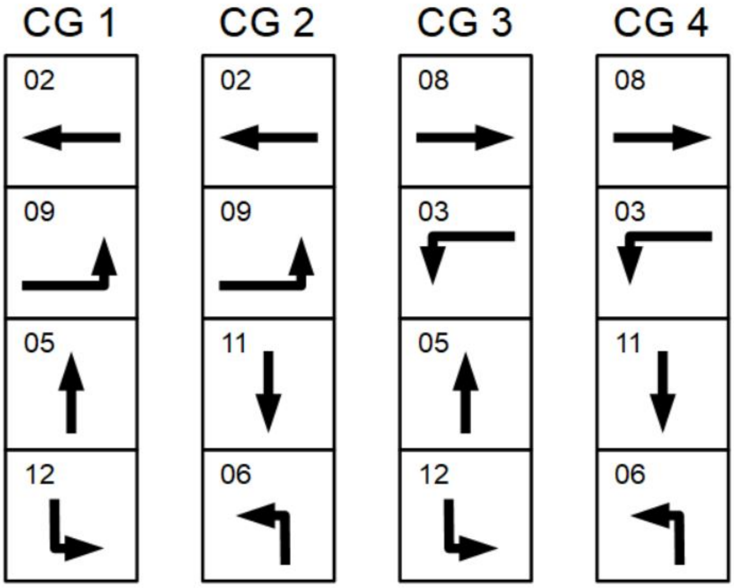


Figure 14: Critical Conflict Groups (CG) included in cycle length calculations

In OC Reduce, the signal's objective is to cycle as quickly as possible. It operates effectively the same as an uncoordinated fully-actuated intersection. The signal continuously calculates the start and end times for the two subsequent phases, taking into account all four of the critical conflict groups and the applicable yellow and red clearance times. The planned green duration for each upcoming phase is consists of the desired green plus OC credits, subject to the constraint that the planned green time must exceed the minimum reduced green rate (80 or 90% of the normal maximum green duration, depending on the scenario). If an active phase in a critical conflict group changes to yellow, that phase's end time in the step raster is changed to the current time. This may result in the start and end times of the subsequent phases shifting earlier, depending on the phases active in other conflict groups.

While in OC Reduce, the controller also calculates the green start and end moments of all phases in the critical conflict groups which would apply if the signal were in OC Extend. If the start time in OC Reduce for both of the current phases in the critical conflict groups are earlier than the start time of those phase in OC Extend, then the signal has caught up enough time and can switch to OC Extend.

In OC Extend, the start and end moments are calculated working backwards from the In Sync end time of the current phase, using the normal pattern maximum green for each phase, as well as any remaining positive OC credits. Once all positive OC credits are redeemed for all phases, the mode is allowed to change to In Sync if appropriate. This occurs if the planned end of the current phases are within two seconds of their end times in the background (In Sync) timing plan. The two seconds of leeway are required because the OC Extend times may differ from the In Sync by up to two seconds due to rounding, since clearance times are calculated in half-seconds but the step raster only operates in full seconds.

If Offset Correction is initiated, the status always starts in OC Reduce, and then switches to OC Extend once it is ahead of the In Sync timings. However, if the TSP action resulted in a net reduction of the cycle length (and OC Reduce is thus unnecessary), OC Reduce will only last one clock cycle of the controller (0.1 seconds) before switching to OC Extend.

4.4 Controller Testing

Throughout the development of the new TSP systems, each component was extensively testing by running the controller in a test environment (shown in figure 15) and simulating detector calls.

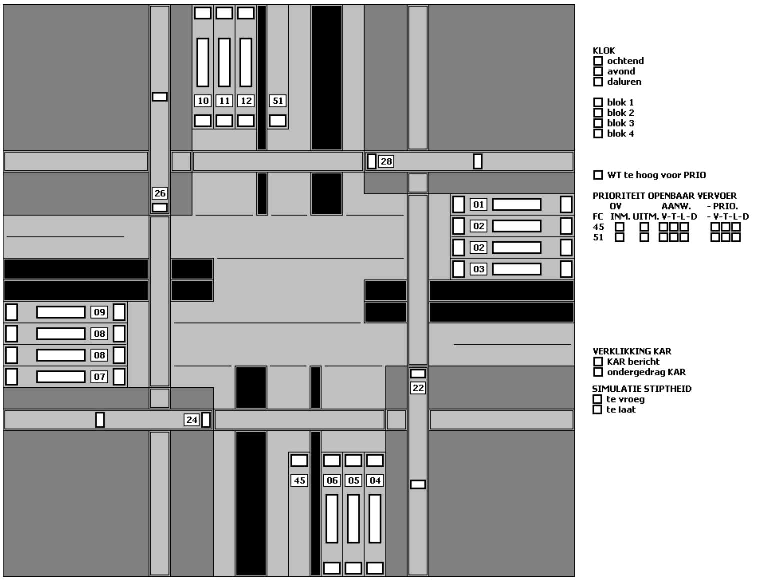


Figure 15: Test environment user interface

Each detector in the model is represented by a button which can be activated and deactivated manually or can be triggered automatically at random intervals to represent a particular traffic volume. TSP calls are generated manually by activating the check-in and check-out detectors.

Prior to each test, internal variables are selected which will be continuously displayed throughout the test. In addition to the particular variables which are most relevant to the module being tested, most tests included the intersection's current clock time, OC status and the maximum allowable cycle length increase. They also included each phase's start and end times, maximum green durations, desired green durations, current offset correction credits, and if there is currently a TSP request, the estimated number of seconds until a bus arrives at the stopbar.

Throughout the test, the controller's operation is intermittently paused and the desired values for some relevant internal variables are manually calculated to ensure that the currently displayed values are correct.

Countless tests were conducted throughout the development of the new TSP systems. For example, the basic functionality of the OC Reduce status was first tested by manually setting the OC status to OC Reduce and verifying that the correct phase durations were being served. Then various scenarios were tested such as phases ending early due to lack of demand (and thereby changing the planned start and end times for subsequent phases), and again the displayed values were verified against manual calculations. Later once all four OC statuses were operating correctly on their own, testing was done to ensure the controller was switching between them at the correct moments.

Finally, once the new controller was complete it was also run in Vissim with a variety of parameters and carefully observed to ensure that it was behaving as intended. Unfortunately due to the way the software used to connect the controller to the Vissim environment, it was not possible to see monitor the status of internal variables while the controller is running in Vissim. Phase durations could be monitored using the clock in Vissim, but the decision logic within the controller could not be checked. For example, in Vissim it is possible to identify that phase insertions were denied for an approaching bus, but it is not necessarily clear why the controller denied the TSP request. In some cases it was necessary to replicate the scenario in the test environment in order to identify the internal variable statuses which led up to an undesired behaviour.

5 Model Design

To evaluate the performance of the new controller relative to the base controller, a microsimulation model was created with a simple fictional network of 3 intersections.

5.1 Model run characteristics

Each model run in Vissim depends on a *random seed* which is the basis of the pseudo-random variations such as the moments at which vehicles are generated, the types of vehicles, and the parameters that the 'drivers' of those vehicles have (PTV Group, n.d.). For each of the six scenarios described in subsection 3.6, model is run 10 times, each with a different random seed. For each scenario, random seeds 40 through 49 are used. For example, in all scenarios, run number 3 uses seed 42. This means that all vehicles in the network are generated at the same moments for a given run number in each of the six different scenarios. This consistency helps isolate the effects of the differing traffic signal controllers, and enables the use of paired T-tests to evaluate the significance of differences in metrics.

Each simulation run consists of 2 hours of data collection according to the model's clock. As described in subsection 5.4 the tested network has a signal cycle length of 120 seconds, so 60 signal cycles are observed per model run. Prior to the start of data collection, there is a 15-minute 'warm-up' period to allow traffic to enter the network and for traffic conditions to stabilise. These are the standard simulation durations used by Movensis for Vissim studies.

5.2 Network Design

The tested network consists of three coordinated intersections spaced 150 and 400 metres apart (centre-to-centre). At the middle intersection, a bus route travels perpendicularly to the coordinated directions in a dedicated busway in the centre of the street. The middle intersection has the heaviest traffic volumes of the three intersections, and is thus critical for the network capacity. This will help make the capacity effects of TSP interventions more apparent.

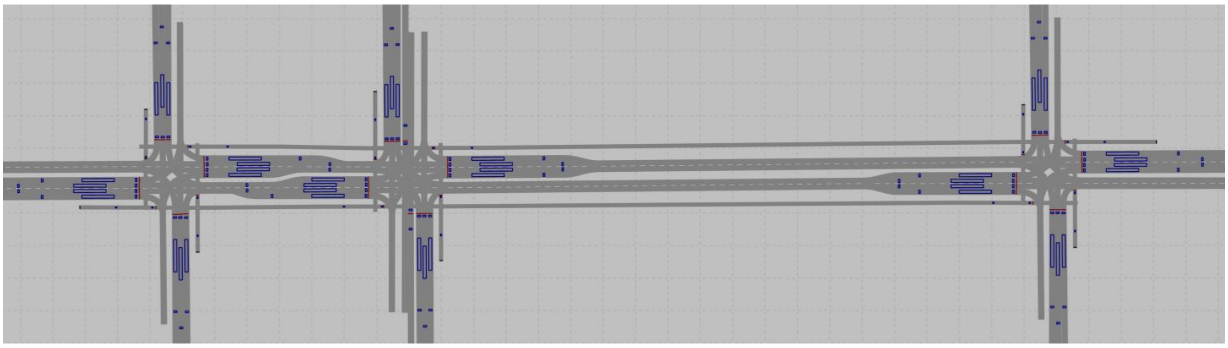


Figure 16: Tested network, as modelled in Vissim

A network of three intersections is the smallest practical network which provides platoons of traffic in both coordinated directions approaching the central intersection. A small network is preferred in order to minimise the work required to create signal timings and set up the network to be modeled.

Spacing the intersections unevenly allows the effects of different intersection spacings to be observed without the need to make separate models to test multiple intersection spacings. To ensure that the results for the different-length road segments are comparable, the model is as symmetrical as possible. Traffic volumes and lane configurations at the two outer intersections are identical to each other, but rotated 180 degrees. The traffic volumes and lane configurations at the central intersection are also symmetrical on opposite approaches. For this reason, the busway is located in the centre of the street rather than on one side or the other.

Making the public transport line perpendicular to the coordinated directions means that the coordinated direction is negatively affected by TSP actions, and stands to benefit from green time compensation during offset correction. It also avoids the complications in making TSP action decisions which occur when the public transport vehicle travels in a coordinated direction. Namely, the travel time savings of TSP actions at an upstream intersection can be enhanced or nullified depending on its offset relative to the downstream intersection. In some cases, a TSP action could have no net effect, if the public transport vehicle will end up waiting for the same start of green at the downstream intersection. While in others, the TSP action could have extra benefits where it allows the transit vehicle to catch a green light at the downstream intersection that it would otherwise miss (Zeng et al., 2021).

Priority interventions only affect the intersection where the public transport vehicle is approaching, not the other intersections in the coordinated group. In a small signal group such as the proposed 3-signal network, it may actually be better to vary the cycle length simultaneously at all three intersections as part of the priority intervention, so that the signals always remain in sync with each other and no offset correction is required. However, that method is not proposed for this study because it does not scale well to larger networks, particularly if there are multiple intersections equipped with TSP. The researchers of this study are more interested in systems which could work well with large networks. The effectiveness of systems which vary the entire network cycle length during TSP interventions could be an area for future research.

5.2.1 Intersection layout

The layout of the central intersection of the model can affect how well a priority system works there. For example, if there are very long crosswalks, it may be physically impossible in for buses to pass through without delay in some circumstances, because of the time required to clear pedestrians out of the intersection. The primary objectives for the intersection designs is to create a simple intersection with a realistic layout. The characteristics of the intersections are summarised in table 9 below.

Table 9: Intersection characteristics

Item	Type	Reason
Permissive conflicts	no	simplicity
Stopbar detectors	all directions	Realism
Setback detectors	all directions	Realism
Faraway detectors	where feasible	Realism
Bus ROW	dedicated	simplicity
Bus stops	far-side	simplicity
Speed Limit	all 50 km/h	realism
Bike crossings	single-stage	simplicity
Pedestrian crossings	no	simplicity

The bus route is placed in a dedicated busway so that it is not affected by general traffic conditions. This avoids the need for the TSP system to estimate the time required to clear the queue waiting at the signal ahead of the bus and thus enables the TSP system to be simpler. Similarly, there are no bus stops upstream of the intersection, to avoid the need to estimate dwell times at stops.

All roads have a speed limit of 50 km/h, which is the standard speed for through-traffic routes in urban areas in the Netherlands (ANWB, n.d.).

To ensure that it is always possible to give a movement a green light when there is an opportunity to do so, movements are controlled by separate fully-protected signal phases. For the general motor traffic phases, each approach lane uses the standard Dutch detector setup illustrated in figure 5. This consists of a short stopbar detector and 20-metre long setback detector. Where space is available, each motor vehicle lane also has a short far-away detector, located 70 metres upstream of the stopbar for through movements, or 60 metres upstream for turning movements. This setup with fully-protected signal phases for all movements and 3 vehicle detectors per lane is common for intersections along major arterial roads in the Netherlands, and the placement is consistent with standard practice.

For bicycle phases, each approach lane has a short stopbar detector, and a short far-away detector located 25 metres upstream of the stopbar. Bus phases are actuated primarily by the TSP check-in loop located 200 metres in advance of the intersection, and the check-out loop. To allow the signal to display green as the bus is approaching if the signal happens to already be in north-south green, a far away loop is located 75 metres before the stop line.

Ideally, the intersections would include multiple-stage pedestrian crossings, as is the standard along arterial roads in the Netherlands. In practice the multiple stages are often "coupled", which effectively creates a green wave for pedestrians to cross the street without getting stopped in the median. But in order to the avoid the complexity which comes with coupling phases together, pedestrian phases are entirely omitted from the model.

Single-stage one-way bicycle crossings are included in the model on all intersection arms.

Turning lanes have a desired tangent length of 70 metres, which provides just enough space to fit the far away loop located 60 metres upstream of the stop bar. All lanes achieve this desired storage length except for phase 03 (westbound left) at intersection 1 and phase 09 at intersection 02, which are located back-to-back in the short roadway segment between those two intersections. As a result, those phases do not have enough room for a far away loop, and they both commonly experience queues spilling back into the adjacent through traffic lane.

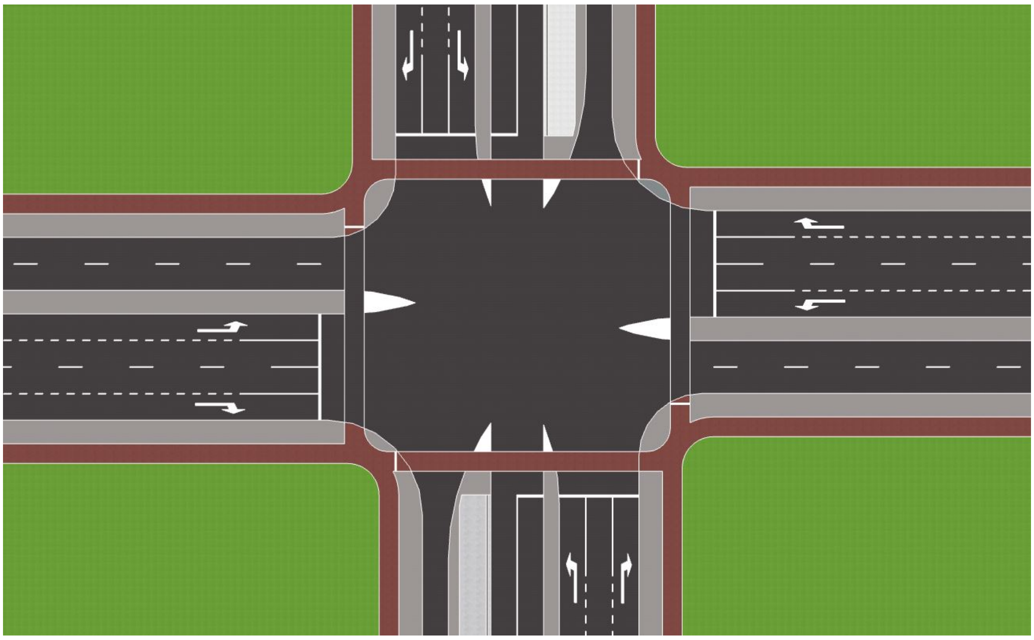


Figure 17: Intersection configuration - studied intersection

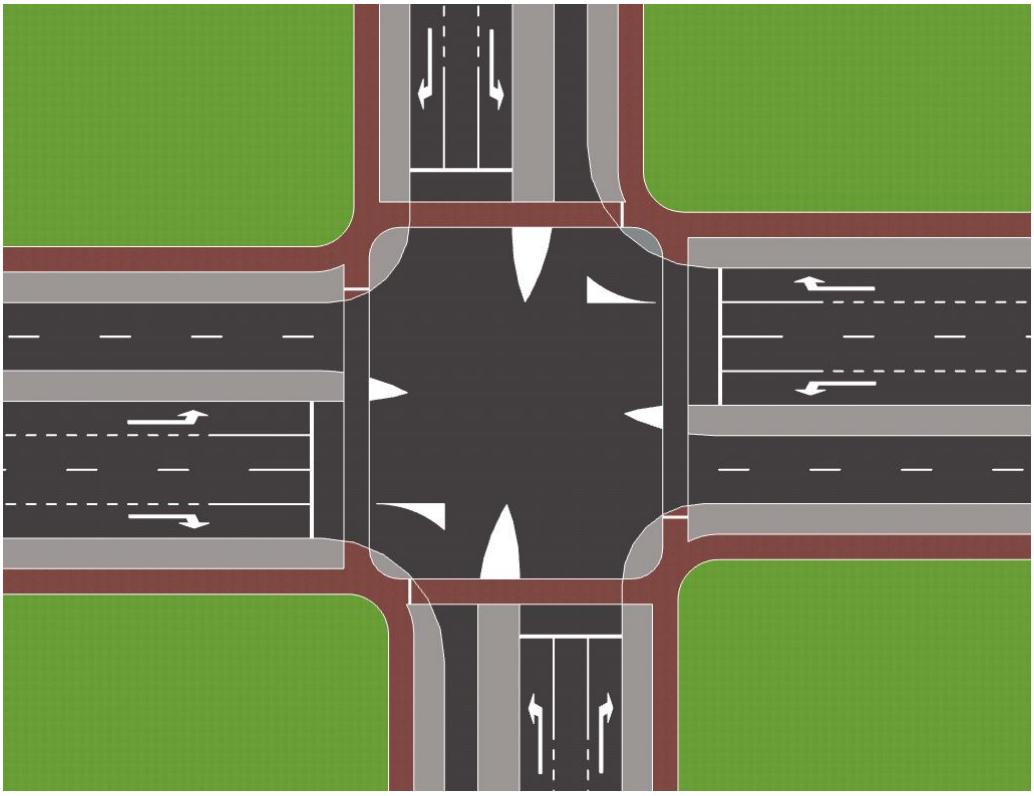


Figure 18: Intersection configuration - other intersections in model

5.3 TSP Parameters

The priority actions available in the new controller truncation, green extension and phase insertion. The base case signal controller also has these three options, as well as the option of "doorsnijden" (phase slicing), in which a phase is truncated, but then returns following an inserted transit phase. This feature is not enabled during testing, because it was not programmed in the new controller.

As explained in subsubsection 3.8.3, the TSP systems operate with two-level headway-based conditional priority, where buses with below-average headways are considered "early", and the remaining buses are considered "late". Given that the planned bus frequency is 12 buses per hour per direction, the headway threshold is 300 seconds (5 minutes).

With the new controller, the parameter for the maximum number of cycles of offset correction is set to 2. A value of 1 would not provide informative study results, because it would act nearly identically to a fixed-cycle system. A value of 3 is rejected because with the high bus frequencies in the model it is expected that the longer duration of offset correction would cause excessive disruption to coordination. With a value of 2, the largest possible disruption is that two maximum-disruption interventions (causing 2 cycles out of sync each) occur 2 cycles apart from each other, causing the signal to be out-of-sync for 4 consecutive cycles.

5.4 Signal Timings

Since the network is fictional, there are no real-world traffic volumes around which to develop optimal signal timings. Instead, realistic signal timings are generated to reflect the types of intersection (major or minor), and then traffic volumes are selected to reflect those signal timings.

To ensure that results are comparable between the road segments between intersections, all signal timings and all traffic volumes are symmetric for opposing directions.

5.4.1 Yellow warning intervals

Because the modeled intersections are all on level ground, the duration for the yellow warning intervals for each phase are simply the default durations from the Dutch CROW traffic signal timing manual (CROW, 2014). These are listed in table 10 - note that the speed limit is 50 km/h on all roads in the model.

Table 10: Yellow warning durations from CROW traffic signal handbook

Phase Type	Yellow
50 km/h through	3.5 s
Turning movements	3.0 s
Bicycle	2.0 s
Bus	2.0 s

5.4.2 Red clearance intervals

As described in subsubsection 3.8.1, red clearance intervals in the Netherlands are calculated per pair of phases. It would therefore be impractical to calculate all clearance times manually. Instead, the program OTTO was used to automatically calculate clearance times as per the CROW manual, using figure 17 and figure 18 for the intersection geometries.

5.4.3 Green timings

Since the objective is to test the traffic signal controllers' performance under nearly-saturated traffic conditions, a network cycle length of 120 seconds is used. According to observations by Movensis consultants, 120 seconds is a common upper limit for permitted cycle lengths among road authorities in the Netherlands. Increasing the cycle length is a way to increase the overall vehicle throughput of an intersection (Wolput et al., 2016), so in situations such as modelled scenario where demand is approaching capacity, intersections will commonly operate at the maximum cycle length permitted in the jurisdiction.

For reference purposes, signal phases are numbered as per the Dutch standard phase numbering. The phase numbers used in the model are illustrated in figure 19 for intersection 2. Intersections 1 and 3 are identical except that they lack the bus phases 45 and 51.

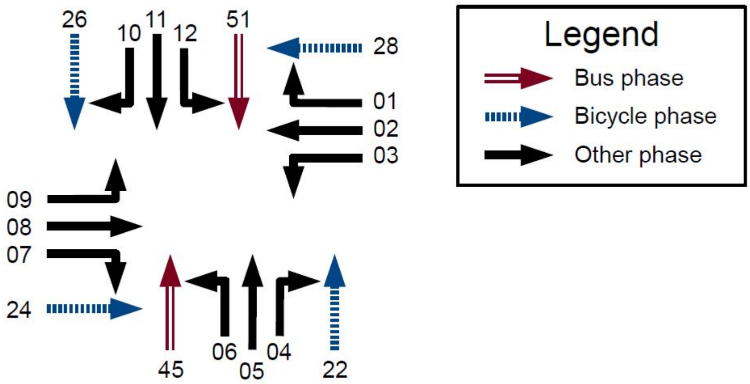


Figure 19: Phase numbers: Intersection 2

Where it is necessary to distinguish between phases at different intersections, the intersection number is added as a prefix. For example phase 103 is phase 03 at intersection 1.

Since intersection 2 is the capacity critical intersection, the timings are first locally optimised for that intersection, and then the timings at intersections 1 and 3 are generated such that a green wave is created in both directions.

The optimal phase order at intersection 2 is to use lagging left turns in all directions, in order to minimise the clearance time per cycle in the critical conflict groups. With the resulting clearance durations, the maximum green times for each phase in the critical conflict groups were selected as listed in table 11. All phase durations are identical for corresponding phases in opposite directions, to avoid creating a confounding variable which would affecting the validity of comparisons of traffic performance in opposite directions.

Table 11: Signal timing of critical phases at intersection 2

Phases	**Max Green**	**Yellow**	**Red Clear***
02/08 East-West Through	43 s	3.5 s	0.5 s
03/09 East-West Lefts	15 s	3.0 s	0.5 s
05/11 North-South Through	32 s	3.5 s	0.5 s
06/12 North-South Lefts	15 s	3.0 s	0.5 s

**Largest red clearance time relative to subsequent phases in critical conflict groups*

The east-west phases are given the longest green durations because they are the coordinated phases, and are intended to represent the primary traffic movements in the model. Maximum green durations of 15 seconds for left turn phases were selected to reflect the researchers' previous experience with busy left turn phases at comparable major intersections. The north-south movements at intersection 2 are intended to represent a relatively major busy road, so a duration of 32 seconds was selected - significantly longer than the left turn phases, but shorter than the east-west coordinated phases.

The signal coordination though the network is set up such that green waves in both directions pass simultaneously through the centre intersection (intersection 2), as per the locally-optimal phase structure indicated in table 11. The green periods for the coordinated phases at the three intersections are shown for a single cycle in figure 20.

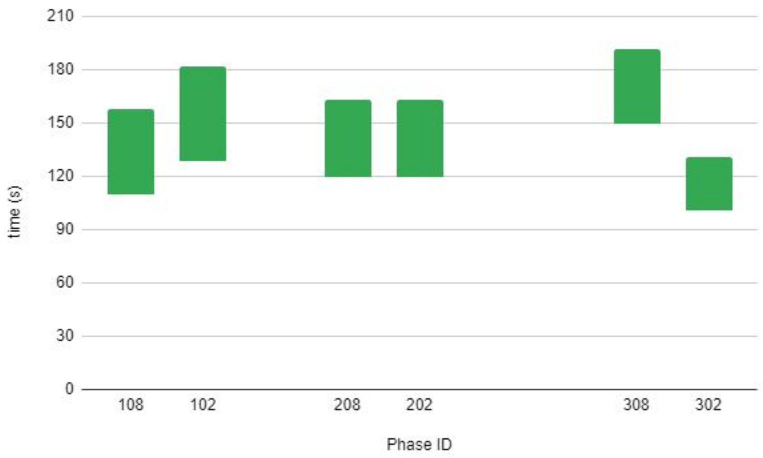


Figure 20: Green periods for coordinated phases

The phase structure at each intersection is displayed in figure 21. For a more detailed illustration of the signal timings, the step rasters for each intersection are also illustrated in Appendix B.

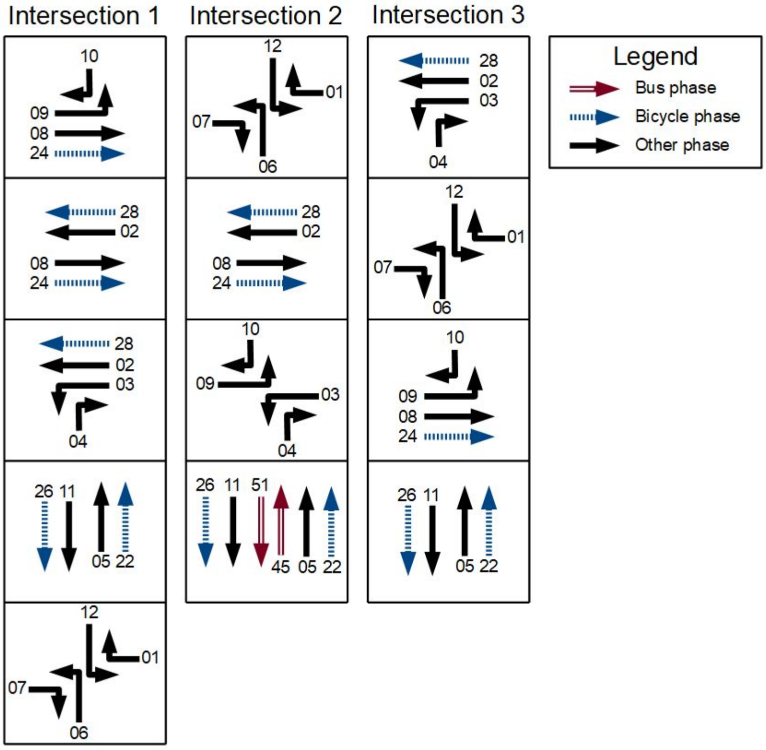


Figure 21: Phase structures at each intersection

At intersection 1, the green bands partially overlap as visible in figure 20, so the intersection is set up with lead-lag phasing, with leading left turns heading eastbound toward intersection 2, and lagging left turns arriving westbound from intersection 2. The remaining two sets of phases are arranged with lagging left turns, to minimise the clearance times in the critical conflict groups. To minimise the queuing in the short space between intersections 1 and 2, phase 08 is only permitted to be green during its primary realisation period.

Because the volume of straight-through traffic is higher at intersection 1 than at intersection 2, the durations of the coordinated phases are also longer. The precise duration and offsets of the coordinated phases were fine-tuned based on observation of the Vissim model.

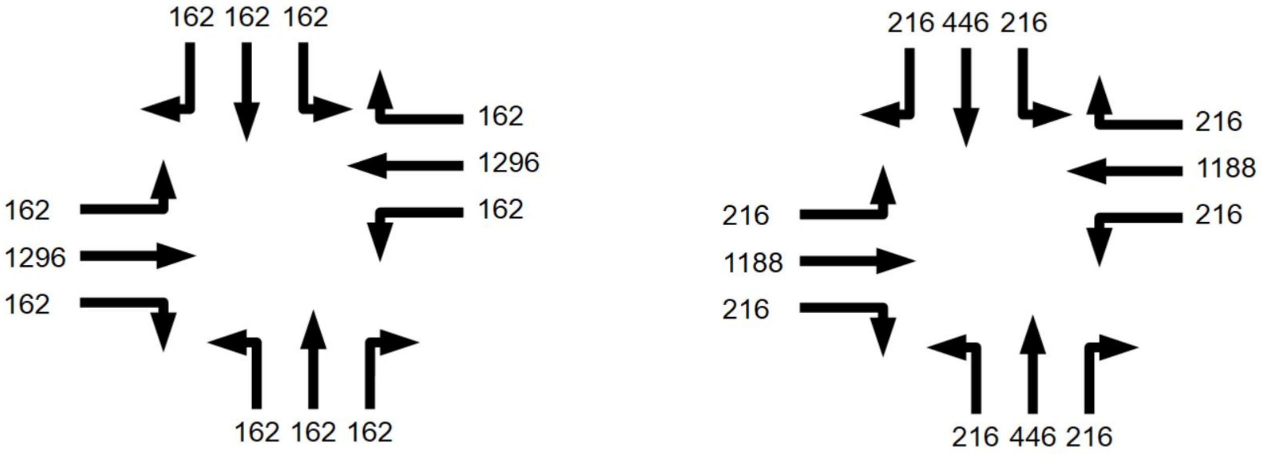
The phase structure at intersection 3 is unusual, because the green waves in the two directions arrive at the intersection during opposite parts of the cycle. The eastbound coordinated phase (08) is paired with the eastbound left turn phase (09), and the westbound coordinated phase (02) is paired with the westbound left turn phase. However, since the left turn volumes are relatively low, in practice much of the time during the coordinated phases is used by alternative realisations of the opposite coordinated phase. For this reason - as well as the need to fit the 2 other sets of phases in the cycle - the primary realisation periods for the coordinated phases at intersection 3 are *not* longer than those at intersection 2.

5.5 Traffic volumes

The desired traffic situation to study is where traffic volumes are approaching capacity. The intent is to see how the controllers cope with the disruptions to general traffic caused by TSP interventions. With lower levels of saturation, the difference between controllers would likely not be visible, since the green times would tend to be determined by vehicle actuation rather than by the step rasters.

The traffic volumes used in the model are indicated in figure 22. The volumes were determined by starting with an 80% v/c ratio at intersection 2 with the signal timings in table 11. The volumes were iteratively increased with the Base-90 controller until all phases failed to clear the queue during some but not most cycles.

All movements with private motor traffic are composed of 98% cars and 2% heavy vehicles.



(a) Intersections 1 and 3

(b) Intersection 2

Figure 22: Motor traffic volumes (vehicles per hour) per intersection

The volume-to-capacity (v/c) ratios of the phases in the critical conflict groups are approximately 0.9, as summarised in table 12. The saturation flow rates were based on the default values from the Highway Capacity Manual (Transportation Research Board, 2000), summarised in table 13. The effective green time (g_e) for each phase is one second longer than the displayed green time, since the duration of yellow in which vehicles still enter the intersection is one second longer than the time lost at the start of the phase as drivers react to the green light and start moving. The ratio of effective green (g_e) to cycle length (C) is multiplied by the saturation flow rates to estimate the actual capacity per approach.

Table 12: Saturation rates of phases in critical conflict groups

Phases	**Sat. Flow**	g_e/C	**Capacity**	**Volume**	v/c
EW Thru (02/08)	3800 pcu/h	0.37	1393 pcu/h	1212 pcu/h	0.87
EW Left (03/09)	1805 pcu/h	0.13	241 pcu/h	220 pcu/h	0.92
NS Thru (05/11)	1900 pcu/h	0.28	523 pcu/h	455 pcu/h	0.87
NS Left (06/12)	1805 pcu/h	0.13	241 pcu/h	220 pcu/h	0.92

Table 13: Assumed values for capacity calculations, from HCM2000

Item	**Value**
Base Saturation Flow	1900 pcu/h
Left Turn adjustment	0.95
Start-up loss	2 s
Yellow used	3 s

The above calculations are based on the green times specified in the step raster. However, since the intersection is fully-actuated, the actual green durations vary from one cycle to another. The calculated v/c ratios for left turn movements are slightly higher than for through lanes because with the selected phase order (with lagging left turns) they are more likely to receive extra time leftover from other phases. In order for a left turn phase to receive extra time, only the opposing through movement needs to end early due to lack of demand. In contrast, in order for a through movement phase to start early, *both* of the perpendicular left turn phases need to end early.

5.6 Bus line implementation

As described in subsection 3.7, the bus service in the model consists of a north-south bus line at intersection 2 with a frequency of 12 buses per hour in each direction. To avoid bias, the arrival of buses should be evenly distributed throughout the signal cycle. Having a disproportionate number of buses arriving during a particular portion of the cycle would over represent the TSP action which is best-suited to reduce delays during that portion of the cycle, and by extension controllers which are better suited for that TSP action. The arrival rate of buses should also be inconsistent, to test the controllers' ability to accommodate TSP requests at varying intervals from each other, including simultaneously.

Given these requirements, buses in the model cannot simply be generated at regular intervals every 5 minutes, instead there needs to be some randomisation in the headways. Within Vissim the two ways this could be done. The simpler option is to use built-in vehicle generation, which would generate buses at random intervals which average to 5 minutes. For each of the 10 model runs per scenario, a different set of pseudo-random headways would be automatically generated by Vissim using the run's random seed. The alternative option is to manually input a timetable which will generate buses at predetermined moments. In the latter case, the timetable would need to be designed separately (for example, using a spreadsheet) to introduce variability which includes inconsistent rates of TSP calls at the subject intersection while still evenly distributing bus arrivals throughout the signal cycle. A single timetable would be used for all 10 runs of each scenario, since manually changing the timetable per run would make running the total 60 runs of the study impractical.

The method selected to generate buses is the built-in pseudo-random vehicle generation of Vissim. Although it will produce a larger variation in headways than would be typical of a single bus line, the vastly simpler implementation enables much more of the project's time to be devoted to be other aspects of the study.

6 Results

The design of the study included 6 scenarios, with 3 signal TSP Systems each tested with two different levels of flexibility in green times. The scenarios are described in subsection 3.6 and the summary table 7 is reproduced below for convenience:

Table 14: Characteristics of the modeled scenarios

Scenario Name	**Base-80**	**Base-90**	**ParFlex-80**	**ParFlex-90**	**ResFlex-80**	**ResFlex-90**
TSP System	Base	Base	New	New	New	New
Traffic Response	N/A	N/A	Off	Off	On	On
Min Maximum Green	80%	90%	80%	90%	80%	90%
Min Green before truncate	80%	90%	80%	90%	80%	90%
Desired Green during OC	N/A	N/A	90%	95%	Varies	Varies

Scenarios named "-80" test traffic signal timings with "high flexibility" in the amount by which green times can be varied. Scenarios marked "-90" test low-flexibility signal timings.

The primary objective of any TSP system is to reduce the travel time for the prioritised public transport vehicles. But this alone does not necessarily indicate the performance TSP system, as any TSP system can trade public transport delay off with delay for other road users by increasing or decreasing the restrictions on priority interventions. The performance for other road users is thus also measured, in terms of delay and number of stops. Finally the performance for public transport and other road users are compared to determine the efficiency of the studied TSP systems.

6.1 Bus performance

The key metrics for TSP performance from the perspective of bus service are the bus delay, and the effects on bus service regularity.

6.1.1 Bus Delay

All of the tested scenarios used two-level headway-based conditional priority with a threshold of 300 seconds (5 minutes), which is the average headway for a service with 12 buses per hour. Buses with a headway less than 300 seconds are only eligible for green extensions, while buses with a headway greater than 300 seconds are eligible for green extensions, phase insertions and red truncations. As a result, the delay results will vary significantly from one bus to another, depending on its headway.

For the sake of simplicity, buses with a headway greater than 300 seconds are referred to as "late" in this paper, and buses with a headway less than 300 seconds are referred to as "early". Late buses would need to be sped up in order to restore a regular service, while early buses would need to be slowed down. In this respect, buses with long headways are treated similarly to buses on a timetabled route which are behind schedule. This simple method for determining which buses need to be accelerated is the typical way headway-based conditional priority is implemented in practice (Hounsell and Shrestha, 2012).

In figure 23, the travel time approaching the stop line is plotted for all 480 buses in the 10 runs of Scenario 1 (Base TSP system, 80% flexibility) relative to their headways. The two directions of bus traffic are not distinguished, since the roadway layouts are perfectly symmetrical, with a free-flow travel time of 18 seconds in both directions.

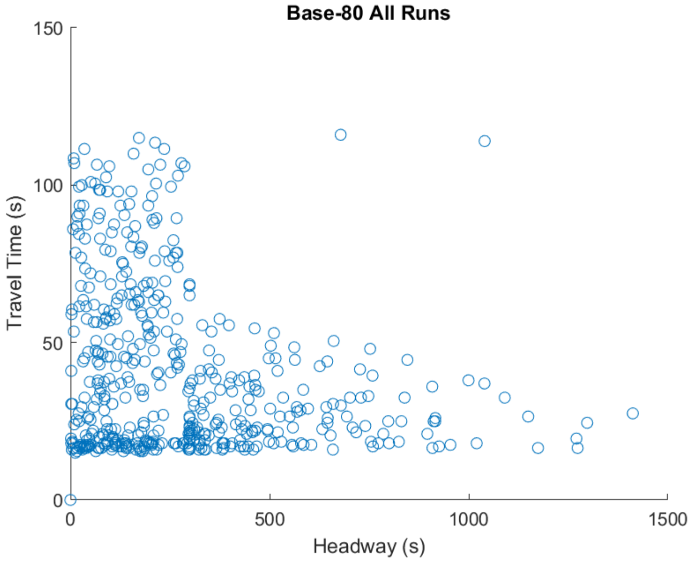


Figure 23: Travel time relative to headway, Base TSP system, high flexibility

The ideal situation from a public transport operator's perspective is that the delay experienced by buses would decrease as buses' headways increase - helping them to catch up to the target headway and provide a more evenly-spaced service. The results in Scenario 1 are close to that ideal. There is a clear distinction between buses with a headway above and below 300 seconds. Early buses take up to 120 seconds to cover the 260 metres approaching the intersection, which represents up to around 100 seconds of delay relative to the 18-second free-flow travel time. This delay occurs because those buses can only proceed during the normal signal stage for north-south traffic. Buses may thus need to wait through the entire duration of the other phases (approximately 90-100 seconds, depending on vehicle actuation) before they get a green light.

Late buses, on the other hand, rarely have a delay greater than 30 seconds. Those buses can insert additional north-south signal phases into the cycle, so as long as the TSP request meets the requirements for a phase insertion which were determined by the signal controller, the bus will never need to wait longer than the duration of one phase. The longest phases in the study intersection are the coordinated phases 02 and 08, which have a normal green duration of 43 seconds, which can be truncated to 34 seconds in Scenario 1 (and in other scenarios with high flexibility). Including the 3 seconds of yellow and 1 second of red clearance, the worst case scenario is that the bus needs to wait approximately 38 seconds for a green light.

Scenario 2 includes the same TSP system, but now with a lower flexibility to reduce phase durations. The relationship between headway and level of priority is now much less distinct, as illustrated in figure 24.

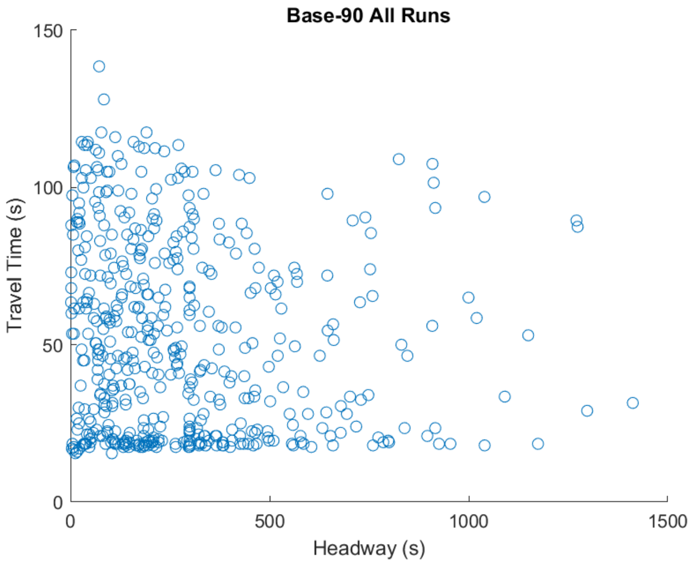
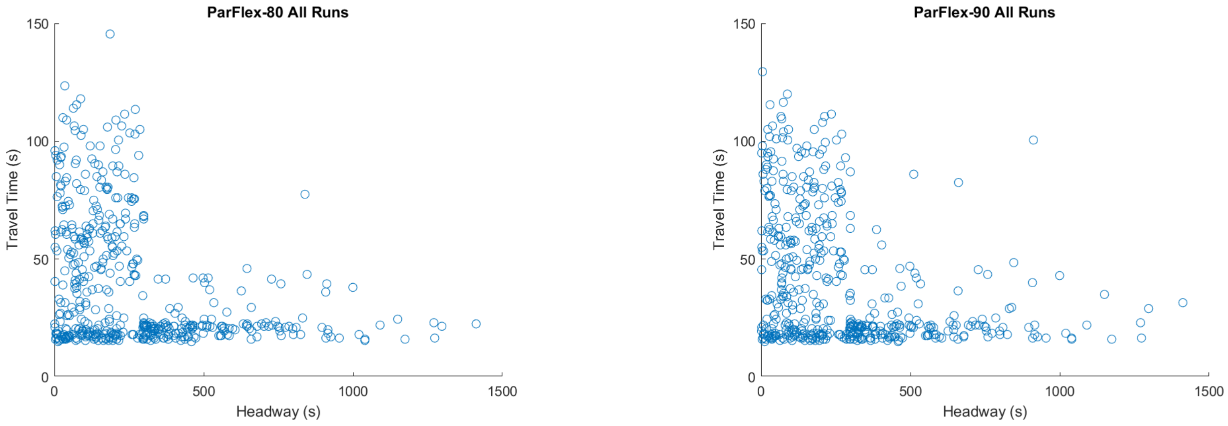


Figure 24: Travel time relative to headway, Base TSP system, low flexibility

Many late buses experience delays as long as 90 seconds even though they have a headway greater than 300 seconds and are therefore eligible for phase insertions. This occurs because the Base TSP system has a fixed cycle length and in this scenario it is not possible to reduce the maximum green durations for other phases sufficiently to fit an additional north-south signal phase in the cycle. An additional phase can only be inserted if other phases happen to end early due to lack of demand. Late buses do still experience lower delay than early buses, due to the occasional possibility of phase insertions, and the phase truncations which are always possible.

With flexible-cycle systems, the cycle length is less of a constraint for TSP actions. With the new parametric flexible-cycle TSP system (ParFlex), the effect of reducing the flexibility in the background signal cycle is much less severe, as illustrated in figure 25.



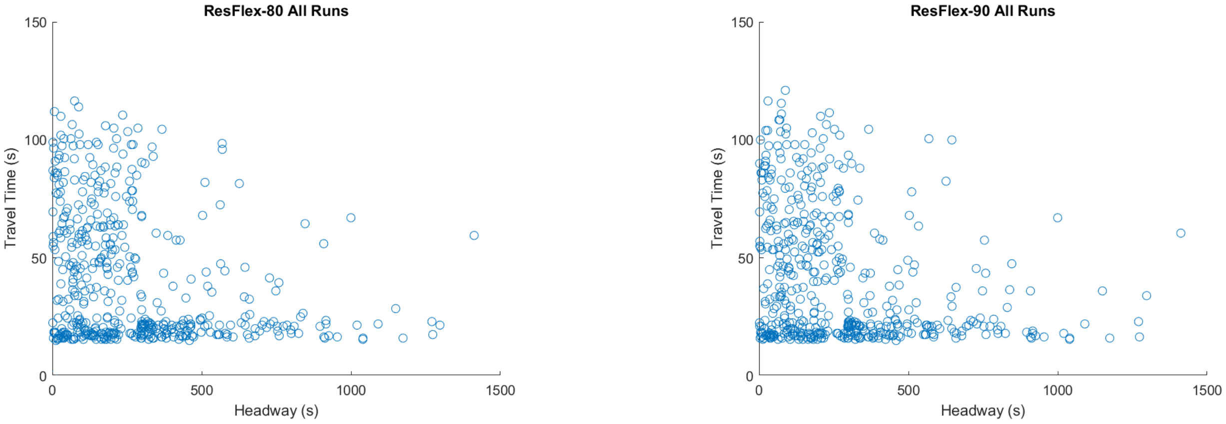
(a) High flexibility

(b) Low flexibility

Figure 25: Travel time relative to headway, ParFlex system

In this scenario, the only requirement for implementing a given TSP action is that it be possible to recover the time within two cycles. In both the low-flexibility and high-flexibility scenarios, the desired green durations during offset correction are predetermined, and in both cases it is possible to recover more than the duration of an inserted phase (approximately 10 seconds) over the course of two cycles. The only exception is if the TSP request occurs while the signal is still in offset correction from a previous intervention. In that case the combined offset impacts of an inserted phase and the pre-existing offset error may be too much to recover within two cycles, in which case a phase insertion would be denied. That scenario explains why a one late had a travel time of more than 50 seconds in the high-flexibility scenario, and several buses had a travel time of more than 50 seconds in the low-flexibility scenario. With lower flexibility, the signal will take longer to complete offset correction, and there is therefore a slightly higher chance that a bus arrives during offset correction from a previous TSP intervention. Bus delays will also be higher because of the stricter limits on phase truncation.

The responsive-parameter flexible-cycle TSP system (ResFlex) implements stricter requirements on priority requests, because the desired green times are constantly updated based on the current traffic volume. As the traffic volumes increase, the amount of time which can be recovered during offset correction will decrease, reducing the maximum durations for green extensions and phase insertions. Both scenarios using the ResFlex system had a larger number of late buses experiencing long delays than either of the ParFlex scenarios, as illustrated in figure 26.

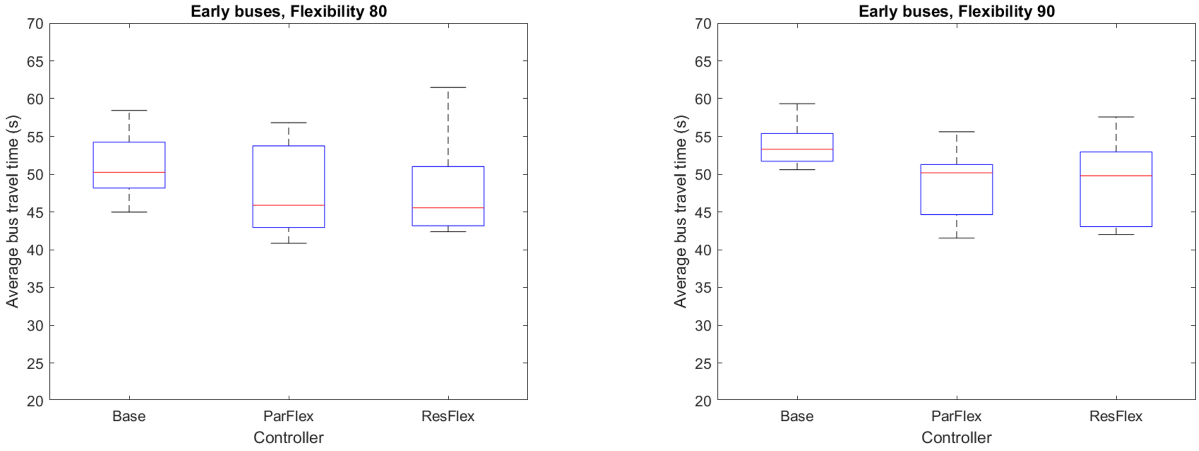


(a) High flexibility

(b) Low flexibility

Figure 26: Travel time relative to headway, ResFlex system

The delay for early buses in each scenario is illustrated in figure 27 as a box plot, where each box displays the distribution of average bus delays in the 10 simulation runs of the scenario.



(a) High-flexibility scenarios

(b) Low-flexibility scenarios

Figure 27: Bus travel times per scenario - early buses

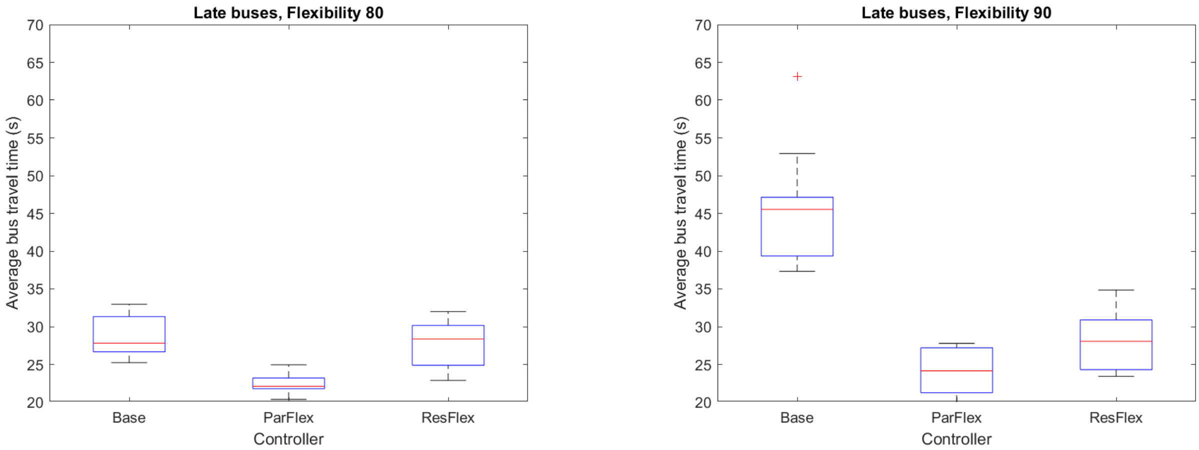
Table 15: Average delay per scenario: Early buses

	High-flexibility			Low-flexibility		
	Base	ParFlex	ResFlex	Base	ParFlex	ResFlex
Average bus delay	32.81 s	29.52 s	29.76 s	35.72 s	30.75 s	30.54 s
Change relative to Base	-	-10%	-9%	-	-14%	-15%
Paired T-test P-value	-	0.01	0.04	-	0.00	0.00
T-test significant @95%	-	Yes	Yes	-	Yes	Yes

With all TSP systems, early buses are not eligible for phase insertions, and must pass through during the primary realisation of the north-south green, a green extension of that green, or an alternative realisation of the north-south phases. However, due to the large traffic volumes in the tested scenarios, alternative realisations of the north-south phases at intersection 2 are very rare.

Regardless of flexibility, the delay for early buses is lower for the new TSP systems than for the Base system. Part of this difference is due to a difference in the way the bus phases are displayed. With the new systems, an ordinary phase call is sent along with the TSP request, so if the signal happens to already be in north-south green, the bus phase will change to green immediately. But with the Base system, the light does not change to green until a few seconds before the expected arrival, or when the bus reaches the far away detector 75 metres (6 seconds) from the stopbar. Since the light is not green while the buses are approaching, it is not eligible for a green extension to allow the bus to pass through after the normal end of the phase.

For scenarios with a high degree of flexibility, the flexible-cycle systems reduce the average delay for early buses by approximately three seconds, and with low flexibility, the difference is five seconds. The greater difference between systems in the low-flexibility scenarios is due to the limited availability of green extensions within the fixed cycle length of the Base system. The maximum duration of a green extension is limited to the amount that the other conflicting phases can be reduced, which is 8 seconds in those scenarios. With the flexible-cycle systems, the maximum green extension is roughly twice as long, because the time can be recovered over the course of two cycles, including from the phases parallel to the priority phase.



(a) High-flexibility scenarios

(b) Low-flexibility scenarios

Figure 28: Bus travel times per scenario - late buses

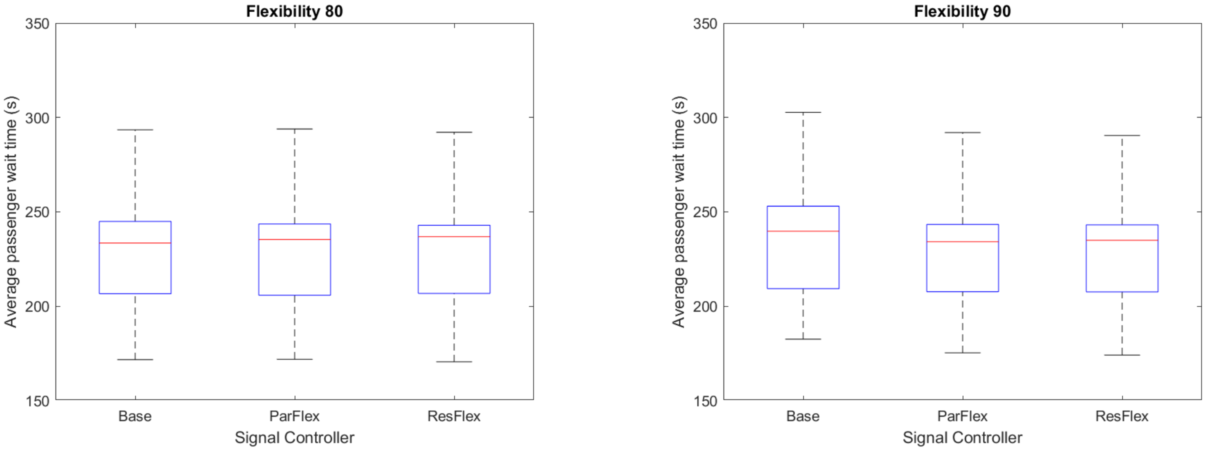
The average delay for late buses is calculated by subtracting the 18-second freeflow travel time from the observed average travel times. The resulting delays are displayed in table 16.

Table 16: Average delay per scenario: Late buses

	High-flexibility			Low-flexibility		
	Base	ParFlex	ResFlex	Base	ParFlex	ResFlex
Average bus delay	10.74 s	4.39 s	9.90 s	27.97 s	6.22 s	10.47 s
Change relative to Base	-	-59%	-8%	-	-78%	-63%
Paired T-test P-value	-	0.00	0.54	-	0.00	0.00
T-test significant @95%	-	Yes	No	-	Yes	Yes

6.1.2 Bus Reliability

The wait time for passengers is calculated for a bus stop located immediately downstream of the intersection, assuming a constant rate of passenger arrivals at the stop. The average waits for each scenario are consolidated for both directions of bus service, and displayed in figure 29.



(a) High-flexibility scenarios

(b) Low-flexibility scenarios

Figure 29: Average passenger waiting time at a stop downstream of the intersection

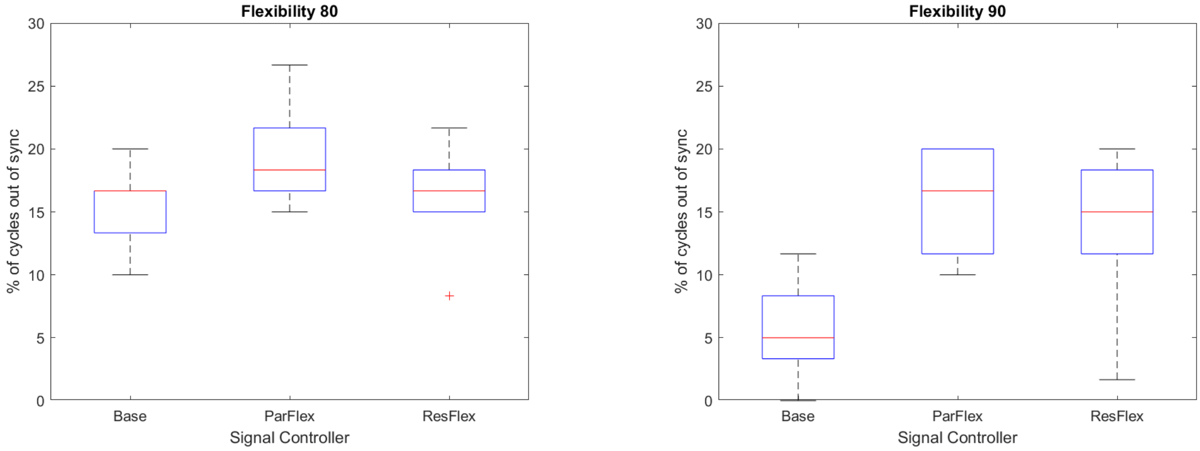
If the bus service were perfectly evenly spaced, the average wait would be 150 seconds (half of the 300-second headway). The much higher average waits observed in figure 29 are to be expected given that the buses in the model arrive at random intervals. Among the high-flexibility scenarios, there is no noticeable difference at all between the scenarios. For the low-flexibility scenarios, the flexible-cycle systems produce a slightly lower wait time, but the difference is not significant. The lack of significant difference can be attributed to the relative magnitudes of variability of bus headway, and the potential time savings through TSP. The range of delay for buses is generally between 0 and 100 seconds, but due to the random bus generation, many buses had headways hundreds of seconds larger than the target headway. Even in a best-case scenario where all early buses pass through without delay and all late buses experience 100 seconds of delay, a single traffic signal would not be able to produce an evenly-spaced bus service.

6.2 Coordination

The tested TSP systems' ability to maintain coordination are measured by their ability to maintain the programmed offset, the number of stops for vehicles in the coordinated directions, and the length of queues on intersection approaches with constrained storage space.

6.2.1 Ability to maintain coordination

In each scenario, the controller's ability to maintain the programmed start and end times for the coordinated phases is measured by the proportion of cycles where the one or both of the coordinated phases begin later than the intended coordinated moment. The percentage of late coordinated phase starts for each scenario are shown in figure 30.

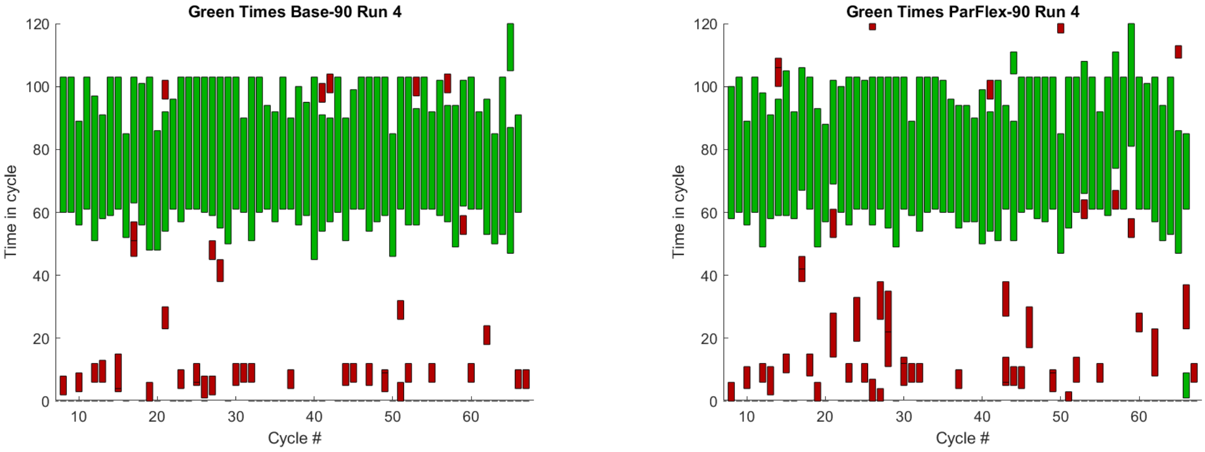


(a) High-flexibility scenarios (b) Low-flexibility scenarios

Figure 30: Proportion of cycles deviating from programmed offset

As expected, for a given level of flexibility, the scenarios with the Base system have considerably lower rate of late coordinated phase starts than the new flexible-cycle system. But despite having a fixed cycle length, some model runs with the Base system had as many of 20% of cycles where the coordinated phases started later than planned. This is due to TSP phases which were inserted between the previous set of phases (north-south left turns) and the start of the east-west coordinated phases, which necessitates taking time away from the coordinated phase. With lower flexibility, the number of late starts is significantly lower for the Base system, because is generally less likely that it will be possible to accommodate an inserted phase within the fixed cycle length.

As an illustration of the behaviour between the flexible-cycle systems and the Base system, the green moments for coordinated phase 02 are shown in green in figure 31 for model run number 4 for the low-flexibility scenarios with the Base TSP system and the parameteric flexible-cycle TSP system (ParFlex). The green moments for the bus phases (45 and 51) are shown in red.



(a) Fixed-cycle TSP system (b) ParFlex TSP system

Figure 31: Green moments for phase 02, low flexibility scenarios, run 4

The Base system has a fixed cycle length, and the green moments follow typical CRSV halfstarre behaviour for most cycles. The coordinated start time for the phase is at 60 seconds in these diagrams, but in a quarter of the cycles, the green actually started earlier, thanks to previous phases ending early due to lack of demand. Phase 02 always remains green until at least time 85, because it is programmed to remain green for that period regardless of vehicle demand. After time 85, it can end early if there is no longer any approaching traffic detected. The programmed end time for the phase is at 103 seconds in these diagrams, and if there is demand for the subsequent phase (eastbound left turns), phase 02 must end even if there is still approaching traffic detected.

Despite the fixed cycle length which prevents the timings from drifting away from coordination, there were two cycles in which phase 02 started later than its programmed start time. Both of those instances were due to TSP phase insertions immediately prior to the start of phase 02. In cycle number 18, both bus phases were inserted starting around time 50. Although their greens ended prior to second 60, their yellow and red clearance times resulted in phase 02 beginning 4 seconds late. In cycle number 59, a bus phase was inserted shortly before the start of phase 02, which resulted in phase 02 starting two seconds late.

With the flexible-cycle systems, there are more situations in which TSP can insert a bus phase outside of its primary realisation period. In this run, there were five cycles where phase 02 started significantly later than its normal start time. Four of those were due to phase insertions around the normal start time of phase 02 (in cycles 22, 53, 57 and 59), while the fifth (in cycle 18) was due to a pair of bus phases which had been inserted earlier in the cycle from which the signal had not yet recovered. The behaviour of the controller in cycle 59 is unexpected, where phase 09 is served prior to phase 02 rather than after it, resulting in phase 02 being much later in the cycle than normal. There are several such unexpected behaviours in the new flexible-cycle systems, and with continued development, the controller's ability to maintain the desired coordinated phase start and end times could be further improved.

In addition to the cycles where the start time of phase 02 is significantly late or early, there are numerous cycles where it starts a few second late or early as the controller has not quite returned in sync following an earlier TSP intervention.

6.2.2 Motor vehicle queues

In networks with closely-spaced signals, one of the reasons signal coordination is implemented is to manage queuing. If an upstream signal provides a green light which does not align with a downstream green light, a queue may build up between the intersections which exceeds the storage length. To identify such situations in the model, the maximum queue lengths for each 2.5 minute measurement period were compared to the available storage for the applicable lane(s). The combined results for all runs for each scenario are summarised in table 17.

Table 17: Percentage of measurements where a queue exceeded storage capacity

	Phase											
	West (short) link						**East (long) link**					
	101	**102**	**103**	**207**	**208**	**209**	**201**	**202**	**203**	**307**	**308**	**309**
Base-80	0%	6%	30%	0%	20%	68%	2%	11%	21%	0%	0%	2%
ParFlex-80	0%	8%	28%	0%	22%	71%	2%	3%	24%	0%	0%	1%
ResFlex-80	0%	6%	29%	0%	21%	69%	2%	4%	23%	0%	0%	2%
Base-90	0%	6%	28%	0%	13%	70%	1%	1%	20%	0%	0%	2%
ParFlex-90	0%	5%	28%	0%	22%	70%	3%	3%	21%	0%	0%	1%
ResFlex-90	0%	5%	28%	0%	19%	70%	2%	6%	21%	0%	0%	3%

The phases which most commonly overflow their storage are phase 103 (westbound left turn) at intersection 1, and phases 208 (eastbound through), 209 (eastbound left turn), 202 (westbound through) and 203 (westbound left turn) at intersection 2. The most common overflow by far is for phase 09 at intersection 2, which was expected. The storage length is severely limited due to the limited space between the intersections, and the turning traffic volume is relatively high given that intersection 2 represents the intersection of two major roads. This situation was knowingly included in the model to test how the TSP systems would affect situations with suboptimal storage lengths. In many cycles, the queue does briefly spill into the through traffic lanes near the end of the through signal phase, immediately before the start of green for phase 09, as seen in figure 32.

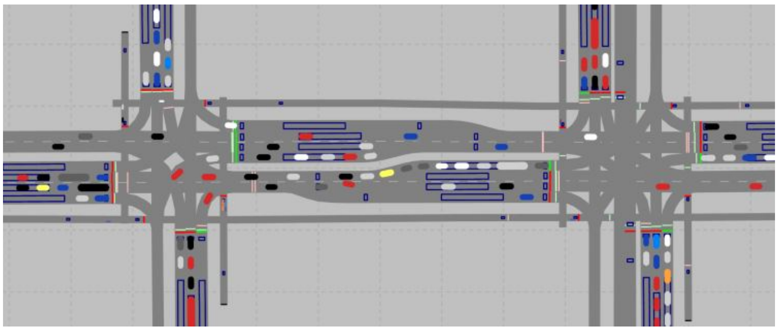
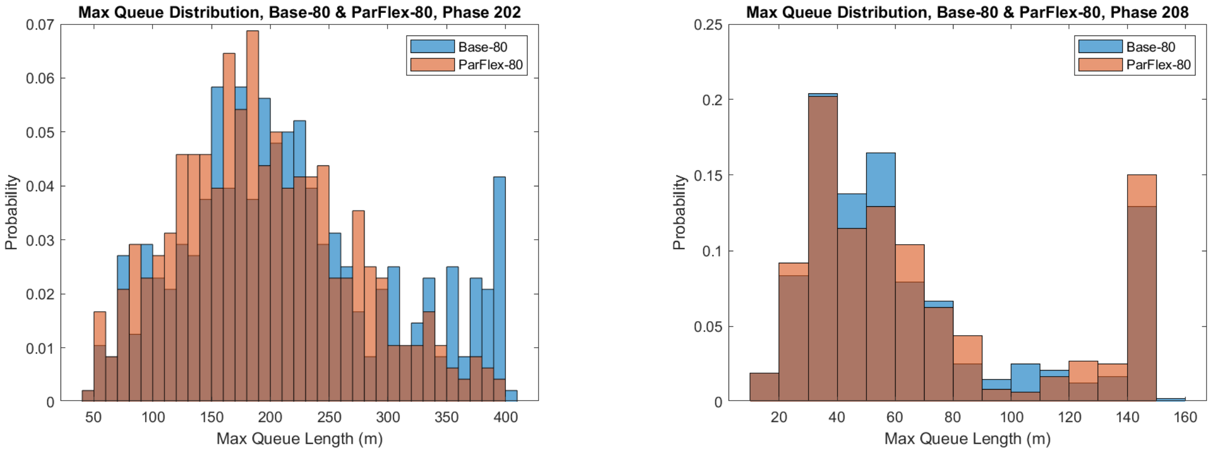


Figure 32: Queues spilling back from phases 103 (westbound left turn) and 209 (eastbound left turn)

It is notable that with a high level of flexibility, the flexible-cycle TSP systems result in an increase in queue overflows for phase 208, but a significant reduction for phase 202. This is counterintuitive considering that the TSP system will tend to allocate the same amount of time to each.

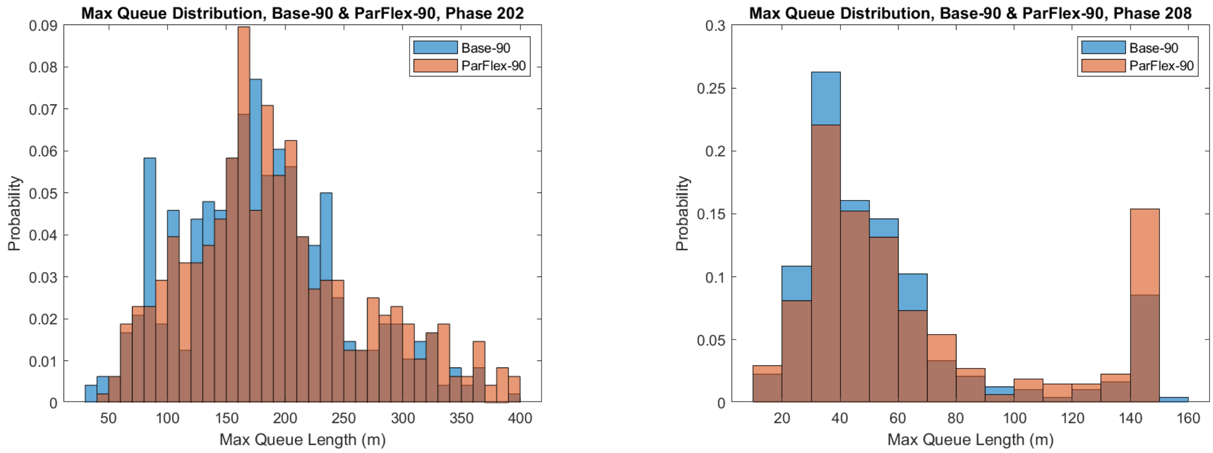
The inherent differences between those phases relate primarily to their distances from the previous coordinated signals. The approach link for phase 208 is very short (105 m), and the green durations and green moments of the upstream signal (phase 108) are deliberately limited to manage the amount of vehicle traffic arriving at intersection 2 during a red signal. Phase 202, on the other hand has a long (355 m) approach link, and the upstream signal does not strictly limit the volumes of traffic approaching during the red light at intersection 2. The distributions of queue lengths are illustrated in figure 33 for scenarios with a high intersection flexibility and figure 34 for scenarios with a low intersection flexibility.



(a) Phase 02 (westbound through)

(b) Phase 08 (eastbound through)

Figure 33: Distribution of queue lengths, Base and ParFlex systems, high flexibility



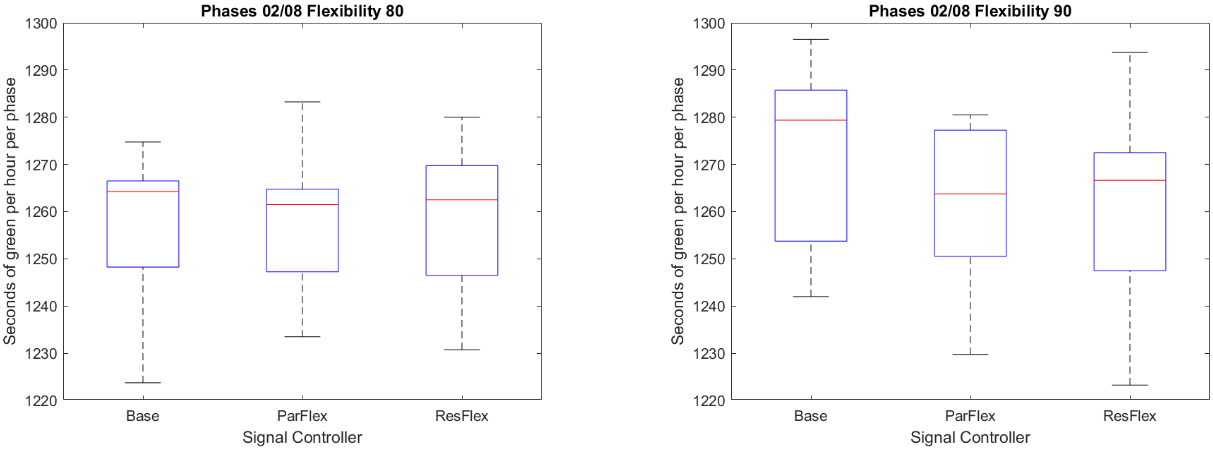
(a) Phase 02 (westbound through)

(b) Phase 08 (eastbound through)

Figure 34: Distribution of queue lengths, Base and ParFlex systems, low flexibility

With a high level of flexibility in green times, the parametric flexible-cycle (ParFlex) system produces fewer long queues for phase 02 than the Base system. This could be explained by the effect of the OC Credits, which will tend to compensate the coordinated phases for time they lose due to truncations and phase insertions, unlike in the Base system. The amount of displayed green time for the coordinated phases is slightly higher in scenario 3 than in scenario 1 (see figure 35) but the difference is not enough to definitively explain the reduction in queues. Note that scale in the figure does not begin at 0 - the difference between phases is minimal.

Meanwhile although phase 08 would benefit from increased green times following TSP actions, it is also highly sensitive to variations in the offset, due to the extremely close spacing from intersection 1. That may explain why the number of long queues slightly increases for phase 08, contrary to phase 02.



(a) High flexibility

(b) Low flexibility

Figure 35: Amount of green time served in coordinated phases

In the scenarios with a low degree of flexibility, the flexible-cycle systems produce longer queues for phase 202 than the Base system, which is the opposite effect observed in the high-flexibility scenarios. There are two main factors which contribute to this effect. In the low-flexibility scenarios, the flexible-cycle systems both provide a far higher level of priority to buses than the Base system, including far more TSP-inserted phases. Furthermore, although the flexible-cycle systems always redeem all *positive* OC credits during offset correction, they does not necessarily redeem all negative OC credits. If a phase's green duration can only be reduced by a small amount - which is often the case in the low-flexibility scenarios - the controller will reclaim time where possible, and delete any remaining negative OC credits once the controller returns in sync. This behaviour is also visible in the green times served in the coordinated phases (figure 35. With high flexibility, the amount of green time is consistent among the systems, but with low flexibility the Base system serves more time in the coordinated phases than the flexible-cycle systems.

6.2.3 Number of stops

Deviating the subject intersection from its normal offset will generally disrupt the green waves which were set up as part of the signal timing plan. As a result, it is to be expected that flexible-cycle systems have a larger number of stops for the coordinated motor traffic streams than fixed-cycle systems. The average number of stops per vehicle per scenario are listed in table 18 for each direction along the two road segments connecting the three intersections. The segment between intersections 1 and 2 is referred to as the "short link", while the segment between intersections 2 and 3 is referred to as the "long link". In each segment, the number of stops is only recorded for vehicles travelling straight through the intersection.

Table 18: Average stops per vehicle, per scenario

	High Flexibility			Low Flexibility		
	Base	ParFlex	ResFlex	Base	ParFlex	ResFlex
Long link, westbound	1.48	1.29	1.30	1.17	1.27	1.25
Short link, westbound	0.05	0.06	0.06	0.06	0.05	0.06
Short link, eastbound	0.24	0.26	0.26	0.19	0.25	0.24
Long link, eastbound	0.18	0.22	0.20	0.19	0.19	0.19

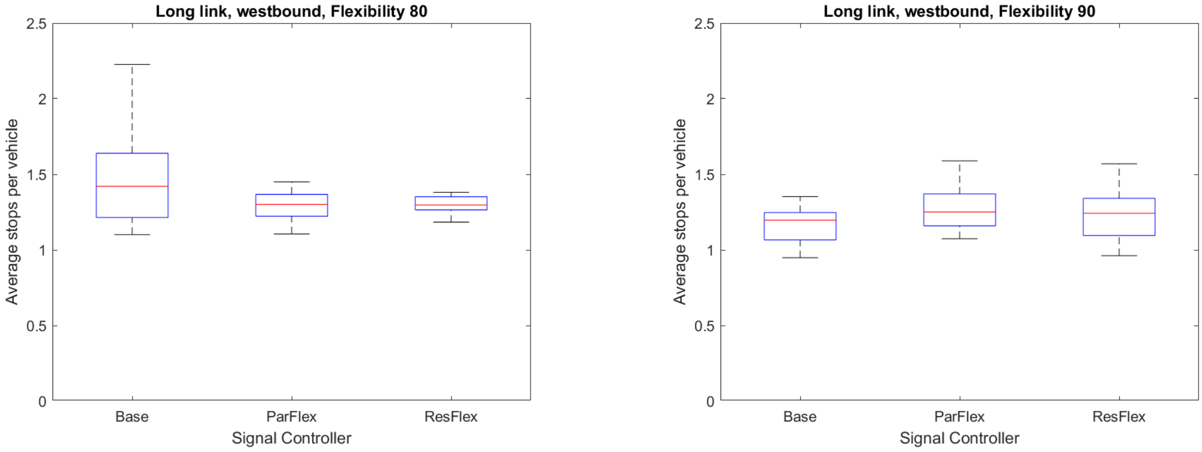
Table 19: Average stops per vehicle per scenario: Short link, eastbound

	High-flexibility			Low-flexibility		
	Base	ParFlex	ResFlex	Base	ParFlex	ResFlex
Average stops per vehicle	0.24	0.26	0.26	0.19	0.25	0.25
Change relative to Base	-	7%	5%	-	28%	23%
Paired T-test P-value	-	0.21	0.28	-	0.00	0.00
T-test significant @95%	-	No	No	-	Yes	Yes

Table 20: Average stops per vehicle per scenario: Long link, westbound

	High-flexibility			Low-flexibility		
	Base	ParFlex	ResFlex	Base	ParFlex	ResFlex
Average stops per vehicle	1.48	1.29	1.30	1.17	1.27	1.25
Change relative to Base	-	-13%	-12%	-	9%	7%
Paired T-test P-value	-	0.11	0.09	-	0.10	0.01
T-test significant @95%	-	No	No	-	No	Yes

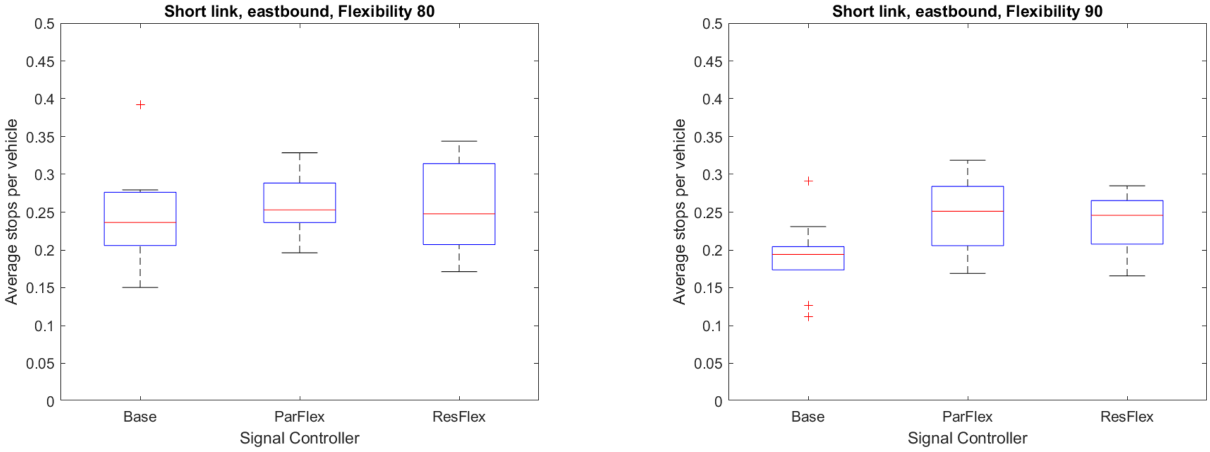
There is only one segment where there was a significant difference in the number of stops between the different systems: the long link, westbound approaching intersection 2. In the high-flexibility scenarios, the flexible-cycle systems had significantly fewer stops than the Base system, which is illustrated in figure 36. The reduced number of stops can be explained by the shorter queues, as discussed in subsubsection 6.2.2.



(a) High flexibility (b) Low flexibility

Figure 36: Average number of stops per vehicle, long link westbound

But even with the flexible-cycle systems, the number of stops in that segment is still very large - around 1.3 stops per vehicle. Vehicles are unable to travel uninterrupted along the green wave, because the entire duration of the westbound green at intersection 2 is spent clearing the queue. In the eastbound direction, queues approaching intersection 2 are minimised by restricting phase 08 at intersection 1 to its primary realisation period - it cannot receive extra time left over from other movements. This method is very effective, as visible in figure 37, where the average number of stops per vehicle is approximately 0.25 in all scenarios. Note that the scale is different than in figure 36. This indicates that at least 75% of vehicles passed through without stopping, even though phase 208 is only green 36% of the time. But this strategy would not work westbound due to the longer distance between intersections.



(a) High flexibility (b) Low flexibility

Figure 37: Average number of stops per vehicle, short link eastbound

The volume of through traffic westbound at intersection 3 is greater than the through traffic at intersection 2, requiring a longer green. As a result, it is not possible to accommodate the entire platoon departing intersection 3, and a portion of the platoon will need to be stopped and accommodated the following cycle. This is also true eastbound, but in that case a larger proportion of vehicles are able to pass through intersection 2 without stopping, thanks to the shorter distance between intersections, and thus less platoon dispersion. The platoon therefore makes more efficient use of the green time than would be possible if westbound traffic were held back at intersection 3 until the programmed primary realisation period for phase 02.

6.3 Delay for other road users

To observe the general effects of the systems' different TSP and Offset Correction strategies on road users other than transit users, the non-transit phases are grouped into four types based on how they are affected by TSP actions:

- Coordinated phases are the eastbound and westbound through phases (08 and 02) at all intersections. These phases are negatively affected by TSP actions, and are also negatively affected by variations in the relative timing between the intersections.
- Priority phases are the northbound and southbound through phases (05 and 11) at intersection 2 only. These phases are positively affected by TSP actions, and are not affected by variations in the relative timing between intersections.
- Bicycle phases are all bicycle phases. These phases may be positively or negatively impacted by TSP actions. They are combined into a single measurement to see if there is an overall effect on cyclists.
- Other phases are the remaining phases at all intersections. These are the all turning phases at all intersections, and the north-south through phases at intersections 1 and 3. These phases are negatively affected or not affected by TSP actions, and may be positively, negatively or not affected by variations in the relative timing between the intersections.

The average vehicle-delay per phase type is summarised in table 21, using a weighed average across phases based on the number of vehicles. Because each phase type only includes a single vehicle composition, the average vehicle-delay is also equal to the average person-delay.

Table 21: Average vehicle-delay per phase type - general traffic

	High flexibility			**Low flexibility**		
Delay	**Base**	**ParFlex**	**ResFlex**	**Base**	**ParFlex**	**ResFlex**
Coordinated Phases	28.0 s	26.6 s	26.4 s	24.8 s	25.9 s	26.6 s
Priority Phases	47.4 s	50.0 s	50.4 s	47.6 s	51.7 s	50.7 s
Bicycle Phases	35.0 s	34.7 s	34.5 s	34.8 s	35.5 s	35.3 s
Other Phases	40.0 s	40.0 s	40.2 s	39.1 s	40.1 s	40.0 s

All of the scenarios with a flexible-cycle system had a significant increase delay for motor vehicles parallel to the bus movements (table 22). This is by design, since the OC credits will always favour directions which conflict with the priority movement and is thus negatively affected by the TSP intervention.

Table 22: Average vehicle-delay: phases parallel to bus movements

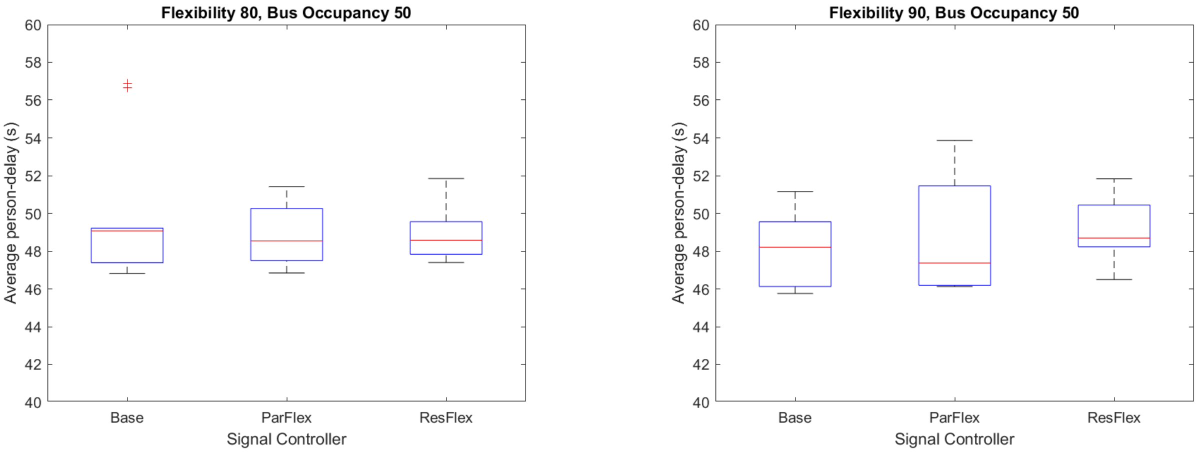
	High-flexibility			Low-flexibility		
	Base	ParFlex	ResFlex	Base	ParFlex	ResFlex
Average vehicle-delay	47.4	50.0	50.4	47.6	51.7	50.7
Change relative to Base	-	6%	6%	-	9%	7%
Paired T-test P-value	-	0.00	0.00	-	0.00	0.01
T-test significant @95%	-	Yes	Yes	-	Yes	Yes

There is no significant change in delay for bicycles between any of the scenarios. This is to be expected, because the primary variable affecting the amount of green time available for the bicycle signals is the volume of conflicting right-turning traffic - which does not vary between scenarios. Although the flexible-cycle scenarios do permit longer cycle lengths than the constant 120-second cycles of the fixed-cycle scenarios, the flexible-cycle systems always maintain the same average cycle length. So increased bicycle delay which occurs during cycles with longer cycle length will be partially offset by the reduced cycle length and delay during offset correction.

6.4 TSP System efficiency

Given that buses have a disproportionately high occupancy rate, reductions in delay for buses will have a disproportionate effect on the overall person-delay. As described in subsubsection 3.5.7, the calculated person-delay includes delays to late buses, but not early buses. To test the sensitivity of overall person-delay to the weighting of delays to late buses, the occupancy rate for buses is varied from 50 to 150 passengers per bus.

Using an occupancy rate of 50 passengers - which corresponds to a relatively full standard bus, the overall person delay per scenario is shown in figure 38.



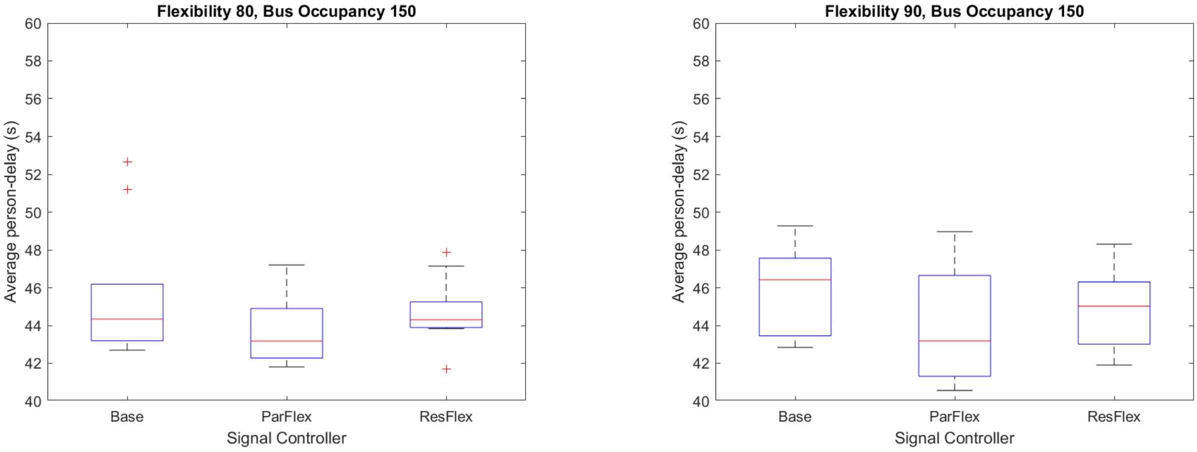
(a) High flexibility (b) Low flexibility

Figure 38: Average network person-delay, bus occupancy 50 people

There is no significant difference between the Base and flexible-cycle systems in any of the scenarios. This indicates that the significant reductions in delay to buses (subsubsection 6.1.1) are offset by the significant increases to delays to other road users (subsection 6.3).

With a bus occupancy rate of 100 passengers, the results are similar to with 50 passengers. No scenarios include any significant differences in overall person-delay.

An occupancy of 150 passengers artificially increases the weighting of delays to late buses above the actual number of passengers to reflect secondary delays and/or the road authority's desire to provide a higher level of priority to public transport. The average person-delay per scenario is illustrated in figure 39 and listed in table 23.



(a) High flexibility (b) Low flexibility

Figure 39: Average network person-delay, bus occupancy 150 people

Table 23: Average person-delay, bus occupancy 150 people

	High-flexibility			Low-flexibility		
	Base	ParFlex	ResFlex	Base	ParFlex	ResFlex
Average person-delay	45.6	43.9	44.7	45.9	43.9	44.9
Change relative to Base	-	-4%	-2%	-	-4%	-2%
Paired T-test P-value	-	0.11	0.41	-	0.01	0.25
T-test significant @95%	-	No	No	-	Yes	No

There continues to be no significant difference in person-delay for the high-flexibility scenarios, but with low-flexibility scenarios, the ParFlex system results in significantly lower person-delay than the Base system. This indicates that the time savings for passengers in late buses (78% reduction in delay) is sufficient to outweigh the increase in delay to other road users.

However, although the improvement for the ParFlex system is significant, the responsive-parameter ResFlex system does not show a significant improvement, despite generally providing lower impacts to other road users. This suggests that it is actually less efficient than the parametric TSP system.

7 Conclusions

7.1 Key Findings

The most notable findings of the research are broken down into two categories: answers to the research questions (subsection 1.3), and other unexpected findings.

7.1.1 Effects of relaxing cycle length during TSP

The primary research question was to determine how relaxing the cycle length constraint during and after TSP interventions affects the performance of a TSP system.

As expected, allowing the cycle length to vary during and after TSP interventions allows TSP systems to provide a higher level of priority to public transport vehicles than would be possible with a fixed cycle length.

Conversely, allowing the cycle length to vary also allows the signal's offset to vary, negatively impacting coordination with adjacent intersections. Along the link connecting the two closely-spaced intersections in the study, the variable cycle-length controllers had either no effect or negative effects on vehicles' delay and number of stops. Where intersections are closely spaced, the effect of varying the cycle length can be either positive or negative to conflicting traffic. Although the coordination is negatively impacted, the greater ability to redistribute green time to compensate for TSP impacts can reduce queue lengths and thereby reduce delay and number of stops for conflicting vehicles.

These effects are described in greater detail in the subsequent sections, which answer the research subquestions.

7.1.2 Effects on public transport performance

The first research subquestion was to determine how varying the cycle length affects signals' ability to promptly serve public transport vehicles.

Temporarily lifting the constraint of cycle length during TSP interventions and the subsequent offset correction can allow the delay for the prioritised public transport vehicles to be significantly reduced. At intersections where there is a high level of flexibility to shorten phase durations, the delay for prioritised vehicles can be reduced by as much as 59% compared to a fixed-cycle TSP system with the same TSP parameters. At intersections with a low level of flexibility to shorten phase durations, the difference is larger, with as much as 78% lower delay than an equivalent fixed-cycle system. Although it is typically possible for any given TSP to trade bus delays off against delays for other users, the level of priority provided by the flexible-cycle systems is beyond the limits of what could be provided with a fixed-cycle system. The fixed cycle length makes it impossible to implement many of the TSP actions which were implemented by the flexible-cycle controllers. Typically the level of priority at a given intersection can be increased by permitting more types of TSP action, and by increasing the permitted duration of those actions. But the TSP parameters used in the study already permitted a full range of actions to reduce delay (green extension, red truncation and phase insertion), and the durations of those actions in the Base system were solely limited by the cycle length and the minimum actuated green durations dictated by the scenario (either 80% or 90% of the normal maximum green time). It would have been possible to increase the level of priority in the Base system by reducing the minimum actuated green duration, but those minimums were included in the study to represent real-life constraints such as long pedestrian clearance times which cannot be shortened.

The study tested a headway-based conditional TSP system, which in theory should improve service regularity by disproportionately reducing delay for buses which have a longer headway. However, the study was unable to identify any significant difference in service regularity, due to the large variation in headways among the arriving buses, and the fact that there was only a single TSP-equipped intersection along the bus route in the model.

7.1.3 Effects on signal coordination

The second research subquestion was to determine how varying the cycle length affects the benefits of signal coordination for vehicles.

The new flexible-cycle controllers tested in the study aimed to return in sync within two cycles of offset correction. With that objective, the New controllers were able to maintain the programmed green band for at least 75% of cycles in sync in all model runs, with an average rate of 80% to 85% depending on the scenario (see subsubsection 6.2.1). This indicates that in situations such as the one tested where the adjacent intersections have a fixed cycle length, a green wave can be provided for the vast majority of cycles even with a flexible-cycle TSP system and as many as 12 buses per hour per direction.

However, the cycles in which coordination was disrupted did have negative effects on traffic performance, particularly on the shorter (150 metre) roadway link. With a low degree of flexibility in green times (and thus a low level of bus priority with a fixed cycle length), the flexible-cycle system significantly increased the number of cycles where a queue exceeded storage, and significantly increased the number of stops per vehicle by 23% to 28% for eastbound vehicles on the short link.

7.1.4 Situations where flexible-cycle TSP is beneficial

Given the conditions in which the flexible-cycle systems performed the best relative to the Base controller, flexible-cycle TSP systems appear to be most suitable at intersections with relatively long distances to adjacent signals, and they provide the largest improvement to public transport service at intersections where the constraint of a cycle length prevents the controller from implementing some TSP actions.

The study results indicate that a flexible-cycle TSP system can significantly reduce public transport delay without significantly increasing vehicle delay, number of stops or queue lengths under the following conditions:

Table 24: Criteria in which flexible-cycle TSP systems are most beneficial

Criterion	**Ideal condition**
Network control type	Fixed-cycle coordinated
Public transport direction	Perpendicular to coordination
Public transport frequency	Up to 12 buses per hour per direction
Priority requests	50% of buses granted high priority
Intersection spacing	400+ metres
Minimum green times	constrained
Traffic volume	v/c up to 0.9

The study's findings are limited to the network configuration used in the model, which is a fixed-cycle coordinated network with public transport travelling perpendicular to the coordinated directions. In networks where the public transport line operates along the coordinated direction, adjacent intersections may both have variable offsets, unlike this study where only a single intersection varied while both adjacent intersections maintained a fixed cycle length. The situation with buses travelling along the coordinated direction has already been studied by Zeng et al. (2021), who found that removing the constraint of cycle length vastly improves the system performance for their novel system which determines TSP actions based on the estimated progression of the prioritised vehicle through the coordinated network.

Even with the relatively high frequency of bus service in the model, with 12 buses per hour in each direction, signal coordination was maintained during at least 75% of cycles (subsubsection 6.2.1). This is partly due to the use of headway-based conditional priority, which only grants a high level of priority to roughly half of buses. Similar results can also be expected with schedule-based priority, as long as only half of buses are granted strong priority.

The roadway link with a length of 150 metres experienced a significant increase in vehicle stops and queues exceeding storage, while the link with a length of 400 metres did not significantly worsen in any measure. As a result, it appears that link lengths greater than 400 metres are suitable for flexible-cycle TSP systems. It is possible that link lengths between 150 metres and 400 metres would be acceptable, but that is not possible to discern from this study.

The most dramatic results in the study occurred in the scenarios with a low level of flexibility to reduce phase durations. This is largely because TSP with a fixed cycle length has a low effectiveness in that situation. With a higher level of flexibility there is less difference in performance between the fixed-cycle and flexible-cycle system, because many more TSP actions are possible despite the fixed cycle length. In such situations, the choice between fixed-cycle and flexible-cycle systems becomes a matter of preference: fixed-cycle systems maintain the coordinated timings more consistently (e.g. maintaining green waves), but the flexible-cycle systems distribute can reduce queues by redistributing green time to compensate for TSP actions.

The study included only a single level of traffic demand, which is a nearly-saturated condition with a volume-to-capacity (v/c) ratio of approximately 0.9. This results in some severe queuing at the critical intersection regardless of the TSP system being used. The results of the study therefore apply primarily to networks with very heavy traffic demand. In networks where the critical intersection has a lower level of saturation, both fixed- and flexible-cycle TSP systems will provide better traffic performance. The responsive-parameter TSP systems will also provide lower delays for buses, since the constraints on offset correction will automatically reduce along with the vehicle volumes, and by extension so will the constraints on TSP actions.

The net result of implementing flexible-cycle TSP in these circumstances is that the public transport delay can be significantly reduced without any increase in overall network person-delay, and without any significant increase in delay or number of stops for the coordinated movements.

7.1.5 Other findings

In addition to the answers to the research question and subquestions, the study produced an unexpected result, which was that the responsive-parameter flexible-cycle TSP system (ResFlex) parameters failed to outperform the ParFlex system. This is surprising because the ParFlex version uses a much more simplistic approach for calculating desired green times and OC credits which lacks the real-time observations of traffic volumes and green time utilisation which are used in the ResFlex system.

Across the scenarios the ResFlex system displayed superior performance for other traffic, including lower delays, shorter queues and fewer stops. But this difference can be explained entirely by the lower level of priority it provides to buses. If the reduced impacts were due to an improvement in TSP system efficiency there would be a corresponding improvement in person-delay, but instead the opposite is true. Regardless of the assumed bus occupancy rate, the ResFlex system produced *higher* person-delay than the ParFlex systen.

A possible explanation of the lack of performance for the more advanced version is that the system is based on assumptions such as saturation flow rates and green utilisation rates which were not calibrated to reflect the conditions in the model. So although the more intricate calculations provided a more precise estimate of the optimal green times, the lack of calibration meant that those estimates were no more accurate than the parametric version. The limitations in the implementation of the responsive-parameter version of the controller are further discussed in subsection 7.2.

7.2 Study Limitations

The results of the study come with caveats resulting from the way the study was conducted. These are broken down into limitations by design, and by implementation. Limitations by design consist primarily of the scope of the study, while limitations in implementation result from decisions made while developing the new controllers and the microsimulation model.

7.2.1 Limitations by design

The use of microsimulation provided a very detailed look into the behaviour of the traffic signal controllers and traffic system, but since each scenario is set up manually, the number of variables which could be effectively studied was limited. For practicality reasons, only a single level of traffic volume was used, which was selected to approach the capacity of the critical intersection in the model. But with lower traffic volumes, the constraint of cycle length will have a different effect. It is expected that with lower traffic volumes there would be less difference between fixed- and flexible-cycle TSP systems, since phases are more likely to end before their maximum green duration, which results in extra time being available for other purposes such as inserting TSP phases or providing extra time to phases where queues have built up.

The study only included a single signalised intersection along the bus route in the model. It is possible that with a larger string of intersections with conditional TSP, the cumulative effects of headway-based conditional TSP interventions on wait times would be able to produce a significant improvement in service regularity.

The study was conducted with Dutch traffic signal controllers, which use extensive vehicle actuation, including multiple detectors in each lane, and no fixed-time phases. Vehicle actuation acts as a passive form of compensation which may already provide some of the benefits of a flexible-cycle TSP system. If a movement has received more time than necessary thanks to a parallel TSP intervention, the queue for that phase will tend to decrease, increasing the likelihood that it ends prior to its maximum green time and thus makes extra time available for other movements. This means that the study's results are primarily applicable in jurisdictions where coordinated signals include significant vehicle actuation as in the Netherlands. In systems where some phases are fixed-time, as common in North America, there may have been a larger net benefit of actively compensating for green time redistribution.

7.2.2 Limitations due to study execution

As noted in subsubsection 6.1.1, there is a slight discrepancy in the way in which bus signals behave between the Base and New controllers. In the Base controller, the signal remains red at least until the bus reaches the vehicle detector 75 metres upstream of the stop line, whereas in the New controllers, the light could turn green as soon as the TSP call is submitted, 200 metres upstream of the stop line. The result is that green extensions are virtually impossible in the Base controller, since there is no green light to extend as the bus is approaching. This primarily affects early buses, since green extensions are the only TSP action permitted for them. This is not an issue for late buses, since they can insert transit phases outside of their normal primary realisation periods, including immediately following the normal period, which achieves the same effect as a green extension. Given that the person-delay calculations only included delay from late buses, the effect of this difference in behaviour between the Base and New controllers is expected to be minimal.

As noted in subsection 5.6, buses are generated at random intervals which average to 5 minutes per direction. This produces an extremely irregular bus service, far more so than such a bus service would typically be, especially with a dedicated busway. A more regular service would have a slightly different effect on the traffic signal operations, with a lower likelihood of multiple buses passing through during the same cycle. This might reduce the delay per bus with the New controllers, since a bus arriving during another TSP intervention or offset correction is more likely to have TSP actions blocked due to the combined impacts being too much to recover within 2 cycles.

Despite the extensive testing regime described in subsection 4.4 in which countless bugs were identified and resolved, the new flexible-cycle TSP systems still contain some bugs. One such bug is noted in subsubsection 6.2.1, where the coordinated starts significantly later than intended, due to alternative realisations being permitted during the portion of the cycle normally reserved for the coordinated phases. An even more severe instance of this bug, or possibly a different bug, is that bus phases inserted around the normal end of the coordinated phases' primary realisation periods sometimes cause the signal to deviate exceedingly far from the coordinated timings. For example, in cycle 46 in figure 40, a bus phase is inserted around time 90, truncating the end of phase 02 and itself ending at time 100. All of the time truncated from phase 02 is returned as OC credits at a 1:1 rate, so the cycle length increase of this action should be equivalent to the duration of the inserted phase (roughly 12 seconds including yellow and clearance times). Yet the following cycle, phase 02 begins 32 seconds late, which is far more than expected. It then takes another cycle of offset correction before returning in sync. The cause of this excessive disruption is not known and would require further investigation.

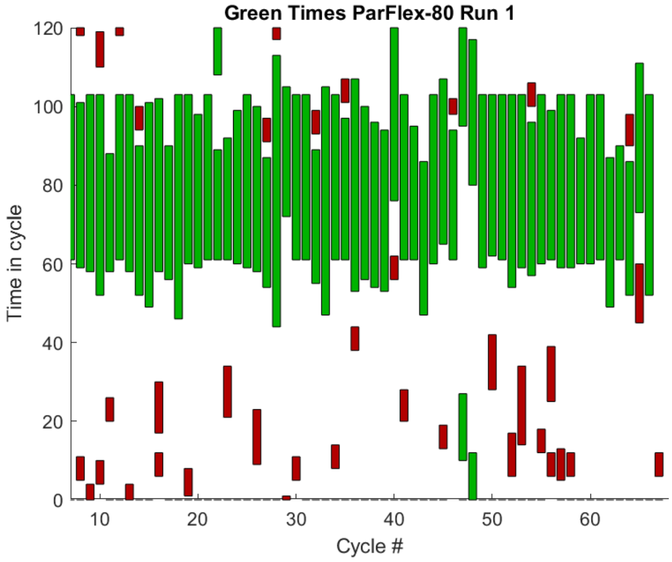


Figure 40: Green moments, Parametric Flexible-cycle TSP System, high flexibility

The effect of these bugs on the study's results will be artificially poor performance for the flexible-cycle TSP systems. The bugs result in larger-than-necessary deviations from coordination for the coordinated phases. These larger deviations will result in longer queues due to vehicles arriving at intersection 2 during a red light, and will also lengthen the duration of offset correction. It is possible that the bugs also cause an incorrect green time distribution between phases. That in turn would positively affect some phases and negatively affect others. It is extremely that the bugs would produce an improvement in overall person-delay, because they will produce more arbitrary phase durations than the background signal timings or the modified timings which should have been produced by the Offset Correction mechanism.

The lack of benefits of the more advanced responsive-parameter version of the new TSP system over the simpler parametric version can be traced back to limitations in the setup of the ResFlex controller. In that more advanced controller, the desired green times and earned offset correction credits are calculated based on proven traffic engineering formulas, but doing so involves many assumptions which need to be validated. For example, the calculation for desired green time makes use of a saturation flow rate, which is calculated according to empirical formulas. But if the actual saturation flow rate differs from the assumed value, the controller will systematically overestimate or underestimate the amount of green time required to accommodate the measured traffic demand.

Ideally, when implementing such a traffic-responsive system, the assumptions such as saturation flow rate and green utilisation for inserted phases would be validated based on the actual conditions, either using a manual study to calculate values, or by using the vehicle detectors to automatically calculate those values in real time. It seems that the ResFlex system achieves a 'worst of both worlds' situation where it is more complicated than the ParFlex system, but not complicated enough to provide the more accurately-tailored green times which should theoretically be possible given the greater amount of traffic information being considered.

Compounding the limitations of the new controller itself, was a lack of internal status monitoring while running the studied scenarios in the Vissim model. During development of the new controller, its functionality was extensively tested in a simple test environment where detector calls could be placed. This testing checked functions such as correctly counting vehicles using the detectors, the assigning correct green durations given the internal variable statuses, and approving the correct forms of TSP action for a given set of conditions. But due to the indirect way in which the controller was connected to Vissim, there was no way of monitoring those variables and conditions in the final model runs, which makes it very difficult to explain or troubleshoot unexpected behaviours. This is less of a problem for the ParFlex system, because its behaviour is primarily based on conditions which are visible in the simulation, namely the green durations, signal offset and detector statuses.

If the new TSP systems were to be implemented on-street, the TSP modules would need to be updated to transmit key variables to the controller's user interface, as with other modules in the controller. In particular, the current maximum TSP extension duration for each TSP phase would be very helpful in troubleshooting the controller's decision to permit or deny TSP actions. It would also be helpful to transmit the current desired green times, planned green times and OC credits.

7.3 Practical Application

Although the criteria listed in table 24 may appear to be highly restrictive, there are groups of signals in the Netherlands which meet all of them. At these locations, implementing a flexible-cycle TSP system would allow the delay for public transport vehicles to be significantly reduced, with only modest impacts on other road users.

For example, the Western Ring Road (Westelijke Randweg) of Haarlem uses a CRSV Halfstarre control system with a fixed cycle length, has intersections more than 700 metres apart, has heavy traffic volumes and as a result relatively little flexibilty in green times, and has three bus routes crossing it at Zijlweg with a combined peak period frequency of 6 buses per hour per direction with irregular scheduled headways.

7.4 Future Research

The study provided a detailed look at one traffic situation in which a flexible-cycle TSP system could be beneficial, and found that such a system would indeed be useful if a road authority wishes to increase the level of priority for public transport beyond the limits of a fixed-cycle system, without increasing the overall delay for road users at the intersection.

Having now demonstrated that the theory of flexible-cycle systems is effective in at least one situation, the next steps for research would be to determine how the effectiveness varies across different types of networks and traffic situations. In the network and traffic conditions of the study, the flexible-cycle systems worked well with an intersection separation of 400 metres - in some cases reducing reducing delay not only for buses but also for the coordinated traffic directions. But the flexible-cycle systems did not work well with a separation of 150 metres - consistently reducing performance for other traffic. Future study could determine with more precision where between these two distances the flexible-cycle systems start to work poorly.

This study tested a constant nearly-saturated traffic volume, but future studies could also examine how flexible-cycle TSP systems compare to fixed-cycle systems with lower traffic demand, as well as with varying traffic demand. In a study where the vehicle volumes change over the course of the study period, the responsive-parameter flexible-cycle system may display advantages which were not visible in this study with a steady traffic volume. Its ability to vary the level of priority based on real-time traffic volumes may allow the traffic performance to remain acceptable while providing as much priority as possible.

A Appendix: Terminology

A.1 Signal System Components

Road Authority	*wegbeheerder*
Road User	*weggebruiker*
Stop line	*stopstreep*
Control System	*verkeersinstallatie*
Controller	*regelcomputer*
Cabinet	*kast*
Stopbar loop	*koplus*
Setback loop	*lange lus*
Faraway loop	*verweg lus*
Critical Intersection	*maatgevende kruispunt*

A.2 Signal Timing Elements

Control Structure	*regelstructuur*
Cycle Length	*cyclustijd*
Stage	*blok*
Phase	*fasecyclus*
Movement	*richting*
Minimum Green	*vastgroentijd*
Maximum Green	*maximumgroentijd*
Yellow Change Interval	*geeltijd*
Red Clearance Interval	*ontruimingstijd*
Passage Time	*hiaattijd*
Clock Cycle	*machineslag*

A.3 Signal Control Types

Vehicle-actuated	*voertuigafhankelijk*
Fixed-time	*star*
Traffic-adaptive	*verkeersafhankelijk*
Traffic-responsive	*programmakeuses op basis van intensiteiten*
Coordinated	*gecoördineerd*

A.4 TSP Elements

TSP Request	*prioriteitsaanvraag*
TSP Intervention	*prioriteitsingreep*
TSP Action	*prioriteitsacties*

A.5 TSP Actions

Extension	*verlengen*
Truncation	*afkappen / afbreken*
Insertion	*prioriteits realisatie*
Rotation	*aangepaste blokindeling*

B Appendix: Intersection Parameters and Variables

B.1 Step Rasters

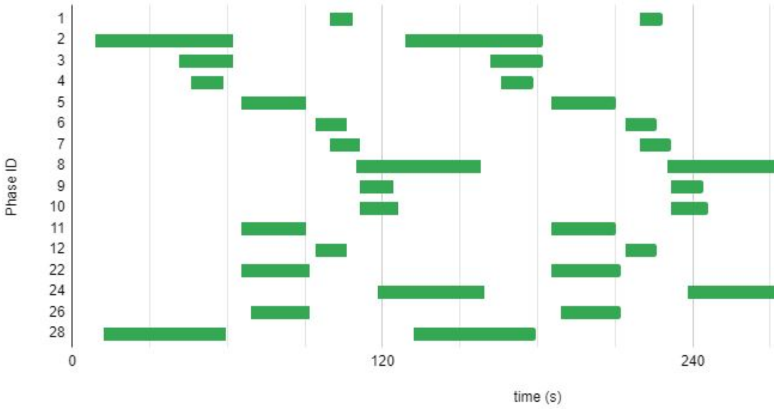


Figure 41: Step raster, intersection 1

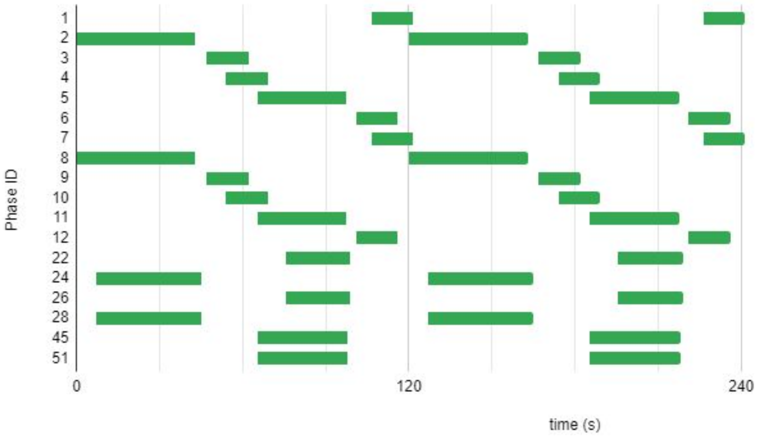


Figure 42: Step raster, Intersection 2

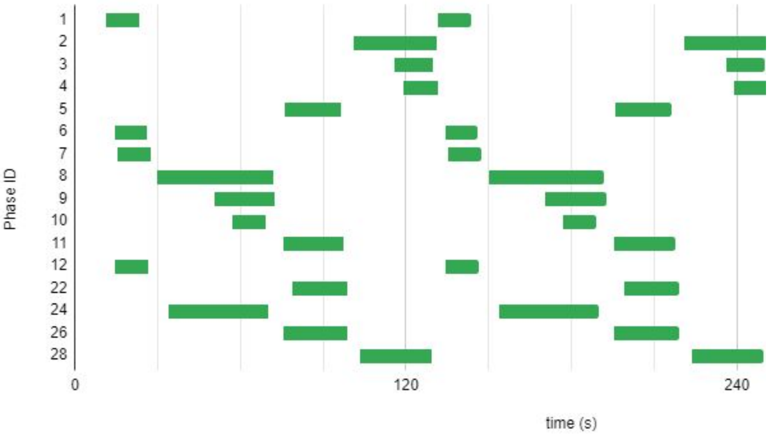


Figure 43: Step raster, intersection 3

B.2 Internal variables

The **offset error** "eOC" for the intersection is the amount of time that the signal is currently deviating from the normal coordinated operation. This represents the amount of time which needs to be recovered through offset correction. This is updated once per cycle when the synchronisation pulse is received from the master controller. A positive value indicates that that the signal is later than in its intended offset (so the cycle length should be decreased), and a negative value indicates that the signal is early (so the cycle length should be increased).

The **spare green** "$tRdc_i$" for phase *i* is the basic amount of time by which a phase can be reduced during the offset correction process. It is defined as the lesser of the reduction based on the minimum reduced green time and the reduction based on the capacity requirement.

$$tRdc_i = gMax_i - gRdc_i \tag{B.1}$$

where:

$gMax_i$ = *maximum green time for phase i, in seconds*
$gRdc_i$ = *minimum reduced green time for phase i, in seconds*

It is worth noting that the phases served during a TSP action can be reduced by more than the spare green during Offset Correction, since they can be further reduced by the utilised portion of TSP extension or insertion time.

The **OC adjustment time** "tOC_{ik}" for a phase *i* during a cycle *k* is the amount by which the phase's maximum green time is adjusted during offset correction. A negative adjustment value produces a shorter than normal green, and a positive value produces a longer than normal green. It is calculated as the difference between the *required green time* and the normal *maximum green time*, plus the *green credit*. If the *minimum green time* prevents a negative *green credit* from being fully redeemed, the remaining green credit is carried over to the next cycle. For a phase i and offset correction cycle K,

let OCC_{ik} be the remaining *green credit* in cycle k

min of $gMin_i$ and

$$tOC_{ik} = gReq_i - gMax_i + OCC_{ik} \tag{B.2}$$

subject to the constraint that $gReq_i + tOC_{ik} > gRdc_i$.

where:

$gMax$ = *maximum green time (from background timing plan), in seconds*
$gReq$ = *actual green time served, in seconds*
OCC = *green credit, in seconds*

B.2.1 Per-phase parameters for general traffic

The **saturation flow rate**, "S_i" for a given phase *i* is the maximum possible rate of traffic flow in passenger car units per hour (PCU/h). It is used to calculate the capacity of that phase, which in turn affects the amount the green duration can be reduced during offset correction. The flow rate could be determined in real time using the vehicle detectors at the intersection, but since the controller does not have a way of knowing if the maximum observed flow rate actually represents a saturated condition, this could result in significantly underestimated saturation flow rates for phases with relatively low demand. As a result, the saturation flow rate should be calculated separately and entered as a static value.

If the capacity for a particular phase is not at all a concern (i.e. it is not a phase which determines the cycle length requirements), an arbitrary large value (e.g. 9999) can be used to effectively bypass the minimum capacity requirements. In that case, the phase adjustments would be limited only by the fixed parameters such as the minimum green durations and minimum time before truncation. This could be helpful for movements which often receive extra time beyond their programmed maximum green time due to unused time from other phases. For example, a right-turn phase may receive extra green time if there happens to be no pedestrian call on the conflicting crosswalk. Since the system's calculation for the minimum capacity required doesn't account for this extra time, it could identify a capacity constraint for a movement which is actually has capacity to spare. This would cause TSP actions to be unnecessarily restricted.

The **observed traffic volume**, "$vObs_i$", for a given phase *i* is the average number of vehicles per hour (veh/h) observed using a phase. It is continually updated based on the observed number of vehicles crossing the stopbar loop(s) over the last 15 minutes. A 15-minute observation period is consistent with the standards set out by the Highway Capacity Manual.

The **percentage heavy vehicles**, "$\%HGV_i$" for a given phase *i* is needed in order to convert the observed traffic volume in vehicles per hour (veh/h) to an effective traffic volume in passenger car units per hour (PCU/h), so the volumes can be compared with the saturation flow rate, which is provided in passenger car units per hour. Since the intersection itself has no way of distinguishing between heavy vehicles and passenger cars, this must be manually entered.

B.2.2 Intersection-wide parameters for general traffic

The **maximum level of saturation**, "$satMax$" at the intersection should be kept below a certain threshold, which can be adjustable by the road authority. This threshold ranges from 0 to 1, and represents the maximum volume-to-capacity (v/c) ratio at the intersection as a whole. This value is lower than the theoretical capacity (i.e. v/c = 1.0) because the theoretical capacity is determined based on average volumes measured over a 15-minute period. In practice, during roughly half of cycles, the actual traffic volumes will be higher than the assumed volumes due to random variations in the vehicle arrival rate.

Zeng et al. (2021) used volume-to-capacity ratios around 0.90 to 0.95 as part of their calculation for allowable phase reductions during TSP actions or offset. To start with a priority level as high as possible, a value of 0.95 will be used as a starting point. If the priority system appears to be causing capacity issues at the intersection as a whole, this value can be reduced to further restrict the TSP actions. If only particular phases are observed to have capacity issues, then the assumed saturation flow rate for those particular phases should instead be reduced.

The **maximum cycles of offset correction**, "nOC" is the maximum number of cycles that offset correction is permitted to last following a TSP intervention. The magnitude of TSP actions is restricted such that the amount of time added to the cycle can be recovered within this number of cycles. This is the primary parameter which determines the level of coordination disruption in most cases. A value of 2 cycles will be used as a starting point.

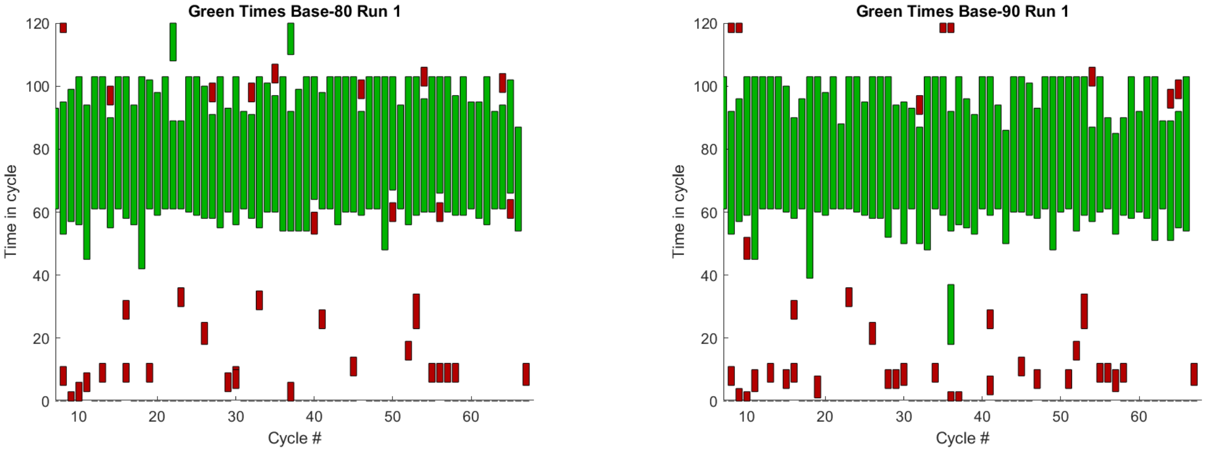
B.2.3 Per-phase TSP parameters

The **maximum green extension**, "$tExtMax_i$" is a hard upper limit for the amount a phase *i* can be extended above its normal maximum green time. The parameter is included because many jurisdictions have policies regarding the maximum amount that TSP may extend a green light Huang et al. (2012). Ideally this value should be as high as practical, since the system can identify on its own when long green extensions would result in unacceptable coordination or capacity impacts.

The **minimum time before truncation**, "$\%gMin$" parameter represents the percentage of a phase which must elapse before the TSP system may truncate the green. This is the standard method of restricting truncations at Dutch intersections. The minimum time before truncation does not consider the current traffic volumes and capacities for the phase because all time truncated from a phase will be returned to the same phase during the offset correction process, thereby resulting in no net impact on capacity.

The **minimum reduced green time**, "$\%gRdc$" is the smallest maximum green value that the TSP system is permitted to use for phase *i* during offset correction.

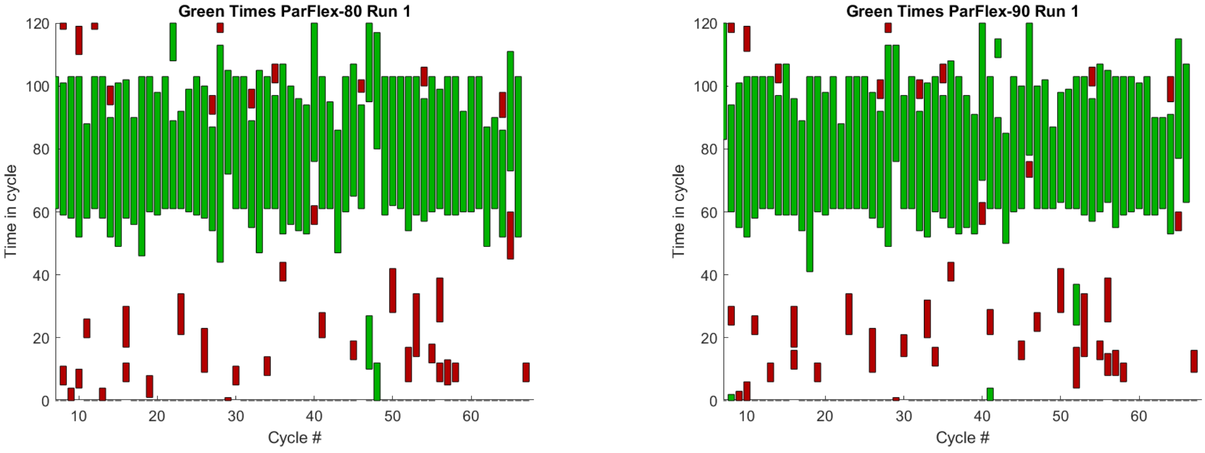
C Appendix: Additional Results



(a) High Flexibility

(b) Low Flexibility

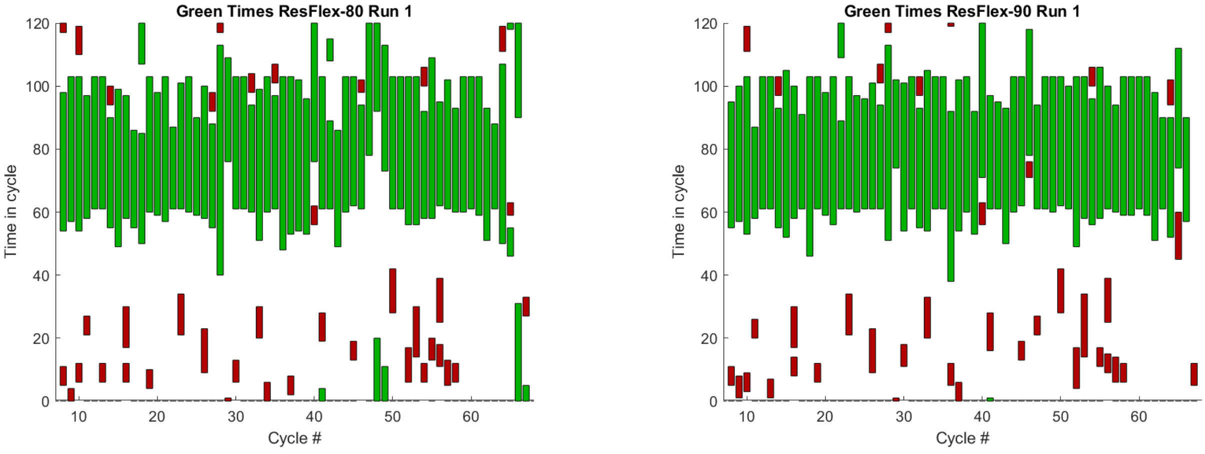
Figure 44: Green moments for phase 02, Base system, run 1



(a) High Flexibility

(b) Low Flexibility

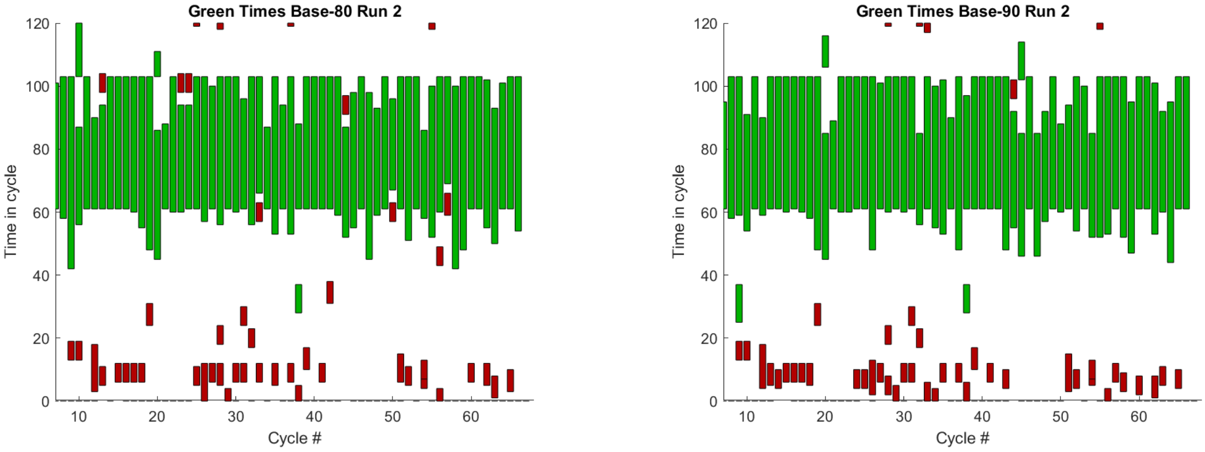
Figure 45: Green moments for phase 02, ParFlex system, run 1



(a) High Flexibility

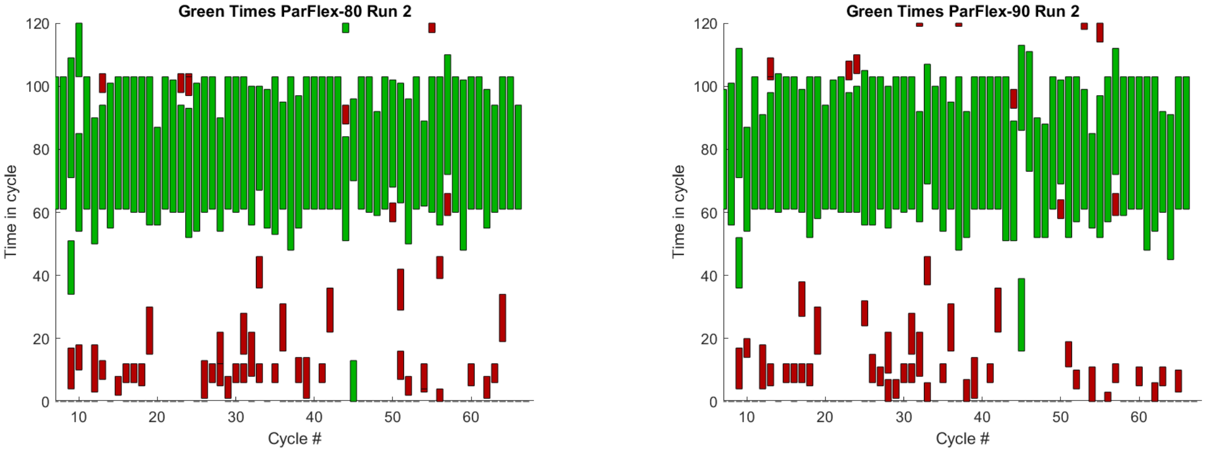
(b) Low Flexibility

Figure 46: Green moments for phase 02, ResFlex system, run 1



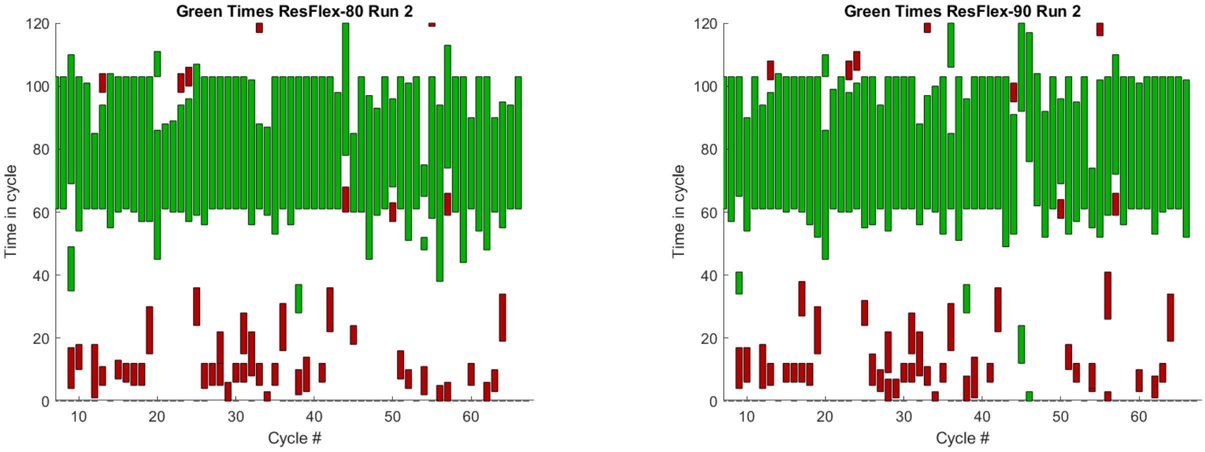
(a) High Flexibility (b) Low Flexibility

Figure 47: Green moments for phase 02, Base system, run 2



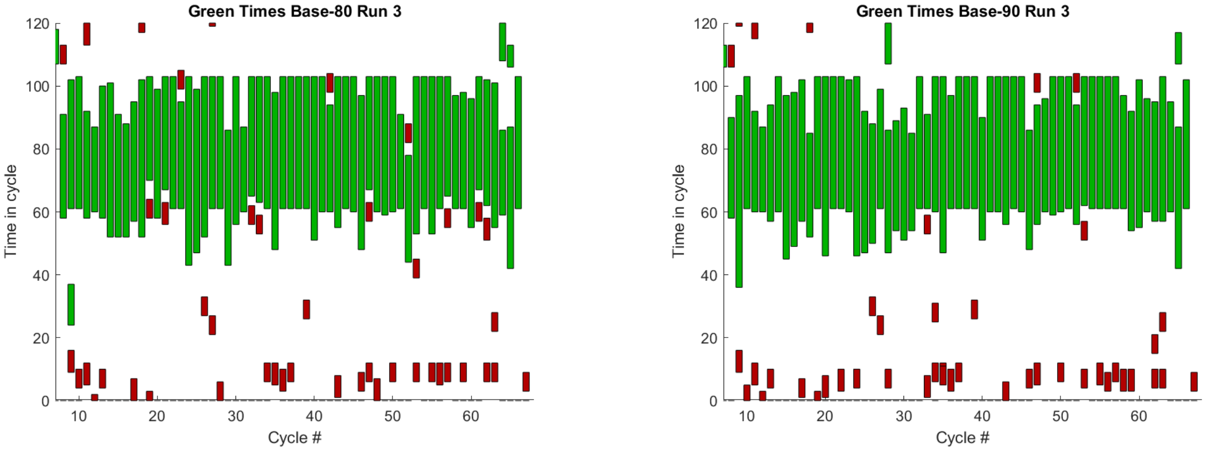
(a) High Flexibility (b) Low Flexibility

Figure 48: Green moments for phase 02, ParFlex system, run 2



(a) High Flexibility (b) Low Flexibility

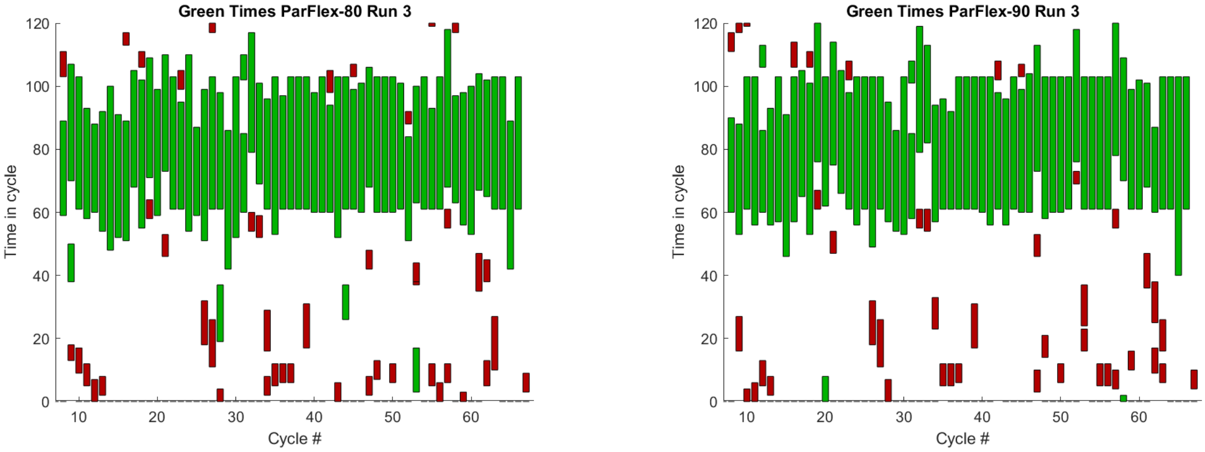
Figure 49: Green moments for phase 02, ResFlex system, run 2



(a) High Flexibility

(b) Low Flexibility

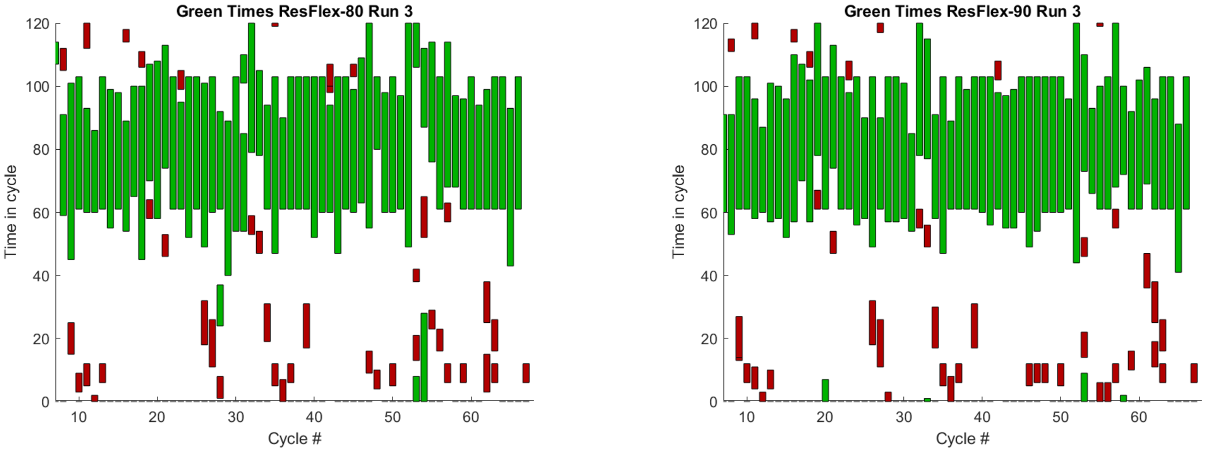
Figure 50: Green moments for phase 02, Base system, run 3



(a) High Flexibility

(b) Low Flexibility

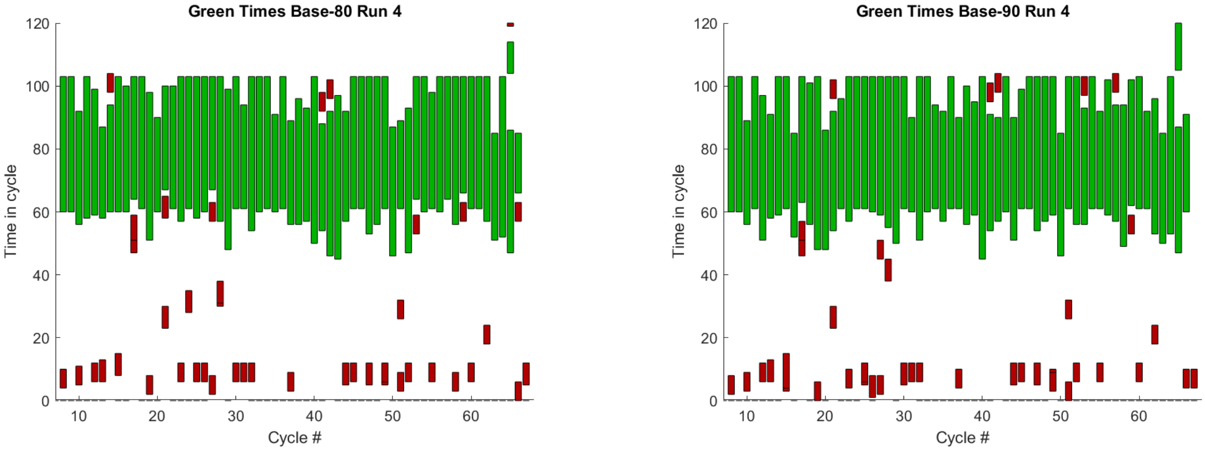
Figure 51: Green moments for phase 02, ParFlex system, run 3



(a) High Flexibility

(b) Low Flexibility

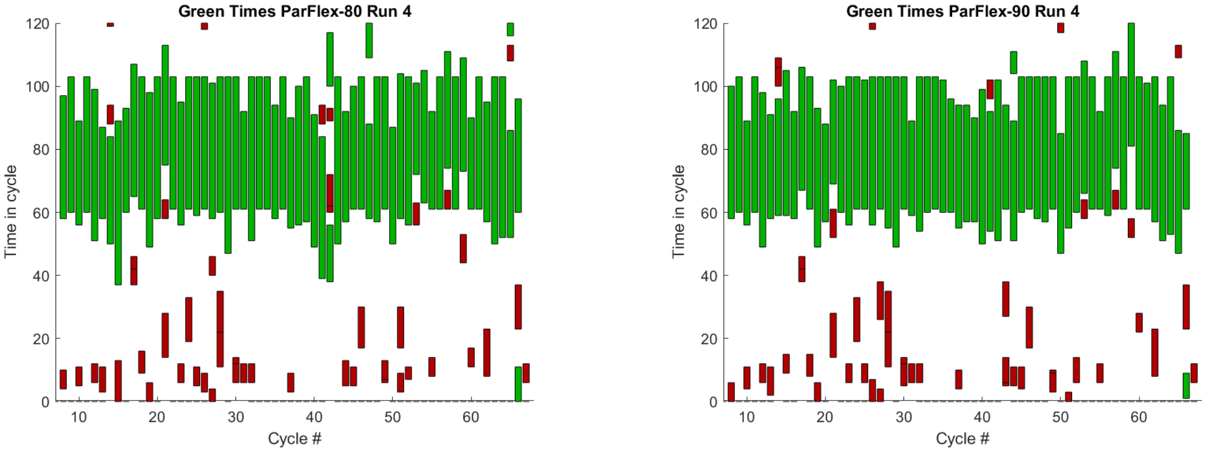
Figure 52: Green moments for phase 02, ResFlex system, run 3



(a) High Flexibility

(b) Low Flexibility

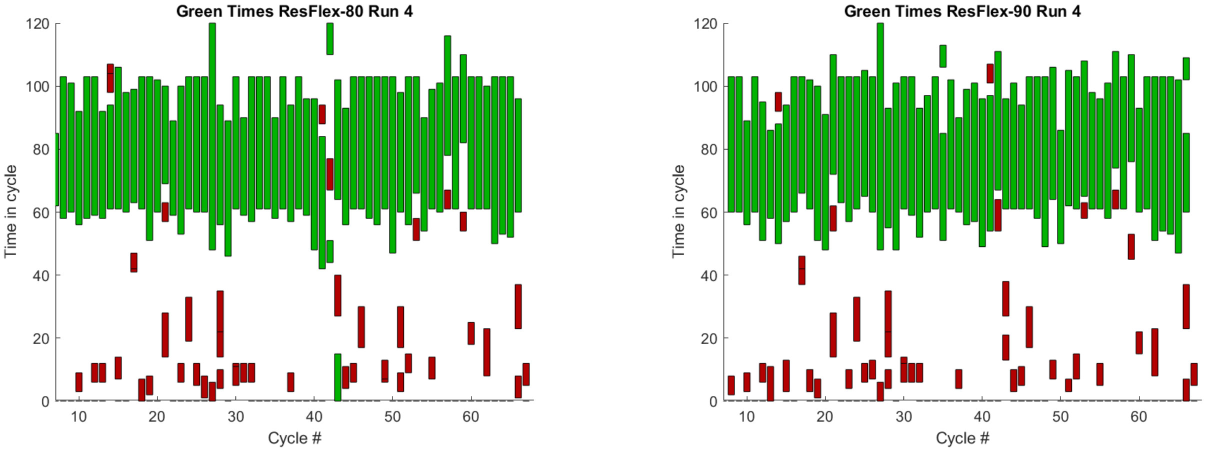
Figure 53: Green moments for phase 02, Base system, run 4



(a) High Flexibility

(b) Low Flexibility

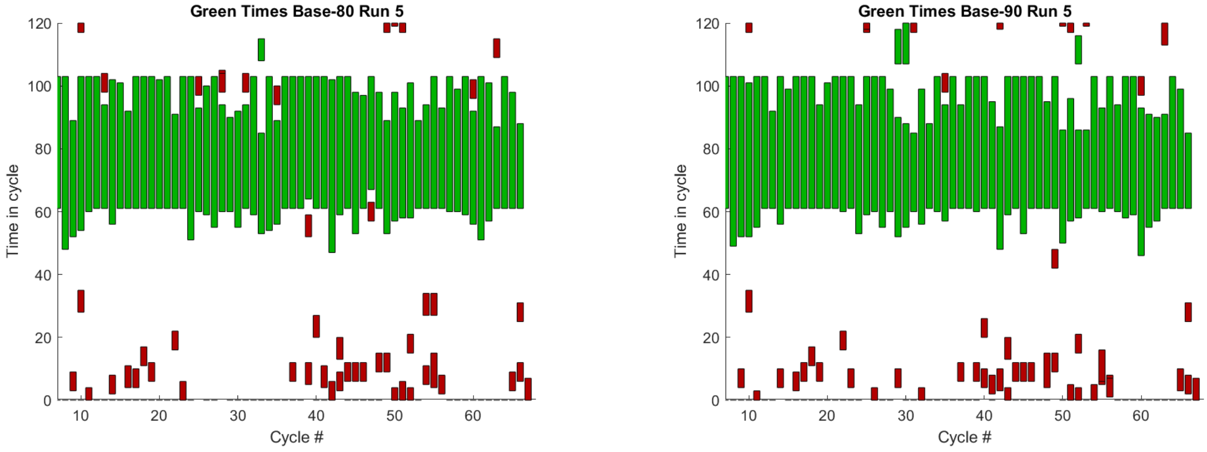
Figure 54: Green moments for phase 02, ParFlex system, run 4



(a) High Flexibility

(b) Low Flexibility

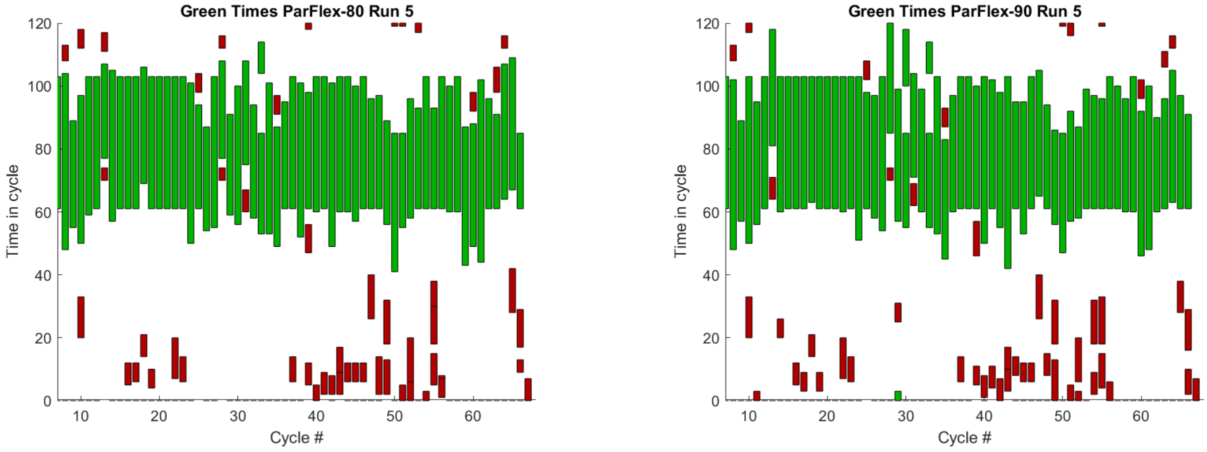
Figure 55: Green moments for phase 02, ResFlex system, run 4



(a) High Flexibility

(b) Low Flexibility

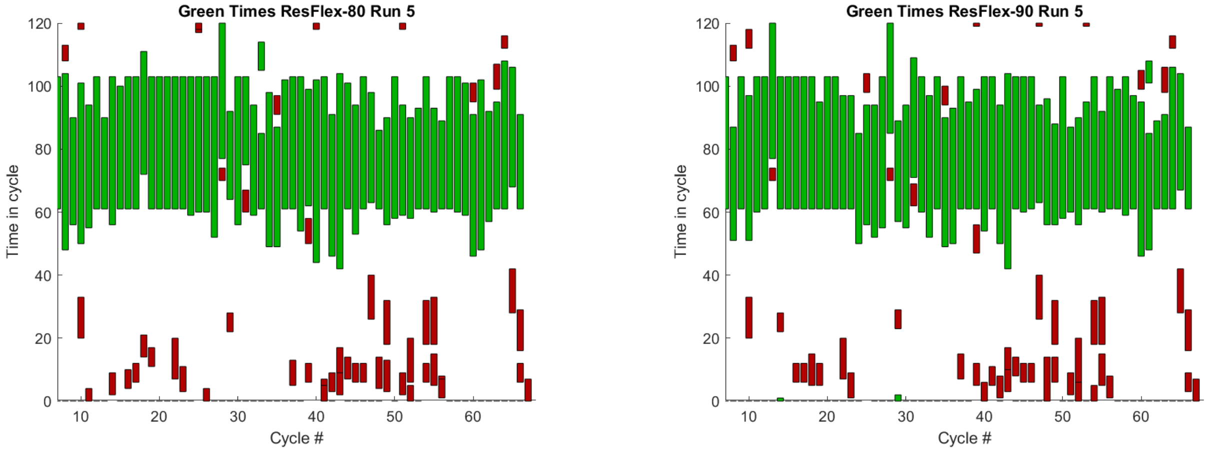
Figure 56: Green moments for phase 02, Base system, run 5



(a) High Flexibility

(b) Low Flexibility

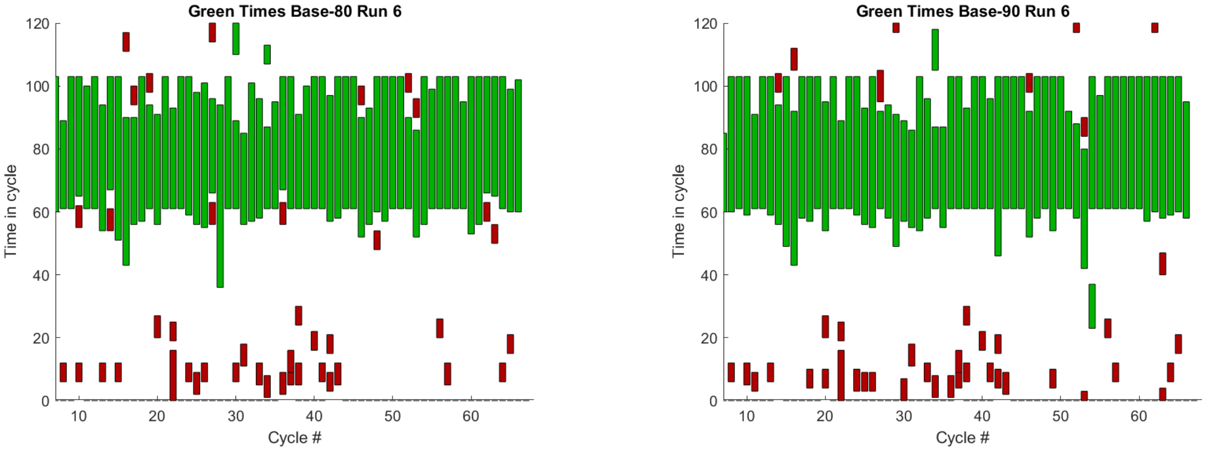
Figure 57: Green moments for phase 02, ParFlex system, run 5



(a) High Flexibility

(b) Low Flexibility

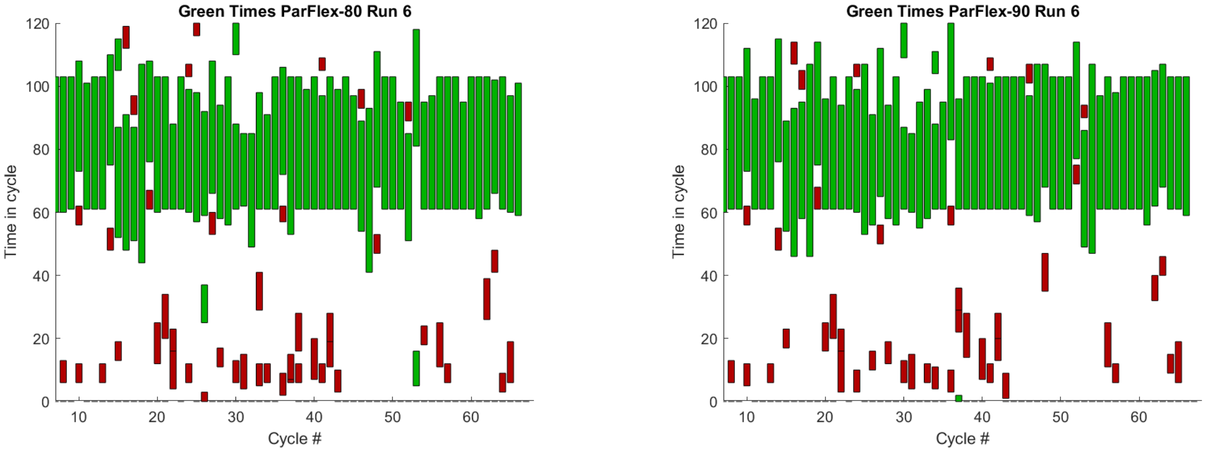
Figure 58: Green moments for phase 02, ResFlex system, run 5



(a) High Flexibility

(b) Low Flexibility

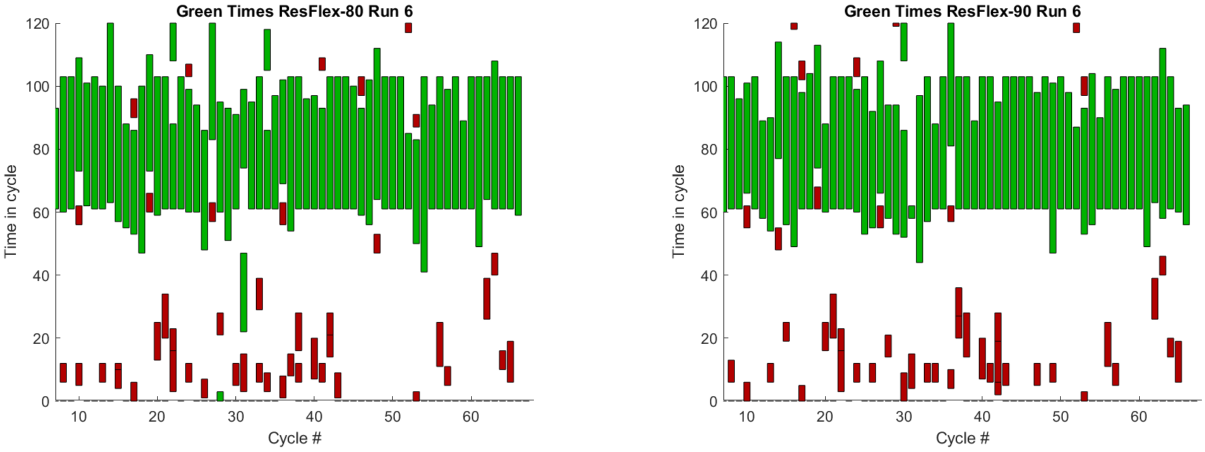
Figure 59: Green moments for phase 02, Base system, run 6



(a) High Flexibility

(b) Low Flexibility

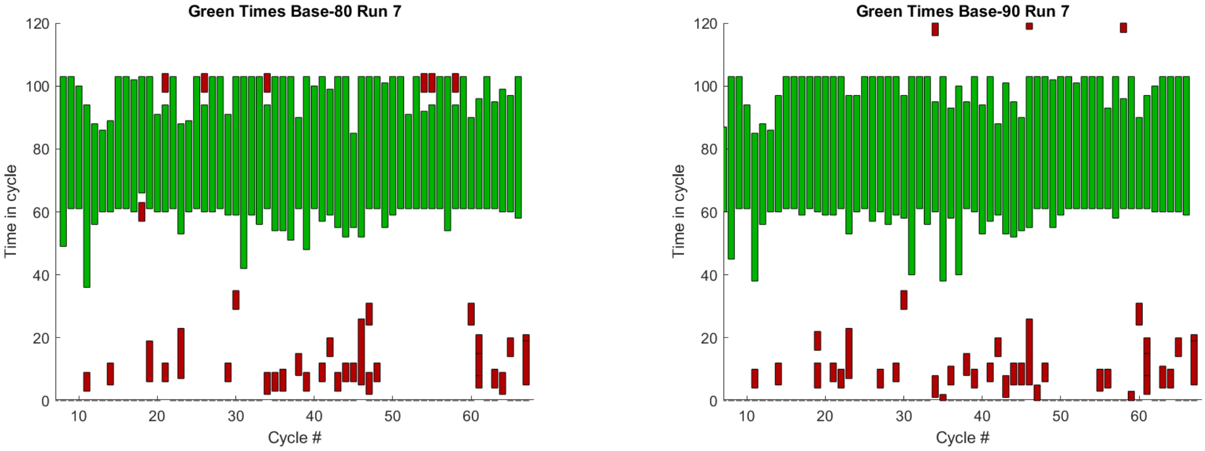
Figure 60: Green moments for phase 02, ParFlex system, run 6



(a) High Flexibility

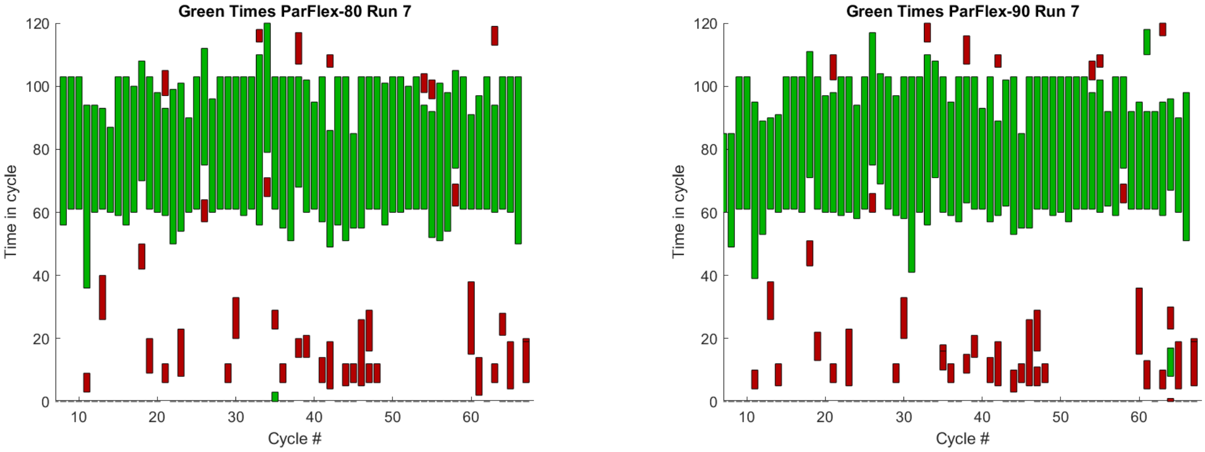
(b) Low Flexibility

Figure 61: Green moments for phase 02, ResFlex system, run 6



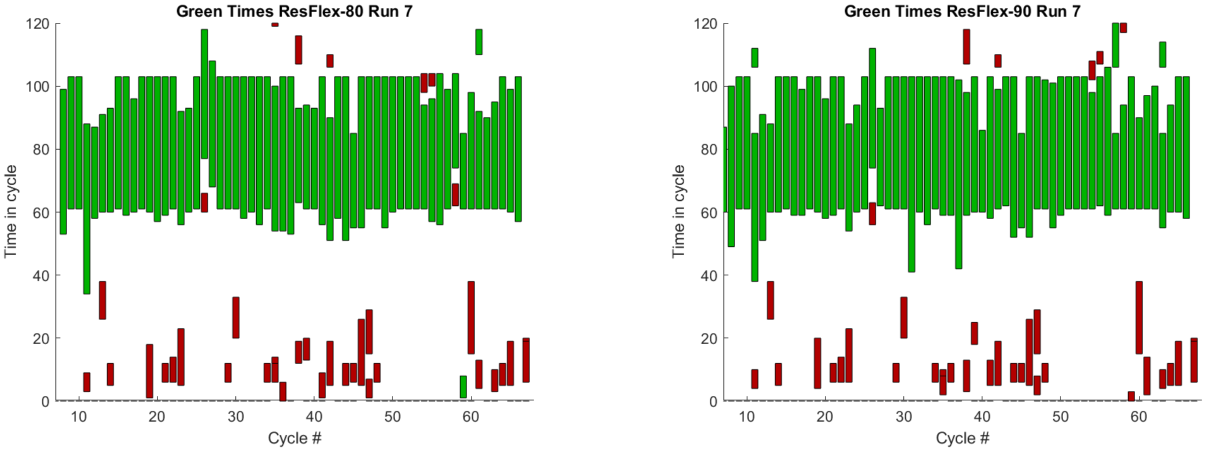
(a) High Flexibility (b) Low Flexibility

Figure 62: Green moments for phase 02, Base system, run 7



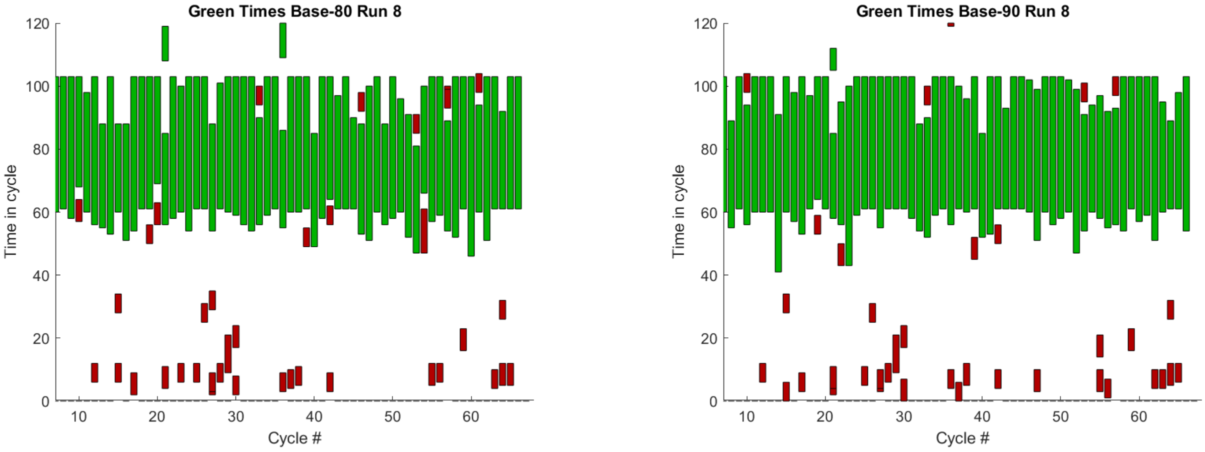
(a) High Flexibility (b) Low Flexibility

Figure 63: Green moments for phase 02, ParFlex system, run 7



(a) High Flexibility (b) Low Flexibility

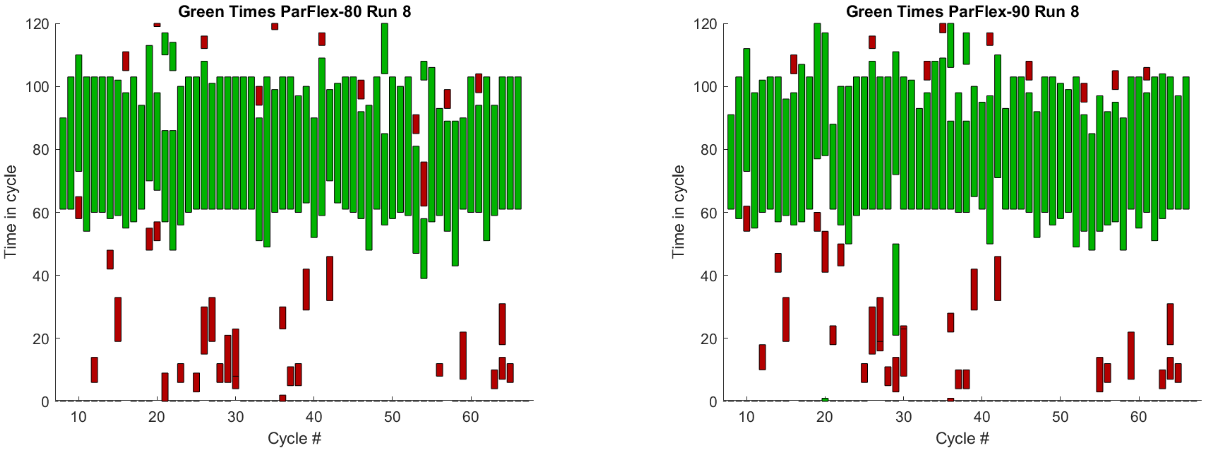
Figure 64: Green moments for phase 02, ResFlex system, run 7



(a) High Flexibility

(b) Low Flexibility

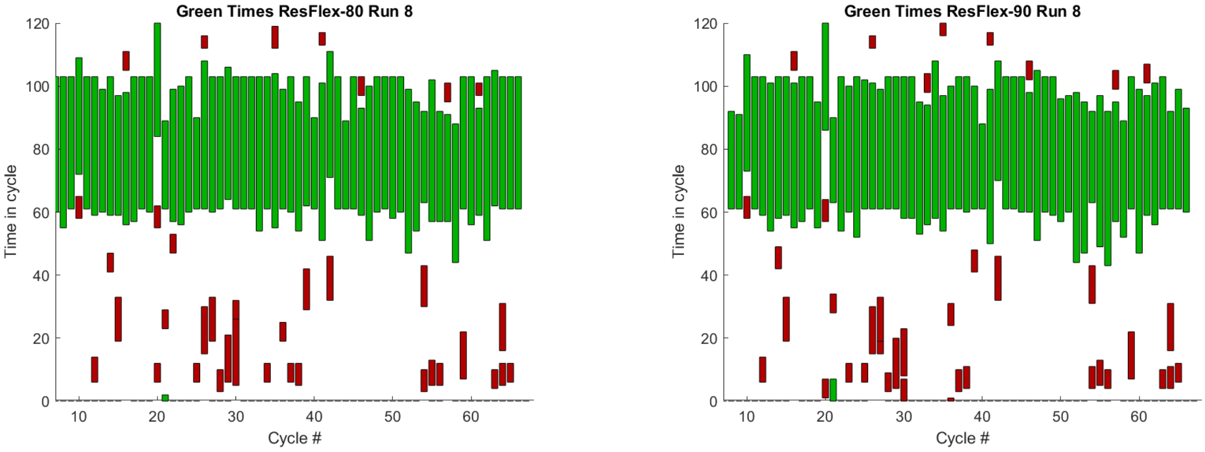
Figure 65: Green moments for phase 02, Base system, run 8



(a) High Flexibility

(b) Low Flexibility

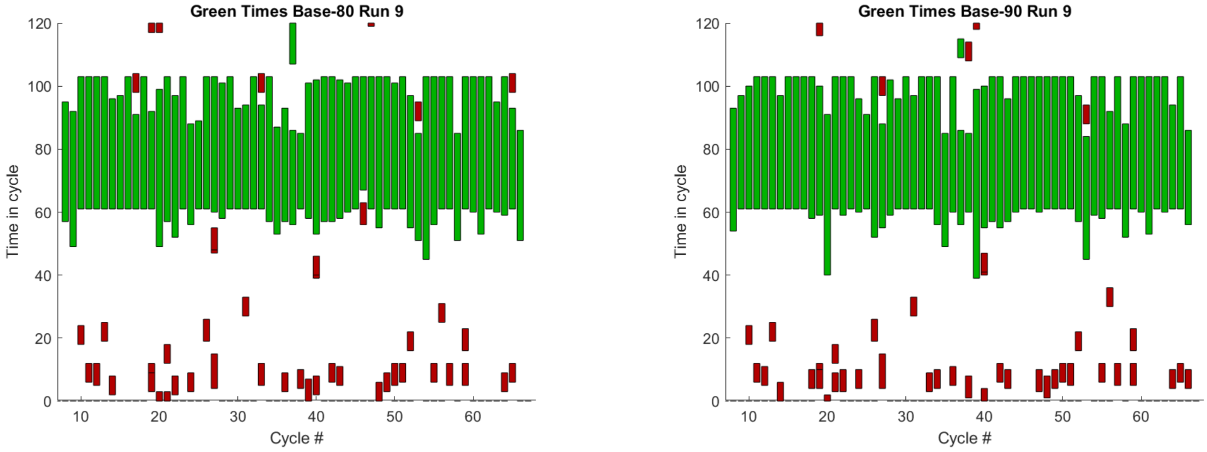
Figure 66: Green moments for phase 02, ParFlex system, run 8



(a) High Flexibility

(b) Low Flexibility

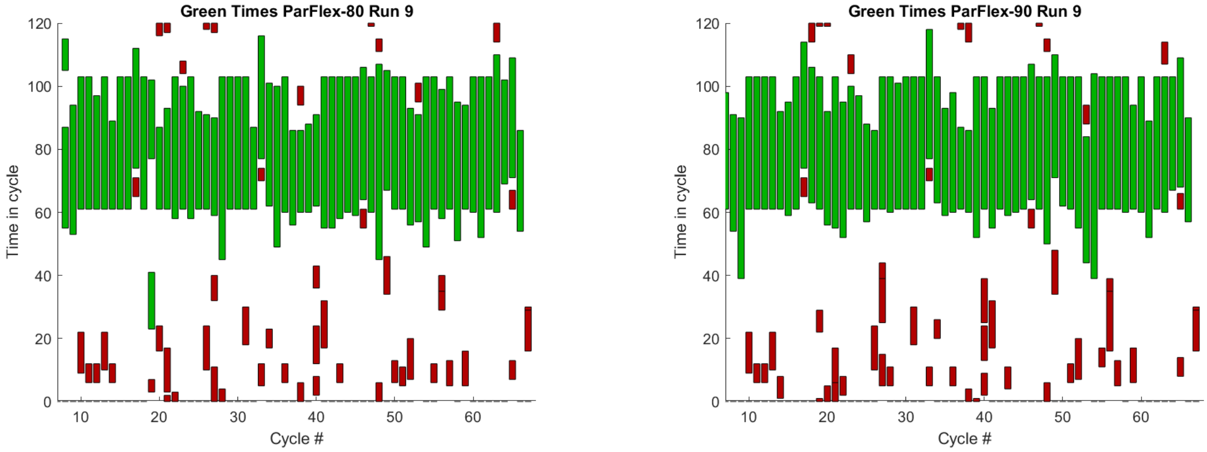
Figure 67: Green moments for phase 02, ResFlex system, run 8



(a) High Flexibility

(b) Low Flexibility

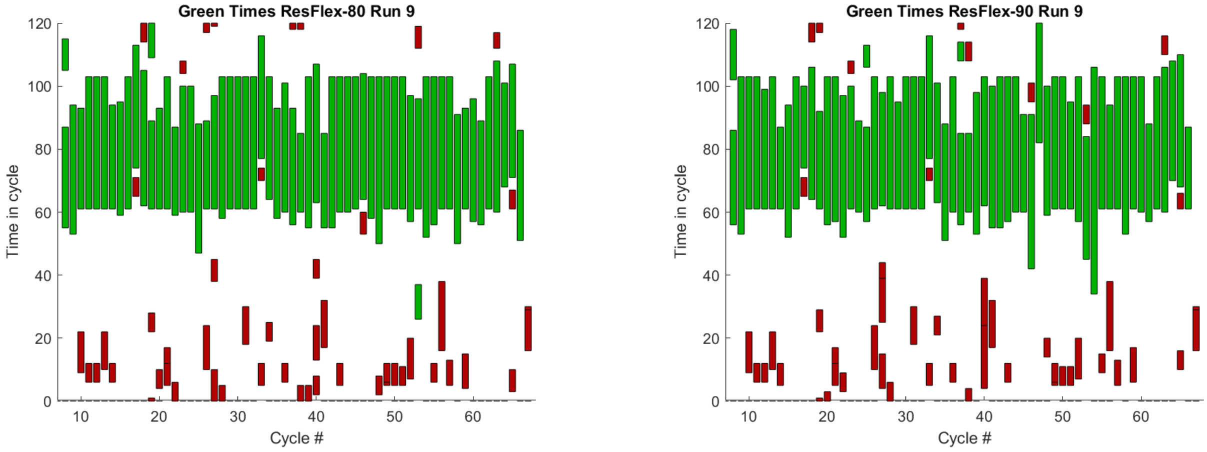
Figure 68: Green moments for phase 02, Base system, run 9



(a) High Flexibility

(b) Low Flexibility

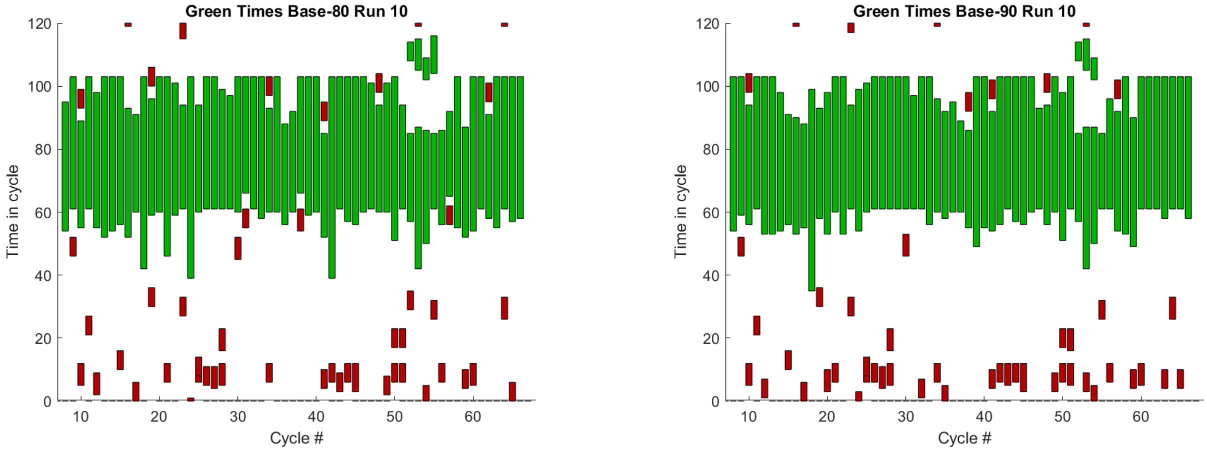
Figure 69: Green moments for phase 02, ParFlex system, run 9



(a) High Flexibility

(b) Low Flexibility

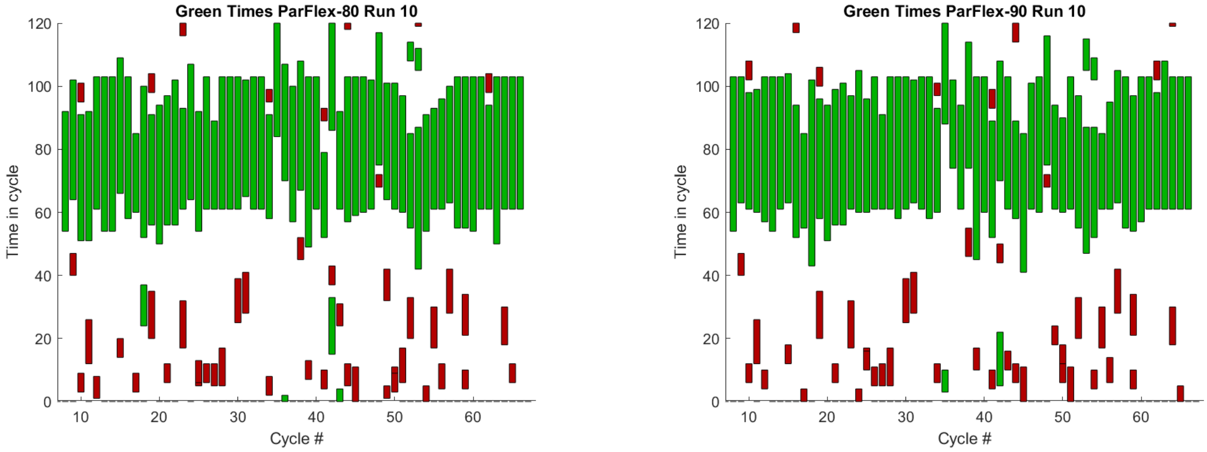
Figure 70: Green moments for phase 02, ResFlex system, run 9



(a) High Flexibility

(b) Low Flexibility

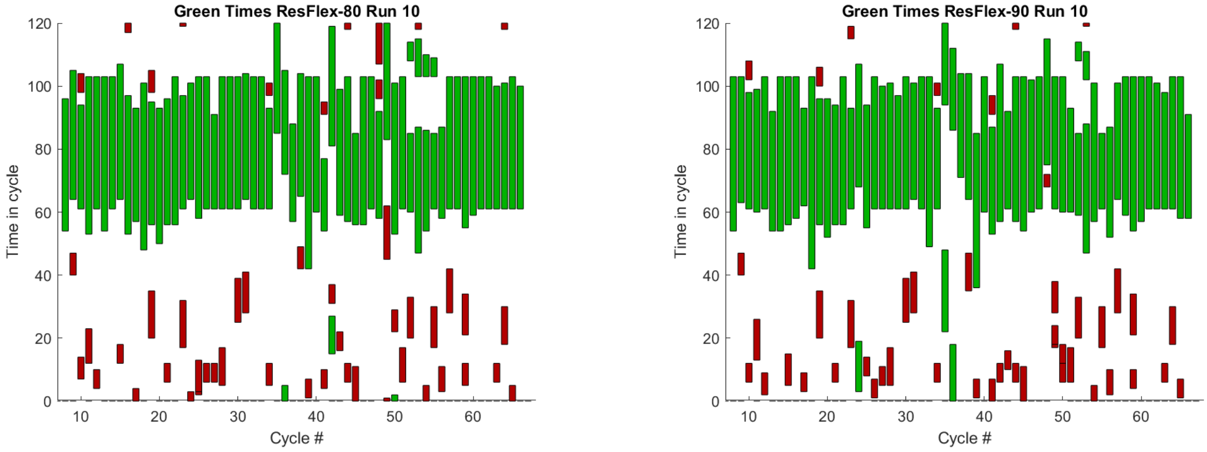
Figure 71: Green moments for phase 02, Base system, run 10



(a) High Flexibility

(b) Low Flexibility

Figure 72: Green moments for phase 02, ParFlex system, run 10



(a) High Flexibility

(b) Low Flexibility

Figure 73: Green moments for phase 02, ResFlex system, run 10

Bibliography

ANWB. Snelheid in het verkeer: de regels. Website, n.d.

Chih-Sheng Chou and Andrew Nichols. Characterizing emergency vehicle preemption operation with high-resolution traffic signal event data. *Transportation Research Reccord*, 2557:108–121, 2016.

City of Toronto. *Traffic Signal Operations Policies and Strategies*, 2015.

CROW. *Korte Afstand Radio voor prioriteit bij verkeerslichten*. Kennisplatform Verkeer en Vervoer, 2010.

CROW. *Handboek Verkeerslichtenreglingen*. CROW343, 2014.

Wael Ekeila, Tarek Sayed, and Mohamed El Esawey. Development of dynamic transit signal priority strategy. *Transportation Research Record*, 2111:1–9, 2009.

European Environment Agency. Occupancy rates of passenger vehicles. Indicator Assessment, 2010.

FHWA. *Traffic Signal Timing Manual*. U.S. Department of Transportation Federal Highway Administration, 2008.

Peter Furth and Theo Muller. Conditional bus priority at signalized intersections: Better service with less traffic disruption. *Transportation Research Record*, 1731:23–29, 2000.

Peter Furth, Burak Cesme, and Tarannum Rima. Signal priority near major bus terminal. *Transportation Research Record*, 2192:89–96, 2010.

Nick Hounsell and Birendra Shrestha. A new approach for cooperative bus priority at traffic signals. *Transactions on Intelligent Transportation Systems*, 13(1):6–14, 2012.

Jia Hu, Byungkyu Brian Park, and Young-Jae Lee. Transit signal priority accommodating conflicting requests under connected vehicles technology. *Transportation Research Part C*, 69:173–192, 2016.

Stephen Huang, Amer Shalaby, and Rajnath Bissessar. Impacts of transit signal priority on signal coordination. *Transportation Research Record*, 2311:29–43, 2012.

Georgios Laskaris, Marcin Seredynski, and Francesco Viti. Enhancing bus holding control using cooperative its. *Transactions on Intelligent Transportation Systems*, 21(4):1767–1778, 2020.

Jiali Li, Yugang Liu, and Bin Chen. Bus priority signal control considering delays of passengers and pedestraians of adjacent intersections. *Journal of Advanced Transportation*, 2020:1–12, 2020.

Meng Li, Yafeng Yin, Wei-Bin Zhang, Kun Zhou, and Hideki Nakamura. Modeling and implementation of adaptive transit signal priority on actuated control systems. *Computer-Aided Civil and Infrastructure Engineering*, 26:270–284, 2011.

Chen-Fu Liao and Gary Davis. Simulation study of bus signal priority strategy: Taking advantage of gps, avl and wireless communications. *Transportation Research Record*, 2034:82–91, 2007.

Yongjie Lin, Xianfeng Yang, Gang-Len Chang, and Nan Zou. Transit priority strategies for multiple routes under headway-based operations. *Transportation Research Record*, 2356:33–43, 2013.

Haiqing Liu, Laxmisha Rai, and Jianping Xing. Transit signal priority controlling method considering non-transit traffic benefits and coordinated ring states for multi-rings timing plan at isolated intersections. *Transactions on Intelligent Transportation Systems*, 22(2):913–936, 2021.

Zhenyu Mei, Zhen Tan, Wei Zhang, and Dinhai Wang. Simulation analysis of traffic signal control and transit signal priority strategies under arterial coordination conditions. *Simulation: Transactions of the Society for Modeling and Simulation International*, 95(1):51–64, 2019.

Mercedes-Benz Buses. Citaro ngt technische gegevens. Brochure, 2020.

Milos Mladenovic and Montasir Abbas. Multi-scale integration of petri net modeling and software-in-the-loop simulation for the novel approach to assessment of operational capabilities in advanced transportation controllers. *14th International IEEE Conference on ITS*, 2011.

NACTO. *Transit Street Design Guide*. National Association of City Transportation Officials, 2016.

Jon Obenberger and John Collura. Methodology to assess traffic signal transition strategies for exit preemption control. *Transportation Research Record*, 2035:158–168, 2007.

Rita Peñabaena-Niebles, Victor Cantillo, Sandra Cuentas, and José Luis Moura. Transition between signal timing plans: A methodological and analytical overview. *Transportation Reviews*, 34(3):356–378, 2014.

PTV Group. Effects of changing random seed. Website FAQ, n.d.

Marcin Seredynski, Francesco Viti, and Djamel Khadraoui. Signal phase and timing for cooperative public transport priority measures. *22nd ITS World Congress, Bordeaux, France 2015*, ITS-2664, 2015.

Marcin Seredynski, Georgios Laskaris, and Francesco Viti. Analysis of cooperative bus priority at traffic signals. *Transactions on Intelligent Transportation Systems*, 21:1929–1940, 2020.

Amer Shalaby, Jinwoo Lee, John Greenough, Stanley Hung, and Michael Bowie. Development, evaluation and selection of advanced transit signal priority concept directions. *Journal of Public Transportation*, 9(5): 97–120, 2006.

Alexander Skabardonis and Christofa Eleni. Impact of transit signal priority on level of service at signalized intersections. *6th International Symposium on Highway Capacity and Quality of Service*, 2011.

Courtney Slavin, Wei Feng, Miguel Figliozzi, and Peter Koonce. Statistical study of the impact of adaptive traffic signal control on traffic and transit performance. *Transportation Research Record*, 2356:117–126, 2013.

Transportation Research Board. *Highway Capacity Manual*. HCM2000, 2000.

Long Tien Truong, Graham Currie, Mark Wallace, Chris De Gruyter, and Kun An. Coordinated transit signal priority model considering stochastic bus arrival time. *Transactions on Intelligent Transportation Systems*, 20(4):1269–1277, 2019.

Johan Wahlstedt. Impacts of bus priority in coordinated traffic signals. *6th International Symposium on Highway Capacity and Quality of Service*, 2011.

Bart Wolput, Eleni Christofa, and Chris Tampère. Optimal cycle length formulas for intersections with or without transit signal priority. *Transportation Research Record*, 2558:78–91, 2016.

Xiaosi Zeng, Yunlong Zhang, and Kai Yin. Route-based transit signal priority using connected vehicle technology to promote bus schedule adherence. *Transactions on Intelligent Transportation Systems*, 22(2):1174–1184, 2021.

Milan Zlatkovic, Peter Martin, and Aleksandar Stevanovic. Predictive priority for light rail transit. *Transportation Research Record*, 2259:168–178, 2011.

Milan Zlatkovic, Peter Martin, and Ivana Tasic. Implementation of transit signal priority and predictive priority strategies in asc/3 software-in-the-loop simulation. *Advances in Artificial Transportation Systems and Simulation*, pages 203–218, 2015.