applied sciences

Article

Research on Optimizing Low-Saturation Intersection Signals with Consideration for Both Efficiency and Fairness

Lingxiang Zhu [1], Lujing Yu [2] and Liang Zou [2,*]

[1] College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China
[2] College of Civil and Traffic Engineering, Shenzhen University, Shenzhen 518069, China
* Correspondence: zouliang@szu.edu.cn; Tel.: +86-135-5488-3640

Abstract: In response to the fairness issue arising from the unequal delay of vehicles in different phases at intersections and considering the actual situation of small and variable delays for vehicles in low-saturation intersection phases, this paper proposes the concept of "sacrificing efficiency for fairness". Firstly, the universality of unfair delay phenomena at intersection phases is explained, especially at low-saturation intersections where the fluctuation in phase delays is 1.87 times higher than at other intersections. Then, a fairness evaluation index is constructed using information entropy, and the feasibility of the proposed approach is demonstrated. Subsequently, a signal optimization model that balances efficiency and fairness is proposed. Finally, the proposed model is validated through case studies, showing that it not only simultaneously considers efficiency and fairness but also has minimal impact on efficiency. Moreover, the changes to timing schemes in the efficiency model are much smaller compared to the model that only considers fairness. Sensitivity analysis reveals that the model performs better under low-saturation intersection conditions.

Keywords: information entropy; conversion rate; average vehicle delay; saturation; sensitivity analysis

check for updates

Citation: Zhu, L.; Yu, L.; Zou, L. Research on Optimizing Low-Saturation Intersection Signals with Consideration for Both Efficiency and Fairness. *Appl. Sci.* **2024**, *14*, 8047. https://doi.org/10.3390/app14178047

Academic Editors: Gianluigi Ferrari and Francesco Zirilli

Received: 26 June 2024
Revised: 21 August 2024
Accepted: 28 August 2024
Published: 8 September 2024

1. Introduction

In recent years, China's comprehensive strength has significantly increased, and socialist development has become more democratized, leading to a higher happiness index among the people. However, there are still areas that require improvement. At the 19th National Congress of the Communist Party of China, General Secretary Xi Jinping pointed out, "Socialism with Chinese characteristics has entered a new era, and the principal contradiction in our society has evolved into the contradiction between unmet growing needs for a better life and unbalanced and inadequate development" [1]. With the stable development and improvement of China's social productivity, imbalances in social development have emerged as a major constraint to people's pursuit of a better life. Throughout history, social fairness has always been a common aspiration of humanity, and achieving social fairness is one of the greatest demands for social development. To meet the people's aspirations for a better life, the central leadership of the Party attaches great importance to the issue of social fairness, repeatedly emphasizing its importance and stressing the need to focus more on social fairness and prioritize people's well-being in the construction process across various domains. Transportation is an integral component of socioeconomic development and, to some extent, reflects social fairness [2]. Improving transportation fairness can contribute to enhancing social fairness. Currently, in various aspects of our daily lives, we have adopted strategies to improve transportation fairness. For instance, there is a strong emphasis on giving priority to public transportation. Measures such as creating dedicated bus lanes and HOV lanes aim to enhance the efficiency and environment of bus travel, thus reducing the commuting disparity among different income groups. To put people first, facilities such as tactile paving on pedestrian walkways and barrier-free elevators at subway transfer stations are installed to meet the travel needs of special populations. These are all effective strategies for enhancing both transportation fairness and social fairness.

Urban road intersections serve as crucial nodes in urban road networks, with signal control playing a pivotal role. However, they also represent bottlenecks that cause traffic delays. According to the "2020 Second Quarter Analysis Report on Traffic in Major Cities in China" released by Gaode Map, the average delay at intersections during peak hours in major Chinese cities exceeds 30 s per vehicle, with Shenzhen recording the highest delay at 39.22 s per vehicle. In urban signalized intersections, the most direct manifestation of fairness is the consistency of average delay across different phases. However, currently, most signal timing plans are designed with the objective of minimizing overall delay for vehicles without considering the unfairness caused by differences in average delay among phases before establishing the optimization model. Equal average delay among phases is only achieved when the saturation levels of each phase are equal. The distribution of urban traffic volume over time is uneven, with significantly lower traffic during off-peak hours compared to rush hours. For example, during off-peak hours, the hourly traffic volume in cities like Shenzhen and Guangzhou is only one-third of that during rush hours. Therefore, sacrificing intersection delay during off-peak hours to achieve fairness is feasible.

Considering the sustained attention to transportation fairness in society and the actual situation of low saturation at intersections during off-peak periods, this paper proposes a signal timing optimization approach for low-saturation intersections, sacrificing some delay to achieve fairness. Through an in-depth analysis of intersection efficiency and fairness, this paper first illustrates the universality of fairness issues in intersection phase delay. Next, the feasibility of the proposed approach is verified. Furthermore, a signal timing optimization model that considers both efficiency and fairness is constructed. Finally, the effectiveness of the model is validated using case studies, and sensitivity analysis is conducted on intersection saturation. The Technical roadmap is shown in detail in Figure 1. The first section of the paper summarizes the current research status in related fields domestically and outlines the main research content. The second section delves into the fairness of each phase of the Webster model from both theoretical and simulation perspectives. The third section introduces a fairness evaluation function based on information entropy, upon which a signal optimization model for low-saturation intersections considering delay and fairness is built. In the fourth section, the effectiveness of the model is analyzed using case studies, and sensitivity analysis is conducted on intersection saturation.

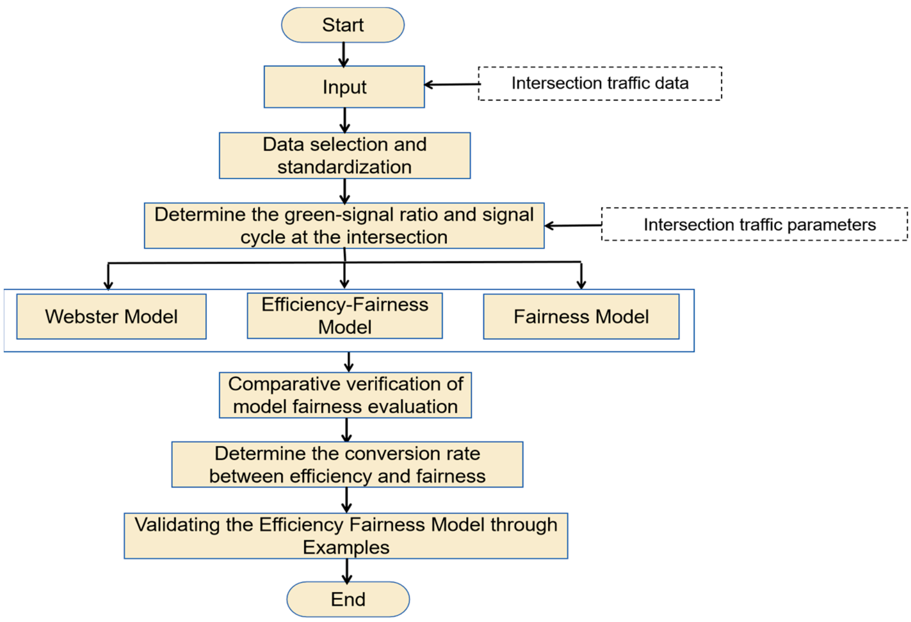

Figure 1. Technical roadmap.

2. Literature Review

2.1. General Traffic Signal Timing Optimization Problem Framework

In typical traffic signal optimization problems, the goal is to optimize the traffic flow at intersections by adjusting signal timing. The signal control parameters include the green light time, saturation rate, and signal period for each phase, which affect the efficiency of traffic flow. The optimization problem of traffic signals generally selects the minimum total delay time within the region as the objective function of the optimization model. Currently, the models used to calculate delay at signalized intersections mainly include the Webster delay model [3], the ARRB model [4], and the HCM2000 model [5]. The following is the general framework of this optimization problem.

(1) Objective function:

The average delay time of vehicles at intersections can be expressed as Equation (1):

$$d = \frac{C(1-\lambda)2}{2(1-\lambda \cdot x)} + \frac{x^2}{2q(1-x)} \tag{1}$$

Among them:

$$\lambda = \frac{g}{C}, x = \frac{q}{S} \tag{2}$$

C is the signal period, λ is the green signal ratio, x is the saturation rate, q is the actual traffic flow at the intersection, g is the green light time, and S is the intersection saturation flow rate.

Therefore, the objective function for minimizing the total delay time in the region is expressed as Equation (3): [6]

$$Delay = \min \sum_{k=1}^{n} \sum_{i=1}^{m} \left[\frac{C\left(1-\lambda_i^k\right)^2}{2\left(1-\lambda_i^k \cdot x_i^k\right)} + \frac{\left(x_i^k\right)^2}{2q_i^n\left(1-x_i^k\right)} \right] \cdot q_i^k \tag{3}$$

In the formula, n is the number of intersections, and m is the number of phases.

(2) Constraint condition

According to the general situation, it is stipulated that there are shortest and longest green light durations for each phase in the optimization process, ensuring that each phase can obtain the right-of-way within one cycle [7]. The constraints for solving are:

1. Minimum Green Light Time Constraints:

The green light time for each phase cannot be less than a minimum time to ensure the safe passage of vehicles:

$$g_i \geq g_{\min}, \forall i \tag{4}$$

Among them, $g_{\min}$ is the minimum green light time for each signal stage.

2. Cycle Length Constraints:

$$C_{\min} \leq C \leq C_{\max} \tag{5}$$

Among them, $C_{\max}$ and $C_{\min}$ are the maximum and minimum periods of the intersection.

(3) Mathematical optimization model

The mathematical model framework for optimizing the timing of general traffic signals can be represented by the following Equation (6) and Constraint (7):

$$Delay = \min \sum_{k=1}^{n} \sum_{i=1}^{m} \left[\frac{C\left(1-\lambda_i^k\right)^2}{2\left(1-\lambda_i^k \cdot x_i^k\right)} + \frac{\left(x_i^k\right)^2}{2q_i^n\left(1-x_i^k\right)} \right] \cdot q_i^k \tag{6}$$

$$s.t.\begin{cases} \lambda = \frac{g}{C} \\ x = \frac{q}{S} \\ C_{\min} \le C \le C_{\max} \\ g_i \ge g_{\min}, \forall i \end{cases} \tag{7}$$

At the same time, research on signal timing optimization at intersections has primarily focused on minimizing delay, with the most classic being the Webster timing optimization scheme [8]. In recent years, on the one hand, scholars have innovated signal timing optimization schemes based on considering delay. On the other hand, foreign scholars have also been continuously exploring the consideration of fairness in the signal optimization process.

2.2. Traffic Signal Timing Optimization

Signal timing control schemes directly impact the operational effectiveness of intersections. In foreign research, many scholars [9–11] often construct mathematical function models targeting one or more parameters in the evaluation criteria for signal timing optimization. They utilize methods like genetic algorithms to solve the established models, followed by simulation. Results indicate that the proposed schemes can reduce intersection delay. Weal et al. [12] formulated an optimization model targeting the minimization of average delay at signalized intersections. Murat [13] and Schmoecker [14] proposed a multi-objective control model for single intersections based on fuzzy logic methods, selecting multiple performance indicators as optimization objectives. Chen [15] introduced a multi-objective optimization and decision-making method to optimize signal timing at individual intersections. Additionally, Deng et al. [16] applied data fusion technology to propose a new multi-objective signal control parameter optimization model for urban intersections. Wei J et al. [17] proposed an optimization method for traffic intersection signal control based on an adaptive artificial fish swarm algorithm, which reduces average delay, parking frequency, and travel time. Himabindu et al. [18] proposed a model-based demand response traffic control system that significantly reduces intersection delays and can adapt to different traffic demands. He et al. [19] proposed an optimization control model for signal parameters on main traffic roads, which optimizes signal cycles, green light times, and phase offsets.

There has been a significant amount of research in China, as well. Scholars [20–22] are constructing signal timing models for intersections, aiming to optimize multiple objectives by targeting one or more parameters in the intersection evaluation criteria and assigning different weights to these criteria based on varying traffic flow conditions. Li Xun et al. [23] investigated signal control issues at urban arterial road intersections, focusing on multiple intersections and establishing signal control function models. The results of model solutions effectively reduced average delay per vehicle and improved intersection capacity. Li Juan et al. [24] proposed a signal timing optimization model aimed at minimizing average delay per person at intersections based on total delay for motor vehicles, non-motor vehicles, and pedestrians, while considering the differences in two crossing modes for non-motor vehicles.

2.3. Transportation Fairness

The currently recognized concept of transportation fairness originates from the project report of the International Joint Highway Research in 1994: "Transportation fairness refers to the allocation of costs and benefits generated by a policy, typically considering various demographic groups" [25,26]. Litman [27] provided a comprehensive overview of this concept, suggesting that transportation fairness should include horizontal fairness, vertical fairness considering different classes and incomes, and vertical fairness considering differences in transportation abilities and needs.

Scholars have not only researched the influencing factors of transportation fairness but also explored the fairness issues between different modes of transportation. For instance, Lu Dandan [28] et al. analyzed factors such as residents' travel efficiency, road infrastructure quality, travel costs, and the impact of transportation facilities on the environment, summarizing the impact of these factors on transportation fairness. Scholars have also studied various approaches to measuring transportation fairness. For example, Delbose and Currie utilized mathematical models such as Lorenz curves and the Gini coefficient to investigate transportation fairness issues in the Melbourne area.

As the principle of people-oriented development gains increasing recognition, scholars have also studied the application of fairness in signal timing optimization research. Zhi Chun [29] developed a heuristic solving algorithm that combines penalty functions and simulated annealing methods, incorporating both environmental and fairness objectives into traffic signal timing problems by maximizing traffic capacity and minimizing traffic emissions. Ozgur Baskan [30] proposed a heuristic solving algorithm based on harmony search and penalty function methods, optimizing traffic signal timing schemes in urban road networks by considering traffic capacity and fairness constraints. Liang Zheng [31] proposed a dual-objective signal timing simulation optimization model based on uncertainty by balancing the Atkinson index (evaluating transportation fairness) and average travel time (evaluating transportation efficiency). With the widespread application of information entropy in engineering, technology, and socioeconomics, explorations in the field of transportation have also been conducted. Shi Jing [32] and others proposed a transportation fairness evaluation method considering regional fairness and fairness of benefit attribution to different groups based on the Wilson entropy model. Lv Bin [33] and others designed a phase difference optimization algorithm for the line control system using information entropy theory and multi-attribute decision-making methods, with travel time, vehicle delay, and queue length as evaluation indicators.

2.4. Summary

(1) Existing signal timing methods often optimize for one or several objectives and construct optimization functions, primarily focusing on delay as the target, with relatively few studies considering both delay and fairness.
(2) Information entropy is widely used in various fields, but there are relatively few studies applying it to signal timing optimization.
(3) Currently, research on fairness in the transportation field is relatively broad, with increasing attention being paid to fairness in signal control. However, there are relatively few studies on fairness regarding delay fairness for each phase.

In order to comprehensively consider the various objectives in signal timing optimization for low-saturation intersections and to reflect fairness, this paper proposes incorporating the differences in phase delay fairness into the optimization objectives. It utilizes information entropy and the Webster delay model as the basis to construct a fairness evaluation function, further establishing a signal timing optimization model considering both delay and fairness. The results validate the feasibility of sacrificing delay for fairness in this paper's approach. Finally, a multi-objective model incorporating delay, fairness, and emissions is constructed, and the model is verified through case studies, with sensitivity analysis conducted on intersection saturation.

3. Intersection Delay Fairness Analysis

Based on the actual conditions of the research object, this section first identifies Webster as the delay model. Subsequently, it analyzes the fairness of the Webster model from both theoretical and empirical perspectives. Finally, it utilizes information entropy to design a fairness evaluation function for intersection delay.

3.1. Delay Model Fairness Analysis

3.1.1. Synthetic Sample Generation

(1) Basic situation of the intersection

The study focuses on a typical four-lane intersection with east–west and north–south directions, each with separate left-turn lanes. Right-turn movements are not considered in this analysis.

(2) Traffic flow setting

Based on the "Urban Road Capacity Table of Various Levels" in China and empirical data, the saturation capacity for each lane is set at 1200 pcu/h/lane, which is appropriate for urban trunk roads. The traffic volume for each lane is randomly generated within the range of [0, 1200] pcu/h.

(3) Phase setting

Considering the crossover practical situation, this paper chooses the classical opponent four-phase, and the following Figure 2 gives the schematic diagram.

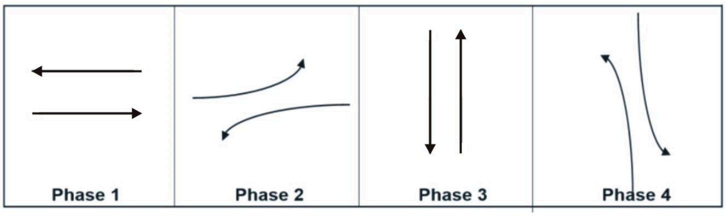

Figure 2. Classic opposite four-phase indication.

(4) Arithmetic sample generation

Using Python 3.10.5, 2500 sets of traffic flow data are randomly generated for each lane. Focusing on low-saturation intersections, 362 datasets with intersection saturation between 0.0 and 0.8 are selected for analysis.

(5) Sample description

In this case study, the total saturation degree is set between 0.05 and 0.8 with a gradient interval of 0.015. The total intersection traffic flow ranges from 600 to 5600 vehicles per hour (veh/h) with an interval of 100 veh/h, generating 2500 sets of basic cycle data. These data sets are traversed to calculate the coefficient of variation for average vehicle delay per phase and phase saturation degree.

Since the selected valid datasets are randomly generated, the distribution is relatively uniform. The average intersection traffic volume is 8448 pcu/h, with a maximum of 14,976 pcu/h and a minimum of 960 pcu/h; the corresponding average intersection saturation is 0.45, with a maximum of 0.78 and a minimum of 0.07, used for further analysis.

(6) Sample of generating code logic

Therefore, this paper uses a Python program to validate the model by employing the built-in scipy.optimize.minimize library to solve the constrained nonlinear objective function. The specific programming logic is as follows:

1. Since the research subject of this chapter is low-saturation urban intersections, the total intersection saturation range is set between 0.05 and 0.8 with a gradient interval of 0.015; the total traffic flow ranges from 600 veh/h to 5600 veh/h with a gradient interval of 100 veh/h. These are paired to form a set of basic data, totaling 2500 sets.
2. Under the constraints of not exceeding the phase saturation flow and meeting the specified intersection saturation and total traffic flow, random values are assigned to the actual flow of each phase. The generated phase flow values are used for model traversal and solution.

The sample plays a crucial role in validating our theory and approach. The detailed generation and solution of this instance can be found in the pseudo-code for instance generation and model solution in Supplementary Materials. Subsequently, we will delve deeper into the analysis of this instance. The sample features are shown in Tables 1 and 2.

Table 1. Intersection non-peak hour traffic flow.

Entrance	East Entrance			West Entrance			South Entrance			North Entrance		
	Left	**Straight**	**Right**	**Left**	**Straight**	**Right**	**Left**	**Straight**	**Right**	**Left**	**Straight**	**Right**
Number of lanes	1	2	1	1	2	1	1	2	1	1	2	1
Saturated traffic flow(veh/h)	1990	6828	2155	1990	6828	2155	2010	6780	2168	2020	6652	2168

Table 2. Vehicle driving parameters and signal parameters.

Type	Name	Symbol	Value	Unit
Vehicle Parameters	Free Flow Speed	v_f	11.1	m/s
	Average Deceleration	a_d	−2.5	m/s^2
	Average Acceleration	a_a	1.5	m/s^2
Signal Parameters	Number of Phases	M	4	
	Saturation Flow Rate (Phase 1)	s_1	6652	veh/h
	Saturation Flow Rate (Phase 2)	s_2	2010	veh/h
	Saturation Flow Rate (Phase 3)	s_3	6379	veh/h
	Saturation Flow Rate (Phase 4)	s_4	2005	veh/h
	Start-Up Loss Time	l	3	s
	Yellow Light Time	Y_A	3	s
	Green Light Interval Time	I	3	s

3.1.2. Theory Analysis

(1) Prevalence of unfairness

Since the study is on low-saturation intersections, $0 < x_i < x < 1$; therefore, Equation (8) < 0 is constant. From this, we can get that the phase vehicle average delay is monotonically decreasing, so there is only one case where each phase saturation is equal, i.e., $x_1 = x_2 = \cdots = x_i = \cdots = x_m = \frac{x}{m}$. The conclusion can be illustrated that signalized intersections designed based on the Webster model can only appear in specific cases where the phase vehicle average delays are equal, i.e., the phenomenon of absolute fairness, whereas the phenomenon of inequality is universal.

$$\frac{d}{dx_i}(d_i) = \frac{C(1-x_i)\left(\frac{x_i}{x}-1\right)}{2x\left(1-\frac{{x_i}^2}{x}\right)^2} \tag{8}$$

(2) The less saturated, the less fair

It is not difficult to find that the numerator part of Equation (8) monotonically decreases with x, so the denominator part can be simplified to the following Formula (9). Due to $0 < x_i < x < 1$, it can be known as monotonically increasing with x, so Equation (8) monotonically decreases with x, i.e., the smaller the degree of saturation, the greater the change in the phase car average and the more unfair.

$$x + \frac{x_i^4}{x} - 2x_i^2 \tag{9}$$

3.1.3. Sample Analysis

To further analyze, this study takes a four-phase single-point intersection as a sample and randomly generates 362 sets of effective basic data for different intersections. Using the Webster model, the cycle length, green time ratio, and phase-average delay of each intersection in each group are calculated. Considering that the coefficient of variation can eliminate the influence of measurement scales and dimensions, it is chosen as a parameter to measure the degree of difference in data, such as phase-average delay.

(1) Prevalence of unfairness

To measure the volatility of the dataset, the ratio of the standard deviation to the mean of the data is defined as the coefficient of variation, which is expressed as Equations (10)–(12). Here, the delay coefficient of variation and phase saturation coefficient of variation for each group of phase vehicles were calculated separately.

$$CV = \frac{\sigma}{\mu} \tag{10}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{11}$$

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{12}$$

The results indicate that unfairness is indeed widespread. The coefficient of variation for phase delay among the 362 intersection groups is greater than zero, with a mean of 0.37. Moreover, there is a roughly proportional relationship between the coefficient of variation for phase saturation and the coefficient of variation for phase-average delay. In other words, the greater the difference in phase saturation, the worse the fairness of phase-average delay. The scatter plot below illustrates the relationship between the coefficient of variation for phase saturation and the coefficient of variation for delay, as shown in Figure 3:

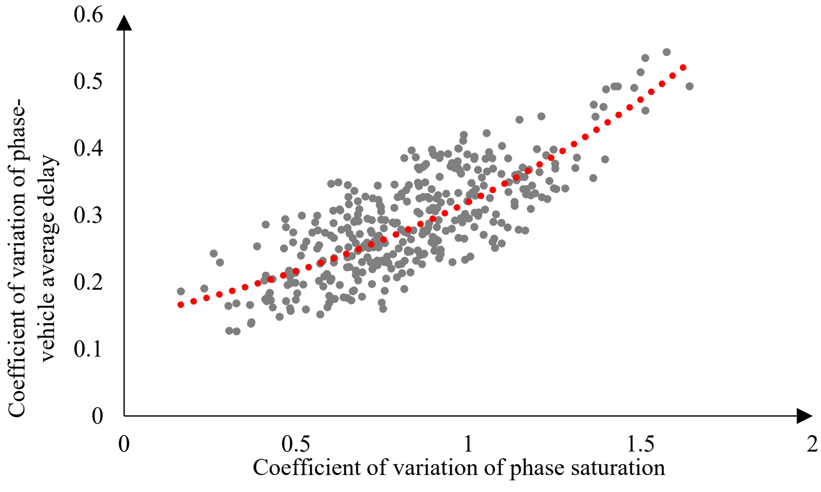

Figure 3. Phase delay coefficient of variation scatter.

(2) The less saturated, the less fair

By dividing the 362 sets of data into three saturation intervals, it is observed that the lower the intersection saturation, the greater the coefficient of variation of phase-average delay. The average coefficient of variation for saturation in the [0.0–0.2] range is 0.56, which is 1.87 times that of the other three saturation groups. This indicates a more severe unfairness phenomenon. Figure 4 below presents the curves of different saturation coefficients of variation.

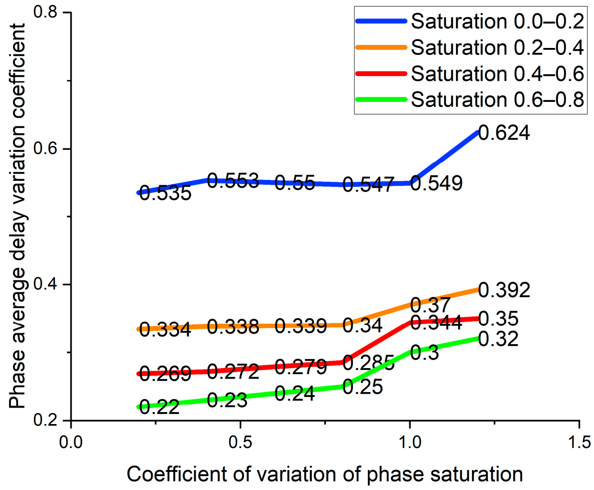

Figure 4. Phase-average delay coefficient of four saturation rates.

Combining the above analysis, it can be inferred that under low saturation conditions, with lower traffic volume and poorer fairness, there is a larger space for adjusting traffic efficiency and a greater necessity for improving the fairness of phase-average delay.

3.2. Delay Model Determination

Currently, the models used to calculate delay at signalized intersections mainly include the Webster delay model [3], the ARRB model [4], the HCM1985 model [5], and the HCM2000 model [6]. In estimating delay at low-saturation signalized intersections, these models are generally consistent because the latter three are derived from the Webster delay model and are more widely applicable. However, as saturation increases, the trend of consistency gradually weakens. When saturation is below 0.8, the relative error percentage of the average delay per vehicle obtained from the Webster model compared to simulation models falls within the range of 0 to 30% for over 95% of cases, significantly better than the other three delay models. Since this study focuses on low-saturation conditions during off-peak periods, the Webster model is chosen as the delay calculation model. The phase-average delay in the Webster model consists of three parts: uniform delay, random delay, and delay correction, as shown in Equation (13):

$$d_i = \frac{C(1-\lambda_i)^2}{2(1-\lambda_i x_i)} + \frac{{x_i}^2}{2q_i(1-x_i)} - 0.65\left(\frac{C}{{q_i}^2}\right)^{\frac{1}{3}} {x_i}^{(2+5\lambda_i)} \quad (13)$$

In the equation, d_i represents the phase i average delay; λ_i represents the phase i green time ratio; C represents the signal cycle length; and q_i represents the flow of the phase i. In the equation of $\lambda_i = \frac{x_i}{x}$, x_i represents the phase i saturation and x represents the total intersection saturation.

Since the last two terms in Equation (13) are much smaller compared to the first term, they can usually be ignored in analysis. Therefore, Equation (13) can be simplified to Equation (14):

$$d_i = \frac{C(1-\lambda_i)^2}{2(1-\lambda_i x_i)} \quad (14)$$

Then substituting $\lambda_i = \frac{x_i}{x}$ into the above equation gives, as expressed in Equation (15):

$$d_i = \frac{C(1-\frac{x_i}{x})^2}{2(1-\frac{{x_i}^2}{x})} \quad (15)$$

3.3. Delay Model Fairness Evaluation

Cross-intersection delay fairness mainly manifests in the consistency of average delay across phases. Descriptive parameters typically include variance, standard deviation, and, to eliminate the influence of scales and dimensions, the coefficient of variation. However, the ranges of these parameters are uncertain, posing difficulties in effectively integrating multiple objectives. Information entropy can also describe data consistency, and for specific problems, its range of values is fixed. Therefore, this study selects information entropy as the descriptive parameter for cross-intersection delay fairness.

3.3.1. Information Entropy

In 1948, Shannon introduced the concept of "information entropy," addressing the quantification issue of information. For an uncertain system Y, if its source symbols have n possible values with corresponding probabilities $P_1, \ldots, P_i, \ldots, P_n$, and each occurrence of values is independent of others, then the average uncertainty of the source should be the statistical mean (E) of the individual symbol uncertainties ($-\log P_i$), known as information entropy and denoted as Equation (16):

$$H(X) = E(-\log P_i) = -\sum_{i=1}^{n} P_i \log P_i \tag{16}$$

In Equation (16), the base of the logarithm is not specified and is typically chosen as 2, e, or 10. Different bases represent information units differently. In this paper, the base e is selected.

Information entropy can be used to evaluate the equilibrium of a system. The closer the individuals are to each other, the less significant the differences and the larger the information entropy, indicating a more balanced system. Conversely, smaller entropy values imply greater system uncertainty and higher information content. When $\exists P_i = 1$, entropy is minimal; when and only when $P_i = \frac{1}{n}$, entropy is maximal. For a four-phase delay problem, this is calculated as approximately $-\sum_{i=1}^{4} \frac{1}{4} ln \frac{1}{4} \approx 1.386$.

3.3.2. Fairness Evaluation Index

The fairness index for the average vehicle delay per phase refers to the concept where, in a delay system with different phases, the number of information sources equals the number of phases. The probability associated with each phase's delay is the proportion of that phase's delay to the total delay across all phases; the larger the proportion value, the greater the fairness it represents. Therefore, we need the number of information sources and the probability associated with each source.

Therefore, we use the theory of information entropy mentioned above to measure the fairness index of phase vehicle average delay in the system, as represented by Equation (17):

$$H(k) = -\sum_{i=1}^{m} k_i ln k_i \tag{17}$$

In Equation (17), k_i represents the ratio of the average delay of each phase to the sum of the average delays of all phases, $k_i = \frac{d_i}{\sum_{i=1}^{m} d_i}$; and m represents the number of phases. When and only when $k_1 = k_2 = \cdots = k_m$, the evaluation index is the maximum of $-\sum_{i=1}^{m} \frac{1}{m} ln \frac{1}{m}$.

This model is only suitable for situations with small non-peak delays, sacrificing total delays for fairness, so no relevant research has been conducted on long delays.

4. Efficiency and Fairness Signal Optimization Model

4.1. Feasibility Analysis

Considering equity will have an impact on delays and will require sacrificing delays for fairness, but is the sacrifice worth it? This section discusses this issue by analyzing the delay-to-fairness conversion rate.

The delay-to-fairness conversion rate represents the ratio between the gain in fairness and the delay sacrifice. Here, the fairness gain is measured by the proportion of fairness improvement, while the delay sacrifice is quantified by the proportion of the increase in delay. This is specified in Equation (18) below:

$$I_{cr} = \frac{(H - H')/H'}{(D - D')/D'} \tag{18}$$

In Equation (18), D' and H' denote the average intersection vehicle delays calculated using the Webster model and the corresponding fairness evaluation indexes; D and H denote the average intersection vehicle delays and fairness indexes, respectively.

4.1.1. The Delay-to-Fairness Conversion Rate

The delay-to-fairness conversion rate refers to the proportion of fairness improvement in the average delay of vehicles in each phase of the intersection that sacrifices the total delay of the intersection. When this conversion rate is maximum, it can maximize the fairness improvement obtained by sacrificing the total delay of the intersection.

In the context discussed in this paper, it can be considered that sacrificing delay for equivalent or greater benefits is worthwhile or feasible; otherwise, it is not feasible. The paper proposes using the delay-to-fairness conversion rate to measure the feasibility of sacrificing delay for fairness. The delay-to-fairness conversion rate represents the ratio between the benefits of fairness and the sacrifice of delay. If this ratio is greater than 1, it indicates feasibility; if less than 1, it indicates infeasibility. The fairness improvement ratio is used to represent the benefits of fairness, while the delay increase ratio represents the sacrifice of delay. The baseline fairness and delay are respectively referenced to the fairness evaluation index and total delay obtained from the signal timing scheme calculated based on the Webster model.

4.1.2. Feasibility Discussion

In order to realize the feasibility discussion, this section first gives the conversion rate calculation method and provides an in-depth discussion of the results, confirming the feasibility of livestock delays in exchange for fairness.

(1) Calculation method

The first step involves calculating the average delay per vehicle D' and the corresponding fairness evaluation index H' based on the intersection data using the Webster model.

In the second step, building upon the classical signal timing optimization model, the objective function is replaced with the fairness evaluation index H. Additionally, to further explore delay sacrifice, a constraint on delay sacrifice is added to the model, as shown in Constraint (19):

The feasibility condition for optimizing the intersection signal timing scheme of the model is to increase the fairness of the intersection to varying degrees. It is necessary to appropriately relax the total delay value of the intersection output by the model to represent the applicable conditions of the comprehensive model.

$$(w + 5\%) \cdot D' \geq D \geq w \cdot D' \tag{19}$$

The third step involves using the results from the second step to calculate fairness indexes and total delay for each set of intersection data; w represents the degree of relaxation of the total delay value at the intersection, and its value range is $(100\%, 105\%, 110\%, 115\%, 120\%, 125\%)$. With total delay sacrifice levels ranging from [100%,105%], [105%, 110%], [110%, 115%], [115%, 120%], [120%, 125%] and [125%, 130%]. Based on the results obtained in the first step, the conversion rates are then calculated accordingly.

(2) Results

After conducting the calculations six times for the 362 sets of data, the results indicate that the overall conversion rate is not favorable, with a mean value of only 0.69. As the sacrifice level increases, the average conversion rate decreases at an average rate of 29%. However, at a delay sacrifice level of [100%, 105%], the average conversion rate is 1.78, exceeding 100%. For other sacrifice levels, the average conversion rate is only 0.48, with [125%, 130%] being as low as 0.28. This suggests that sacrificing delay for fairness is feasible when the delay sacrifice is small, indicating that this approach has minimal impact on delay, with an average impact of only 2.5%.

Regarding saturation, as saturation increases, the average conversion rate gradually decreases. The saturation range of [0, 0.2] has the highest average conversion rate of 1.25, which remains the highest across all sacrifice levels. At a sacrifice level of [100%, 105%], it even reaches 3.08. In contrast, the saturation range [0.6, 0.8] has an average conversion rate of only 0.46. This indicates that sacrificing delay for fairness is more effective in low-saturation scenarios. The figure below illustrates the conversion rate curves for different saturation and sacrifice levels.

In conclusion, sacrificing delay for fairness is feasible and has minimal impact on delay, making it more suitable for low-saturation intersections. The purpose is to demonstrate the fair conversion rate achieved through sacrificing delay. Initially, when the total delay time sacrificed is relatively short, the conversion rate is high, indicating a significant fairness effect. However, as the total delay time sacrificed increases, the fairness effect obtained becomes less apparent, as shown in Figure 5.

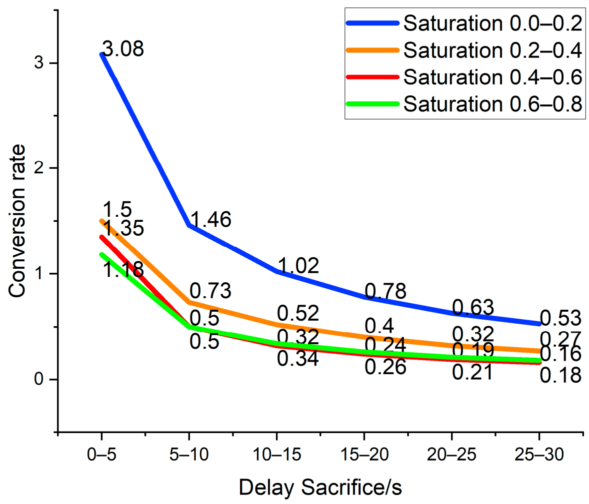

Figure 5. Four-saturation delay sacrifice conversion rates.

4.2. Optimization Model Construction

Considering the idea of feasibility analysis in Section 4.1, in order to realize the organic combination of fairness and efficiency, this paper proposes to take the conversion rate as the objective function and at the same time take the degree of delay sacrifice as the constraint.

4.2.1. Objective Function

To simultaneously consider the efficiency and fairness of signalized intersections, it is necessary to organically combine the Webster intersection total delay model with the phase-average delay fairness evaluation function to construct a dual objective comprehensive function. Based on the conclusion of the feasibility analysis of the comprehensive model in the previous section, a variable n is introduced into the comprehensive function, where n takes any positive integer, to adjust the weight ratio of total delay and fairness at the intersection in the comprehensive function. The final decision on the comprehensive examination model is as shown in Equation (20):

$$\max \frac{H(k)}{\sqrt[n]{D}} \tag{20}$$

4.2.2. Optimization Model

In this paper, based on the classical signal timing optimization model, the objective function is replaced with the one proposed in the previous section, and the constraints on the conversion rate and the degree of delay sacrifice are added. Thus, the signal optimization model considering fairness is obtained with the following Equation (20) as the objective function and Equation (21) as the constraints, where the first and second in Equation (21) are the phase green time length constraints; the third is the cycle length constraint; the fourth is the conversion rate constraint, which is required to be greater than or equal to 1; and the fifth is the delay sacrifice degree constraint.

$$s.t.\begin{cases} 5 < g_i < (C-L) \\ g_i > (C-L)(q_i/s_i) \\ 15m \leq C \leq 220 \\ \frac{(H-H')/H'}{(D-D')/D'} \geq 1 \\ 105\% \cdot D' \geq D \geq D' \\ i = 1, 2 \cdots m \end{cases} \tag{21}$$

In Equation (20), $D = \frac{\sum_{i=1}^{m} q_i \cdot d_i}{\sum_{i=1}^{m} q_i}$, $H = -\sum_{i=1}^{m} k_i lnk_i$, and $k_i = \frac{d_i}{\sum_{i=1}^{m} d_i}$. In Equation (21), g_i represents the phase i green time; $g_i = (C-L) \cdot \lambda_i$; s_i represents the phase i saturation flow; L represents lost time; and w represents delay sacrifice, and its value is greater than 1.

5. Model Sample Validation

To validate the model, this section first adopts the traditional Webster model as the efficiency model and constructs the fairness model with the fairness evaluation index as the target number. This section conducts a comprehensive comparative analysis of the proposed efficiency–fairness trade-off model, fairness model, and efficiency model from two aspects: the effectiveness of the model and the sensitivity of parameters. Metrics such as delay, fairness evaluation index, cycle change rate, and green ratio change rate are utilized for the comparison and analysis.

5.1. Fairness Model

To further validate the model, a signal optimization model focusing solely on fairness was constructed based on the model from the previous section. Specifically, the objective function was replaced with the fairness evaluation index H, and the fourth conversion rate constraint was removed. Equation (23) represents the constraint.

$$\max \frac{H(k)}{\sqrt[n]{D}} \tag{22}$$

$$s.t.\begin{cases} 5 < g_i < (C-L) \\ g_i > (C-L)(q_i/s_i) \\ 15m \leq C \leq 220 \\ 105\% \cdot D' \geq D \geq D' \\ i = 1, 2 \cdots m \end{cases} \tag{23}$$

5.2. Validity and Sensitivity Analysis

This section will analyze the validity of the model in terms of both model effectiveness and impact on the efficiency model, as well as perform a sensitivity analysis for intersection saturation.

5.2.1. Comparative Analyses of Validity

(1) Fairness

In terms of the model performance, overall, the fairness model performed the best, followed by the efficiency–fairness model, and the efficiency model performed the worst. Furthermore, the difference between the efficiency–fairness model and the fairness model was significantly smaller than that between the efficiency–fairness model and the efficiency model, with mean fairness evaluation index values of 1.37, 1.32, and 1.21, respectively, as shown in Figure 6.

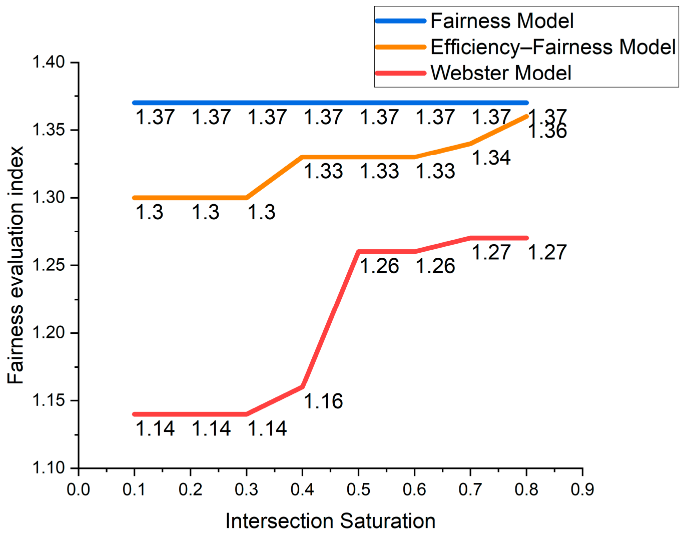

Figure 6. Three-model fairness evaluation index.

Regarding saturation, under low saturation conditions, the efficiency–fairness model showed better performance. The mean improvement in the fairness evaluation index for saturation levels between 0.1 and 0.4 was 0.16, while for saturation levels between 0.5 and 0.8, it was only 0.075. In terms of trend, higher saturation levels correlated with higher fairness evaluation index values. At saturation levels of 0.1 and 0.8, the mean fairness evaluation index values were 1.27 and 1.33, respectively. Furthermore, the higher the saturation level, the smaller the differences between the models. At saturation levels of 0.1 and 0.8, the mean differences between the models were 0.115 and 0.03, respectively, especially notable between the efficiency–fairness model and the efficiency model, with differences of 0.16 and 0.09 at saturation levels of 0.1 and 0.8, respectively. The following figure illustrates the fairness evaluation index curves for each model.

(2) Efficiency

In terms of the models, overall, the efficiency model performed the best, followed by the efficiency–fairness model, and the fairness model performed the worst. Furthermore, the difference between the efficiency–fairness model and the efficiency model was significantly smaller than the difference between the efficiency–fairness model and the fairness model. The mean vehicle delay values were 13.72, 14.08, and 35.12, respectively, which are shown in Figure 7.

Regarding saturation, the saturation level had little impact on the efficiency–fairness model. As for the trend, higher saturation levels correlated with greater vehicle delays. At saturation levels of 0.1 and 0.8, the mean vehicle delays were 18.4 and 24.98, respectively.

Additionally, as saturation levels increased, the differences between the models did not change significantly, with a mean difference of 10.7. At saturation levels of 0.1 and 0.8, the differences were 10.16 and 11.12, respectively. The following figure illustrates the delay curves for each model.

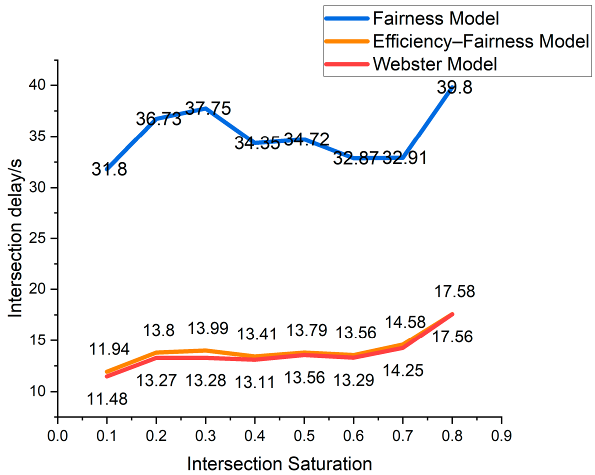

Figure 7. Three-model intersection delay.

(3) Conversion rate

In terms of the models, overall, the efficiency–fairness model significantly outperformed the fairness model, with mean conversion rates of 0.09 and 9.6, respectively, a difference exceeding 100 times. Furthermore, the efficiency–fairness model ranged from a minimum of 2.37 to a maximum of 0.14 for the fairness model, representing a difference of over 16 times, which is shown in Figure 8.

Regarding saturation, under low saturation conditions, the efficiency–fairness model exhibited better performance. Excluding the outlier at saturation level 0.8, the mean conversion rates for saturation levels 0.1 to 0.4 were 4.0, greater than the rates for saturation levels 0.5 to 0.7, which were 2.76. With changes in saturation levels, the two models showed different trends. While the fairness model exhibited a decreasing conversion rate with increasing saturation levels, the efficiency–fairness model showed fluctuations in its conversion rate with changes in saturation levels without a clear trend. The figure below illustrates the conversion rate curves for both models.

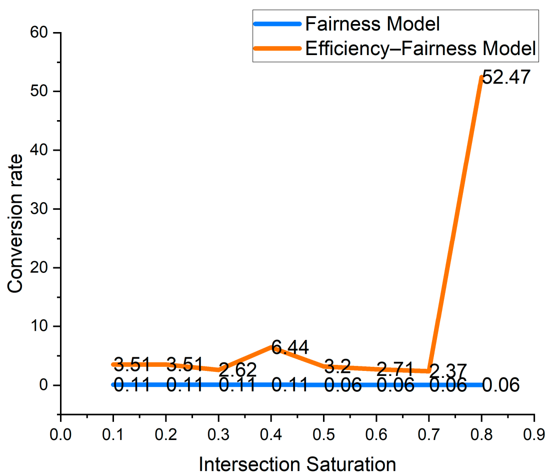

Figure 8. Two-model conversion rate.

5.2.2. Comparative Analyses of Fluctuations

This section analyzes the fluctuations of the two models relative to the efficiency model based on changes in cycle length and green time ratio.

(1) Cycle Length

Overall, the cycle lengths mostly increased. Concerning the models, the efficiency–fairness model significantly outperformed the fairness model, with mean change ratios of 0.013 and 0.88, respectively, representing a difference exceeding 67 times. Furthermore, the efficiency–fairness model ranged from a maximum of 0.04 to a minimum of 0.62 for the fairness model, a difference exceeding 15 times, which is shown in Figure 9.

Regarding saturation levels, under low saturation conditions, the efficiency–fairness model exhibited better performance. The mean change ratio for saturation levels 0.1 to 0.4 was 0.0035, whereas for saturation levels 0.5 to 0.8, it was as high as 0.026. In terms of trends, with changes in saturation levels, the two models showed different trends. While the fairness model initially exhibited an increasing and then decreasing trend in change ratio and increasing saturation levels with significant fluctuations, the efficiency–fairness model showed fluctuations without a clear trend with changes in saturation levels. The figure below illustrates the change ratio curves for cycle lengths for both models.

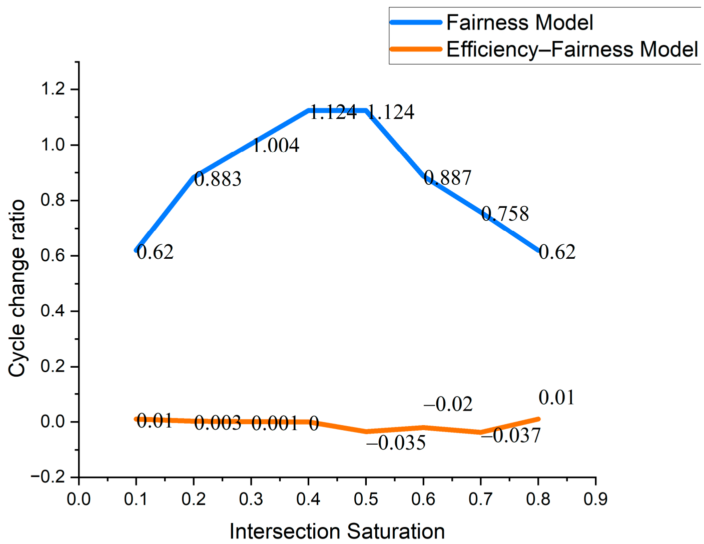

Figure 9. Two–model cycle change ratio.

(2) Green time ratio

Additionally, to visually represent the changes in phase green time ratio for these two models relative to the efficiency model, this section selected the two phases with the lowest and highest average delay in the efficiency model for comparison.

Overall, for phases with low green time ratios, the ratios further decreased, while for phases with high green time ratios, the ratios increased. Concerning the models, the efficiency–fairness model generally outperformed the fairness model. The mean change ratios for phases with low green time ratios were −0.48 and −0.51 for the efficiency–fairness and fairness models, respectively, while for phases with high green time ratios, the mean change ratios were 0.57 and 1.27, respectively. Only for phases with low green time ratios and saturation levels of 0.7 and 0.8 did the efficiency–fairness model slightly outperform the fairness model.

Regarding saturation levels, although the efficiency–fairness model exhibited slightly higher change ratios under low saturation conditions, its fluctuation was significantly smaller than under high saturation conditions. For saturation levels of 0.1 to 0.4, the mean change ratios for both low and high green time ratios were 0.47 and 0.7, respectively, with standard deviations of 0.01 and 0.08. For saturation levels of 0.5 to 0.8, the mean change ratios were 0.484 and 0.446, respectively, with standard deviations of 0.19 and 0.43. In terms of trends, the change rate for phases with low saturation levels increased with saturation levels, while for phases with high saturation levels, the change rate decreased with saturation levels. The figures below illustrate the change ratio curves for green time ratios for both models, which are shown in Figures 10 and 11.

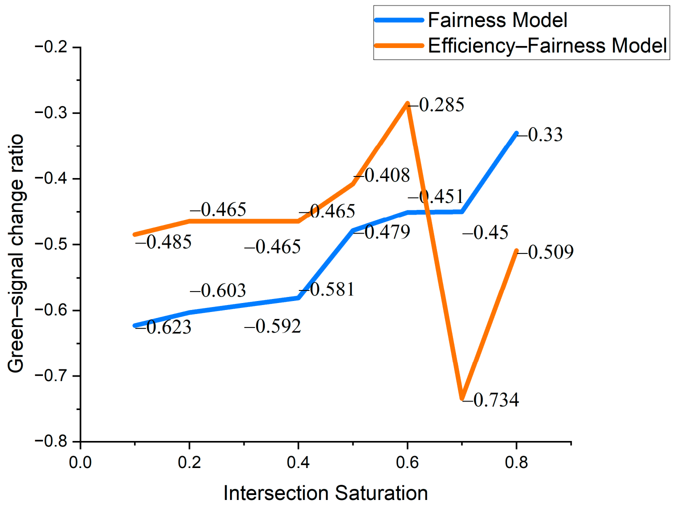

Figure 10. Two–model green-signal change ratio with the lowest average delay.

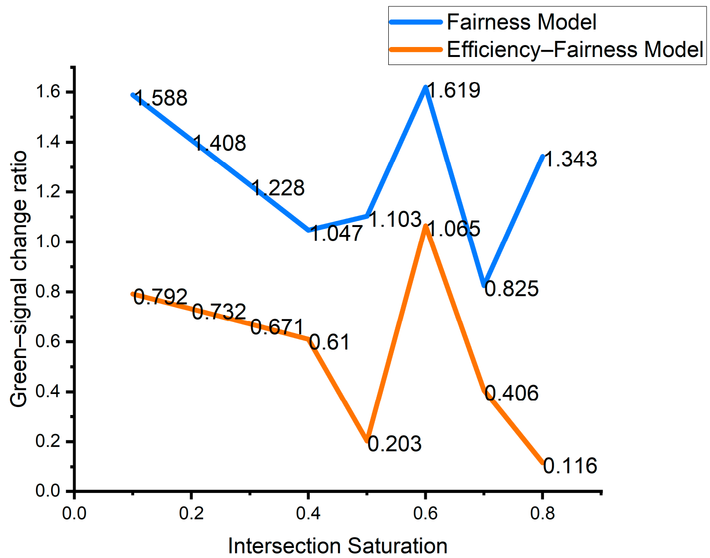

Figure 11. Two-model green-signal change ratio with the highest average delay.

In summary, the efficiency and fairness model proposed in this paper not only balances efficiency and fairness simultaneously but also has minimal impact on efficiency. Furthermore, the changes to the timing schemes in the efficiency model are much smaller compared to the fairness model. Therefore, it can be concluded that the proposed model in this paper is valid and effective. Sensitivity analysis indicates that the efficiency and fairness model is more effective in low saturation conditions.

6. Conclusions

To further advance fairness-related research in the transportation field, this paper addresses the issue of fairness in the average delay per vehicle at low-saturation intersection phases. It proposes a strategy of sacrificing efficiency for fairness. Initially, it constructs a fairness evaluation metric for intersection phase delay using information entropy. Then, it validates the feasibility of this approach based on simulated data. Subsequently, it introduces the concept and calculation formula for the efficiency–fairness conversion rate and uses it to develop a signal optimization model that balances efficiency and fairness. Finally, the proposed model is validated using simulated data, showing that it not only achieves a balance between efficiency and fairness but also has minimal impact on efficiency compared to fairness-oriented models. Sensitivity analysis reveals that the model is only applicable to situations where intersection density and overall delay are not high. The model has a high conversion rate of fairness with minimal delay sacrifice and can improve traffic fairness with a minimal increase in overall delay.

The research scope of this paper is limited to low-saturation and single-point intersections. As a result, the signal timing optimization model lacks universality and applicability. In future research, it will be extended to high-saturation arterials and regional network studies to enable broader research and application of the fairness of intersection emissions and average vehicle delay per phase. Additionally, in the effectiveness analysis, the influence of cycle length and green split ratio on the output of evaluation indicators is analyzed. In future research, not only should comparisons and analyses be made on the evaluation output indicator results, but also more attention should be paid to the influence of changing factors on the output of evaluation indicators to identify the causes of results and the internal connections between influencing factors from the source.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app14178047/s1.

Author Contributions: Conceptualization, L.Z. (Liang Zou) and L.Y.; methodology, L.Y.; software, L.Y. and L.Z. (Lingxiang Zhu); validation, L.Y.; formal analysis, L.Z. (Lingxiang Zhu); writing—original draft preparation, L.Y.; writing—review and editing, L.Z. (Liang Zou). All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Shenzhen Science and Technology Plan Project (No. KJZD20230923115223047) and the Shenzhen Higher Education Stable Support Plan Project (No. 20231123103157001).

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Xi, J. *Secure a Decisive Victory in Building a Moderately Prosperous Society in All Respects and Strive for the Great Success of Socialism with Chinese Characteristics for a New Era: Report to the 19th National Congress of the Communist Party of China*; People's Publishing House: Beijing, China, 2017; pp. 11–12.
2. Zhang, Z. Research on Equity Measurement of Transportation Development in China from the Perspective of Equalization of Public Services. Master's Thesis, Beijing Jiaotong University, Beijing, China, 2019.
3. Webster, F.V. *Traffic Signal Settings*; Road Research Technical Paper; TRID: Washington, DC, USA, 1958; Volume 39.
4. Akcelik, R. *Traffic Signals: Capacity and Timing Analysis*; TRID: Washington, DC, USA, 1981.
5. Wisconsin Department of Natural Resources Council. *Transportation Research Board (TRB). Highway Capacity Manual, 2000*; Wisconsin Department of Natural Resources Council: Madison, WI, USA, 2000.
6. Liu, C.; Ren, Y.Y.; Bi, X.J. Optimization of regional traffic signal timing based on improved firefly algorithm. *Control Decis.* **2020**, *35*, 2829–2834. [CrossRef]
7. Rahmi, A. *Guide to Traffic Engineering Practice: Part 7: Traffic Signals*; Austroads Ltd.: Sydney, Australia, 2003.
8. Wang, P.; Yang, Q. Genetic Algorithms Based Traffic Signal Optimization at a Congested Intersection. *Appl. Mech. Mater.* **2012**, *209–211*, 814–817. [CrossRef]
9. Kesur, K.B. *Multiobjective Optimization of Delay and Stops in Traffic Signal Networks*; Elsevier: Amsterdam, The Netherlands, 2013.
10. Ceder, A.; Reshetnik, I. An algorithm to minimize queues at signalized intersections. *J. Oper. Res. Soc.* **2001**, *52*, 615–622. [CrossRef]
11. Liu, H.; Balke, K.N.; Lin, W.H. A reverse causal-effect modeling approach for signal control of an oversaturated intersection. *Transp. Res. Part C Emerg. Technol.* **2008**, *16*, 742–754. [CrossRef]
12. Alhajyaseen, W.K.M.; Najjar, M.; Ratrout, N.T.; Assi, K. The effectiveness of applying dynamic lane assignment at all approaches of signalized intersection. *Case Stud. Transp. Policy* **2017**, *5*, 65–72. [CrossRef]

13. Murat, Y.S.; Gedizlioglu, E. A fuzzy logic multi-phased signal control model for isolated junctions. *Transp. Res. Part C* **2005**, *13*, 19–36. [CrossRef]
14. Schmoecker, J.D.; Ahuja, S.; Bell, M. Multi-objective signal control of urban junctions—Framework and a London case study. *Transp. Res. Part C Emerg. Technol.* **2008**, *16*, 454–470. [CrossRef]
15. Chen, J.; Jia, M. Robust Multi-objective Optimization and Decision for Signal Timing at Single Intersection. In Proceedings of the 2016 International Conference on Automatic Control and Information Engineering, Hong Kong, China, 22–23 October 2016.
16. Deng, W.; Jia, L.M.; Dong, H.H. Multi-Objective Optimization Model for Urban Traffic Intersection Control Based on Data Fusion. *Adv. Mater. Res.* **2011**, *156–157*, 505–510. [CrossRef]
17. Wei, J.; Ju, Y. Research on optimization method for traffic signal control at intersections in smart cities based on adaptive artificial fish swarm algorithm. *Heliyon* **2024**, *10*, e30657. [CrossRef]
18. Himabindu, M.; Lelitha, V.; Rama, B.C. A probe-based demand responsive signal control for isolated intersections under mixed traffic conditions. *Transp. Lett.* **2024**, *16*, 117–130.
19. He, R.; Zhao, H.; Ye, Z.; Yang, L. Time optimization of urban arterial coordinated control based on multidimensional neighbourhood artificial bee colony algorithm. *Transp. Plan. Technol.* **2024**, *47*, 113–145. [CrossRef]
20. Gao, Y.; Xu, L.; Hu, H.; Han, H.; Yang, X. Multi-objective optimization method for fixed-cycle signal control at intersections. *China, J. Highw. Transp.* **2011**, *24*, 82–88.
21. Liu, Q.; Lu, B.; Ma, Q. Research on multi-objective optimization of signal control at signalized intersections. *J. Traffic Sci. Eng.* **2014**, *16*, 47–50.
22. Wang, Y.; Jiang, X.; Guo, W. Research on optimization method of signalized intersection cycle length considering travel cost. *J. Chang. Univ. (Nat. Sci. Ed.)* **2015**, *35*, 104–107.
23. Li, X.; Liu, Y.; Zhou, J.; Liu, W.; Hong, L. Traffic signal timing optimization model based on improved genetic algorithm. *Ind. Instrum. Autom.* **2017**, *5*, 125–130.
24. Li, J.; Jia, T. Signal timing optimization at signalized intersections based on per capita delay. *J. Highw. Traffic Sci. Technol.* **2021**, *38*, 134–141.
25. Cai, S. Traffic network design model and application based on traffic fairness. Master's Thesis, Huazhong University of Science and Technology, Wuhan, China, 2016.
26. Zhu, Z. Research on Multi-Mode Traffic Congestion Charging Methods Considering Fairness. Ph.D. Thesis, Southeast University, Nanjing, China, 2016.
27. Litman, T. Evaluating transportation equity—methods for incorporating distributional impacts into transport planning. *Costs* **2005**, *8*, 1–30.
28. Lu, D.; Zhang, S.; Guo, M. Analysis and Countermeasures of Urban Traffic Fairness. *J. Transp. Sci. Eng.* **2008**, *32*, 103–105.
29. Zhi, C.L.; Xiao, Y.G. Traffic signal timing problems with environmental and equity considerations. *J. Adv. Transp.* **2014**, *48*, 1066–1086.
30. Baskan, O.; Ozan, C. Reserve Capacity Model for Optimizing Traffic Signal Timings with an Equity Constraint. In *Highway Engineering*; Intechopen: Rijeka, Croatia, 2017.
31. Zheng, L.; Xue, X.; Xu, C.; Ran, B. A stochastic simulation-based optimization method for equitable and efficient network-wide signal timing under uncertainties. *Transp. Res. Part B* **2019**, *122*, 287–308. [CrossRef]
32. Shi, J.; Yang, L.; Ying, X.; Huang, Q. Quantitative Evaluation Model of Traffic Fairness Based on Wilson Entropy Distribution Assumption. *J. Wuhan Univ. Technol. (Transp. Sci. Eng.)* **2008**, *32*, 1–4.
33. Lv, B.; Niu, H. Optimization Method of Phase Difference in Urban Traffic Line Control System. *J. Transp. Eng.* **2011**, *11*, 103–107.