Mikolášek
European Transport Research Review (2024) 16:57
https://doi.org/10.1186/s12544-024-00672-y

European Transport Research Review

ORIGINAL PAPER

Efficiency of dynamic traffic signal control in work zones with shuttle operation: theory and practice

Igor Mikolášek[1*]

Abstract

Portable traffic signals with fixed-time signal plans are a common type of traffic control at work zones with shuttle traffic. The most-used alternatives are flagging and intelligent transport systems with traffic-actuated signals. These can provide more efficient traffic control, but existing policies often do little to encourage their practical application. This paper provides a clear and accessible overview of shuttle operations and a comparison of the main signal control types while addressing some knowledge gaps, such as whole-day operation efficiency. The relations between different variables of the signal plan and traffic flow are derived to build a theoretical framework and models for both deterministic and random arrivals to estimate delays and to find the optimal signal plan setting for a wide range of circumstances. The hypothetical scenarios are supported by a case study. Traffic and signal control data from several construction phases of a work zone with shuttle operation were gathered, processed, analysed, and used to compare different control scenarios. The results provide solid evidence for the efficiency of the dynamic systems. The magnitude of the difference is heavily affected by circumstances. The efficiency of the green signal almost doubled with the dynamic control. The case study also revealed a severe impact of road conditions (milling) on the work zone capacity. Several standards and policies are proposed based on the findings to encourage wider and more efficient adoption of traffic-actuated signals in work zones.

Keywords One-lane two-way traffic, Portable traffic signals, Flagging, Saturated flow, Highway capacity, Temporary traffic management

1 Introduction

Highway capacity significantly affects the level of service of a road network. Maintenance and reconstruction work zones (WZ) create significant bottlenecks causing recurrent traffic flow (TF) congestions due to capacity reduction. Since WZs are an inevitable part a road's lifecycle, research is focused on understanding the properties of the WZs and mitigating their impact on the TF. Papers address TF behaviour and modelling at bottlenecks [1–4], WZ capacity estimation [5–8], safety [9, 10], or ITS solutions improving the capacity and reducing travel times [11–13]. However, these studies focus on motorway WZs with their high traffic flows, yet the issues and TF behaviour significantly differ from shuttle operation. Limited research focuses on two-lane highways with shuttle operation (Fig. 1), most of which address flagging [14, 15]. However, given the extent of the two-lane road network, e.g. 85% of US [16] and over 95% of Czech [17] road network are rural paved highways, optimizing traffic control in shuttle operated WZs could be equally beneficial, especially where fixed-time traffic signals predominate.

Many studies on this topic refer to two older papers that formulated delay estimation equations for signalized intersections [18, 19]. These models can be applied

*Correspondence:
Igor Mikolášek
igor.mikolasek@cdv.cz
[1] Transport Research Centre CDV, Lisenska 33a, 636 00 Brno, Czech Republic

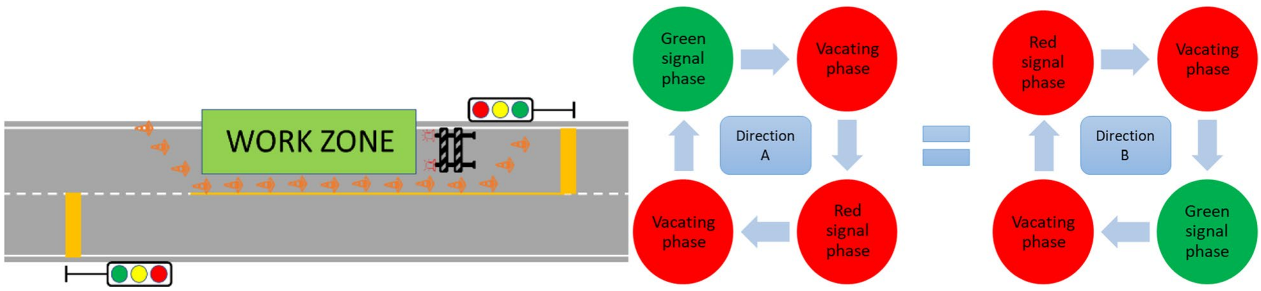


Fig. 1 Idealized scheme of a typical work zone with signalized one-lane two-direction shuttle operation on a two-lane highway and its schematic signal plan

as shuttle operation is principally the same as an intersection of two one-way streets. Webster formulated equations for determining optimal signal settings based on simulation, which form the bases of many formulas in national design standards all over the world, while Newell used fluid (deterministic queuing) and diffusion approximations of traffic flow. He stated that the green signal should switch promptly (except for extreme cases of asymmetric demand) once the TF intensity drops below saturated (queue discharge) flow and, for nearly saturated conditions, the average delay is three times smaller with traffic-actuated (dynamic) control at isolated intersections. The traffic-actuated signals perform even better when considering stochastic and evolving arrivals thanks to their flexibility. Son [20] extended Newell's models to include WZ passage. Other papers on modelling delay include a Monte Carlo simulation study with an estimated statistical distribution of vehicle delay [21] or a comparison of four different types of models at intersections [22].

Others combined these models with cost functions of vehicle time, WZ set-up, and operation and construction costs to find optimal signal timing and length of WZs based on traffic demand [23, 24]. However, implementation of this methodology would require accurate input data for the construction and time costs, and motivation or enforcing policies for the contractors and/or authorities to utilize them.

The paper emerges from a study conducted for the Road and Motorway Directorate of the Czech Republic (ŘSD ČR) which was seeking justification for mandating dynamic traffic control systems on first-class roads in Czechia. Other Czech Road administration offices rarely mandate the use of dynamic control, too. While rooted in a specific national context, the findings of this research carry broader relevance as common usage of suboptimal control systems is not unique to Czechia. For example, part 6 (temporary traffic control) and associated sections of MUTCD [25] does not even explicitly recognize dynamic traffic control signals as a separate type of traffic control for shuttle operation, unlike e.g. pilot car or flag transfer methods. Flagging, common in the USA, can mimic dynamic control, based on how the flaggers allocate the right of way, but the cost of human resources often makes it more expensive. There are likely two main reasons for limited adoption of dynamic control. Firstly, there's a lack of awareness among the public, including administration offices and contractors, regarding the significant time-saving potential of dynamic systems in specific situations. Secondly, dynamic systems tend to be costlier, making them less attractive to contractors without incentives or obligations.

To address these issues, this paper provides a comprehensive assessment of the differences between fixed-time and dynamic control systems and offers policy suggestions for local governments and administration offices to promote the use of more efficient traffic control systems. It presents a broader but accessible analytical examination of the relationships between signal plan (SP) settings, WZ capacity, and total delay that allows for an easy understanding of the relations and core principles and makes the topic more comprehensible for practitioners and policymakers, yet it expands the state-of-the-art by addressing some of the knowledge gaps such as modelling of delay caused by random arrivals with dynamic control or capacity definition, variable relations, and whole-day operation, previously pointed out by Zhu et al. [26]. The major deficiency of the existing literature is the focus on specific aspects, often with very complex modelling tools, while lacking clear and comprehensible assessment of the multidimensional dependencies of different variables that affect the outcomes. Most papers also consider only peak hour, rather than whole-day WZ operation. Following the theoretical background, including modelling of a hypothetical WZ with varying parameters, a practical part presents a case study where TF data were measured over several weeks with different traffic control conditions. Some observations feed back to the model settings and the TF data are used to estimate delay using the derived models, while bringing additional insight into the efficiency of dynamic control in practice and into the effects of road conditions and saturated flow on the overall capacity. Based on the findings, the concluding section proposes several policy recommendations and best practices which can serve as a basis for other countries considering policy changes, given the continued prevalence of fixed-time signal control across the world.

2 Methodology

2.1 Fundamental relations

The elementary characteristic of any WZ is its capacity C. For shuttle operation, it is practical to define it as the combined sum of both directions, as they share the same infrastructure:

$$C = \left(1 - \left(T_{V,sum}/T_C\right)\right) \cdot I_{sat} = \frac{T_{G,sum} \cdot I_{sat}}{T_C} \tag{1}$$

where $T_{V,sum}$ represents the total time needed for both directions to vacate (clear) the WZ per cycle, including yellow signals and time lost to acceleration/deceleration, T_C is the total cycle time, I_{sat} is the intensity of saturated flow on the single lane, which is different from the capacity. Capacity can also be deducted from the duration of green signals once defined. In that case, capacity can be also defined separately per direction.

The ratio $T_{V,sum}/T_C$ is a key parameter – low $T_{V,sum}$ is essential for high capacity, making long WZs inefficient with inherently low capacity. Figure 2 illustrates the relationship between capacity, $T_{V,sum}$, and T_C. Capacity initially grows rapidly with increasing green signal duration, but experiences diminishing gains as T_G and T_C increase, asymptotically approaching the saturated flow I_{sat}. These diminishing returns are especially significant for the lower, favourable values of $T_{V,sum}$. Variation in I_{sat} only alters the magnitude. In case of dynamic control, capacity fluctuates from cycle to cycle due to dynamic changes in T_C (and potentially $T_{V,sum}$). Conversely, capacity remains constant with fixed-time control, but its effective utilisation suffers if demand drops.

While I_{sat} and C are often measured in passenger car equivalents (PCE) as vehicle properties affect these variables, using PCE is nonsensical when dealing with waiting time and delays. This conflict introduces accuracy challenges not discussed in most of existing literature. When estimating delays, it is more pragmatic to use vehicles (per hour) even to gauge saturated flow and capacity for delay estimation, but the proportion of slow and long vehicles when measuring the saturated flow should mirror their representation at the assessed WZ. However, if the only concern is design of a SP, it is more accurate to use pc/h.

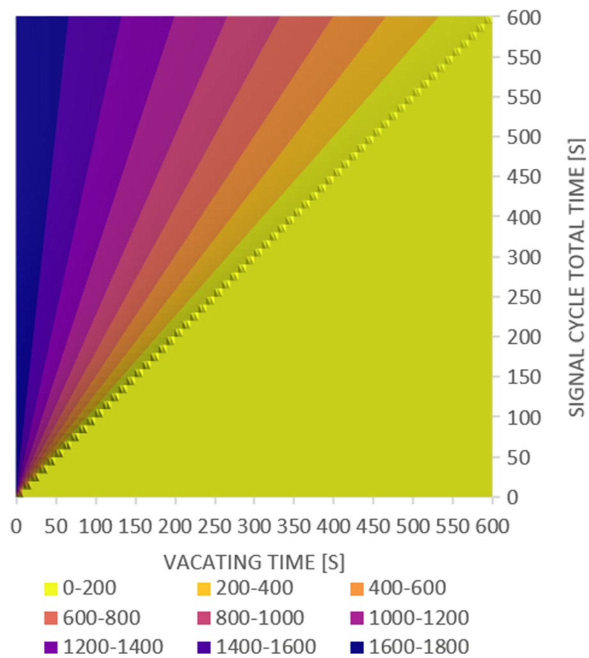


Fig. 2 Total capacity [veh/h] for I_{sat} = 1800 veh/h and varying vacating (sum per cycle) and total cycle time

It is considered unacceptable for the capacity to be exceeded in the following derivations and modelling (i.e., oversaturation is not allowed and considered) as that leads to accumulation of queuing vehicles and rapidly growing delays (see e.g. [22]). The red signal time for each cycle and direction T_R is given by:

$$\begin{aligned} T_{R,A/B} &= T_{V,sum} + T_{G,B/A} = T_C - T_{G,A/B} \\ T_{R,A/B} &= T_{V,sum} + \frac{\left(T_C - T_{V,sum}\right)}{2} \quad if \quad T_{G,A} = T_{G,B} \end{aligned} \tag{2}$$

where T_G is the green signal duration and A and B denote directions. Extending the green signal duration in one direction prolongs the red signal duration, and thus the delay, in the other direction (in fact, the delay grows with the square of T_R, see Eq. (5)). Therefore, it is essential to minimize its duration while ensuring adequate capacity.

Assuming undersaturated conditions and known, homogenous traffic flow (deterministic arrivals) and saturated flow, the optimal total cycle time (without reserves) can be determined using Eq. (3) for a given demand or TF intensity I.

$$T_C = \frac{I_{sat} \cdot T_{V,sum}}{I_{sat} - I_{crit}} \quad where \quad I_{crit} = I_{A,max} + I_{B,max} \tag{3}$$

In practical scenarios involving whole-day operation, the overall capacity of a fixed-time SP must accommodate I_{crit} – the sum of peak flows in individual directions $I_{A,max}$ and $I_{B,max}$ over the whole operation interval – which exceeds the combined peak flow in both directions. This is particularly significant when the morning and afternoon peaks occur in opposite directions but is relevant even when scheduled fixed-time plans with different SP for different times of day are used.

The green signal time is then assigned to individual directions based on the demand in each direction using Eq. (4), which defines the length of the red signals (Eq. (2)). Non-integer values must be rounded up, possibly extending T_C (or the result of Eq. (3) can be rounded up to nearest even number in advance).

$$T_{G,A/B} = \left(T_C - T_{V,sum}\right) \cdot \left(I_{A,max/B,max}/I_{crit}\right) = T_{G,sum} \cdot \left(I_{A,max/B,max}/I_{crit}\right) \quad (4)$$

For modelling dynamic control, the T_C, T_R, and T_G must be established separately for each cycle or interval with constant (average) demand using current values of I_A and I_B ($I_{crit}=I_{sum}=I_A+I_B$). Additionally, T_G (and by extension T_C and T_R) must be extended with the duration of the detection window for incoming vehicles before the green signal is switched to red. Roughly 5 s for T_G (10 s for T_C assuming two approaches) is recommended.

In practice, the clearance time is mostly determined by the WZ length and the driving speed of a slow vehicle within and kept fixed, leaving just the green signal time in each direction to be tampered with. There are currently no systems known to the author that can dynamically adjust the clearance time.

Once the SP is defined, the total delay T_D per direction for intervals t with constant demand can be estimated ($t=1$ h in this paper). These can be summed over the whole day or used to calculate average vehicle delay. Unless specified otherwise, namely in Eqs. (11)–(13), delay is estimated separately for each direction and variables I, T_R, and T_G are substituted into the formulas for individual directions as relevant. If queue dissipation was disregarded:

$$T_D = N_{veh} \cdot \frac{T_R}{2} \cdot \frac{t}{T_C} = I \cdot \frac{T_R^2}{2} \cdot \frac{t}{T_C} \quad (5)$$

where $N_{veh} = T_R \cdot I$ is the number of vehicles queuing at the entrance to the WZ at the end of the red signal in each cycle. The second version of Eq. (5) illustrates the quadratic impact of the red signal duration.

However, the queue dissipation cannot be omitted as it has significant impact on the overall delay, especially when the degree of saturation X (Eq. (6)) is close or even equal to one, as theoretically optimal, as even some of the vehicles arriving during the green phase are delayed when they arrive to the gradually shortening queue.

$$X = \frac{I \cdot T_C}{I_{sat} \cdot T_G} \quad (6)$$

Equation (7), derived from first (deterministic) term of Webster's equation [19] can be used to correct for this. The added term $1/(1 - I/I_{sat})$ reflects the additional delay due to the queue dissipation.

$$T_D = \frac{T_R^2 \cdot I \cdot t}{2 \cdot T_C \cdot (1 - \frac{I}{I_{sat}})} \quad (7)$$

There are two main issues stemming from the assumption of uniform arrivals, which is obviously violated in practice. Firstly, the irregular arrivals would cause immediate oversaturation (unless dynamic control is used), resulting in additional delays. To avoid the oversaturation, a capacity reserve is recommended to allow for some fluctuation in demand (and saturated flow). Zhu et al. [26] recommend 20–30% or a minimum of 200 veh/h capacity reserve, which minimizes the average delay. This can be included in the SP design by multiplying I_{crit} in Eq. (3) by a coefficient $r \in < 1.2, 1.3 >$ and then checking whether the partial capacity is at least 100 veh/h above the peak demand in each direction. Secondly, to model the additional delay due to the random arrivals with fixed-time SP, the second term of Webster's equation (Eq. (8), where $t_{avg,add}$ is the added average delay per vehicle – in hours, if I and I_{sat} are in veh/h) can be included in the delay calculations:

$$t_{avg,add} = \frac{X^2}{2 \cdot I \cdot (1 - X)} = \frac{I \cdot T_C^2}{2 \cdot I_{sat} \cdot T_G \cdot (I_{sat} \cdot T_G - I \cdot T_C)} \quad (8)$$

The overall estimated delay per interval t with constant average demand I is then given by Eq. (9). As all delays are calculated in seconds, $t_{avg,add}$ must be transformed from hours to seconds. Then:

$$T_{D,+rnd,fix} = I \cdot t \cdot \left[\frac{T_R^2}{2 \cdot T_C \cdot \left(1 - \frac{I}{I_{sat}}\right)} + \frac{I \cdot T_C^2 \cdot 3600}{2 \cdot I_{sat} \cdot T_G \cdot (I_{sat} \cdot T_G - I \cdot T_C)} \right] \quad (9)$$

However, Webster's formula overestimates the delay caused by random arrivals, especially during times of high saturation [22] (it in fact grows to infinity for $X=1$), which is likely why Webster's formula for SP design leads to unreasonably long cycles for highly saturated intersections [27]. On the other hand, the effect of random arrivals is negligible for $I < 0.6\ C$ [22]. Therefore, the estimated delay can be considered to be the average of values given by Eqs. (7) and (9), i.e. the added delay from Eq. (8) is taken at half the value, resulting in Eq. (10). This average roughly corresponds to values given by the US Highway Capacity Manual [28] for values of X around 0.7–0.8 [22, Fig. 11], which are assumed to be optimal for fixed-time control, as discussed above.

$$T_{D,+rnd/2,fix} = I \cdot t \cdot \left[\frac{T_R^2}{2 \cdot T_C \cdot \left(1 - \frac{I}{I_{sat}}\right)} + \frac{I \cdot T_C^2 \cdot 3600}{4 \cdot I_{sat} \cdot T_G \cdot (I_{sat} \cdot T_G - I \cdot T_C)} \right] \tag{10}$$

Note that Eqs. (8)–(10) must be adjusted if dynamic control is used given its ability to adapt to traffic conditions. Specifically, the added delay must be considered per both directions together, and the capacity is given by the maximal allowed total cycle duration. Therefore:

$$X_{dyn} = \frac{I_{sum} \cdot T_{C,max}}{I_{sat} \cdot T_{G,sum,max}}, \text{ where } T_{G,sum,max} = T_{C,max} - T_{V,sum} \tag{11}$$

and

$$t_{avg,add,dyn} = \frac{X_{dyn}^2}{2 \cdot I_{sum} \cdot (1 - X_{dyn})} = \frac{I_{sum} \cdot T_{C,max}^2}{2 \cdot I_{sat} \cdot T_{G,sum,max} \cdot (I_{sat} \cdot T_{G,sum,max} - I_{sum} \cdot T_{C,max})} \tag{12}$$

Same as with the fixed signal plan, $t_{avg,add,dyn}$ can be assumed at half its value to compensate for the overestimation given by the original Webster's formula, thus:

$$T_{D,+rnd/2,dyn} = \frac{T_{R,A}^2 \cdot I_A \cdot t}{2 \cdot T_C \cdot (1 - \frac{I_A}{I_{sat}})} + \frac{T_{R,B}^2 \cdot I_B \cdot t}{2 \cdot T_C \cdot (1 - \frac{I_B}{I_{sat}})} + \frac{I_{sum} \cdot T_{C,max}^2 \cdot 3600 \cdot I_{sum} \cdot t}{4 \cdot I_{sat} \cdot T_{G,sum,max} \cdot (I_{sat} \cdot T_{G,sum,max} - I_{sum} \cdot T_{C,max})} \tag{13}$$

In practice, $T_{C,max}$ for dynamic control is usually given by $T_{G,A,max} + T_{G,B,max} + T_{V,sum}$, where $T_{G,A,max} = T_{G,B,max}$, while the model assumes that $T_{G,A} + T_{G,B} \le T_{G,sum,max}$ instead of $T_{G,A} \le T_{G,A,max}$ and $T_{G,B} \le T_{G,B,max}$, which can lead to some minor error, but it is considered negligible, especially in the context of accuracy of the Webster's model. In fact, the whole added delay due to random arrivals with dynamic control should be negligible in most cases where fixed-time signals can be operated with sufficient capacity reserve to avoid oversaturation.

2.2 Inevitable delay with optimal control

Initially, the minimal total delay over an hour of isolated operation was estimated based on the assumption of known, homogenous (deterministic) traffic demand in each direction, and an optimal SP for that. The model applies Eq. (7) to different combinations of demand in both directions, up to 1800 veh/h in total. For each combination, the optimal SP is selected and the total delay for $t = 1$ h is calculated. Different values of $T_{V,sum}$ (40, 120, 300, and 500 s) and I_{sat} (2000, 1600, 1200, and 800 veh/h) were used to assess different WZ lengths and saturated flows. Furthermore, three distinct maximal total cycle times $T_{C,max}$ (300, 600, and 900 s) were used to evaluate their impact on the total delay.

In practice, dynamic traffic control can reasonably fulfil the assumptions by continuously detecting traffic and allowing the SP to adapt. However, fixed-time control would not be able to adapt to the random variance, leading to oversaturation as discussed in the previous section. Moreover, even the actual saturated flow is neither known, nor constant, making such theoretically optimal setting of the fixed-time SP impossible in practice.

2.3 Comparison of fixed-time and dynamic signal control in a hypothetical work zone

This section addresses whole-day operation and compares the dynamic and fixed-time control in a wide range of hypothetical situations. Known, homogenous traffic demand and saturated flow are still assumed for both control methods (i.e., there is no capacity reserve and Eq. (7) is used to estimate delay), allowing the fixed-time control to precisely match critical demand. A time-series of hourly volumes is used to define the stepwise constant demand (Fig. 3). The demand profile comes from a traffic survey conducted by [29] on road I/9 near Líbeznice, Czechia. The profile reflects the average workday demand on a surveyed road profile, marked by a pronounced asymmetry due to morning and afternoon commuting peaks to and from Prague. A symmetric demand profile was created by splitting the sum evenly between both directions for each hour.

Four scenarios were devised for comparison, differentiated by demand symmetry and control type:

1. Fixed-time SP, asymmetric demand: SP tailored to 1650 veh/h demand with a 51:49 T_G ratio based on peak demand in each direction (840 and 810 veh/h,

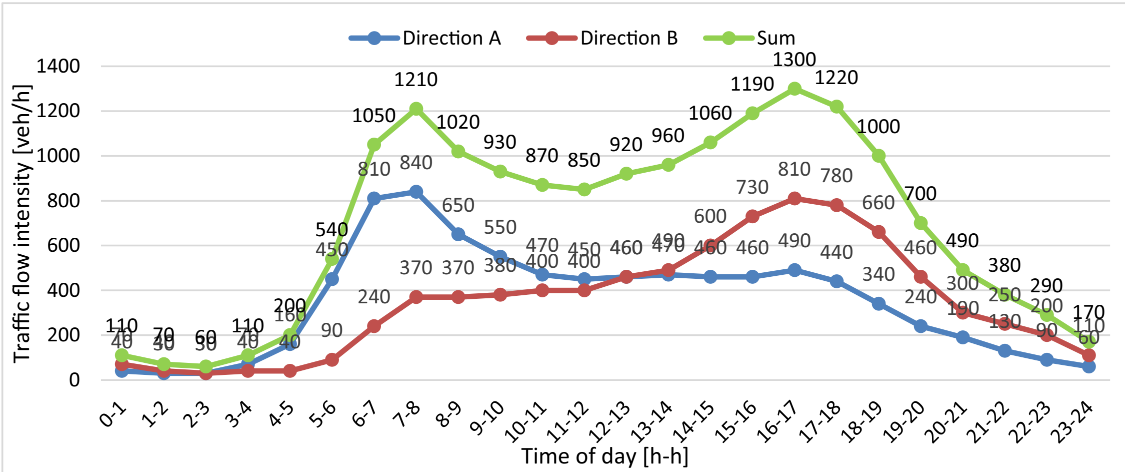


Fig. 3 Traffic flow profile in a hypothetical work zone

respectively). With $T_{V,sum}=40$ s and $I_{sat}=1800$ veh/h, this results in $T_C=480$ s and $C=1650$ veh/h.

2. Fixed-time SP, symmetric demand: SP adjusted for 1300 veh/h with 50:50 T_G ratio based on peak demand in each direction (650 veh/h both). With $T_{V,sum}=40$ s and $I_{sat}=1800$ veh/h, this results in $T_C=144$ s and $C=1300$ veh/h.
3. Dynamic SP, asymmetric demand: T_C and T_G vary hourly based on the demand sum and ratio. At 8–9 AM, total demand is 1020 veh/h (650 + 370) with a demand and T_G ratio 64:36, resulting in $T_C=94$ s, and $C=1034$ veh/h.
4. Dynamic SP, symmetric demand: principally identical to the previous case, demand and T_G ratio fixed at 50:50.

For each case, T_C was adjusted in 2 s increments, capped at 480 s (8 min). Every scenario was tested with two clearance times $T_{V,sum}$ (40 and 120 s) and three I_{sat} values (1800, 1500, and 1200 veh/h, representing the range of realistic values). If the demand exceeded capacity for given parameters, the control was assumed unsuitable.

2.4 Case study measurements

A field test was conducted on road I/56 near Ludgeřovice in a roughly 100 m-long WZ during bridge repairs. A non-intrusive microwave traffic detector was installed in the WZ on a signpost closer to one end of the WZ. The trial was divided into three periods:

- Period 1: Traffic was regulated using traffic lights with fixed-time SP set up by a contractor, as is common practice in Czechia. The road surface was milled, causing slow driving due to road steps and reducing saturated flow.
- Period 2: The WZ was re-ordered to utilize the finished half of the road. The traffic lights were replaced with a dynamic signal control system, although it still used fixed-time SP due to the significant change in road surface to gather data comparable with Period 3.
- Period 3: The dynamic control system was switched to dynamic mode. It was configured to turn off the green signal after 5 s without any detected incoming vehicle, with a minimal green signal duration of 30 s (not recommended). The clearance time per direction was reduced from 50 to 30 s to deter vehicles from driving on the red signal due to excessive clearance time.

Each measurement period lasted for about 13 days in May and June 2020. Raw vehicle data were processed to isolate individual cycles and directions. For the fixed-time plan, vehicle transits were aligned with the SP sequence by shifting the SP to match the start of each green phase. For dynamic control, the separation relied on the 30 s minimal green signal time, 30 s clearance time, and identification of sizeable gaps between vehicles. This makes the direction assignment unreliable but does not affect the performed analyses meaningfully.

Histograms of saturated flow were created for all three periods, considering only green phases with at least 10 consecutive vehicles with headways under 5 s at the phase start. Saturated flow in each phase was calculated as the average TF intensity between the first and last vehicle until a headway of over 4 s. The histograms were compared among the three periods. The efficiency of both control systems (static and dynamic) was evaluated similarly, with the 1. period assessed for comparison, too. Only phases with at least 10 (and 4 for comparison) recorded vehicles were considered. Vehicle counts per green phase were transformed into hourly volume and plotted in histograms. Since green phase start and end times for dynamic control weren't precisely known, its duration was measured from the first to last vehicle for both control types to make the results comparable. Average intensity over the whole assumed green phase duration was also estimated but must be interpreted cautiously. For the dynamic control, the green phase duration was considered from the first to the last vehicle plus 5 s or a minimum of 30 s (not for delay modelling as 30 s minimal green duration is highly inefficient). If the minimal green signal duration was set to 5 s as suggested, the average intensity could not drop below the equivalent of 240 veh/h, assuming only one vehicle passing during a 15 s phase (5 s before first detection, 5 s continuous detection during arrival, 5 s waiting).

Finally, the recorded TF intensities were used to model total delay similarly to the hypothetical WZ. Equation (10) or (13) were used to estimate the delay in this case to reflect the additional delay caused by random arrivals. Average demand profiles for weekdays and weekends were created from collected data (Fig. 4). Three values of saturated flow (1200, 1500, and 1800 veh/h) were used to calculate capacity and estimate delays. All six cases were modelled with the contractor-set

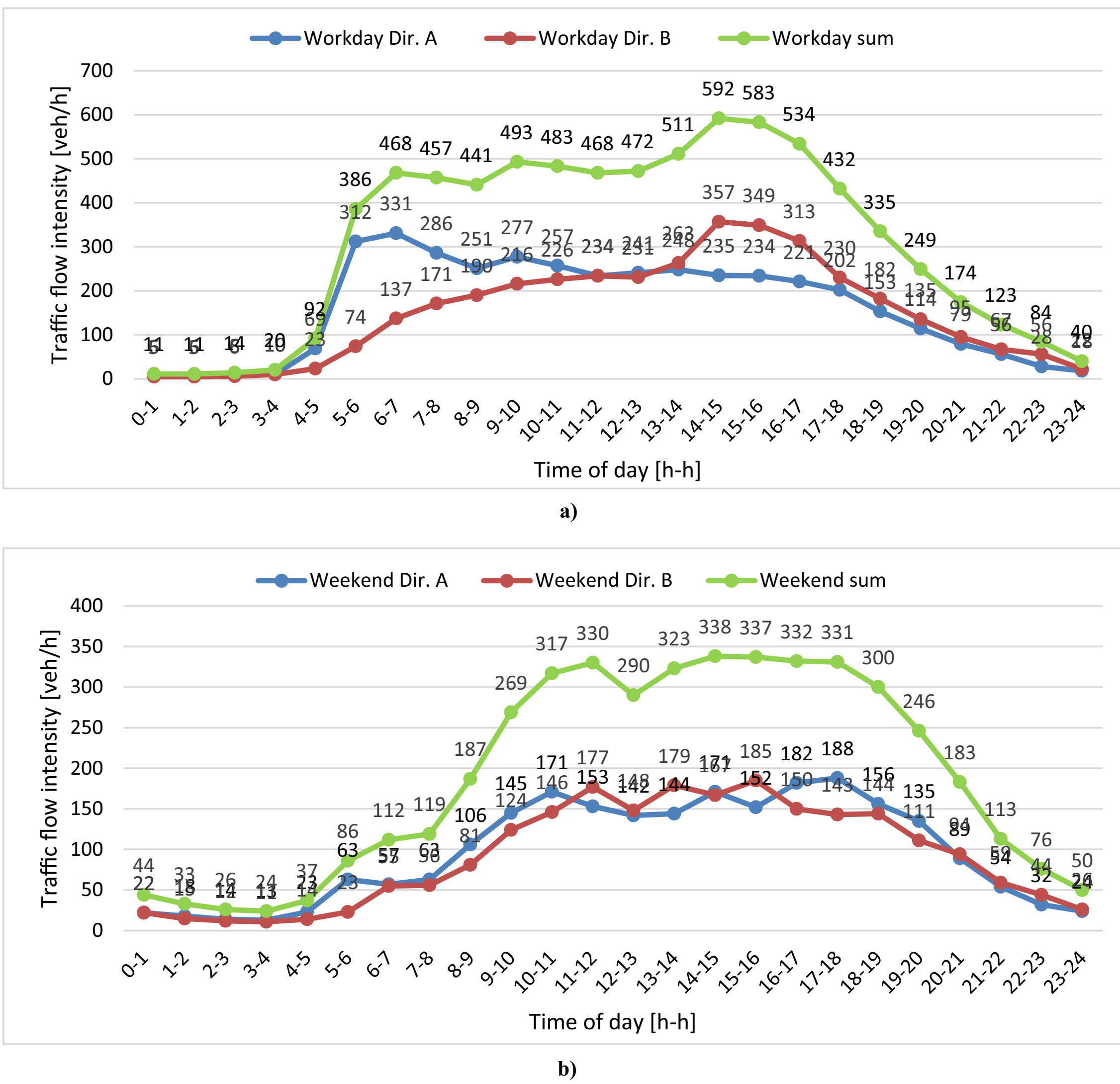


Fig. 4 Average traffic flow demand in the case study work zone; **a** workday **b** weekend

fixed-time SP, with an optimized fixed-time SP given known (estimated) traffic demand profile and saturated flow, and with a dynamic control. Both fixed-time control types were assumed not to change their signal plans for the weekend unless stated otherwise.

For the real SP model, the SP cycle duration T_C was 8:30 min (510 s), the sum of clearance times $T_{V,sum}$ was 100 s and the green signal ratio was 50:50. For optimal fixed-time control, $T_{C,max}$ was also set to 8:30 min and for dynamic control $T_{C,max}$ was set to T_C as required for the fixed-time control under the same circumstance. $T_{V,sum}$ was modelled as both 60 and 100 s for optimal fixed-time and dynamic control. The green signal ratios were 48:52 for the fixed-time control and varied for the dynamic control.

To better reflect reality, 20% capacity reserve was required for the optimal fixed-time control, while for the dynamic control, the minimal possible T_C was extended by 10 s (5 per direction) to mimic the 5 s vehicle detection window before switching to red signal (as well as to secure the minimal 5 s green signal duration as recommended).

3 Results and discussion

This section presents and interprets the results of observations and calculations that compare dynamic and fixed-time control in one-lane two-way WZs. The results are divided into three main subsections (Inevitable delay with optimal control, Hypothetical work zone, Case study), plus one with resulting recommendations.

3.1 Inevitable delay with optimal control

Only seven out of the possible 48 combinations of $T_{C,max}$, $T_{V,sum}$, and I_{sat} considered are presented to maintain conciseness while retaining all the key insights. For instance, limiting the total cycle time just cuts the delay graph off at lower TF intensity (compare Fig. 5a and b), so only the longest maximal cycle duration can be displayed without losing any information. The comparison also illustrates why, besides driver's acceptability and queue overspilling issues, the total cycle duration should be limited – 10.4% increase in capacity results in maximal delay growing by 305.1% (40.34 vs. 123.08 h/h). In the extreme, increasing the capacity (by increasing $T_{C,max}$) to match the demand can potentially increase delay to infinity while technically maintaining under-saturated conditions if $I_{sum} \leq I_{sat}$. As also argued in the Methodology section, Dion [22] considers Webster's stochastic model unsuitable for near-saturated conditions as the estimated delay grows to infinity, but that that could in principle happen even without considering random arrivals if the total cycle time was not limited.

The graph is truncated similarly when reduced I_{sat} limits the available capacity, but then the shape changes, too (Fig. 5b–d). The delay grows with decreasing I_{sat} for any demand combinations. The graphs also underscore that asymmetric demand can be very beneficial under optimal control, possibly in contrast to intuitions based on fixed-time control experience.

Figure 5e and f represent a longer WZ with $T_{V,sum} = 120$ s. The cases for $I_{sat} = 1600$ and 1200 veh/h are shown. Compare with the respective graphs from Fig. 5c and d to see the effect of increased clearance time. Besides the slight decrease in capacity, the delay increases markedly. Finally, Fig. 5g illustrates the effect of $T_{V,sum} = 300$ s.

3.2 Hypothetical work zone

This section presents results of the estimated total daily delay in a hypothetical WZ with either optimal fixed-time or dynamic signal control, assuming known, piecewise constant, homogenous (deterministic) demand and saturated flow. Average delay per vehicle can be obtained in combination with the intensities from Fig. 3.

Figure 6 illustrates the difference between the two control systems throughout the day with asymmetric demand. The largest savings with dynamic control are achieved in peak hours due to the highly asymmetrical peaks in both directions. This supports the earlier statement – asymmetric demand is beneficial if the SP is dynamically adjusted. On the contrary, it is detrimental for fixed-time control if the peaks occur at different times in each direction as it must provide capacity sufficient for their sum, making it even less efficient.

Tables 1 and 2 compile comprehensive results for the hypothetical WZ, considering varying saturated flow, demand, clearance time, and control types. With asymmetric demand (Table 1) and fixed-time control, capacity fell short of the relatively high traffic demand (16,700 veh/day) most of the time with T_C restricted to

(See figure on next page.)

Fig. 5 Minimal delay per hour for different combinations of TF intensity and WZ parameters: **a** $I_{sat} = 2000$ veh/h, $T_{V,sum} = 40$ s, $T_{C,max} = 300$ s → $C_{max} = 1730$ veh/h, $T_{D,max} = 40.34$ h/h, **b** $I_{sat} = 2000$ veh/h, $T_{V,sum} = 40$ s, $T_{C,max} = 900$ s → $C_{max} = 1910$ veh/h, $T_{D,max} = 123.08$ h/h, **c** $I_{sat} = 1600$ veh/h, $T_{V,sum} = 40$ s, $T_{C,max} = 900$ s → $C_{max} = 1520$ veh/h, $T_{D,max} = 88.67$ h/h, **d** $I_{sat} = 1200$ veh/h, $T_{V,sum} = 40$ s, $T_{C,max} = 900$ s → $C_{max} = 1140$ veh/h, $T_{D,max} = 66.50$ h/h, **e** $I_{sat} = 1600$ veh/h, $T_{V,sum} = 120$ s, $T_{C,max} = 900$ s → $C_{max} = 1380$ veh/h, $T_{D,max} = 95.24$ h/h, **f** $I_{sat} = 1200$ veh/h, $T_{V,sum} = 120$ s, $T_{C,max} = 900$ s → $C_{max} = 1040$ veh/h, $T_{D,max} = 73.67$ h/h, **g** $I_{sat} = 1600$ veh/h, $T_{V,sum} = 300$ s, $T_{C,max} = 900$ s → $C_{max} = 1060$ veh/h, $T_{D,max} = 87.57$ h/h, **h** common colour scale for graphs a)-g), values in [h/h] (hours of delay per hour of operation)

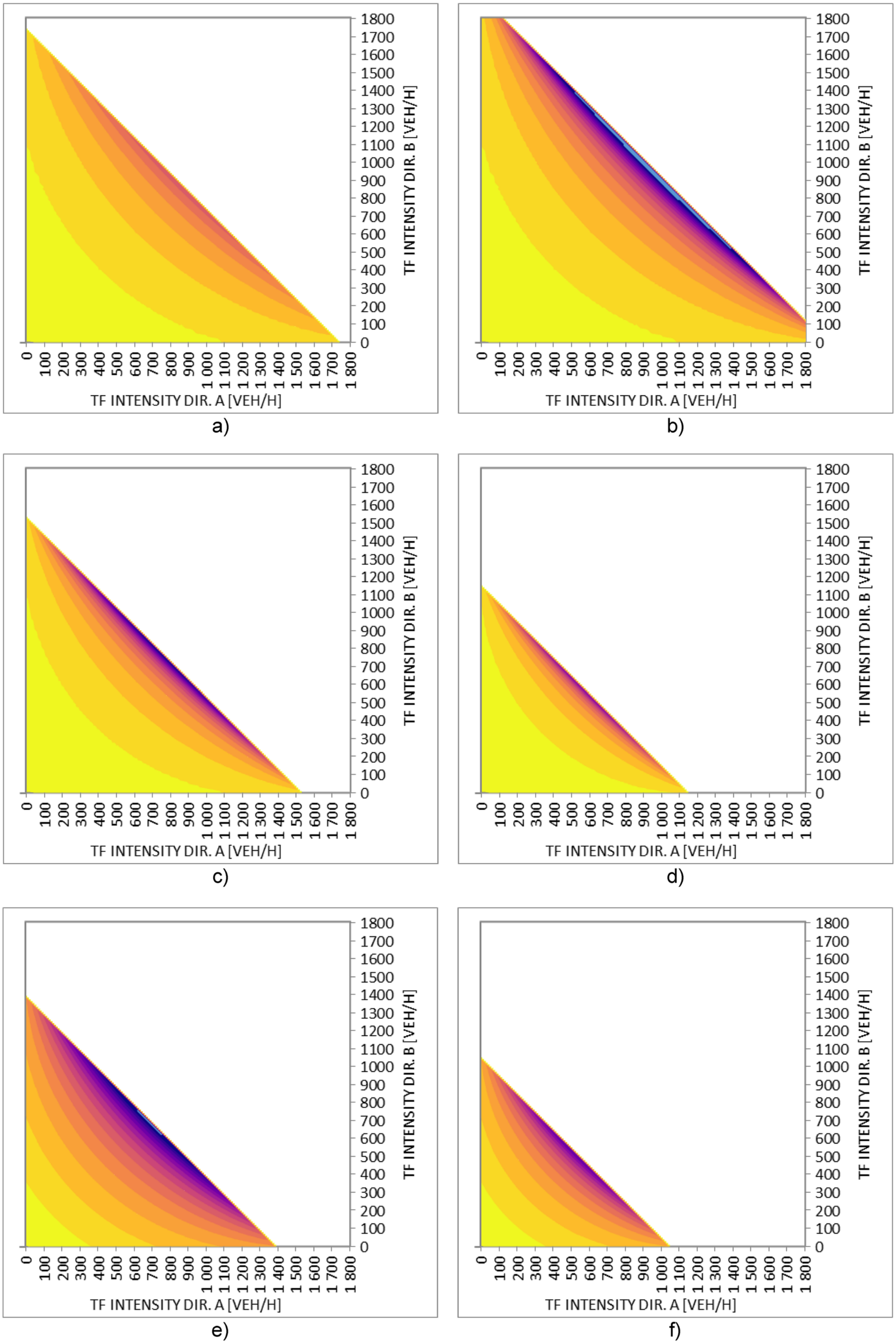
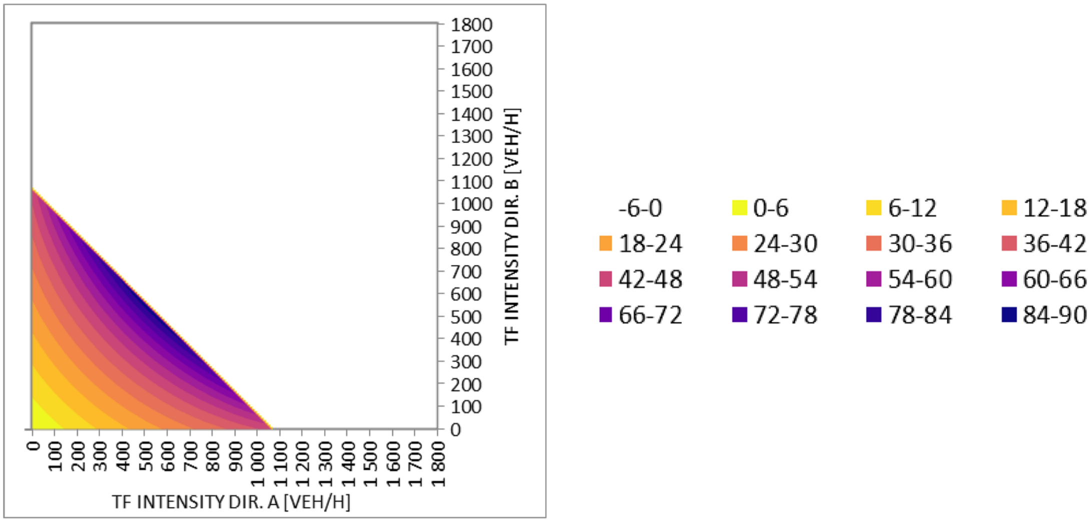


Fig. 5 (See legend on previous page.)



g) h)

Fig. 5 continued

8 min. Even dynamic control struggled to manage the full demand, especially with extended clearance times. However, in some cases, the control type was the difference between oversaturated and undersaturated conditions, where oversaturation could lead potentially up to double-digit multiples of lost time under specific circumstances. Where both control types maintained undersaturated conditions, the difference reached up to 213.6% (with dynamic control as the base). Reduced demand could be managed in undersaturated conditions under all scenarios with variable delays.

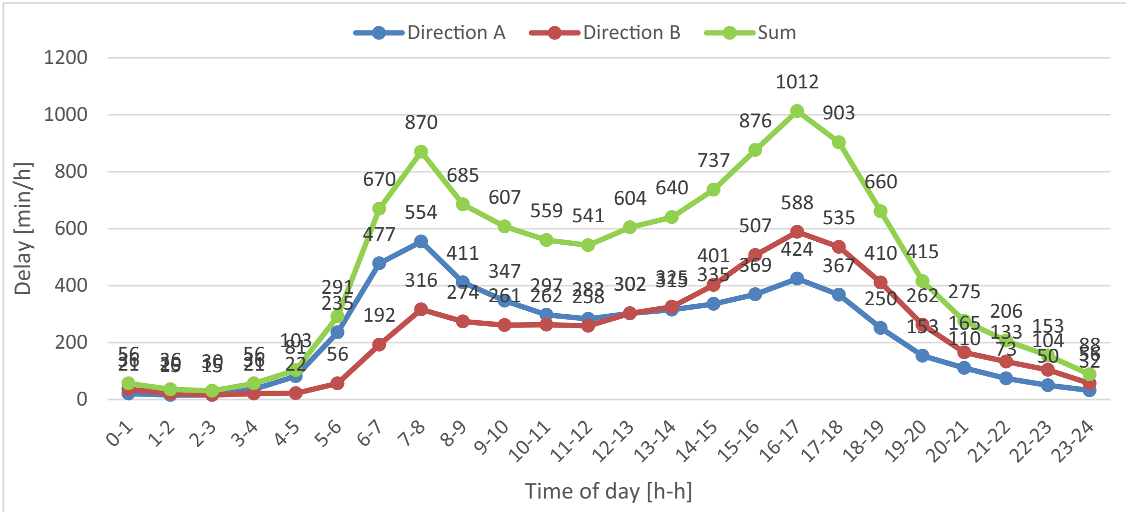
With symmetric demand (Table 2), the capacity of the fixed-time control sufficed in more cases as the capacity could be set for the overall peak demand, reducing the delay by ca 5–60%. Conversely, dynamic control could not fully capitalize on efficiently distributing green signal time to suit current demand, leading to a slight increase in delay (ca 1–4%). However, it still outperformed the fixed-time control in all cases. Overall, the closer the demand is to the potential capacity, i.e. saturated flow, the bigger the difference between the two control types.

All presented results emphasize that both delay itself and the disparity between the two control systems exhibit highly nonlinear behaviour, subject to circumstances. Saturated flow intensity, WZ length (clearance time), and demand (both magnitude and symmetry) all exert substantial influences and interact significantly. Nevertheless, there is compelling evidence supporting the efficiency of the dynamic approach. While the benefits appear marginal in some cases, they can be significant in others, often preventing oversaturation. Even when fixed-time control can provide adequate capacity, the disparity in total delay can be substantial (up to single digit multiples).

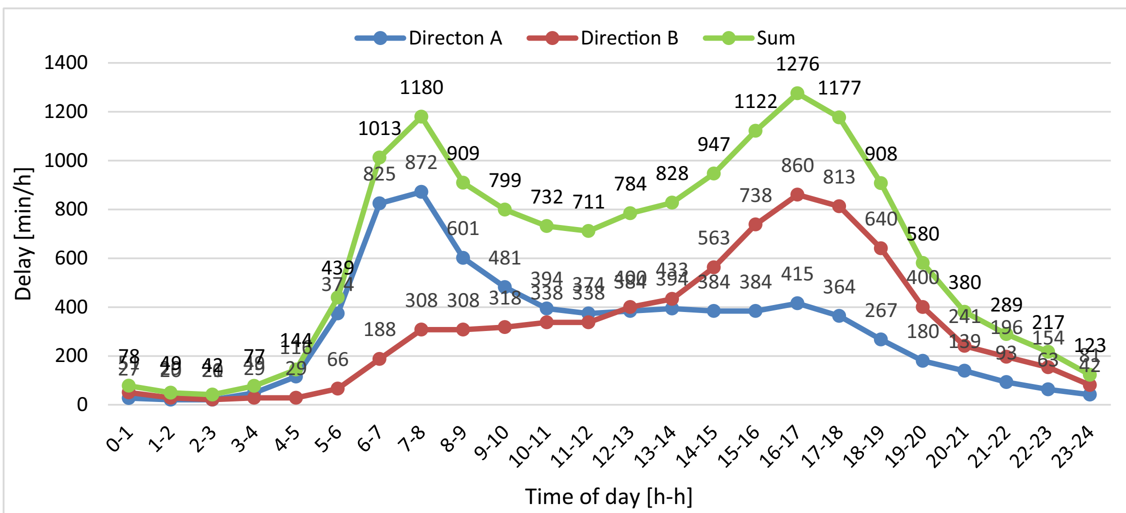
3.3 Case study

Figure 7a shows the distribution of saturated flow in the three measurement periods. The milled road surface in period 1 markedly reduced saturated flow in the WZ, reducing it by 30% (mean dropped approximately from 1660 to 1180 veh/h). There is a slight statistically significant difference between the distributions in the 2nd and 3rd periods (mean 1720 veh/h), possibly caused by higher share of passenger vehicles in the traffic flow due to the easing of covid-19 restrictions during the measurements.

Figure 7b presents the distribution of average flow during a green phase, measured between the first and the last recorded vehicle in phases with at least 10 passing vehicles. The mean values for periods 1–3 are approximately 590, 730, and 1410 veh/h, respectively. Dynamic control roughly doubled the average intensity during a green phase, i.e. the efficiency of the green signal. The ability to dynamically adjust the green signal and utilize its duration more efficiently is pivotal as it reduces the red signal duration in the opposite direction and, consequently, minimizes delays. Furthermore, this effect compounds over time, as the shorter red signals in the opposite direction result in fewer queuing vehicles, leading to even



a)



b)

Fig. 6 Modelled hourly time loss in a hypothetical work zone with **a** dynamic control **b** fixed-time control; I_{sat} = 1200 veh/h, $T_{V,sum}$ = 120 s, and 50% asymmetric demand

shorter green signal durations. The average TF intensity distributions for all periods with at least four vehicles are displayed in Fig. 8 for comparison. The distributions have heavier left tails as that is where most of the added phases with 4–9 vehicles are situated.

Because the duration of green phase is considered only from the first to last recorded vehicle (see Methodology section), the actual average intensity may have been even lower. In period 3, due to the extended minimal green duration, up to 20 s may be omitted. However, these results are representative for a shorter minimal green duration, such as the recommended 5 s. In periods 1 and 2 the green phase can be reduced even much further if no vehicles passed in the later part of the phase. If all recorded phases are included, a whole quarter of them has average TF intensity below 100 veh/h in these periods. These do not contribute too much to the total delay as only a few vehicles are waiting but are particularly frustrating for the waiting drivers.

Figure 9 combines the histograms of saturated and average flow from Fig. 7a, b, highlighting the (in)efficiency of the control as the effect of the road surface on capacity is already represented in the shifted distribution of the saturated flow. There is a notable difference in the distributions for both periods with fixed-time control. In

Table 1 Comparison of optimal fixed-time and dynamic signal control in a hypothetical work zone with asymmetric demand under different circumstances. Dash indicates insufficient capacity

Control	Delay for $T_{V,sum}$ = 40/120 s [h/day]	
***Isat* = 1800 veh/h**	**100% demand**	**50% demand**
Fixed-time optimal	468.7/–	59.6/178.9
Dynamic optimal	149.4/447.5	54.4/162.9
Difference	319.2/– (+213.6/– %)	5.3/15.9 (+9.7/+9.8%)
***Isat* = 1500 veh/h**	**100% demand**	**50% demand**
Fixed-time optimal	–/–	66.1/197.7
Dynamic optimal	197.9/–	56.8/170.3
Difference	–/– (–/–%)	9.3/27.4 (+16.3/+16.1%)
***Isat* = 1200 veh/h**	**100% demand**	**50% demand**
Fixed-time optimal	–/–	82.2/246.7
Dynamic optimal	–/–	61.6/184.5
Difference	–/–(–/–%)	20.7/62.2 (+33.5/+33.7%)

Table 2 Comparison of optimal fixed-time and dynamic signal control in a hypothetical work zone with symmetric demand under different circumstances. Dash indicates insufficient capacity

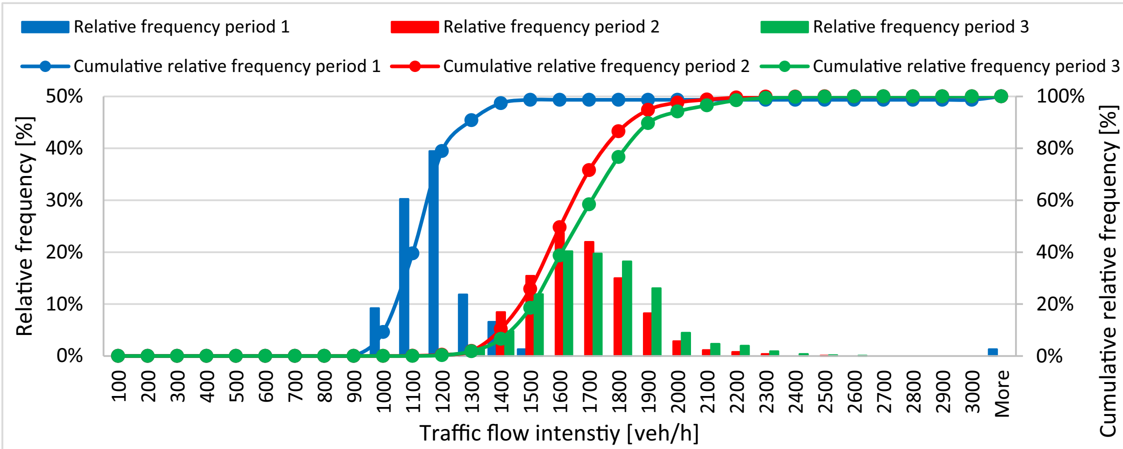
Control	Delay for $T_{V,sum}$ = 40/120 s [h/day]	
***Isat* = 1800 veh/h**	**100% demand**	**50% demand**
Fixed-time optimal	186.9/560.7	56.5/168.8
Dynamic optimal	154.5/462.5	55.2/165.1
Difference	32.4/98.2 (+20.9/+21.2%)	1.4/3.7 (+2.5/+2.2%)
***Isat* = 1500 veh/h**	**100% demand**	**50% demand**
Fixed-time optimal	332.1/–	60.1/179.5
Dynamic optimal	207.2/–	57.8/173.1
Difference	124.8/–(+60.2/–%)	2.3/6.4 (+4/+3.7%)
***Isat* = 1200 veh/h**	**100% demand**	**50% demand**
Fixed-time optimal	–/–	67.6/202.1
Dynamic optimal	–/–	63/188.6
Difference	–/–(–/–%)	4.6/13.5 (+7.3/+7.2%)

contrast, there is a much smaller shift for period 3 with the dynamic control, demonstrating its ability to make the most of the available capacity (saturated flow) despite the unnecessarily long minimal green signal duration.

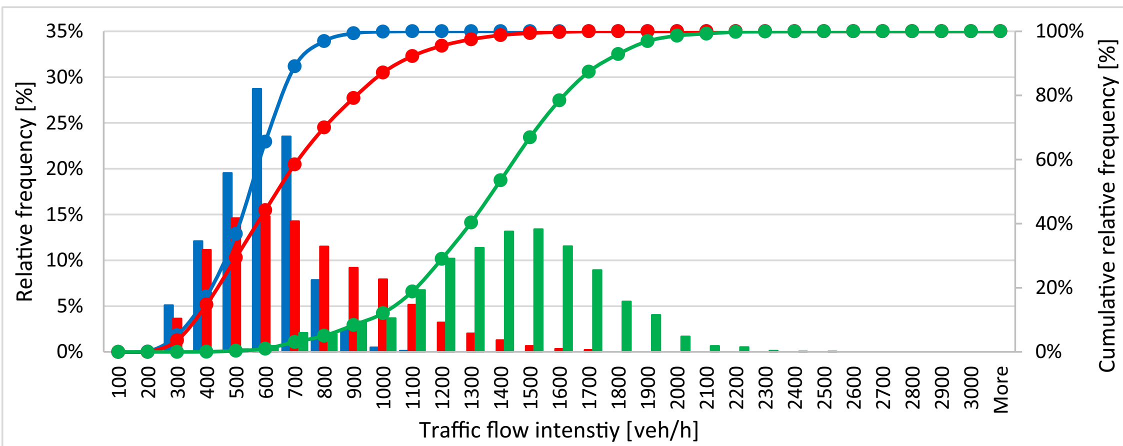
This is further illustrated in Fig. 10. At the top, the fixed-time SP is displayed with a red line. The sections at the top and bottom represent the green signal phases in each direction, while the part in the middle represents the clearance phases. There were prolonged instances with no vehicles traversing in one direction during green phase, while many vehicles accumulated in the other direction. The resulting queue then took most of the green phase to dissipate. Sometimes, vehicles drove through the red signal, either following vehicles after the signal turned red or at the end of the clearance period as vehicles stopped in the opposite direction could be seen. The WZ occasionally become briefly oversaturated in period 1 due to the milled surface despite the reduced traffic due to covid-19 restrictions. The dynamic control instead cut off the green signal if no arriving vehicle was detected over 5 s. The rapid switching of the right of way resulted in reduction of both the queuing vehicles and their waiting times, markedly reducing total delay. The progression of the right of way does not display the clearance phases, since the SP for dynamic control is only estimated from gaps between vehicle groups and known control rules.

The recorded traffic flow data were used to model the total delay per workday and weekend day depending on saturated flow and type of signal control (Table 3). The real fixed-time control assumes clearance time and SP as set up by the contractor. Comparing the results with an optimized control puts the earlier results into a different perspective. The difference between optimal fixed-time and dynamic control is not major in most conditions (10.2–36.7%). However, real-world fixed-time control rarely operates optimally, and the difference grew up to 79.4% with the real control setup. Additionally, no fixed-time control could sustain undersaturated conditions in the peak hours if saturated flow were below ca 900 veh/h (the recommended capacity reserve was not achievable for I_{sat} < ca 1050 veh/h) while the dynamic control would be able to manage the traffic even with saturated flow as low as 800 veh/h.

Table 4 provides the outcomes of the same models but with the $T_{V,sum}$ = 60 s for the two optimized control cases as this was well sufficient to clear the WZ. This reduction more than doubles the relative difference compared to the contractor configuration (up to 196.4%). It would also allow the optimal fixed-time control to maintain undersaturated conditions even with I_{sat} = 800 veh/h (950 with the reserve). The differences between optimal fixed-time and dynamic traffic control are 10.7–39.0%, so very similar in relative terms as with the longer clearance time but are nearly halved in absolute terms. This is in contrast with [14] who claim that longer WZs with longer clearance times offer less room for improvement. However, the conclusions that dynamic control outperforms all other control strategies are aligned. The absolute differences are unsurprisingly lower for weekend due to lower demand, but even the relative difference is lower between optimal fixed-time and dynamic control. This is likely



a)



b)

Fig. 7 Distributions of **a** saturated flow and **b** average traffic flow intensity between the first and the last vehicle in a green signal phase. Only phases with 10 + vehicles traversing are considered. Period 1 – milled road surface, fixed-time control; period 2 – new tarmac, fixed-time control; period 3 – new tarmac, dynamic control

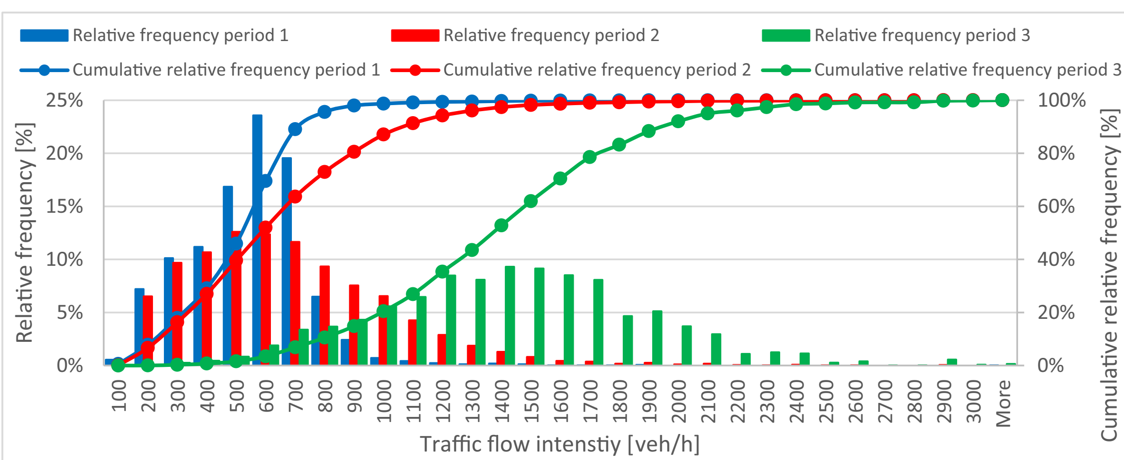


Fig. 8 Distribution of average traffic flow intensity between the first and the last vehicle in a green signal phase. All phases with 4 + vehicles are considered. Period 1 – milled road surface, fixed-time control; period 2 – new tarmac, fixed-time control; period 3 – new tarmac, dynamic control

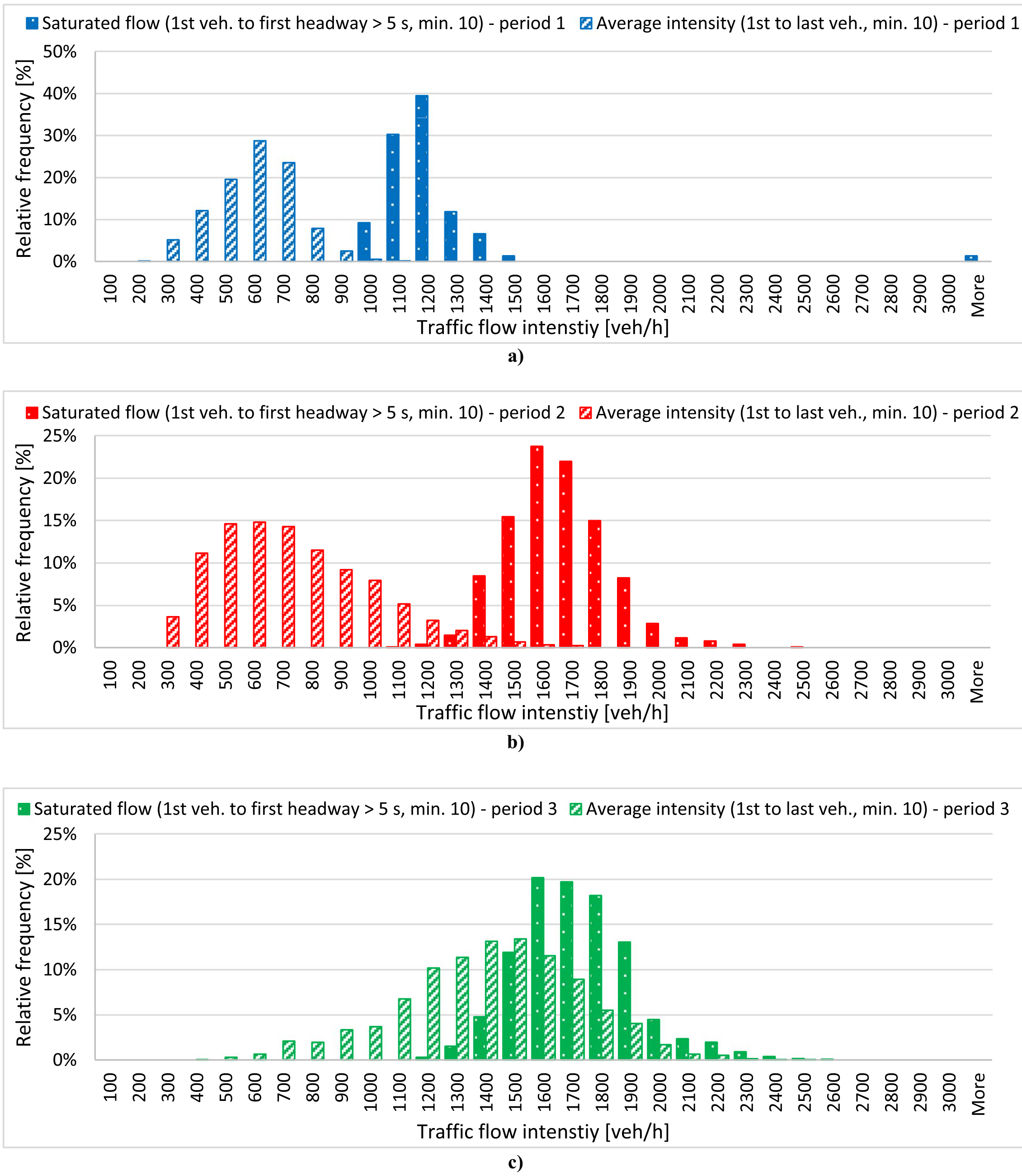


Fig. 9 Comparison of the distributions of saturated flow and actual average traffic flow during a green phase for **a** period 1 – milled road surface, fixed-time control; **b** period 2 – new tarmac, fixed-time control; **c** period 3 – new tarmac, dynamic control

because much of the delay is unavoidable in these cases and to lesser extent owing to the limited effect of random arrivals thanks to the large capacity reserve.

The results again highlight that the effect of efficient control steeply grows as the peak demand approaches the saturated flow. On the other hand, the potential savings and differences are much lower on weekends due to lower traffic. Reduced I_{sat} affects the total delay with fixed-time control (once it is set) only due to queues taking longer to dissipate and possibly due to the random arrivals (depending on capacity reserve) if the capacity is sufficient with the impact being fairly limited in the tested case (240.7 vs. 219.0 h/day or 10% for I_{sat} reduction from 1800 to 1200 veh/h). The optimal fixed-time control is affected much more (116.6 vs. 87.1 or 34%) as T_C is extended non-linearly (see Fig. 2) compared to I_{sat}

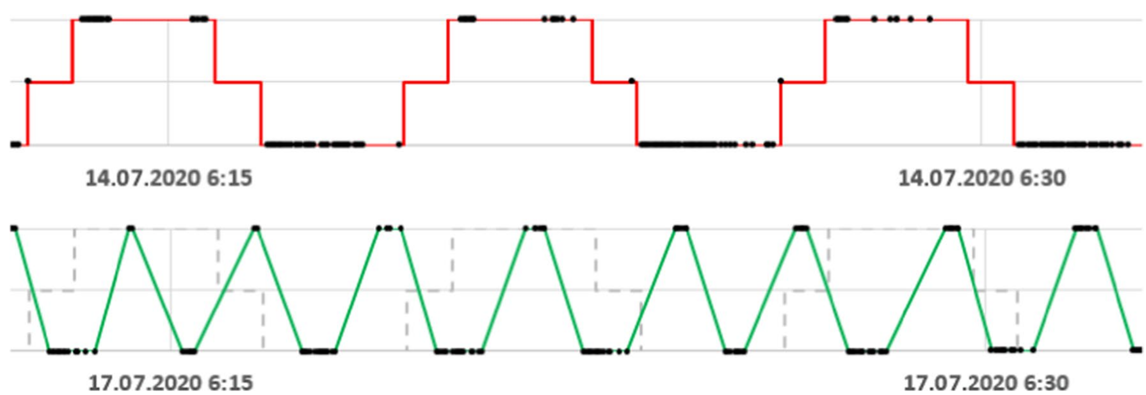


Fig. 10 Right-of-way progression and vehicle detection (black dots) during morning peak hours with fixed-time (period 2 – red) and dynamic control (period 3 – green)

Table 3 Expected delay with different saturated flows and signal control; clearance time $T_{V,sum} = 100$ s

Control	Delay [h/day]/[s/veh]	
***Isat* = 1800 veh/h**	**Workday**	**Weekend**
Real fixed-time	219/106	123.4/99
Optimal fixed-time	139.9/67	75.8/61
Optimal dynamic	122.7/59	68.8/55
Real fixed-time vs. dynamic	96.3 (+78.5%)	54.6 (+79.4%)
***Isat* = 1500 veh/h**	**Workday**	**Weekend**
Real fixed-time	226.9/109	125.6/100
Optimal fixed-time	153.2/74	82.1/66
Optimal dynamic	127.9/62	70.3/56
Real fixed-time vs. dynamic	99 (+77.4%)	55.4 (+78.8%)
Optimal fixed-time vs. dynamic	25.3 (+19.8%)	11.8 (+16.8%)
***Isat* = 1200 veh/h**	**Workday**	**Weekend**
Real fixed-time	240.7/116	129.3/103
Optimal fixed-time	188.4/91	99.1/79
Optimal dynamic	137.8/66	72.7/58
Real fixed-time vs. dynamic	102.9 (+74.7%)	56.6 (+77.8%)
Optimal fixed-time vs. dynamic	50.6 (+36.7%)	26.4 (+36.4%)

Table 4 Expected delay with different saturated flows and signal control; clearance time $T_{V,sum}$ optimized to 60 s ($T_{V,sum}$ is still 100 s for the real fixed-time control)

Control	Delay [h/day]/[s/veh]	
***Isat* = 1800 veh/h**	**Workday**	**Weekend**
Real fixed-time	219/106	123.4/99
Optimal fixed-time	87.1/42	46.1/37
Optimal dynamic	74.5/36	41.6/33
Real fixed-time vs. dynamic	144.5 (+193.9%)	81.8 (+196.4%)
Optimal fixed-time vs. dynamic	12.6 (+16.9%)	4.5 (+10.7%)
***Isat* = 1500 veh/h**	**Workday**	**Weekend**
Real fixed-time	226.9/109	125.6/100
Optimal fixed-time	95.2/46	49.8/40
Optimal dynamic	77.8/37	42.6/34
Real fixed-time vs. dynamic	149.1 (+191.7%)	83.1 (+195%)
Optimal fixed-time vs. dynamic	17.4 (+22.4%)	7.2 (+16.8%)
***Isat* = 1200 veh/h**	**Workday**	**Weekend**
Real fixed-time	240.7/116	129.3/103
Optimal fixed-time	116.6/56	60.2/48
Optimal dynamic	83.8/40	44.1/35
Real fixed-time vs. dynamic	156.9 (+187.1%)	85.2 (+193.4%)
Optimal fixed-time vs. dynamic	32.7 (+39%)	16.1 (+36.5%)

when it drops closer to the traffic demand to maintain the required capacity reserve. While the elongation of T_C is also true for dynamic control, the impact is much lesser as it does not need the capacity reserve (74.5 vs. 83.8 h/day or 12.5%).

Figure 11 illustrates the progression of average vehicle delay with optimal fixed-time and dynamic control. Fixed-time control displays much higher average waiting times, especially during the peak hours, as it is negatively affected by the asymmetrical demand, along with the increased demand. The dynamic control, on the other hand, benefits from the asymmetry and especially the morning peak is not even noticeable due to

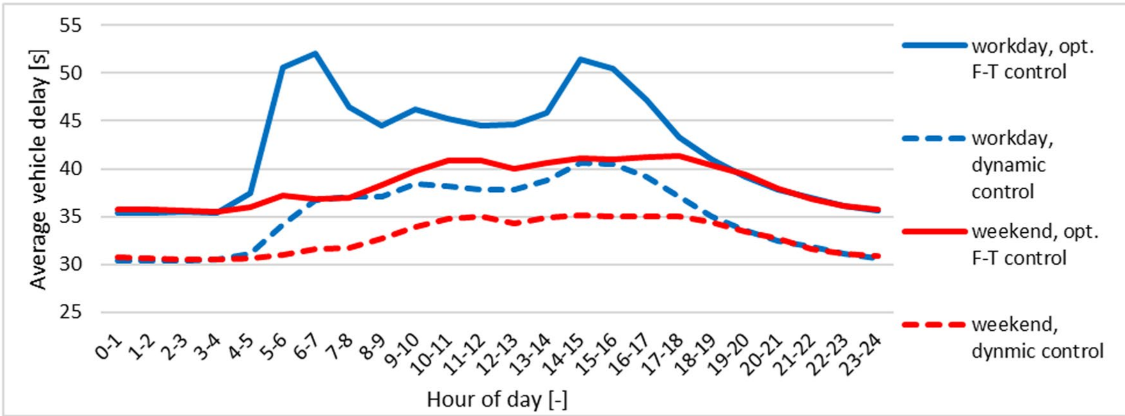


Fig. 11 Estimated average delay per vehicle during a day. $T_{V,sum}=60$ s, $I_{sat}=1500$ veh/h

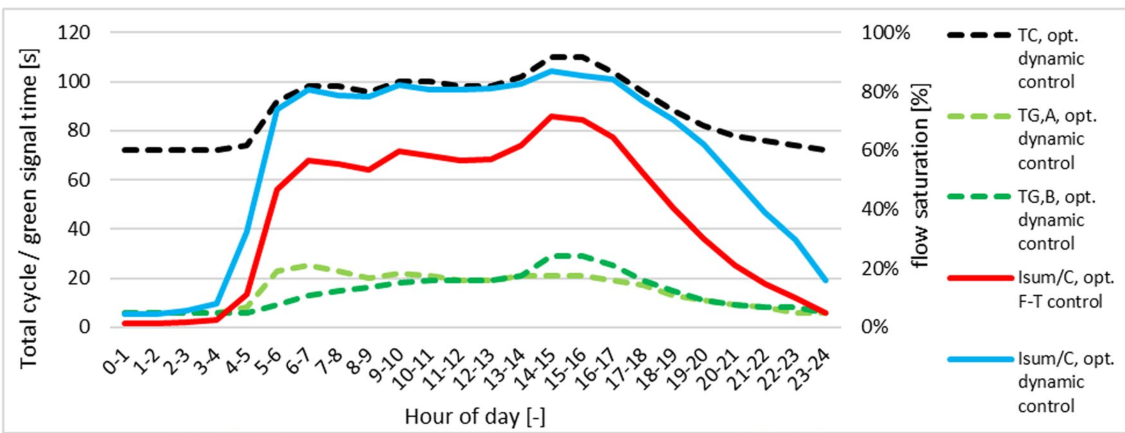


Fig. 12 Signal control parameters throughout the day with fixed-time or dynamic control

the low demand in one direction. Even in situations where the overall benefit of dynamic control is smaller, it avoids excess waiting when no vehicles are arriving from the opposite direction. This mitigates drivers' frustration and consequent aggressive driving, ultimately lowering the risk of accidents [30]. For comparison, the average waiting time with the real signal plan is 91–116 s (not shown for better clarity).

Figure 12 presents additional information about the models depicted in Fig. 11. Saturation (I_{sum}/C) for both control types for workday are shown along with SP data for dynamic control. Since $T_C=134$ s, $T_{G,A}=36$ s, and $T_{G,B}=38$ s are constant for fixed-time control, they are not displayed. The saturation is about 80% for most of the day with dynamic control despite the long 30 s minimal green signal time, while it stays around 50–60% even with optimally set up fixed-time control due to the asymmetrical demand.

Some other tests were performed to analyse how scheduling could improve performance of the fixed-time control. Separate signal plans were set for morning (midnight to midday) and afternoon (midday to midnight). This shown some interesting results based on I_{sat}, T_{Vsum}, and required capacity reserve. At best, with 30% required capacity reserve, low saturated flow (1200 veh/h) and long clearance time (50 s per direction), this reduced the delay by roughly 20% compared to having a single SP per whole day. However, with the capacity reserve reduced to 20%, the time savings dropped to just 12% as the efficiency of the unscheduled control improved much more. With that and saturated flow increased to 1800 veh/h, the delay was virtually equal in both cases. And with the clearance time reduced to 30 s per direction, the scheduled SP actually got worse due to the suboptimal capacity reserves in peak hours as the minimal recommended reserve was not enforced. Similar results occurred when the signal plan was set separately for day (5–20 h) and night (20–5 h) or for workday and weekend, except the suboptimal capacity reserves got much more pronounced due to the low demand in night and/or weekend. This highlighted the need for a minimal capacity reserve of 200 veh/h (or 100 per direction).

This indicates that where the conditions are not favourable, it is suitable to have timetable with different signal plans if fixed-time control is used, but where the conditions are favourable, the savings will be relatively low, although this is also true for dynamic control. Overall, scheduled fixed-time control sits somewhere between the unscheduled and dynamic control, but only as long as the recommended minimal capacity reserve is maintained.

3.4 Recommendations

Several policy and standards recommendations can be deduced from the findings regarding different control types, evaluation of real-world operation, or measurements of saturated flow. These recommendations primarily focus on broader adoption of more efficient traffic control systems, optimizing work zone designs to enhance capacity, and establishing standards for setting up different types of traffic control, leading to reduced delays and improved safety.

- Mandatory dynamic control on all major roads (arterials, I. class roads etc.).
- Recommended/mandatory dynamic control elsewhere if green signal time in either direction exceeds certain thresholds, as this variable encompasses all the other variables (I, I_{sat}, $T_{V,sum}$, and consequently T_R and T_C). Suggested thresholds: 30 s for recommendation (time saved per day in order of hours), 60 s for strong recommendation (order of tens of hours per day), 120 s for obligation (up to hundreds of hours saved per day).
- Flagging can be used instead of dynamic control (flaggers should be educated on how to assign the right of way efficiently), but it is usually more expensive in developed countries and poses treat to the flaggers. Suitable for short term work zones or temporarily during irregular operation etc.
- Different timetable for morning and afternoon to cater for morning and afternoon peaks in opposite directions if fixed-time control is used, possibly for weekends, too.
- Design fixed-time control for critical demand (sum of peak flows in individual directions within the fixed SP period) and with sufficient capacity reserve of 20–30% (depending on accuracy of demand and saturated flow estimates) and at least 100 veh/h per direction to avoid temporary oversaturation.
- Acceptable total cycle time may vary based on region and circumstances roughly from 5 to 10 min.
- Short detection window before switching the green signal to red for dynamic control should be set depending on detection location and type such that roughly 5 s gap between two consecutive approaching vehicles should lead to the end of green signal with enough time for the second vehicle to stop safely at the stop line on the red signal.
- Maximal green signal duration for dynamic control set to the value needed for fixed-time control to allow for sufficient flexibility but to avoid unreasonably long green signal in cases where vehicles arrive continuously with gaps below 5 s but not at saturated flow. If fixed-time control would lead to excessive cycle time, limit green signal per direction to $(T_{C,max}-T_{V,sum})/2$.
- Short minimal green signal duration, 5 s recommended. Where relevant due to particularly low demand, constant red with green-on-demand may be used in either or both directions (reliable detection is crucial). Suitable especially for rarely used side entrances into the WZ (e.g. with only public transport allowed) where full-fledged signalling is not suitable due to elongation of total cycle time.
- Set safe but not excessively long clearance times, especially with good visibility into the work zone.
- Use dynamic control whenever significant excessive traffic, e.g. due to holidays or local events, is expected during the work zone duration as it can adapt to the increased (and often highly asymmetric) demand. Possibly extend maximal green signal time temporarily for the event duration.
- Design work zones and work timetables such that the saturated flow is not reduced due to milling which severely reduces saturated flow and available capacity, especially when lacking it. Other circumstances such as narrow lane, chicane, or steep incline can adversely affect the saturated flow, too.
- Avoid oversaturation if possible. If needed, consider re-routing part of the traffic (e.g. one direction to remove the shuttle operation), splitting the work zone in two (may decrease total delay even with the extended duration of roadworks), closing the section entirely to shorten work zone duration, or similar measures to minimize lost time.

4 Conclusions

The study analytically derives the relations between signal plan settings, overall capacity, and total delay per time, under varying traffic demand conditions, using some simplifying assumptions (deterministic, homogeneous traffic flow). The effect of random arrivals is then added based on Webster's formula. The models are used to estimate total delay per day for different scenarios and control strategies. The results are supported with outcomes of a case study, bridging the gap between theory and practice. The results clearly show that dynamic control is more efficient than fixed-time control. However, it also showed huge potential for time savings if the fixed-time control is set up well. Still, given the little extra cost of the dynamic systems compared to the fixed-time ones and to the overall costs of maintenance and reconstruction works, it should be used as much as possible, especially on busy roads. The total time savings are highly dependent on circumstances with the difference ranging from a few to hundreds of percents or hours of delay per day even with optimal setting and when both control types can maintain undersaturated conditions. In practice, especially fixed-time control is not set up optimally and is prone to human error and the dynamic control can potentially manage even higher demand without oversaturation. The cost–benefit ratio, based on circumstances, can be up to a hundred in Czechia and the situation is likely similar elsewhere although there is still space for fixed-time control on minor roads with short work zones. Additionally, improving the reliability of travel times and efficiency of the green signal can improve traffic safety by reducing the drivers' frustration. The findings of this study, which are principally consistent with existing literature, were used to formulate several best practices and policy recommendations to support worldwide adoption of more efficient traffic control in work zones with shuttle operation.

Future research and development should focus on estimating saturated flow for different work zone designs to enable accurate signal plan setting and good design choice. This can be helped by adopting certain policies such as avoiding driving over the milled surface if high capacity is required as the case study showed milled surface reduced saturated flow by about 1/3. Another way to improve capacity of a shuttle work zone is developing reliable methods of clearance detection that would allow dynamic adjustments of clearance time.

Abbreviations

ITS	Intelligent transport system
PCE	Passenger car equivalent
SP	Signal plan
TF	Traffic flow
WZ	Work zone

Acknowledgements

The article is partially based on a research project conducted for the Road and Motorway Directorate of the Czech Republic. The author would like to thank those who participated in the original study in any way and anyone who provided him with feedback, including the reviewers.

Authors contribution

The author contibuted with writing the whole paper including all computations and data analysis and visualisation. The signalling and data collection system for the case study section was installed and operated by technical staff as part of the original study for ŘSD ČR.

Funding

This article has been published open access with support of the TRA2024 project funded by the European Union. This article was produced with the financial support of the Ministry of Transport within the programme of long-term conceptual development of research institutions.

Availability of data and materials

Data will be made available upon reasonable request.

Declarations

Competing interest

The author has no competing interests.

Received: 28 November 2023 Accepted: 22 August 2024
Published online: 14 October 2024

References

1. Srivastava, A., & Geroliminis, N. (2013). Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C: Emerging Technologies, 30*, 161–177. https://doi.org/10.1016/j.trc.2013.02.006
2. Ramadan, O. E., & Sisiopiku, V. P. (2018). Modeling Highway Performance under Various Short-Term Work Zone Configurations. *Journal of Transportation Engineering, Part A: Systems*. https://doi.org/10.1061/jtepbs.0000176
3. Newell, G. F. (1993). A simplified theory of kinematic waves in highway traffic, part I: General theory, part II: Queueing at freeway bottlenecks, part III: Multi-destination flows. *Transportation Research Part B, 27*, 281–313.
4. Cicic, M., Mikolasek, I., & Johansson, K. H. (2020). Front tracking transition system model with controlled moving bottlenecks and probabilistic traffic breakdowns. *IFAC-PapersOnLine, 53*, 14990–14996. https://doi.org/10.1016/j.ifacol.2020.12.1997
5. Brilon, W., Geistefeldt, J., & Zurlinden, H. (2007). Implementing the concept of reliability for highway capacity analysis. *Transportation Research Record*. https://doi.org/10.3141/2027-01
6. Elefteriadou, L., Lertworawanich, P. (2003) Defining, measuring and estimating freeway capacity. 82nd Annual Transportation Research Board Meeting, 12–16, Washington DC
7. Al-Kaisy, A., & Hall, F. (2003). Guidelines for estimating capacity at freeway reconstruction zones. *Journal of Transportation Engineering, 129*, 572–577. https://doi.org/10.1061/(ASCE)0733-947X(2003)129:5(572)
8. Asgharzadeh, M., & Kondyli, A. (2018). Comparison of highway capacity estimation methods. *Transportation Research Record, 2672*, 75–84. https://doi.org/10.1177/0361198118777602
9. Garber, N. J., Zhao, M. (2002) Crash characteristics at work zones. Charlottesville
10. Weng, J., Xue, S., Yang, Y., et al. (2015). In-depth analysis of drivers' merging behavior and rear-end crash risks in work zone merging areas. *Accident Analysis and Prevention, 77*, 51–61. https://doi.org/10.1016/j.aap.2015.02.002
11. Papamichail, I., Papageorgiou, M., & Stamatakis, I. (2018). Feedback traffic control at highway work zones using variable speed limits. *IFAC-PapersOnLine, 51*, 329–336. https://doi.org/10.1016/j.ifacol.2018.07.054
12. Tympakianaki, A., Spiliopoulou, A., Kouvelas, A., et al. (2014). Real-time merging traffic control for throughput maximization at motorway work zones. *Transportation Research Part C: Emerging Technologies, 44*, 242–252. https://doi.org/10.1016/j.trc.2014.04.006
13. Yang, X., Lu, Y., & Lin, Y. (2016). Optimal variable speed limit control system for freeway work zone operations. *Journal of Computing in Civil Engineering*. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000610
14. Al-Kaisy, A., & Kerestes, E. (2006). Evaluation of the effectiveness of single-lane two-way traffic control at maintenance and reconstruction

zones. *Canadian Journal of Civil Engineering, 33*, 1217–1226. https://doi.org/10.1139/L06-060
15. Finley, M. D., Songchitruksa, P., Jenkins, J. (2015) Evaluation of alternative methods of temporary traffic control on rural one-lane, Two-way Highways
16. Chien, S., Tang, Y., & Schonfeld, P. (2002). Optimizing work zones for two-lane highway maintenance projects. *Journal of Transportation Engineering, 128*, 145–155. https://doi.org/10.1061/(ASCE)0733-947X(2002)128:2(145)
17. Kastlová O, Houšť R (2020) Transport yearbook Czech Republic 2020. Ministry of Transport, Zlín
18. Newell, G. F. (1969). Properties of vehicle-actuated signals: I. *One-Way Streets. Transportation Science, 3*, 30–52. https://doi.org/10.1287/trsc.3.1.30
19. Webster F V (1958) Traffic signal settings
20. Son, Y. T. (1999). Queueing delay models for two-lane highway work zones. *Transportation Research Part B: Methodological, 33*, 459–471. https://doi.org/10.1016/S0191-2615(98)00043-5
21. Cassidy, M. J., Son, Y. T., & Rosowsky, D. V. (1994). Estimating motorist delay at two-lane highway work zones. *Transportation Research Part A, 28*, 433–444. https://doi.org/10.1016/0965-8564(94)90025-6
22. Dion, F., Rakha, H., & Kang, Y. S. (2004). Comparison of delay estimates at under-saturated and over-saturated pre-timed signalized intersections. *Transportation Research Part B: Methodological, 38*, 99–122. https://doi.org/10.1016/S0191-2615(03)00003-1
23. Ceder, A. (2000). Application of an optimal traffic control during lane closure periods of a two-lane road. *Journal of Advanced Transportation, 34*, 173–190. https://doi.org/10.1002/atr.5670340202
24. Schonfeld, P., & Chien, S. (1999). Optimal work zone lengths for two-lane highways. *Journal of Transportation Engineering, 125*, 21–29. https://doi.org/10.1061/(ASCE)0733-947X(1999)125:1(21)
25. U.S. Department of Transportation FHA (2009) Manual on uniform traffic control devices for streets and highways. 818
26. Zhu, W., Li, Z., Ash, J., et al. (2017). Capacity modeling and control optimization for a two-lane highway lane-closure work zone. *Journal of Transportation Engineering, Part A: Systems*. https://doi.org/10.1061/JTEPBS.0000078
27. Zakariya, A. Y., & Rabia, S. I. (2016). Estimating the minimum delay optimal cycle length based on a time-dependent delay formula. *Alexandria Engineering Journal, 55*, 2509–2514. https://doi.org/10.1016/j.aej.2016.07.029
28. Transportation Research Board (2016) Highway capacity manual—A guide for multimodal mobility analysis, 6th ed. Transportation Research Board
29. AF-CityPlan (2014) Měření intenzit automobilové dopravy na vybraných profilech v okolí obce Líbeznice. Prague
30. Su, Z., Woodman, R., Smyth, J., & Elliott, M. (2023). The relationship between aggressive driving and driver performance: A systematic review with meta-analysis. *Accident Analysis and Prevention, 183*, 106972. https://doi.org/10.1016/j.aap.2023.106972

Publisher's Note