



**HAL**  
open science

# Un protocole de comparaison stochastique des systèmes de gestion du trafic routier

Jules Bompard, Philippe Mathieu, Antoine Nongaillard

## ► To cite this version:

Jules Bompard, Philippe Mathieu, Antoine Nongaillard. Un protocole de comparaison stochastique des systèmes de gestion du trafic routier. Trente-deuxièmes journées francophones sur les systèmes multi-agents (JFSMA), Nov 2024, Cargese (Corse), France. pp.35-44. hal-04777205

**HAL Id: hal-04777205**

**<https://hal.science/hal-04777205v1>**

Submitted on 12 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Un protocole de comparaison stochastique des systèmes de gestion du trafic routier

Jules Bompard<sup>a</sup>  
jules.bompard.etu@univ-lille.fr

Philippe Mathieu<sup>a</sup>  
philippe.mathieu@univ-lille.fr

Antoine Nongillard<sup>a</sup>  
antoine.nongillard@univ-lille.fr

<sup>a</sup>Univ. Lille, CNRS, Centrale Lille, UMR 9189, CRIStAL, F-59000 Lille, France

## Résumé

Depuis quelques années, les zones urbaines sont soumises à des trafics routiers importants, créant des congestions qui impactent négativement l'économie et les émissions de gaz à effet de serre. Pour pallier ce problème, des systèmes de gestion du trafic (SGT) ont été développés. Les évaluations existantes de ces systèmes souffrent de plusieurs biais, réduisant la crédibilité des résultats. Dans cet article, nous développons un protocole stochastique d'évaluation des SGT. Le protocole consiste à simuler un grand nombre de fois chaque système sur un unique réseau routier, en faisant varier aléatoirement les flux de véhicules et les paramètres des systèmes. À partir des temps de trajet moyens et des émissions de CO<sub>2</sub> par trajet, nous développons deux nouvelles métriques d'analyse du potentiel et de la difficulté de paramétrage des systèmes. Avec un exemple d'utilisation de ce protocole, nous démontrons que l'utilisation de simulations stochastiques permet d'obtenir des résultats objectifs, contrairement aux protocoles existants.

**Mots-clés :** Evaluation, Stratégies, Simulation, Trafic routier

## Abstract

In ever-expanding cities, road traffic is intensifying and experiencing slowdowns due to congestion, with negative economic and ecological effects. The development of traffic management systems (TMS) aims to reduce the frequency of congestion, but the biases in their evaluations make it impossible to compare them effectively. In this paper, we present a stochastic protocol for evaluating TMS. We proceed by studying a large number of randomly generated simulations, i.e. where the TMS and flow settings are randomly selected from a set of possible configurations, whose operation prevents existing interpretation biases. We are developing two new evaluation metrics, based on average travel time and average CO<sub>2</sub> emissions per travel, which allow to rank TMS according to their performance. With an application example of this protocol, we de-

monstrate that the use of stochastic simulations makes it possible to obtain objective results, unlike existing protocols.

**Keywords:** Evaluation, Strategies, Simulation, Road traffic

## 1 Introduction

L'expansion et la densification des zones urbaines ont entraîné une croissance exponentielle du secteur des transports, principalement du transport routier, dans les villes. Des réseaux routiers qui fonctionnent bien sont associés à de meilleurs résultats économiques, mais les embouteillages peuvent les rendre contre-productifs. Économiquement, les embouteillages entraînent des pertes financières importantes, comme dans l'Union européenne, où les coûts sont estimés à 1% du PIB chaque année [18]. La congestion du trafic routier entraîne également une augmentation des émissions de gaz à effet de serre [4], ainsi que des gaz toxiques pour l'homme, qui augmentent le risque de maladies respiratoires et cardiovasculaires, ainsi que les problèmes liés à la grossesse et à l'accouchement [17]. L'optimisation du trafic routier est donc devenue un sujet majeur ces dernières années, abordé de nombreuses fois au cours des dernières éditions de cette conférence [27, 20, 13].

Pour éviter de construire de nouvelles infrastructures, aux coûts économiques et écologiques importants, des systèmes de gestion du trafic routier (SGTR) ont été créés dans le but d'optimiser le fonctionnement des infrastructures existantes. Parmi eux, les systèmes de gestion des feux tricolores (SGFT), dont l'objectif est de fluidifier l'écoulement du trafic au niveau des intersections. Dans un réseau routier, chaque carrefour est un agent qui gère, avec son propre comportement, un ensemble de feu de manière autonome. Les SGFT sont donc des systèmes multi-agents dont la fluidité est la propriété émergente [5, 24, 25, 8].

Si beaucoup de ces systèmes donnent des ré-

sultats intéressants pour une topologie de réseau routier et des flux de trafic spécifiques, il est difficile de généraliser leurs performances à tous les environnements possibles. La plupart du temps, les SGTR sont évalués sur la base d'une dizaine de simulations maximum par réseau routier, dont les flux varient. Ces conditions expérimentales ne permettent pas de généraliser les résultats à quelque échelle que ce soit, et conduisent parfois à des résultats contradictoires, comme [30] montrant que le système *Max Pressure* [24] est meilleur que le système *MPLight*, alors que [8] montre le contraire. De plus, les paramètres des différents systèmes étudiés sont rarement transmis, ce qui rend impossible la reproduction des simulations. Afin de généraliser les résultats des protocoles d'évaluation, il est nécessaire de développer un protocole qui prenne en compte ces spécificités.

Dans cet article, nous présentons un protocole stochastique pour l'évaluation de SGFT. Ce protocole consiste à simuler un certain nombre de fois un système sur une typologie de réseau donnée avec des paramètres de flux et de système choisis au hasard, afin d'estimer leurs forces et faiblesses dans toutes les conditions. Les valeurs aléatoires sont choisies dans des fourchettes de valeurs basées sur les mesures du trafic réel, ce qui garantit que chaque simulation est représentative des conditions de trafic réelles. Grâce à cela, notre protocole possède deux propriétés qui le différencient des protocoles existants : 1) les résultats des expériences sont généralisables à toutes les conditions de trafic possibles pour un réseau donné, et 2) les paramètres de simulation peuvent être facilement partagés pour assurer la reproductibilité des expériences.

## 2 Travaux connexes

### 2.1 Systèmes de gestion de feux tricolores

Les systèmes de gestion des feux tricolores sont des systèmes qui gèrent le fonctionnement d'une ou plusieurs intersections, souvent avec un objectif d'optimisation de l'écoulement des flux de véhicules. Si quelques systèmes adoptent des approches centralisées [11], la plupart de ces systèmes adoptent une approche multi-agents, où chaque agent contrôle les feux tricolores d'une intersection. Grâce au développement de capteurs toujours plus performants permettant de capter l'environnement local, les agents d'un SGFT sont dotés d'un comportement réactif pour sélectionner une phase, c'est-à-dire une dis-

tribution des droits de passage à l'intersection pour chaque voie. Plusieurs paradigmes ont été étudiés dans le cadre des SGFT multi-agents : les systèmes à base de règles, les systèmes à base d'optimisation et les systèmes à base d'apprentissage par renforcement.

**Systèmes à base de règles.** Les systèmes basés sur des règles sont définis par un ensemble de règles de gestion, définissant la stratégie de l'agent pour la sélection d'une phase pour une intersection. Parmi eux, les systèmes *Max Pressure* [24], *Self Organized Traffic Lights* (SOTL) [10], et *Longest Queue First* (LQF) [26] sont les plus populaires. Ces systèmes sont faciles à tester. Comme ils ne nécessitent aucun entraînement, ces systèmes sont au maximum de leur capacité dès leur déploiement. Ainsi, toutes les métriques collectées après déploiement sont représentatives des performances réelles du système et peuvent être utilisées pour l'évaluer.

**Systèmes à base d'optimisation.** Les systèmes basés sur l'optimisation implémentent des agents dont la stratégie est défini par un algorithme d'optimisation, dont le résultat permet de sélectionner une phase pour l'intersection. Plusieurs paradigmes d'optimisation ont été étudiés, comme la représentation de l'environnement et l'utilisation de solveurs à logique floue [3, 9], l'utilisation d'algorithmes de recuit simulé [2] ou génétiques [21, 19]. Les systèmes basés sur l'optimisation sont aussi faciles à tester que les systèmes basés sur des règles, car ils ne nécessitent pas non plus d'entraînement avant d'être déployés.

**Systèmes à base d'apprentissage par renforcement.** Les systèmes basés sur l'apprentissage par renforcement implémentent des agents selon un modèle de Markov. Leur comportement est défini par un ensemble d'actions pour chaque état, dont la valeur est estimée par un réseau de neurones ou une matrice, et d'une fonction de sélection d'action [22, 25, 8, 30]. Contrairement aux systèmes basés sur des règles ou sur l'optimisation, ces systèmes nécessitent un entraînement avant d'être déployés, entraînement durant lequel leur comportement est sous optimal. Ainsi, l'évaluation d'un système dont les agents suivent un modèle d'apprentissage par renforcement ne devra être faite que lorsque la phase d'entraînement sera terminée, c'est-à-dire quand chaque agent a atteint sa performance maximale.

## 2.2 Protocoles de test

Pour tester un SGFT, il est d'usage de simuler un réseau routier où est implémenté le système. En effet, l'utilisation d'un simulateur présente les avantages de ne pas perturber le trafic réel et de pouvoir récupérer facilement de nombreuses métriques d'évaluation. Les protocoles d'évaluation existants sont majoritairement des protocoles de comparaison. L'objectif est de comparer, sur un ou plusieurs réseaux routiers, les performances de plusieurs SGFT, dans l'objectif de démontrer la supériorité de l'un d'eux. Cependant, nous avons identifié plusieurs problèmes de conception de ces protocoles qui nous amènent à penser que leurs résultats peuvent être biaisés ou partiels. Il n'est donc pas possible d'en généraliser les résultats, et les protocoles de comparaison ne tendent qu'à prouver l'efficacité d'un SGFT dans des conditions uniques. Dans cette section, nous allons énumérer l'ensemble des problèmes des protocoles de comparaison que nous avons identifiés et expliquer pourquoi ils biaisent les résultats.

**Réseaux idéaux.** Un réseau routier est une combinaison de carrefours, chacun d'eux ayant une forme particulière, définie par le nombre de voies pour chaque entrée. Parmi l'ensemble des combinaisons de carrefours possibles, certaines présentent plus de contraintes à l'écoulement du trafic que d'autres. Pour évaluer objectivement des SGFT, il convient donc d'utiliser un réseau présentant un maximum de ces contraintes, un système efficace avec pouvant également l'être sans, ce qui n'est pas le cas des protocoles existants. Certains articles évaluent les systèmes sur un réseau composé d'un seul carrefour [1, 26]. Or, la complexité du problème d'optimisation des feux de circulation vient du fait que deux carrefours voisins peuvent avoir des intérêts contradictoires. Ainsi, un carrefour optimisé peut bloquer complètement un carrefour voisin, ce qui réduirait les performances globales. Une topologie plus complète est donc nécessaire pour tester un SGFT. Des articles récents ont également évalué leur SGFT sur un réseau dont les routes ont une voie spéciale pour tourner à gauche [25, 8, 14]. Cette topologie est idéale, car un véhicule qui attend pour tourner à gauche ne peut pas ralentir ou arrêter les autres. Comme un grand nombre de réseaux routiers n'ont qu'une voie par approche, aucune généralisation ne peut être faite à partir d'un protocole de comparaison réalisé sur des intersections à deux voies ou plus.

**Trop de métriques.** Le grand nombre de métriques de comparaison des protocoles est également problématique, des études utilisant le temps de parcours moyen [28], le débit [8], le temps d'attente [10], le retard [25], et autres valeurs pour comparer les SGFT. La comparaison des résultats de deux articles en devient plus complexe. En effet, deux articles développant des protocoles évaluant les mêmes SGFT avec différentes métriques peuvent obtenir des résultats différents, accentuant l'incertitude concernant le choix du meilleur système. En outre, certaines études ne comparent leur SGFT qu'avec un seul autre système, comme [1, 3, 22, 10]. Cette méthode n'est pas suffisante pour valider l'efficacité d'un SGFT, surtout si le seul système comparé est celui à temps fixe, dont on connaît la faible efficacité.

**Pas assez de simulations.** Tous les articles cités dans notre étude développent des protocoles basés sur un petit nombre de simulations, c'est-à-dire moins de dix, et sur différents réseaux. Ce nombre soulève des questions quant à la généralisation des résultats, car seul un petit nombre de situations, aussi réalistes soient-elles, sont testées. Un système de transport routier doit être flexible, c'est-à-dire qu'il ne doit pas être efficace uniquement dans certaines conditions, mais dans toutes les conditions réalistes. Par exemple, des études récentes ont testé leur système sur les réseaux de New York, Hangzhou et Jinan, avec des flux mesurés sur le terrain, mais rien ne garantit que leurs résultats soient les mêmes dans d'autres situations [31, 30].

**Paramètres inconnus.** Enfin, le problème majeur des protocoles de comparaison est l'absence de transmission des paramètres de simulation. Ces paramètres sont composés en deux groupes : les paramètres de flux, qui sont les paramètres qui définissent la génération des flux sur le réseau routier, tandis que les paramètres des systèmes sont ceux qui définissent le comportement des agents pour chaque système. Les résultats d'un protocole de comparaison où des agents sont mal paramétrés sont trompeurs, car ils offriront des performances inférieures à leur potentiel réel. Le phénomène est d'autant plus important que le nombre de simulations est petit. Il est aussi important de connaître les paramètres de flux du test d'un SGFT, car ils peuvent grandement faire varier leur performance. Ainsi, la transmission des paramètres de simulation est essentielle pour mieux évaluer la qualité des résultats d'une série de tests. Alors que les paramètres du système développé dans un article sont

logiquement transmis, aucun des articles cités ici ne les fournit.

Aujourd'hui, un consensus a émergé autour du protocole développé par [31], qui est devenu un standard [8, 30, 12]. Les SGFT sont testés sur des réseaux et des flux réels de mégapoles, mesurant le temps de parcours moyen et le débit des véhicules. Cependant, ce protocole n'implique qu'un petit nombre de simulations, et les paramètres des autres systèmes ne sont jamais fournis. Cela conduit à des incohérences dans les résultats de certains articles, comme [15] qui propose un système qui réduit le temps de trajet de 95% par rapport à une stratégie à temps fixe, voire à des résultats contradictoires comme [8] et [30] qui classent différemment le système *Max Pressure* et le système *MPLight*.

## 3 Méthodologie

Déterminer le meilleur SGFT pour tout réseau nous semble être une tâche impossible. En effet, tous les réseaux routiers n'ont pas les mêmes contraintes en matière de circulation (trois voies par entrée, trafic à sens unique, etc.), et un système qui serait performant pour un réseau donné pourrait l'être moins s'il était soumis à des contraintes différentes. On pourrait penser que les SGFT devrait alors être testés sur un ensemble de configurations de carrefours représentatifs de l'ensemble des contraintes de fluidité du trafic, mais la complexité de calcul serait telle que le protocole ne serait pas réalisable. Le protocole que nous développons dans cet article vise, plutôt que de déterminer le meilleur SGFT global, à déterminer le meilleur pour un réseau donné.

### 3.1 Idée générale

Comme le nombre de configurations de simulations est une exponentielle du nombre de paramètres, il est impossible de toutes les tester. Cependant, nous avons vu qu'il est nécessaire de simuler les SGFT dans un grand nombre de situations de trafic pour avoir une idée de la façon dont leur performance évolue en fonction des conditions. Pour surmonter ce problème, nous utilisons un protocole d'évaluation stochastique : nous effectuons  $n$  simulations avec des paramètres choisis au hasard, pour chaque SGFT. Ce choix présente deux avantages : il limite le nombre de simulations nécessaires pour évaluer un système, et garantit que ses performances sont évaluées sur un large ensemble de situations, dé-

pendant de  $n$ . Logiquement, plus  $n$  est grand, plus les résultats sont fiables.

### 3.2 Sélection des paramètres

Si les paramètres d'une simulation sont choisis au hasard, il est nécessaire de définir des plages de valeurs sélectionnables pour chacun d'entre eux. C'est la partie la plus délicate du protocole, car il faut choisir une plage de valeurs suffisamment large pour tester une grande variété de configurations, mais suffisamment petite pour que les résultats empiriques des  $n$  simulations soient représentatifs. Les plages de valeurs ne peuvent être définies qu'à partir d'une connaissance préalable du fonctionnement d'un système, une large fourchette étant préférable dans le cas contraire. Dans tous les cas, cette plage de valeurs doit être partagée afin que le protocole puisse être reproduit et les résultats analysés. L'un des paramètres les plus importants concerne la génération de véhicules. Nous pensons que seuls des flux de véhicules réalistes devraient être simulés. En effet, si des scénarios de trafic irréalistes sont inclus, l'espace d'exploration est élargi, ce qui nécessiterait plus de simulations que nécessaire. S'ils sont connus, les flux de véhicules devraient suivre les mêmes proportions que ceux des routes réelles du réseau, l'exploration consistant à faire varier l'intensité de ces flux.

### 3.3 Adaptations pour l'apprentissage par renforcement

Nous avons vu qu'un SGFT à apprentissage par renforcement devait suivre une phase d'apprentissage avant d'être performant, durant laquelle il ne devait pas être évalué. Dans ce protocole, ces systèmes sont d'abord entraînés sur le même réseau que lors de l'évaluation, pendant une durée fixe, et avec des flux sélectionnés dans les mêmes intervalles et selon la même distribution de probabilité que pour l'évaluation. Les paramètres d'entraînement, tels que la durée d'un épisode ou le nombre d'épisodes, doivent être fournis pour assurer la reproductibilité du protocole.

### 3.4 Métriques

Pour chaque simulation, nous conservons le temps de trajet moyen et les émissions moyennes de CO<sub>2</sub> par trajet. Ces mesures sont les plus pertinentes, car elles sont directement liées aux

questions écologiques et économiques mentionnées en introduction. Les émissions de CO<sub>2</sub> d'un réseau par unité de temps ne peuvent pas être utilisées pour évaluer la fluidité d'un réseau : les véhicules émettent moins de CO<sub>2</sub> lorsqu'ils sont à l'arrêt que lorsqu'ils se déplacent. Par conséquent, un embouteillage émet moins de CO<sub>2</sub> par unité de temps qu'un trafic en mouvement.

Les données recueillies sont représentées sous la forme d'un nuage de points dont l'abscisse est la charge totale du réseau, c'est-à-dire le nombre total de véhicules circulant sur le réseau par heure. Cette représentation fournit une indication visuelle du système qui fonctionne le mieux dans différentes conditions de trafic. Cependant, pour pouvoir classer les systèmes, ces données doivent être agrégées pour former une valeur unique. Nous avons donc développé deux nouvelles méthodes d'agrégation des données : le potentiel et la variance.

**Potentiel.** Le potentiel représente la performance empirique moyenne d'un SGFT réglé optimalement. Il est calculé en sélectionnant, pour chaque intervalle de charge, les meilleurs résultats de simulation, c'est-à-dire le temps de trajet moyen et les émissions de CO<sub>2</sub> les plus faibles, et en en faisant la moyenne. L'intervalle de charge est une segmentation de l'espace des résultats selon la charge de trafic. Chaque intervalle représente un profil de trafic, et catégorise les différentes charges possibles sur le réseau. Plus formellement, le potentiel est décrit par l'équation 1, où  $m$  est la valeur de charge maximale possible, et  $k$  l'intervalle de charge. Chaque intervalle de charge peut être considéré comme un profil de trafic différent (léger, moyen, dense, embouteillage), et la moyenne des meilleurs résultats de ces profils est équivalente au calcul de la meilleure performance empirique d'un SGFT.

$$P = \frac{1}{n/k} \sum_{i=0}^{(n/k)-1} \operatorname{argmin}(ik, ik + k) \quad (1)$$

**Variance.** La variance représente la différence moyenne de performance entre un mauvais réglage et un bon réglage. Un système bien réglé peut avoir de très bonnes performances pour un profil de trafic donné, mais des performances médiocres avec un autre paramétrage. La variance quantifie la conséquence moyenne d'un mauvais réglage. Pour chaque intervalle de charge, nous calculons la moyenne des écarts absolus de chaque intervalle. Formellement, la

variance est définie par l'équation 2, où  $R$  représente l'ensemble des résultats des simulations.

$$D = \frac{1}{n/k} \sum_{i=0}^{(n/k)-1} \left( \frac{1}{\#R} \sum_{r \in R[ik, ik+k]} r - \bar{R} \right) \quad (2)$$

Le potentiel et la variance peuvent facilement être présentés dans un tableau, afin de classer les systèmes testés. Le système qui a le meilleur potentiel, soit la valeur la plus basse, est jugé comme étant le meilleur pour un réseau donné, s'il est bien calibré. D'autre part, la variance peut atténuer l'évaluation d'un système ayant un bon potentiel. S'il est mal paramétré, un système ayant une variance élevée aura des performances très mauvaises, quel que soit son potentiel. Dans tous les cas, il est clair que si un système possède les meilleures valeurs de potentiel et de variance, il est le meilleur système absolu pour le réseau routier testé.

### 3.5 Transmission des paramètres

Avec les protocoles existants, nous devrions transmettre les paramètres précis de chacune des simulations, ce qui serait trop volumineux pour être intégré dans un article. Puisque nous générons des simulations de manière aléatoire à partir d'intervalles de valeurs pour chaque paramètre, il suffit maintenant de transmettre uniquement ces intervalles, ainsi que la distribution de probabilité de sélection, sous forme de tableaux multiples par exemple, qui sont plus concis.

## 4 Démonstration

Pour illustrer ce nouveau protocole, nous menons une série d'expériences afin de comparer cinq SGFT différents. Nous simulons les flux de trafic sur un réseau et des scénarios arbitraires. Comme expliqué précédemment, les résultats présentés ici prouvent uniquement la supériorité d'un système sur un réseau imaginaire, à savoir un réseau en grille 3x3. Les expériences sont menées sur le simulateur SUMO [6], et plus précisément avec le module python `sumo-experiments`<sup>1</sup>. SUMO a été préféré à Cityflow [29], car le projet n'est plus maintenu depuis plusieurs années.

1. <https://github.com/cristal-smac/sumo-experiments>

## 4.1 Paramètres de simulation

Chaque simulation dure 3600 secondes, et est effectuée 500 fois pour chaque SGFT. Ces valeurs ont été choisies pour être suffisamment élevées pour tester une grande partie des flux possibles et des calibrages du système, mais suffisamment basses pour effectuer toutes les simulations en une seule semaine sur un ordinateur personnel. Les paramètres du réseau et des flux sont décrits ci-dessous. Un tableau récapitulatif des paramètres de simulation est fourni à la fin de la section dans la table 1.

**Réseau.** Les simulations sont effectuées sur un réseau en grille 3x3, où chaque intersection est située à 200 mètres des intersections voisines. Toutes les routes ont une voie pour chaque sens de circulation. Le réseau comporte neuf entrées et neuf sorties, situées aux extrémités. Les feux de signalisation ne comportent que deux phases : une pour la direction ouest/est (OE) et une pour la direction nord/sud (NS). Les phases orange durent 3 secondes. À chaque intersection, nous implémentons deux types de détecteurs, pour chaque voie entrante : des détecteurs booléens et des détecteurs numériques. Les détecteurs booléens ont une portée de 20 mètres et renvoient simplement la présence d'un véhicule sous la forme d'un booléen. Les détecteurs numériques ont une portée égale à la longueur de la voie, et renvoient le nombre de véhicules détectés. Les véhicules simulés sont les voitures et comportements par défaut de SUMO.

**Flux.** Plusieurs plages de paramètres doivent être définies pour générer du trafic sur le réseau. Ces plages sont définies en fonction du flux de trafic supposé réel, que nous imaginons ici. La charge du réseau, c'est-à-dire le nombre de véhicules en circulation, varie de 100 à 5000 véhicules par heure. Les flux sont déséquilibrés entre les entrées OE et NS, les flux OE étant toujours plus importants que les autres. Nous définissons donc un paramètre de déséquilibre, compris entre 65 et 80%, qui représente la part de la charge entrant dans le réseau par les entrées OE. La charge est répartie de manière égale sur toutes les entrées du même point cardinal. Les itinéraires possibles des véhicules sont tous les couples entrée/sortie répartis uniformément sur l'ensemble de la charge du réseau. Le déséquilibre et la charge sont sélectionnés avec une distribution uniforme. Pour calculer le potentiel et la variance, nous sélectionnons un intervalle de charge  $k$  de 500 véhicules, ce qui donne 10

tranches de données, soit environ 50 simulations pour chaque SGFT.

TABLE 1 – Paramètres de simulation

Paramètre	Valeur
Nombre de simulations $n$	500
Durée de simulation	3600 secondes
Réseau	Grille 3x3 - 200m entre FT
Portée détecteurs num.	200 mètres
Portée détecteurs bool.	20 mètres
Charge	100 - 5000 véhicules/heure
Déséquilibre	65 - 80 %
Temps feu orange	3 secondes
Intervalle de charge $k$	500 véhicules/heure
Loi de distribution	Uniforme

## 4.2 Systèmes étudiés

Pour cette étude, nous comparons cinq SGFT multi-agents, dont deux systèmes à base de règles et deux systèmes à apprentissage par renforcement. Dans cette section, nous décrivons brièvement chacun d'entre eux et présentons leurs paramètres.

**Système à temps fixe.** Le système à temps fixe [16] est le plus utilisé dans le monde. Chaque intersection a une liste de phases, et chaque phase a une durée. L'agent passe à la phase suivante dès que le temps de la précédente a expiré. Une durée choisie dans une fourchette de 10 à 120 secondes est attribuée à chaque phase. Chaque agent est paramétré de la même manière. La durée de chaque phase est le seul paramètre requis et est résumée dans la table 2.

TABLE 2 – Paramètres du feu à temps fixe

Paramètre	Min	Max	Pas	Unité
Durée de phase	10	120	1	secondes

**Système Max Pressure.** Le système de pression maximale est un système basé sur des règles développé par [24]. Il définit le concept de pression pour une phase, qui est la différence entre le nombre de véhicules se dirigeant vers l'intersection, et le nombre de véhicules se dirigeant vers une intersection voisine, en provenance de cette intersection. À chaque période  $t$ , l'agent calcule la pression de chaque phase et sélectionne celle dont la pression est la plus élevée. La même phase peut être sélectionnée plusieurs fois de suite. En simulation, nous utilisons les détecteurs numériques de chaque intersection pour

calculer la pression, et nous faisons varier la durée d'une période de 10 à 120 secondes. C'est le seul paramètre requis pour ce système, résumé dans la table 3.

TABLE 3 – Paramètres du système *Max Pressure*

Paramètre	Min	Max	Pas	Unité
Durée de période	10	120	1	secondes

**Système SOTL.** Comme le système de pression maximale, le système SOTL est un système basé sur des règles, développé par [10]. L'agent compte et additionne à chaque intervalle de temps, ici une seconde, le nombre de véhicules entrants et en attente sur les voies rouges, c'est-à-dire les voies pour lesquelles le feu est rouge. Lorsque la somme des véhicules est supérieure à un seuil  $x_1$  et qu'il n'y a plus aucun véhicule sur les voies vertes, le carrefour change de phase. Pour éviter que les phases ne durent trop longtemps lorsque des embouteillages se forment, le feu de circulation change également de phase lorsque le nombre de véhicules en attente dépasse à lui seul un seuil  $x_2$ . Comme précédemment, nous utilisons les détecteurs numériques pour compter les véhicules entrants sur chaque voie. De plus, pour éviter que le système ne clignote, lorsque les phases changent trop rapidement durant une forte phase d'embouteillages, nous ajoutons un temps de phase minimum  $t$ . Les plages de paramètres que nous avons utilisées pour cette expérience sont résumées dans la table 4.

TABLE 4 – Paramètres du système SOTL

Paramètre	Min	Max	Pas	Unité
Seuil changement $x_1$	120	600	1	véhs
Seuil passage en force $x_2$	2	$x_1$	1	véhs
Temps minimum $t$	10	120	1	secs

**Systèmes à apprentissage par renforcement.** Pour tester le protocole avec des systèmes nécessitant un entraînement, nous avons développé deux systèmes simples à apprentissage par renforcement. Les deux systèmes utilisent un algorithme de *double deep Q-network* [23] pour calculer la valeur d'un couple état/action, en mettant en œuvre deux réseaux à deux couches avec respectivement 128 et 56 unités réactivées. Le réseau cible est mis à jour tous les  $f$  épisodes d'entraînement. La sélection des actions est effectuée par un algorithme de type  $\epsilon$ -greedy, où  $\epsilon$  diminue à chaque sélection par une multiplication avec un facteur de mise à jour  $k$ , compris

entre 0 et 1, et a une valeur minimale  $\min\text{-}\epsilon$  comprise entre 0 et 0.1. Pour cette expérience, la phase d'entraînement dure six heures, divisée en épisodes d'une durée de  $\delta$ , où les données sont collectées toutes les secondes. La valeur de  $\epsilon$  commence à diminuer uniquement lorsque la phase d'apprentissage atteint le  $d$ -ième épisode. Lors de la mise à jour des poids du réseau neuronal, nous utilisons un *batch* de taille  $s$  pour un nombre d'époques  $e$ , et un facteur déprécié  $\gamma$  pour le calcul des équations de Bellman [7].

Le premier système (RL1) s'inspire du système Intellilight de [25]. L'état d'un agent est représenté par un vecteur contenant les longueurs de files, le nombre de véhicules et les temps d'attente totaux sur chaque voie, ainsi que la phase actuelle. L'agent choisit une action toutes les secondes : soit conserver la phase actuelle, soit passer à la suivante. La récompense de l'algorithme est le délai moyen, c'est-à-dire la différence entre la vitesse des véhicules et la vitesse maximale d'une voie, multipliée par le nombre de véhicules sur chaque voie.

Le second système (RL2) s'inspire des systèmes [25] et [8]. La représentation de l'environnement par l'agent est la même que celle du système RL1, mais l'espace d'actions est différent. À chaque intervalle de temps  $t$ , l'agent sélectionne une phase pour l'intersection, qui peut être la même que la précédente. La récompense à minimiser est la pression de l'intersection.

Toutes les plages de paramètres sont résumées dans la table 5.

TABLE 5 – Paramètres des systèmes à apprentissage par renforcement

Paramètre	Min	Max	Pas	Unité
Valeur de départ $\epsilon$	1	1	aucun	aucune
Coefficient MàJ $k$	0.9	0.999	0.001	aucune
$\min\text{-}\epsilon$	0	0.1	0.01	aucune
Ep. décroissance $d$	20	100	1	aucune
Durée épisode $\delta$	120	600	1	secondes
Taille du batch $s$	20	300	1	données
Nb époques $e$	20	200	1	époques
Facteur déprécié $\gamma$	0.7	0.99	0.01	aucune
MàJ réseau cible $f$	1	10	1	épisodes

## 5 Résultats et discussions

### 5.1 Analyse du classement

Les résultats de la démonstration sont représentés sur la Fig 1. Cette première analyse visuelle

permet de faire plusieurs affirmations. Bien que le système *Max Pressure* ait les meilleures performances moyennes, le système RL1 produit les meilleurs résultats globaux, avec des flux de trafic faibles, mais aussi avec des flux de trafic très élevés. Ce phénomène est la raison exacte pour laquelle nous avons introduit la notion de potentiel. Si nous ne calculions que les performances moyennes de tous les systèmes, le système RL1 aurait une faible évaluation, malgré ses très bonnes performances avec des paramètres précis. Le potentiel met en évidence ces performances, même si la majorité des résultats sont mauvais. Étant donné qu'un réseau routier n'a besoin que d'un seul SGFT avec un seul réglage, le potentiel semble plus juste que le résultat moyen. Mais la disparité des résultats ne doit pas être totalement ignorée. Les systèmes d'apprentissage par renforcement et la stratégie de temps fixe ont des résultats dispersés, ce qui signifie que leur réglage a un impact important sur leurs performances. Ceci contraste avec le système *Max Pressure*, pour lequel les résultats sont beaucoup plus compacts, et montre qu'il est plus résistant aux variations de charge ou de réglages. Le potentiel et la variance de chaque système sont présentés dans la table 6 en termes de temps de trajet moyen.

TABLE 6 – Potentiel et variance des systèmes, calculés à partir du temps de trajet moyen

Système	Potentiel	Variance
Temps fixe	369.22	983.73
<i>Max pressure</i>	137.95	<b>363.58</b>
SOTL	193.6	560.66
RL1	<b>85.53</b>	2533.62
RL2	128.58	1221.5

L'interprétation que l'on peut faire des résultats de ce protocole est que le RL1 est le meilleur système pour une grille 3x3 avec ce type de flux asymétriques, avec un potentiel de temps de trajet moyen de 85 secondes. Cependant, il faut veiller à ce que le système soit très bien réglé. En cas de mauvais réglage, les performances seront pires que celles du système *Max Pressure* dans le pire des cas, bien que son potentiel soit moins bon.

## 5.2 Objectivité des résultats

Le nombre élevé de simulations garantit une meilleure objectivité des résultats que les protocoles existants. Dans les protocoles précédents, seules quelques simulations ont été effectuées,

et les variations des performances en fonction des flux ou des paramètres du système n'ont pas pu être étudiées. Cela a conduit à des résultats biaisés, comme démontré précédemment. Nous pouvons reproduire ce biais avec les données de la Fig 1. Si nous ne sélectionnons qu'un seul profil de trafic pour classer des SGFT, nous pouvons obtenir n'importe quel classement, comme le montre la table 7.

TABLE 7 – Classements des différents systèmes pour des conditions uniques de charge de trafic, exprimées en véhicules/heure

Classement	1768 vehs/h	2422	2883	2907
<b>1er</b>	MP	RL1	SOTL	MP
<b>2ème</b>	RL2	MP	MP	SOTL
<b>3ème</b>	SOTL	RL2	RL1	RL2
<b>4ème</b>	RL1	SOTL	RL2	Fixed
<b>5ème</b>	Fixed	Fixed	Fixed	RL1

Cette contradiction prouve que les protocoles existants ne sont pas pertinents pour évaluer des SGFT. Notre protocole n'est pas soumis à ce biais. Un seul classement peut être produit, et bien que l'aléatoire nous empêche de le définir comme déterministe, ce protocole renvoie presque toujours le même classement pour le même ensemble de paramètres, à condition que  $n$  soit grand.

## Influence des paramètres

Bien que les résultats de ce protocole soient considérés comme objectifs, il ne faut pas négliger l'impact de certains paramètres qui peuvent avoir des conséquences importantes sur les résultats des expériences.

**Nombre de simulations.** Le nombre de simulations a une grande influence sur le classement final du système. Comme dans les protocoles de comparaison existants, un faible nombre de simulations conduit à des résultats non pertinents, puisque n'importe quel résultat peut être obtenu à partir de résultats individuels. Pour garantir l'objectivité du protocole, un nombre élevé de simulations doit être effectué, l'idée étant que plus il est élevé, mieux c'est. Le nombre minimum de simulations dépend de l'espace d'exploration, c'est-à-dire de toutes les combinaisons de débits et de paramètres du système. Selon nous, le strict minimum pour une évaluation objective est de 100 simulations par système.

**Intervalle de charge.** Chaque intervalle représente un profil de trafic, et catégorise les dif-

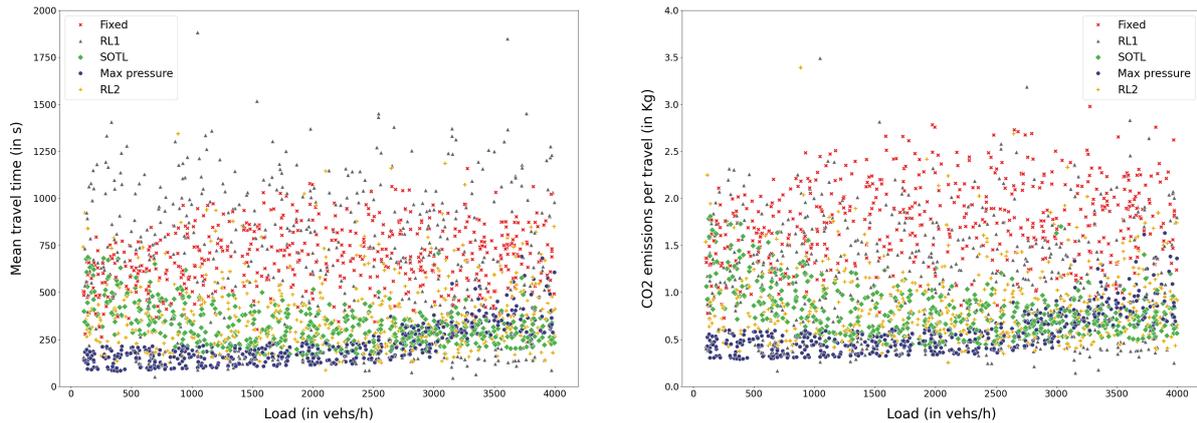


FIGURE 1 – Temps de trajet moyen et émissions de CO2 moyenne par trajet de chaque simulation, en fonction de la charge de trafic et du système utilisé

férentes charges possibles sur le réseau. Ce paramètre doit donc être choisi avec soin. Une valeur faible se traduit par une segmentation fine des résultats de la simulation, qui peut ne pas contenir suffisamment de données pour calculer objectivement les métriques. D'autre part, une tranche de résultats trop large ne définit pas précisément les différents profils de trafic et entraîne une perte d'information. Par conséquent, l'intervalle de charge doit être sélectionné avec une connaissance préalable des profils de trafic du réseau, afin de garantir la précision du potentiel et de la variance.

**Paramètres des systèmes.** La définition des plages de valeurs pour chaque paramètre du système est également un point crucial du protocole. Elle définit les combinaisons de paramètres possibles qui sont simulées et étudiées. Parmi toutes les configurations, on doit trouver les meilleures possibles. Si ce n'est pas le cas, le potentiel réel d'un système ne peut pas être évalué. En cas d'incertitude sur la direction dans laquelle les systèmes doivent être paramétrés, il est préférable d'inclure autant de configurations que possible pour éviter le risque de biais.

## Conclusion

Dans cet article, nous avons présenté un protocole permettant d'évaluer et de classer les systèmes de gestion du trafic routier. L'avantage de ce protocole est qu'il ne souffre pas des biais des protocoles existants. En introduisant deux nouvelles mesures, le potentiel et la variance, ainsi qu'un concept de simulation stochastique, nous sommes en mesure d'évaluer les systèmes de gestion du trafic routier quantitativement et

objectivement. Nous donnons également les limites de notre approche dont l'évaluation dépend de multiples paramètres, et nous donnons quelques lignes directrices pour éviter tout biais. Quoiqu'il en soit, comme ce protocole facilite la transmission de chacun des paramètres sélectionnés, les résultats peuvent être interprétés avec précision et tout relecteur peut détecter des anomalies dans la méthode d'essai. Deux types d'expériences futures pourront utiliser ce travail. Dans le cadre du développement d'un nouveau système, ce protocole pourra être utilisé pour mieux comprendre l'évolution des performances en fonction de ses paramètres, ou de l'évolution des flux de trafic. D'un point de vue plus technique, ce protocole pourra être utilisé par les gestionnaires de réseaux routiers pour sélectionner le système qui fonctionne le mieux pour leur réseau spécifique. Nous pensons que ce protocole a le potentiel pour devenir l'un des protocoles d'évaluation standard à l'avenir.

## Références

- [1] B. Abdulhai, R. Pringle, and G.J. Karakoulas. Reinforcement learning for true adaptive traffic signal control. *Journal of Transportation Engineering*, 129(3) :278–285, 2003.
- [2] H. Amer, N. Salman, M. Hawes, M. Chaqfeh, L. Mi-haylova, and M. Mayfield. An improved simulated annealing technique for enhanced mobility in smart cities. *Sensors*, 16(7) :1013, 2016.
- [3] E. Azimirad, N. Pariz, and M.B.N. Sistani. A novel fuzzy model and control of single intersection at urban traffic network. *IEEE Systems Journal*, 4(1) :107–111, 2010.
- [4] Shashank B., Sudheer B., Rohit, and Munish K. C. Impact of congestion on greenhouse gas emissions for road transport in mumbai metropolitan region. *Transportation Research Procedia*, 25 :3538–3551,

2017. World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016.
- [5] A.L.C. Bazzan. A distributed approach for coordination of traffic signal agents. *Autonomous Agents and Multi-Agent Systems*, 10 :131–164, 2005.
- [6] M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz. Sumo - simulation of urban mobility : an overview. In *Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation*. ThinkMind, 2011.
- [7] R. Bellman. Dynamic programming. *Press Princeton, New Jersey*, 39, 1957.
- [8] C. Chen, H. Wei, N. Xu, G. Zheng, M. Yang, Y. Xiong, K. Xu, and Z. Li. Toward a thousand lights : Decentralized deep reinforcement learning for large-scale traffic signal control. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 3414–3421, 2020.
- [9] M. Collotta, L.L. Bello, and G. Pau. A novel approach for dynamic traffic lights management based on wireless sensor networks and multiple fuzzy logic controllers. *Expert Systems with Applications*, 42(13) :5403–5415, 2015.
- [10] S. Cools, C. Gershenson, and B. D’Hooghe. Self-organizing traffic lights : A realistic simulation. *Advances in applied self-organizing systems*, pages 45–55, 2013.
- [11] B. De Schutter, H. Hellendoorn, A. Hegyi, M. Van den Berg, and S.K. Zegeye. Model-based control of intelligent traffic networks. *Intelligent Infrastructures*, pages 277–310, 2010.
- [12] H. Goel, Y. Zhang, M. Damani, and G. Sartoretti. Socialight : Distributed cooperation learning towards network-wide traffic signal control. *arXiv preprint arXiv :2305.16145*, 2023.
- [13] M. Guériau, F. Armetta, S. Hassas, R. Billot, and N.E. El Faouzi. Contrôle par apprentissage constructiviste et régulation du trafic coopératif. In *Journées Francophones des Systèmes Multi-Agents 2017-Cohésion : fondement ou propriété émergente*, 2017.
- [14] H. Huang, Z. Hu, Z. Lu, and X. Wen. Network-scale traffic signal control via multiagent reinforcement learning with deep spatiotemporal attentive network. *IEEE transactions on cybernetics*, 2021.
- [15] Q. Jiang, M. Qin, S. Shi, W. Sun, and B. Zheng. Multi-agent reinforcement learning for traffic signal control through universal communication method. *arXiv preprint arXiv :2204.12190*, 2022.
- [16] P. Koonce and L. Rodegerdts. Traffic signal timing manual. Technical report, United States. Federal Highway Administration, 2008.
- [17] M. Krzyzanowski, B. Kuna-Dibbert, and J. Schneider. *Health effects of transport-related air pollution*. WHO Regional Office Europe, 2005.
- [18] A. Mačiulis, A.V. Vasiliauskas, and G. Jakubauskas. The impact of transport on the competitiveness of national economy. *Transport*, 24(2) :93–99, 2009.
- [19] T. Mao, A.-S. Mihăiță, F. Chen, and H.L. Vu. Boosted genetic algorithm using machine learning for traffic control optimization. *IEEE Transactions on Intelligent Transportation Systems*, 23(7) :7112–7141, 2021.
- [20] H. Ngo, E. Kaddoum, M.-P. Gleizes, J. Bonnet, and A. Goursolle. P-adrip : un système multi-agent auto-organisateur pour la prévision du trafic routier. In *Journées Francophones sur les Systèmes Multi-Agents : SMA et Smart Cities (JFSMA 2022)*, 2022.
- [21] F. Teklu, A. Sumalee, and D. Watling. A genetic algorithm approach for optimizing traffic control signals considering routing. *Computer-Aided Civil and Infrastructure Engineering*, 22(1) :31–43, 2007.
- [22] E. Van der Pol and F.A. Oliehoek. Coordinated deep reinforcement learners for traffic light control. *Proceedings of learning, inference and control of multi-agent systems (at NIPS 2016)*, 8 :21–38, 2016.
- [23] H. Van Hasselt, A. Guez, and D. Silver. Deep reinforcement learning with double q-learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 30, 2016.
- [24] P. Varaiya. The max-pressure controller for arbitrary networks of signalized intersections. In *Advances in dynamic network modeling in complex transportation systems*, pages 27–66. Springer, 2013.
- [25] H. Wei, G. Zheng, H. Yao, and Z. Li. Intellilight : A reinforcement learning approach for intelligent traffic light control. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2496–2505, 2018.
- [26] R. Wunderlich, C. Liu, I. Elhanany, and T. Urbanik. A novel signal-scheduling algorithm with quality-of-service provisioning for an isolated intersection. *IEEE Transactions on intelligent transportation systems*, 9(3) :536–547, 2008.
- [27] F. Zambanelli. From smart to autonomous (cities, homes, and everything) : Opportunities and challenges. In *SMA et Smart Cities - Trentièmes Journées Francophones sur les Systèmes Multi-Agents, JFSMA 2022, Saint-Etienne, France, June 27-29, 2022*, page 5. Cépaduès, 2022.
- [28] X. Zang, H. Yao, G. Zheng, N. Xu, K. Xu, and Z. Li. Metalight : Value-based meta-reinforcement learning for traffic signal control. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 1153–1160, 2020.
- [29] H. Zhang, S. Feng, C. Liu, Y. Ding, Y. Zhu, Z. Zhou, W. Zhang, Y. Yu, H. Jin, and Z. Li. Cityflow : A multi-agent reinforcement learning environment for large scale city traffic scenario. In *The world wide web conference*, pages 3620–3624, 2019.
- [30] L. Zhang, Q. Wu, J. Shen, L. Lü, B. Du, and J. Wu. Expression might be enough : Representing pressure and demand for reinforcement learning based traffic signal control. In *International Conference on Machine Learning*, pages 26645–26654. PMLR, 2022.
- [31] G. Zheng, Y. Xiong, X. Zang, J. Feng, H. Wei, H. Zhang, Y. Li, K. Xu, and Z. Li. Learning phase competition for traffic signal control. In *Proceedings of the 28th ACM international conference on information and knowledge management*, pages 1963–1972, 2019.