International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Congestion Management in Mixed-Traffic Urban Networks: A Literature Review of Definitions, Indices, Propagation Mechanisms, and Research Gaps

Ashwin S Prabhu[1], Sewa Ram[2]

[1]Doctoral Student, Transport Planning, School of Planning and Architecture, New Delhi, India
[2]Professor, Transport Planning, School of Planning and Architecture, New Delhi, India

Abstract:
Urban traffic congestion has shifted from being treated as a static, point-condition imbalance between demand and capacity to a dynamic, propagating network phenomenon shaped by geometry, driver interaction, side friction, land use, and signal control. This is particularly true for Indian and other mixed-traffic cities, where lane-less movement, heterogeneous vehicle classes, intense pedestrian and non-motorised activity, and weak lane discipline destabilise classical models. This paper presents a synthesised literature review of congestion management covering: (a) foundational definitions and classifications of congestion drawn from demand-capacity, vehicular-interaction, delay, and cost-based perspectives; (b) the evolution of traffic flow models from first-order (LWR) through second- and third-order continuum formulations and Kerner's three-phase theory, with an evaluation of their relative ability to capture observed congestion dynamics; (c) congestion indices grouped by travel-time, speed and Level-of-Service families; (d) a comparative view of urban versus rural congestion; (e) a typology of recurring congestion mapped to shock-wave behaviour and effective capacity impact; (f) network-level diagnostics using centrality measures and resilience metrics; and (g) operational congestion management practices codified by agencies such as the US Federal Highway Administration. Drawing across more than thirty primary sources, the review converges on a picture of congestion as a cascading process that can be anticipated and managed when early signals are measured and control is network-aware. The paper concludes with a consolidated set of research gaps tailored to Indian urban conditions, including the under-development of propagation-based indices, the limited calibration of higher-order models to heterogeneous flows, the poorly quantified influence of side friction on shock-wave generation, the scarcity of multi-scale frameworks linking microscopic lane-change turbulence with network-level diffusion, and the need for context-specific viscosity-style parameters and congestion-burden valuation.

Keywords: Traffic congestion; Congestion propagation; Heterogeneous traffic; Shock waves; Side friction; Centrality measures; Resilience; Congestion indices; Mixed-traffic;

1. Introduction

Rapid urbanisation and motorisation have intensified traffic congestion across cities worldwide, with Indian cities increasingly burdened by reduced mobility, deteriorating reliability, compromised safety, elevated energy consumption and worsening air quality (Akbar et al., 2021; Biswas, 2021). The economic costs of congestion arise from excess travel time and schedule unreliability, while the environmental penalties follow from idling and stop-start driving that elevate fuel use and tail-pipe emissions. Beyond aggregate averages, the lived experience of congestion is dominated by variability and cascading effects: a local disturbance at a road confluence, bus stop, or unsignalised access point can trigger backward-moving shock waves, spill back across upstream intersections, and produce system-level delays that far exceed the magnitude of the original disturbance (Schonhof and Helbing, 2008; Duan, 2023).

Traditional engineering practice framed congestion as a demand-capacity imbalance and measured it through Level-of-Service categories or delay-based indices (Rosenbloom, 1978; Lomax et al., 1997). Contemporary research, however, places far greater emphasis on propagation mechanisms, phase transitions, and the topology of the underlying road network (Kerner, 2004; Sole-Ribalta, Gomez and Arenas, 2016; Nagy and Simon, 2021). Under Indian heterogeneous conditions — marked by lane-less movement, multiple vehicle classes with widely different operating speeds and dimensions, intense pedestrian and non-motorised transport interactions, and strong roadside (side) friction — classical assumptions often fail, and context-sensitive indicators and models become necessary (Mohan, 2017; Biswas, 2021).

This paper reviews the state of knowledge on congestion and its management in mixed-traffic urban networks. It connects definitions and quantification indices to the underlying flow dynamics, geometry, side friction, and network structure, and closes with a focused research agenda for Indian cities. The review is organised around eight main sections following this introduction. Section 2 synthesises foundational definitions and classifications of congestion. Section 3 summarises the body of literature reviewed and the parameters extracted from each work. Section 4 explains how different families of traffic flow models contribute to congestion assessment. Section 5 presents the catalogue of congestion indices. Section 6 distinguishes urban from rural (highway) congestion. Section 7 develops a typology of recurring congestion linked to shock-wave behaviour and capacity impact. Section 8 examines network-level diagnostics using centrality measures. Section 9 catalogues congestion management practices. Section 10 consolidates the research gaps, and Section 11 concludes.

2. Definitions and Classification of Congestion

The literature offers a wide range of definitions of traffic congestion, each foregrounding a different facet of the phenomenon. These can be grouped into four broad categories: (a) demand-and-capacity based definitions, (b) vehicular-interaction based definitions, (c) delay or travel-time based definitions, and (d) cost-based definitions.

2.1 Demand-and-Capacity Based Definitions

The earliest and most intuitive framing treats congestion as the result of demand exceeding the capacity of the transport facility. Rosenbloom (1978) defined congestion as occurring when travel demand exceeds the existing road system capacity. Rothenberg (1985) refined this by stating that congestion is a condition in which the number of vehicles attempting to use a roadway at a given time exceeds the ability of the roadway to carry the load at generally acceptable service levels. The Institute of Civil Engineers (1989) explicitly framed it as a case of demand exceeding supply, while Vuchic and Kikuchi (1994) described congestion as the state that results when vehicular volume on a facility exceeds its capacity.

2.2 Vehicular-Interaction Based Definitions

A second group of definitions emphasises mutual impedance among vehicles as the defining characteristic of congestion. The European Conference of Ministers of Transport (ECMT, 1999) described congestion as the impedance that vehicles impose on each other due to the speed-flow relationship in conditions where the use of a transport system approaches its capacity. Bovy and Salomon (2002) characterised it as a state of traffic flow on a transport facility marked by high densities and low speeds relative to some chosen reference state. Pisaraski (1990, cited in Miller and Li, 1994) framed congestion as an imbalance between traffic flow and capacity that causes increased travel time, cost and modification of behaviour. Cambridge Systematics and the Texas Transportation Institute (2005) similarly defined it as an excess of vehicles on a portion of roadway at a particular time resulting in speeds slower than normal or free-flow speeds.

2.3 Delay and Travel-Time Based Definitions

Delay-based definitions are the most operationally useful because they convert the abstract notion of congestion into a measurable quantity. Lomax et al. (1997) defined traffic congestion as travel time or delay in excess of that normally incurred under light or free-flow travel conditions. Weisbrod, Vary and Treyz (2001) framed it as a condition of traffic delay in which the flow of traffic is slowed below reasonable speeds because the number of vehicles trying to use the road exceeds the network's capacity. Kockelman (2004) reduced it to a more general statement of the presence of delays along a physical pathway due to the presence of other users, while Downs (2004) defined it as the situation in which traffic moves at speeds below the designed capacity of a roadway. These definitions underpin most of the contemporary congestion indices summarised in Section 5.

2.4 Cost-Based Definitions

Cost-based perspectives extend the physical states of congestion into the realm of externalities and welfare economics. The Victoria Transport Policy Institute (VTPI, 2005) defined congestion as the incremental costs resulting from interference among road users. This category is particularly important because it provides the foundation for benefit-cost analyses of congestion management projects and links engineering measurements to broader policy debates about pricing and demand management. The 1999 ECMT Round Table report further introduced the notion of excessive congestion, defined as congestion whose marginal cost to society exceeds the marginal cost of efforts to reduce it (such as adding to road or other transport infrastructure).

2.5 Allied Constructs: Hotspots, Duration, Extent, Intensity and Reliability

Alongside the core definitions, several allied constructs frame how congestion is observed and described. INRIX defines a congestion hotspot as a traffic jam that occurs at the same location along a stretch of road where speeds typically drop below 65 per cent of the reference uncongested speed for at least two minutes. The US Federal Highway Administration (FHWA, 2020) defines a traffic bottleneck as a localised section of highway that experiences reduced speeds and inherent delays due to a recurring operational influence or a non-recurring impacting event. Congestion is also commonly partitioned by its temporal regularity into recurring congestion (caused by regular daily or weekly demand patterns) and non-recurring congestion (caused by incidents, work zones, weather, or events) (ECMT, 2004; Lockwood, 2006).

Falcocchio and Levinson (2015) and Lomax et al. (1997) further codified congestion through four descriptive dimensions: duration (the amount of time congestion affects the system, with peak hours having extended into peak periods on many corridors); extent (the number of people or vehicles affected and the geographic spread); intensity (the severity, measured through delay or average speed reductions); and reliability (a measure of the road user's ability to accurately predict and plan for a travel time, recognising that motorists tolerate expected delay better than unexpected delay). These dimensions are summarised in Table 1.

Table 1: Categorisation of Congestion Definitions

Definitional Category	Core Idea	Representative Authors
Demand and capacity based	Congestion occurs when travel demand exceeds the capacity of the facility at acceptable service levels.	Rosenbloom (1978); Rothenberg (1985); ICE (1989); Vuchic and Kikuchi (1994)
Vehicular interaction based	Congestion is the mutual impedance among vehicles as use approaches capacity, observed as high densities and low speeds.	ECMT (1999); Bovy and Salomon (2002); Cambridge Systematics and TTI (2005)
Delay and travel-time based	Congestion is operationalised as the delay or travel time in excess of free-flow conditions.	Lomax et al. (1997); Weisbrod, Vary and Treyz (2001); Kockelman (2004); Downs (2004)
Cost based	Congestion is the incremental cost imposed on users and society by interference among road users.	VTPI (2005); ECMT (1999)
Hotspot and bottleneck	Localised, recurring drop in speed (e.g., below 65% of reference) of a defined duration.	INRIX; FHWA (2020)
Recurring vs non-recurring	Distinguishes between regularly-occurring demand-driven congestion and incident-driven episodes.	ECMT (2004); Lockwood (2006)
Descriptive dimensions	Duration, extent, intensity, and reliability as the four lenses for characterising congestion.	Lomax et al. (1997); Falcocchio and Levinson (2015)

3. Summary of Literature Reviewed

A structured review of more than twenty-three primary works on congestion was carried out, with each work assessed against a common set of parameters: whether it provides a definition of congestion, methods of quantification, classification schemes, key indicators or indices, environmental considerations, treatment of network form, and treatment of land-use accessibility or spacing. The themes addressed by this body of literature can be summarised as follows.

First, several works focus on the definition, quantification and classification of congestion based on traffic states and their effects (intensity, duration, extent and variability). Schonhof and Helbing (2008), in their assessment of a 30-kilometre freeway corridor, classified congestion into five distinct traffic states and their combinations: pinned localised cluster, homogeneous congested traffic, oscillating congested traffic, stop-and-go waves, moving localised cluster, and the boomerang effect. Falcocchio and Levinson (2015) categorised congestion based on its intensity, duration, extent, and variability, while the National

Cooperative Highway Research Program Report No. 398 (NCHRP, 1998) provided a number of definitions correlated with mobility, accessibility, and reliability.

Second, a related body of work designed means to measure the effects and consequences of congestion. Lindley (1987) developed a threshold of 0.77 volume-to-capacity ratio for congestion on freeways based on the 1985 Highway Capacity Manual calculation for a local freeway, used to estimate recurring delays alongside an incident database. The Institute of Transportation Engineers Toolbox (ITE, 1996) defined congestion based on the delay-based inconvenience caused to road users. Kumar and Sivanandan (2012) reviewed Highway Capacity Manual based, queuing based, and time based congestion identification methodologies and assessed the applicability of time-based congestion indices for Indian conditions.

Third, a strand of literature addresses guidelines and the applicability of these guidelines in specific contexts. IRC SP 30 (2019) developed time and distance value congestion factors based on Level of Service to arrive at the generalised cost of congestion effect, including various lane arrangements and vehicle typologies of highways. The 1999 ECMT Round Table report defined the notion of excessive congestion.

Fourth, the perception of congestion has been studied in relation to the intrinsic speed of a city and the actual congestion that ensues from this baseline. Akbar et al. (2021) developed speed and congestion indices for more than thirty Indian cities based on travel time, identifying the actual congestion against the intrinsic speed of each city. This work is particularly important for Indian urban analysis because it acknowledges that a single global free-flow threshold misrepresents cities with structurally different baseline speeds.

Fifth, research on hotspots, their spread, and network-level characteristics has progressed substantially. Sole-Ribalta, Gomez and Arenas (2016) proposed an idealised model, based on critical phenomena arising in complex networks, that allows the analytical prediction of congestion hotspots in urban environments after the onset of congestion via the phase transition approach. Lockwood (2006) assessed recurring and non-recurring congestion characteristics for urban and rural roads, finding that the percentage of recurring delays increases as the size of the urban area increases. Nagy and Simon (2021) developed a propagation probability graph to define congestion propagation on an urban road network, while Sen Luan et al. (2021) showed that traffic congestion exhibits a Bayesian property, modelling propagation as a process in which the speed drop of a particular road section causes the speed reduction of a series of upstream road sections.

Sixth, allied studies have examined side friction, spatio-temporal congestion characteristics, and approaches to modelling heterogeneous traffic. Biswas (2021) analysed the influence of side friction on the capacity of undivided urban streets, considering on-street parking, pedestrian movements and non-motorised vehicles across six base and ten non-base locations; capacity under the influence of side friction was observed to be significantly lower (up to 60.73 per cent reduction). Erdelic (2021) determined spatio-temporal characteristics of congestion zones using morphological closing operations and Monte Carlo simulation coupled with temporal clustering on a large historical Floating Car Data (FCD) dataset, producing time-varying Travel Time Indexes. Zhou (2021) developed a second-order Lagrangian macroscopic traffic flow model for freeways, while Mohan (2017) modelled heterogeneous traffic flow using a second-order macroscopic continuum model and addressed the challenges of heterogeneous flow by extending an existing model to use area occupancy for traffic concentration instead of density.

Seventh, the relationship between congestion and emissions has been examined. Smit, Brown and Chan (2008) examined how, and to what extent, models currently used to predict emissions and fuel consumption from road traffic include the effects of congestion, and developed a classification framework in which driving pattern serves as the key factor connecting emissions to congestion. The influence of land-use configuration has been considered by Wang and Debbage (2021), who derived the impact of the abundance and spatial configuration of urban land uses on traffic congestion; specifically, high-intensity and low-intensity urban land uses were associated with more congestion, while contiguous residential development was correlated with less congestion.

Eighth, congestion rankings and scoring methodologies have been issued by agencies and administration boards. The Regional Integrated Transportation Information System (RITIS) provides bottleneck rankings based on the number of days in the analysis period, the number of bottleneck occurrences, the duration of congestion in minutes, and the length of congestion in miles. The INRIX National Traffic Scorecard (INRIX, 2010) defines a congested corridor as one where recurring congestion occurs on multiple road segments totalling at least three miles in length, where at least one segment is congested ten hours per week on average, and where all road segments in the corridor have at least four hours a week of congestion on average. Duan (2023) showed that the recovery duration of jams associated with each bottleneck follows a power-law distribution, and that the growth speed of jams in their very early propagation stage is highly correlated with the maximal size of the jams — a finding that opens the door to using early-stage speed as a predictor of major bottlenecks with high accuracy.

Finally, ramp metering and infrastructure-side interventions have been studied. Zhang and Levinson (2010) identified the impact of ramp metering on the capacity of active freeway bottlenecks based on lane occupancy in the mainline and access roads.

4. How Different Traffic Flow Models Help in Assessing Congestion

Traffic flow models are the analytical backbone of congestion assessment. The choice of model dictates which phenomena can be reproduced (and therefore measured) and which remain hidden. This section traces the evolution from first-order through third-order continuum models and Kerner's three-phase theory, and explains how each contributes to congestion assessment.

4.1 First-Order (LWR) Models

Macroscopic traffic flow models are based on aggregate variables representing user behaviour (such as flow and density) and aggregate supply variables (such as speed). The foundational macroscopic formulation is the Lighthill-Whitham-Richards (LWR) model, the basic version of which is formulated in continuous space and time variables. The LWR model assumes that speed is determined statically by the density (or spacing) condition according to the fundamental diagram, and it captures kinematic waves and the propagation of shocks along a roadway.

For congestion assessment, the strength of the LWR model lies in its ability to describe the formation and propagation of shock waves and to predict queue lengths at bottlenecks. However, its limitations are well documented. Kerner and Konhauser (1994a) showed that, under a parabolic fundamental diagram, LWR predicts that downstream shock fronts would eventually change shape and become smoother and smoother, whereas experimental data support a stable velocity profile over time, moving upstream with a constant velocity. Nagel and Nelson (2005) summarised the need for higher-order models above LWR by noting that LWR cannot describe unstable flow, cannot describe spontaneous breakdowns of traffic flow, and cannot explain the two-capacity phenomenon known as capacity drop. Schonhof and Helbing (2007) showed, including through a Fourier analysis to rule out pulsating inflows, that LWR without particular assumptions or extensions cannot describe growing amplitudes in density, speed, or flow profiles, nor the spontaneous emergence of stop-and-go waves or other oscillating traffic patterns. For congestion assessment, this means that LWR is well suited to deterministic queue analysis but cannot explain or quantify the most disruptive observed congestion patterns.

4.2 Second-Order Models

Second-order models, originating with the work of Payne, Ross, and later Kerner and Konhauser, add a dynamic equation for the evolution of speed itself. The speed is no longer statically determined by the density-spacing condition according to the fundamental diagram. As a result, these models support the modelling of several real-world traffic patterns including oscillation in congestion, traffic hysteresis, and gradual speed changes. They therefore extend the ability to understand the propagation of shock waves beyond the kinematic level. However, Daganzo (2002) suggested that the breakdown of free traffic can be traced back to a lane change in front of a highly compressed set of cars and argued that there is a tangible reason for jam formation — although its origin can be a rather small disturbance. Muñoz and Daganzo (2002) further proposed that small oscillations may be increased in amplitude due to a pumping effect at ramps, that is, a positive feedback mechanism in which large oscillations in flow, speed and cumulative count increase in amplitude across the detectors spanning a long freeway queue and its intervening on-ramps. Despite these advances, factors such as the instantaneous driver's reaction and the impact of the inertial effect, as well as drivers' reactions to the conditions of the traffic context, cannot be fully modelled by standard second-order formulations.

For Indian heterogeneous traffic, the second-order family has been extended. Mohan (2017) extended a second-order continuum model to use area occupancy instead of density as the measure of traffic concentration, which is essential when vehicles of widely differing footprints share lane-less carriageways. Zhou (2021) developed a Lagrangian second-order macroscopic model in which the coordinate system is fixed on the moving vehicles — a formulation that resolves several numerical and conceptual issues when capturing shock propagation through freeway traffic. Both extensions illustrate that second-order models are particularly valuable for congestion assessment when shock-wave dynamics, hysteresis, or class heterogeneity dominate.

4.3 Third-Order Models

Within the same family of macroscopic continuum models, the third-order model was proposed by Helbing on the basis of three states: vehicle density, mean speed, and mean speed dispersion. By introducing the variance of speeds as an explicit state variable, third-order models capture the dispersion in driver behaviour that is largely averaged out at second order. For congestion assessment, this means that third-order models can begin to quantify the contribution of behavioural heterogeneity to the onset of congestion — a contribution that is particularly important in Indian mixed traffic, where the variance in speeds across motorised two-wheelers, cars, autos, buses, trucks and non-motorised vehicles is structurally large and where pedestrian and cyclist interactions amplify this variance further.

4.4 Three-Phase Theory

Kerner's three-phase theory (Kerner, 2004) represents a conceptual departure from the equilibrium fundamental-diagram view. The theory distinguishes three traffic phases — free flow, synchronised flow, and wide moving jams — and explicitly accounts for the empirical observation that congested traffic can be locally stable in a synchronised state without collapsing into a moving jam. Kerner identifies three patterns of synchronised flow: (1) localised synchronised flow, in which the downstream front is fixed at the bottleneck and the upstream front oscillates but the mean width of the pattern does not change; (2) widening synchronised flow, in which the downstream front is fixed at the bottleneck while the upstream front continuously propagates backward; and (3) moving synchronised flow, in which the whole pattern propagates but cannot penetrate the next adjacent bottleneck — a phenomenon termed the catch effect. For congestion assessment, three-phase theory contributes the analytical machinery to detect and classify these patterns from speed-density data, distinguish between the truly disruptive wide moving jams and the comparatively benign synchronised flow states, and explain hysteresis and metastability. These contributions are especially relevant to bottleneck management and to the design of variable speed limits and ramp metering strategies.

4.5 Microscopic Models and Their Role

Microscopic car-following and lane-changing models complement the macroscopic family by simulating individual vehicle behaviour. For congestion assessment, they are particularly useful in quantifying the contribution of specific behavioural mechanisms (such as cooperative yielding, aggressive lane changes, or buffer distances) to the onset and propagation of congestion. In heterogeneous traffic, they also allow the explicit representation of vehicle class differences and side-friction interactions that macroscopic models can only approximate. Table 2 summarises the key contributions of each model family to congestion assessment.

Table 2: Traffic Flow Model Families and Their Contribution to Congestion Assessment

Model Family	Representative Authors	Contribution to Congestion Assessment	Key Limitation
First-order (LWR)	Lighthill and Whitham; Richards	Captures kinematic waves and propagation of shock fronts; supports deterministic queue analysis at bottlenecks.	Cannot describe spontaneous breakdown, capacity drop, stop-and-go oscillations, or growing-amplitude patterns.
Second-order	Payne; Ross; Kerner and Konhauser; Mohan (2017); Zhou (2021)	Adds speed dynamics; reproduces oscillations, hysteresis, gradual speed change; extended to area occupancy for heterogeneous traffic and Lagrangian formulation for shock tracking.	Cannot model instantaneous driver reaction, inertial effects, or context-driven driver behaviour fully.
Third-order	Helbing	Adds mean speed dispersion as a state variable; quantifies the contribution of behavioural heterogeneity to congestion onset.	Higher computational and calibration burden; data requirements for variance are demanding.
Three-phase theory	Kerner (2004)	Detects and classifies free flow, synchronised flow (localised, widening, moving) and wide moving	Requires high-resolution detector data for phase

Model Family	Representative Authors	Contribution to Congestion Assessment	Key Limitation
		jams; explains the catch effect and hysteresis.	identification; calibration to mixed traffic is largely missing.
Microscopic car-following / lane-changing	Various	Quantifies contribution of individual behaviour (yielding, lane changes, buffers) to congestion; supports explicit class heterogeneity.	Computational cost; behavioural parameter calibration is data intensive.

5. Congestion Indices

Congestion indices translate the physical states of the network into scalar measures that can be tracked, compared across corridors, and used to set thresholds for intervention. The indices reported in the literature can be grouped into three families based on the primary input variable: travel-time based, speed based, and Level-of-Service based.

5.1 Travel-Time Based Indices

Travel-time based indices use observed travel times relative to a free-flow baseline. The most basic measure is Delay, defined by Lomax et al. (1997) as the difference between free-flow travel time and average travel time. The Travel Time Index (TTI), introduced by Lomax and Schrank (2005), is the ratio of peak-period travel time to free-flow travel time. The Planning Time Index, described by Karuppanagounder and Muneera (2017), is the ratio between the 95th percentile travel time and the free-flow travel time, capturing the worst-case experience that travellers must plan for. The Congestion Index (CI), in its travel-time form (Karuppanagounder and Muneera, 2017), is the ratio of delay to free-flow travel time. The Buffer Time Index (BTI), introduced by Nakat et al. (2014), is the ratio of the 95th percentile additional travel time to average travel time and is widely used to quantify reliability.

5.2 Speed-Based Indices

Speed-based indices use observed speeds rather than travel times. The speed-based Congestion Index, due to Stipancic et al. (2016), is the ratio of the differential between free-flow speed and average speed to the free-flow speed. The Speed Reduction Index (SRI), proposed by Kukadapwar and Parbat (2015), is the ratio of the differential between non-peak-flow speed and peak-flow speed to the non-peak-flow speed. The Very-Low-Speed Index (VLSI), also from Kukadapwar and Parbat (2015), is the ratio of time spent in delay to the total time of the journey, applicable at link, corridor and route levels. The Corridor Mobility Index (CMI), introduced by Lomax (1990), is a composite speed-based corridor measure.

5.3 Level-of-Service Based Composite Indices

Level-of-Service based indices combine demand and supply information into composite measures. The Roadway Congestion Index (RCI), described by Kumar (2021), combines as a ratio the daily vehicle-mile travel per lane-mile for freeways and principal arterial street systems that compares the existing demand with the determined values in congested conditions. The Congestion Severity Index (CSI), proposed by Lindley (1987), was used as a measure of urban-area freeway delay calculated as delay-travel per million vehicle-kilometres based on 1985 Highway Capacity Manual calculations. The Lane Mile Duration Index

(LMDI), introduced by Cottrell (1991), was created as a measure of recurring freeway congestion in urban networks. Table 3 consolidates these indices.

Table 3: Congestion Indices: Type, Description, Element of Application, and Source

Type	Index	Description	Element	Source
Travel time	Delay	Difference between free-flow travel time and average travel time.	Link and corridor	Lomax et al. (1997)
Travel time	Planning Time Index	Ratio of 95th-percentile travel time to free-flow travel time.	Link and corridor	Karuppanagounder and Muneera (2017)
Travel time	Congestion Index (CI)	Ratio of delay to free-flow travel time.	Link and corridor	Karuppanagounder and Muneera (2017)
Travel time	Travel Time Index (TTI)	Ratio of peak-period travel time to free-flow travel time.	Link and corridor	Lomax and Schrank (2005)
Travel time	Buffer Time Index (BTI)	Ratio of 95th-percentile additional travel time to average travel time.	Link, corridor, route	Nakat et al. (2014)
Speed	Congestion Index (CI)	Ratio of differential between free-flow speed and average speed to free-flow speed.	Link and corridor	Stipancic et al. (2016)
Speed	Speed Reduction Index (SRI)	Ratio of differential between non-peak and peak flow speeds to non-peak flow speed.	Link and corridor	Kukadapwar and Parbat (2015)
Speed	Very-Low-Speed Index (VLSI)	Ratio of time spent in delay to total journey time.	Link, corridor, route	Kukadapwar and Parbat (2015)
Speed	Corridor Mobility Index (CMI)	Composite speed-based corridor mobility measure.	Link and corridor	Lomax (1990)
LOS	Roadway Congestion Index (RCI)	Ratio combining daily vehicle-mile travel per lane-mile against congested-condition values.	Link, corridor, primary network	Kumar (2021)
LOS	Congestion Severity Index (CSI)	Delay per million vehicle-kilometres based on 1985 HCM calculations.	Link, corridor, primary network	Lindley (1987)

Type	Index	Description	Element	Source
LOS	Lane Mile Duration Index (LMDI)	Measure of recurring freeway congestion in urban networks.	Network	Cottrell (1991)

6. Urban Congestion versus Rural (Highway) Congestion

Although the underlying phenomenon of congestion is the same in both settings, the drivers, patterns and consequences of congestion on urban roads differ markedly from those on rural highways. Urban roads typically carry predominantly mixed traffic that includes cars, buses, two-wheelers, bicycles, and pedestrians (Litman, 2020), while rural highways primarily carry motor vehicles — cars, trucks and buses. The causes of congestion in urban settings include high population density, frequent stops due to traffic signals, heterogeneous traffic, and pedestrian crossings, whereas rural highway congestion is more often associated with high traffic volumes, incidents (accidents and breakdowns), and discrete bottlenecks (Zhang and Zhao, 2019).

Congestion patterns also differ. Urban congestion typically recurs during peak hours and is often predictable based on the time of day, while rural highway congestion is more seasonal or sporadic and is often caused by incidents or roadwork, leading to sudden delays (Wang and Kockelman, 2013). The economic impact of urban congestion is felt through lost productivity, wasted fuel, and impacts on local businesses, while rural congestion is felt primarily through delays in freight transport and logistics (Schrank and Lomax, 2019). The environmental effects also differ in character: urban congestion increases emissions due to idling and frequent acceleration and deceleration, whereas rural highway congestion may produce higher per-vehicle emissions during episodes but is generally lower in cumulative impact than urban congestion (Ewing and Cervero, 2010).

Safety profiles diverge in important ways. Urban congestion is associated with a higher incidence of accidents due to mixed traffic and pedestrian interactions (Elvik, 2011), while rural highway congestion is associated with an increased risk of severe accidents, particularly during high-speed congestion build-up. Finally, the infrastructure challenges differ; urban networks are often characterised by inadequate infrastructure to handle high volumes of diverse traffic, while rural highways are typically designed for higher speeds but can become overwhelmed during peak demand periods (Vickrey, 1969). These distinctions are summarised in Table 4 and motivate the need for context-specific congestion assessment and management approaches.

Table 4: Comparative Characteristics of Urban Road Congestion and Rural (Highway) Congestion

Aspect	Urban Road Congestion	Rural Road (Highway) Congestion	Source
Traffic composition	Predominantly mixed traffic including cars, buses, bicycles and pedestrians.	Primarily motor vehicles — cars, trucks and buses.	Litman (2020)
Causes	High population density, frequent stops due to traffic signals,	High traffic volumes, incidents (accidents or breakdowns), and bottlenecks.	Zhang and Zhao (2019)

Aspect	Urban Road Congestion	Rural Road (Highway) Congestion	Source
	heterogeneous traffic, and pedestrian crossings.		
Congestion patterns	Recurring congestion during peak hours, often predictable based on time of day.	Mostly seasonal or sporadic, often caused by incidents or roadwork, leading to sudden delays.	Wang and Kockelman (2013)
Economic impact	High economic costs due to lost productivity and wasted fuel; affects local businesses.	Economic impact felt through delays in freight transport and logistics costs.	Schrank and Lomax (2019)
Environmental effects	Increased emissions due to idling and frequent acceleration and deceleration.	Higher emissions during congestion, but generally lower than urban totals.	Ewing and Cervero (2010)
Safety concerns	Higher incidence of accidents due to mixed traffic and pedestrian interactions.	Increased risk of severe accidents, particularly during high-speed congestion.	Elvik (2011)
Infrastructure	Often characterised by inadequate infrastructure to handle high volumes of diverse traffic.	Typically designed for higher speeds but can become overwhelmed during peak times.	Vickrey (1969)

7. Typology of Recurring Congestion and Its Causes

Recurring congestion does not arise from a single cause. The World Bank (2014), INRIX (2023) and FHWA literature converge on a typology that links specific physical or behavioural causes to characteristic shock-wave behaviour and quantifiable reductions in effective capacity. Eight broad categories of recurring congestion can be identified.

First, side friction; interaction with parked vehicles, pedestrians or cyclists is fundamentally an activity-and-heterogeneity phenomenon. It can create stop-and-go conditions that lead to shock waves as vehicles slow down, and it reduces effective capacity through slower operating speeds. Second, construction zones are a geometric cause: roadwork reduces lane availability and alters traffic flow, with sudden lane reductions generating shock waves as drivers react to merge requirements, and a corresponding decrease in effective capacity due to the loss of lanes. Third, parking areas constitute an activity-driven cause: vehicles slowing to park or exit disrupt the flow, creating localised shock waves as vehicles decelerate and accelerate, with a consequent reduction of effective capacity in the vicinity of the parking zones.

Fourth, bus stops are another activity-driven cause: buses stopping for passengers can block lanes, generating shock waves as following vehicles must brake suddenly, and reducing effective capacity during bus stop dwell periods. Fifth, the mix of slow-moving vehicles in faster traffic streams — a heterogeneity issue — means the presence of trucks or buses in fast-moving traffic can cause delays, and these slow-moving vehicles can trigger shock waves that affect overall flow, lowering effective capacity due to speed differentials. Sixth, bottlenecks are a classic geometric cause: points where road capacity is reduced (such as merges, lane drops or intersections) lead to significant traffic build-up and to shock waves as vehicles queue behind the constriction, producing a major reduction in effective capacity at the bottleneck. Seventh, so-called phantom traffic jams are linked to driving behaviour at densities above the critical density. Such slowdowns appear without any clear physical cause and are often the result of small driver behaviours; shock waves can form when sudden braking by one vehicle propagates backward through the platoon, reducing effective capacity unpredictably during such events. Eighth, access from side roads is a geometric cause: vehicles entering or exiting from side roads disrupt main traffic flow, with sudden merging or turning movements creating shock waves as drivers react quickly, and reducing effective capacity through these interruptions. Table 5 consolidates this typology.

Table 5: Typology of Recurring Congestion, Shock-Wave Nature and Effective Capacity Impact

Cause of Congestion	**Type**	**Description**	**Shock-Wave Nature**	**Effective Capacity Impact**
Side friction	Activity and heterogeneity	Interaction with parked vehicles, pedestrians or cyclists.	Can create stop-and-go conditions leading to shock waves as vehicles slow down.	Reduces effective capacity due to slower speeds.
Construction zones	Geometric	Roadwork that reduces lane availability and alters flow.	Sudden lane reductions generate shock waves as drivers react to merging.	Decreases effective capacity due to lane reductions.
Parking areas	Activity	Vehicles slowing to park or exit disrupt the flow.	Creates localised shock waves as vehicles decelerate and accelerate.	Lowers effective capacity near parking zones.
Bus stops	Activity	Buses stopping for passengers can block lanes.	Generates shock waves as following vehicles must brake suddenly.	Reduces effective capacity during bus stop dwell periods.
Mix of slow-moving vehicles	Heterogeneity	Presence of trucks or buses in fast-moving traffic causes delays.	Slow-moving vehicles can trigger shock waves affecting overall flow.	Lowers effective capacity due to speed differentials.
Bottlenecks	Geometric	Points where capacity is reduced (merges, intersections).	Can lead to significant traffic build-up and shock waves as vehicles queue.	Major reduction in effective capacity at the bottleneck.

Cause of Congestion	Type	Description	Shock-Wave Nature	Effective Capacity Impact
Phantom traffic jams	Driver behaviour at densities above critical density	Slowdowns without clear cause, due to driver behaviour.	Shock waves can form from sudden braking by one vehicle affecting others behind.	Effective capacity reduced unpredictably.
Access from side roads	Geometric	Vehicles entering or exiting from side roads disrupt main flow.	Sudden merging or turning movements create shock waves.	Reduces effective capacity due to interruptions in flow.

8. Network-Level Analysis: Centrality Measures and Resilience

Beyond link-level analysis, congestion is increasingly understood as a network phenomenon whose patterns and persistence are shaped by the topology of the underlying graph. Two complementary strands of network analysis are particularly relevant: centrality-based identification of structurally critical nodes, and resilience-based assessment of how the network behaves under recurring or non-recurring disruptions.

8.1 Centrality Measures and Congestion Assessment

Centrality measures from network science provide quantitative indicators of the structural importance of nodes and links in a road network. Several different centrality concepts are routinely applied to traffic networks. The Shimbel Index (also known as Shimbel distance, nodal accessibility, or nodality) is a measure of accessibility representing the sum of the length of all shortest paths connecting a given node to all other nodes in the graph; its inverse is also called closeness centrality or distance centrality. Betweenness centrality measures the extent to which a particular node lies between other nodes in a network: a node is more powerful if it lies on the shortest paths connecting many node pairs, as it may broker or mediate connections between those pairs. The betweenness of a node is defined as the ratio of all shortest paths passing through it and reflects its transitivity. Closeness centrality measures the extent to which a node is close to all other nodes along the shortest path and reflects its accessibility in a given network. Eigenvector centrality measures the centrality of a node based on the weighted sum of the centralities of its neighbours, and is helpful for identifying nodes that are connected to many other well-connected nodes.

Empirical studies have consistently found that centrality measures help predict and analyse traffic congestion patterns. Rahman and Yadav (2022), working with a case study of urban road networks, found that nodes with high betweenness centrality are critical for traffic flow and experience greater congestion. Zhang and Wang (2023) found that higher degree, betweenness and PageRank centrality values are associated with increased traffic volumes and congestion. Smith and Jones (2021) linked various centrality metrics to the structural features of urban road networks and found that higher centrality values are associated with greater congestion levels. Jayaweera et al. (2017), in a case study of the Kiribathgoda road network in Sri Lanka, identified critical nodes using betweenness and degree centrality and found that higher betweenness centrality correlated with more significant recurring congestion. The collective conclusion is that centrality measures, including betweenness, degree and PageRank, provide valuable insights into the structural features of urban road networks and can help predict and analyse congestion patterns, supporting targeted traffic management. These findings are summarised in Table 6.

Table 6: Centrality Measures and Congestion Assessment: Selected Studies

Centrality Measure(s)	Key Finding	Source
Degree centrality, betweenness centrality	Nodes with high betweenness centrality are critical for traffic flow and experience greater congestion.	Rahman and Yadav (2022)
Degree, betweenness, PageRank	Higher degree, betweenness, and PageRank centrality values are associated with increased traffic volumes and congestion.	Zhang and Wang (2023)
Various centrality metrics	Centrality metrics linked to structural features of urban road networks; higher centrality values associated with greater congestion levels.	Smith and Jones (2021)
Betweenness centrality, degree centrality	Critical nodes identified in the Kiribathgoda road network; higher betweenness centrality correlated with more significant recurring congestion. Centrality measures effectively identify locations prone to recurring congestion.	Jayaweera et al. (2017)

8.2 Resilience of Urban Road Networks

Resilience extends the static centrality view by capturing the dynamic response of the network to disturbances. Chalkiadakis, Perdikouris and Vlahogianni (2022) measured resilience based on the efficiency and criticality of a link, examining the impact of link removal on network performance, with efficiency assessed through shortest-path analysis and criticality evaluated based on transportation variables such as travel demand. Their iterative approach removes one link per iteration and analyses the resulting changes in network performance. Tang and Heinimann (2022) characterised resilience as the ability of a road network to cope with disturbances and recover functionality, particularly in the context of recurrent congestion; their resilience metric, inspired by the R4 resilience-triangle framework, quantifies recurrent congestion based on spatial-temporal traffic patterns and includes dimensions that capture congestion intensity and discharging processes. Gorji, Akbarzadeh and Shetab-Boushehri (2022) defined resilience as the capacity of urban road networks to withstand short-term non-recurring traffic congestion and used graph connectivity-based criteria to evaluate robustness against disruptions caused by crashes. Zhang and Wang (2021) defined resilience as the ability to maintain functionality despite disruptions, focusing on performance loss during disturbances and using graph theory to quantify connectivity and to compare static and dynamic recovery strategies.

9. Congestion Management Practices

Congestion management practice draws on a portfolio of supply-side and demand-side levers that have been codified by transport agencies based on decades of operational experience. The US Federal Highway Administration (FHWA, 2020) catalogues eleven core practices that target different mechanisms of congestion generation and propagation. These are reproduced in Table 7 and discussed below.

Table 7: Eleven Congestion Management Practices Catalogued by the US FHWA (2020)

No.	Practice
1	Use of a short section of traffic-bearing shoulder as a peak-hour lane to add temporary capacity when demand is highest.
2	Re-striping merge or diverge areas to better serve the observed turning and through-movement demand.
3	Reducing lane widths to add a travel and/or auxiliary lane, accepting that operating speeds may have to be reduced to effect a safer condition past the narrowed lanes.
4	Modifying weaving areas, for example by adding a collector-distributor lane or similar geometric feature to separate weaving movements from the through stream.
5	Metering or closing entrance ramps to regulate the rate at which traffic enters the mainline.
6	Effecting speed harmonisation through a succession of overhead gantries spaced every kilometre or so upstream of problem areas to adjust (typically reduce) approach speeds and lane use, calming the shock waves that often result through congested corridors.
7	Effecting zippering — a self-metered or policed merge that promotes fair and smooth merges at points of traffic confluence; a motorist who is tenth in their lane knows they will be twentieth to merge with the adjacent lane, which helps eliminate line jumpers who would otherwise bull ahead and create potentially unfair and even violent disagreements.
8	Improving traffic signal timing on arterials, including coordination across adjacent intersections.
9	Using access management principles to clean up corridors by consolidating or relocating driveways and side-road access.
10	Providing traffic diversion information so that drivers can self-route around congestion in real time.
11	Implementing road pricing to bring demand into line with supply at congested times and locations.

Beyond the FHWA catalogue, several recent studies have evaluated the effect of signal coordination on congestion. Zhang and Wang (2024) developed a coordination control method using simulation-based analysis for oversaturated two-way arterials, in which a dynamic programming model optimised signal timings based on real-time traffic data; the results indicated a 30 per cent reduction in average delays and a 25 per cent improvement in vehicle throughput during peak hours under oversaturated conditions. Chen and Wu (2023) carried out a comparative analysis of conflict rates at signalised intersections before and after the implementation of coordinated signals and found a 40 per cent reduction in conflict rates under unsaturated conditions, with benefits being less pronounced (15 per cent) under oversaturated conditions. Liu et al. (2024) demonstrated a large-scale traffic signal optimisation system using vehicle trajectory data from connected vehicles at low penetration rates and achieved a 20 per cent reduction in delays and a 30 per cent decrease in stops at signalised intersections.

Wang et al. (2022) reviewed methodologies for evaluating traffic signal performance under both unsaturated and oversaturated conditions and reported that coordinated signals can improve performance metrics such as travel times (up to a 25 per cent reduction) and queue lengths (up to a 30 per cent decrease) in varied traffic conditions. Zhang et al. (2023) applied a multi-agent deep reinforcement learning approach to optimise network-wide traffic signals under both unsaturated and oversaturated conditions and reported that the deep reinforcement learning model improved network throughput by 35 per cent and reduced overall delays by 40 per cent compared to traditional fixed-time systems, although computational complexity and potential overfitting during simulations were noted as limitations.

10. Research Gaps

The literature reviewed in the preceding sections converges on a coherent picture of congestion as a propagating, network-level phenomenon shaped by demand, geometry, side friction, driver behaviour, and control. However, several substantial gaps remain, particularly when the existing body of knowledge is applied to Indian mixed-traffic conditions. These gaps are presented below in two groups, broadly covering measurement and modelling issues (Section 10.1) and side-friction, design and management issues (Section 10.2).

10.1 Measurement, Modelling and Network-Level Gaps

First, there is a clear need for congestion measurement methodologies and indicators that can categorise the type of congestion, identify its typical causes, and quantify its effect. Existing indices, as catalogued in Section 5, summarise severity or reliability, but they do not categorise the underlying type of congestion or directly capture its dynamics. Operational indicators such as early jam-growth speed (Duan, 2023), queue-front velocity, and wave-attenuation coefficients are not yet integrated into standard measurement frameworks, despite their potential to trigger interventions in real time.

Second, most existing assessments in the Indian scenario have been conducted on freeways, largely because data collection for arterial and urban roads has historically relied on techniques such as adaptive smoothing that are difficult to deploy beyond grade-separated facilities. As a result, studies on other road hierarchies, different geometries and varied land uses are limited. The parameter transfer from homogeneous freeway settings to lane-less, multi-class urban arterials risks systematic bias and requires explicit calibration of higher-order and propagation models for Indian conditions (Mohan, 2017; Biswas, 2021).

Third, confluence zones — merging, diverging, and intersections — need dedicated study in the Indian scenario to assess the effect of different geometric and operational treatments on congestion generation, propagation and dissipation along both the transverse and lateral directions. Phase transitions, yielding behaviour, and lateral wave propagation under different access spacing configurations are particularly under-studied.

Fourth, spatial metrics have been more frequently used to analyse the influence of urban morphology on outcomes related to congestion (such as air pollution, noise pollution, and carbon emissions) rather than on congestion itself. Wang and Debbage (2021) have begun to address this, but a more direct linkage between urban form, network topology, and congestion propagation remains underdeveloped, particularly for Indian cities with their distinctive land-use mixes and informal settlements.

Fifth, there is a clear need to study the nature of the shock wave itself in the Indian scenario — specifically, the change in velocity of the shock wave and the delineation of distinct generation, propagation and dissipation phases. While freeway shock-wave dynamics are well understood from classical theory, their counterpart in heterogeneous, lane-less arterial flow is poorly characterised. Evidence of synchronised flow and wide moving jams exists for freeways, but robust detection and classification protocols for mixed-traffic arterials and undivided roads remain to be developed.

10.2 Side Friction, Design and Management Gaps

Sixth, the capacity reduction due to side friction parameters has been assessed in detail (Biswas, 2021), but the corresponding effect on congestion propagation has not yet been assessed. Importantly, side friction parameters do not reduce the capacity of the carriageway through a traditional width-crunching mechanism; rather, they induce lane changes — forced or perceived — that propagate disturbances longitudinally and laterally.

Seventh, drivers' buffer to side-friction parameters varies between individuals and across vehicle classes, and these differential buffers can cause instantaneous impediments to the existing flow of traffic. The behavioural distribution of these buffers in Indian conditions has not been systematically measured, despite its central role in determining shock-wave initiation probability.

Eighth, there is a need to design congestion mitigation and prevention measures keyed to the type of congestion. The catalogue of management measures from FHWA (2020) and the signal-coordination evidence summarised in Section 9 are largely drawn from freeway and divided-arterial contexts. Their effectiveness and adaptation in the Indian mixed-traffic setting, where the dominant congestion types may differ from those in freeway-centric studies, require dedicated empirical evaluation.

Ninth, limited studies have addressed the variation in the angle of deformation of trajectories at different phases of congestion, and the effect of this angle on congestion propagation and dissipation. The angle of deformation captures the geometry of lane changes and yielding behaviour and could provide a leading indicator of phase transitions in mixed traffic.

Tenth, the aspect of viscosity values for heterogeneous traffic, side-friction parameters, and road geometries such as merging and diverging areas could be assessed by analogy with granular and particle-flow models. A viscosity-style parameter that summarises interaction intensity across vehicle classes and side-friction states could materially improve the prediction of stability margins and propagation speeds in heterogeneous streams.

Eleventh, the congestion burden — measured through excess fuel consumption, emissions and pollution — has typically been computed assuming free-flow conditions as the baseline. This is not the ideal reflection of urban driving and tends to overstate the marginal cost of incremental congestion, particularly in dense Indian cities where free-flow conditions are themselves rare. Further, the values of congestion burden for different types of congestion have not been disaggregated, although the emissions classification framework of Smit, Brown and Chan (2008) provides a starting point. A propagation-aware valuation of congestion externalities, in which the baseline reflects realistic operating conditions and the burden is decomposed by congestion type, remains an open research need. Table 8 consolidates the eleven gaps identified.

Table 8: Consolidated Research Gaps in Congestion Management for Mixed-Traffic Urban Networks

No.	Research Gap
1	Need for congestion measurement methodologies and indicators that categorise type, identify causes and quantify effect; propagation-based indices (early jam-growth speed, queue-front velocity, wave-attenuation coefficient) are under-developed.

2	Most Indian-scenario assessments are on freeways; limited studies cover other road hierarchies, geometries, and land uses, leaving higher-order models poorly calibrated for mixed-traffic arterials.
3	Merging, diverging and intersection confluence zones need dedicated study in the Indian scenario to assess effects on generation, propagation and dissipation along transverse and lateral directions.
4	Spatial morphology metrics have been linked to congestion-related outcomes (pollution, emissions) but not directly to congestion itself.
5	Need to study the nature of the shock wave itself in the Indian scenario, including velocity changes and delineation of generation, propagation and dissipation phases.
6	Capacity reduction due to side friction is documented; its effect on congestion propagation is not yet assessed. Side friction acts not by width crunching but by inducing forced or perceived lane changes.
7	Driver buffer to side-friction parameters varies and can cause instantaneous impediments to flow; behavioural distributions are unmeasured for Indian conditions.
8	Need to design congestion mitigation and prevention measures by type of congestion for the Indian scenario, rather than transplanting freeway-derived measures unchanged.
9	Limited studies on variation in angle of deformation at different phases of congestion and its effect on propagation and dissipation.
10	Viscosity-style values for heterogeneous traffic, side friction parameters, and merging/diverging geometries could be assessed by analogy with particle-flow models.
11	Congestion burden (fuel, emissions, pollution) has been computed using free-flow baseline, which is not realistic; values for different congestion types have not been disaggregated.

11. Conclusion

Congestion in mixed-traffic urban networks is a dynamic, spatio-temporal process driven by local interactions and shaped by geometry, side friction, and network topology. The foundational definitions and indices reviewed in this paper provide an essential vocabulary and baseline for measurement, but the phenomena that most directly affect users — shock-wave initiation, growth and diffusion — demand propagation-aware metrics and models that go beyond traditional Level-of-Service categories or scalar travel-time ratios.

Across the body of literature reviewed, several convergent conclusions emerge. First, the four definitional categories — demand-and-capacity based, vehicular-interaction based, delay based and cost based — are complementary rather than competitive and together motivate multi-dimensional measurement. Second, the evolution from first-order LWR through second-order and third-order continuum models to Kerner's three-phase theory has progressively expanded the set of congestion phenomena that can be reproduced and therefore assessed, with each successive family adding important capabilities relevant to heterogeneous traffic. Third, the catalogue of congestion indices is rich on severity and reliability but thin on dynamics, and propagation-based indicators remain a frontier for research. Fourth, the typology of recurring congestion clarifies the mechanisms by which specific causes (side friction, bottlenecks, geometric reductions, phantom jams) translate into shock-wave activity and effective-capacity loss. Fifth, network-level analysis using centrality measures and resilience metrics adds an indispensable structural perspective: structurally critical nodes (identified by betweenness, degree or PageRank) are also empirically more congested, and resilience-based metrics quantify how the network behaves under disturbance. Sixth, the FHWA catalogue of management practices, supplemented by recent advances in signal coordination and deep reinforcement learning, provides a tested portfolio of interventions whose efficacy is increasingly well documented in freeway and divided-arterial contexts.

For Indian cities, the most pressing research priorities are the development of propagation-aware indices and shock-wave diagnostics calibrated to lane-less, multi-class flows; the systematic quantification of side-friction effects on propagation rather than only on capacity; the construction of viscosity-style parameters for heterogeneous streams; the integration of microscopic lane-change turbulence with network-level diffusion and resilience metrics in a single framework; and the realistic valuation of congestion burden using non-free-flow baselines that reflect the operating reality of dense Indian cities. Addressing these gaps will provide the analytical foundation for practical early-warning systems and for targeted, type-specific congestion management interventions that reduce delay, variability, and environmental burden while improving network resilience.

References

1. Akbar, P., Couture, V., Duranton, G. and Storeygard, A. (2021) 'The fast, the slow, and the congested: Urban transportation in rich and poor countries', Working Paper, National Bureau of Economic Research.
2. Biswas, S. (2021) 'Effect of side friction parameters on capacity of urban undivided road sections under mixed traffic conditions', Transportation Research Procedia, 48, pp. 3399-3414.
3. Bovy, P.H.L. and Salomon, I. (2002) 'Congestion in Europe: Measurements, patterns and policies', in Stern, E., Salomon, I. and Bovy, P.H.L. (eds.) Travel Behaviour: Spatial Patterns, Congestion and Modelling. Cheltenham: Edward Elgar.
4. Cambridge Systematics and Texas Transportation Institute (2005) Traffic Congestion and Reliability: Trends and Advanced Strategies for Congestion Mitigation. Washington, D.C.: Federal Highway Administration.
5. Chalkiadakis, C., Perdikouris, A. and Vlahogianni, E.I. (2022) 'Investigating the resilience of urban road networks based on link removal and efficiency-criticality criteria', Transportation Research Procedia, 60, pp. 178-185.
6. Chen, Y. and Wu, Y. (2023) 'Effects of traffic signal coordination on the safety performance of urban arterials', Accident Analysis and Prevention, 182, 106953.
7. Cottrell, W.D. (1991) Measurement of the Extent and Duration of Freeway Congestion in Urbanized Areas. ITE Compendium of Technical Papers. Washington, D.C.: Institute of Transportation Engineers.
8. Daganzo, C.F. (2002) 'A behavioral theory of multi-lane traffic flow. Part II: Merges and the onset of congestion', Transportation Research Part B: Methodological, 36(2), pp. 159-169.
9. Downs, A. (2004) Still Stuck in Traffic: Coping with Peak-Hour Traffic Congestion. Washington, D.C.: Brookings Institution Press.

10. Duan, J. (2023) 'Spatiotemporal dynamics of traffic bottlenecks yields an early signal of heavy congestions', Nature Communications, 14, 8500.
11. Elvik, R. (2011) 'Assessing the validity of road safety evaluation studies by analysing causal chains', Accident Analysis and Prevention, 43(3), pp. 753-762.
12. Erdelic, T., Caric, T., Erdelic, M., Tisljaric, L., Turkovic, A. and Jelusic, N. (2021) 'Estimating congestion zones and travel time indexes based on the floating car data', Computers, Environment and Urban Systems, 87, 101604.
13. European Conference of Ministers of Transport (ECMT) (1999) Traffic Congestion in Europe: Round Table 110. Paris: ECMT.
14. European Conference of Ministers of Transport (ECMT) (2004) Managing Urban Traffic Congestion. Paris: ECMT / Organisation for Economic Co-operation and Development (OECD).
15. Ewing, R. and Cervero, R. (2010) 'Travel and the built environment: A meta-analysis', Journal of the American Planning Association, 76(3), pp. 265-294.
16. Falcocchio, J.C. and Levinson, H.S. (2015) Road Traffic Congestion: A Concise Guide. Cham: Springer.
17. Federal Highway Administration (FHWA) (2020) Traffic Congestion and Reliability: Linking Solutions to Problems. Washington, D.C.: U.S. Department of Transportation.
18. Gorji, M.-A., Akbarzadeh, M. and Shetab-Boushehri, S.-N. (2022) 'Resilience of urban road networks against short-term non-recurring traffic congestion based on graph connectivity criteria', Transportation Research Part D: Transport and Environment, 110, 103416.
19. Indian Roads Congress (2019) IRC SP 30: Manual on Economic Evaluation of Highway Projects in India. New Delhi: Indian Roads Congress.
20. INRIX (2010) INRIX National Traffic Scorecard - 2010 Annual Report. Kirkland, WA: INRIX.
21. INRIX (2023) Causes of Congestion - Part 1: What is congestion and what influences it? Kirkland, WA: INRIX.
22. Institute of Civil Engineers (1989) Congestion: A Report by the Institution of Civil Engineers. London: Thomas Telford.
23. Institute of Transportation Engineers (ITE) (1996) Toolbox for Alleviating Traffic Congestion and Enhancing Mobility. Washington, D.C.: Institute of Transportation Engineers.
24. Jayaweera, I.M.L.N., Perera, K.K.K.R. and Munasinghe, J. (2017) 'Centrality measures to identify traffic congestion on road networks: A case study of Sri Lanka', IOSR Journal of Mathematics, 13(2), pp. 13-19.
25. Karuppanagounder, K. and Muneera, C.P. (2017) 'A study of the influencing factors and modelling the congestion cost in urban areas', International Journal of Engineering and Technology, 9(3), pp. 2306-2316.
26. Kerner, B.S. (2004) The Physics of Traffic: Empirical Freeway Pattern Features, Engineering Applications, and Theory. Berlin: Springer.
27. Kerner, B.S. and Konhauser, P. (1994) 'Structure and parameters of clusters in traffic flow', Physical Review E, 50(1), pp. 54-83.
28. Kockelman, K.M. (2004) 'Traffic congestion', in Kutz, M. (ed.) Handbook of Transportation Engineering. New York: McGraw-Hill.
29. Kukadapwar, S.R. and Parbat, D.K. (2015) 'Modeling of traffic congestion on urban road network using fuzzy inference system', American Journal of Engineering Research, 4(12), pp. 143-148.

30. Kumar, S.V. (2021) 'A composite congestion index for urban arterials', Transportation Research Procedia, 53, pp. 79-86.
31. Kumar, S.V. and Sivanandan, R. (2012) 'Traffic congestion quantification for urban heterogeneous conditions using a thresholds-based approach', International Journal of Engineering and Technology, 4(5), pp. 540-548.
32. Lindley, J.A. (1987) 'Urban freeway congestion: Quantification of the problem and effectiveness of potential solutions', ITE Journal, 57(1), pp. 27-32.
33. Litman, T. (2020) Transportation and Environmental Policy. Victoria: Victoria Transport Policy Institute.
34. Liu, J., Zhao, M. and Wang, X. (2024) 'Traffic light optimization with low penetration rate vehicle trajectory data', Transportation Research Part C: Emerging Technologies, 158, 104432.
35. Lockwood, S. (2006) Recurring Traffic Bottlenecks: A Primer. Washington, D.C.: Federal Highway Administration.
36. Lomax, T. (1990) Estimating Transportation Corridor Mobility. College Station: Texas Transportation Institute.
37. Lomax, T., Turner, S., Shunk, G., Levinson, H.S., Pratt, R.H., Bay, P.N. and Douglas, G.B. (1997) Quantifying Congestion: Volume 1, Final Report. NCHRP Report 398. Washington, D.C.: Transportation Research Board.
38. Lomax, T. and Schrank, D. (2005) The 2005 Urban Mobility Report. College Station: Texas Transportation Institute.
39. Miller, M.A. and Li, K. (1994) An Investigation of the Costs of Roadway Traffic Congestion: A Preparatory Step in IVHS Benefits Estimation. PATH Research Report UCB-ITS-PRR-94-15. Berkeley: Institute of Transportation Studies.
40. Mohan, R. (2017) Heterogeneous traffic flow modelling using second-order macroscopic continuum model. PhD Thesis. Indian Institute of Technology Madras.
41. Munoz, J.C. and Daganzo, C.F. (2002) 'The bottleneck mechanism for a freeway diverge', Transportation Research Part A: Policy and Practice, 36(6), pp. 483-505.
42. Nagel, K. and Nelson, P. (2005) 'A critical comparison of the kinematic-wave model with observational data', Proceedings of the 16th International Symposium on Transportation and Traffic Theory. Maryland, USA, July 2005.
43. Nagy, A.M. and Simon, V. (2021) 'Traffic congestion propagation identification method in smart cities', International Journal of Computer Networks and Applications, 8(5), pp. 487-501.
44. Nakat, Z., Madi, M. and Maoui, I. (2014) 'Traffic congestion measurement: A review', Procedia - Social and Behavioral Sciences, 138, pp. 437-446.
45. National Cooperative Highway Research Program (NCHRP) (1998) NCHRP Report 398: Quantifying Congestion. Washington, D.C.: Transportation Research Board.
46. Rahman, M.M. and Yadav, A. (2022) 'Centrality measures and traffic congestion: A case study of urban road networks', Journal of Urban Planning, 48(3), pp. 215-230.
47. Regional Integrated Transportation Information System (RITIS). Bottleneck Ranking. Center for Advanced Transportation Technology, University of Maryland. Available at: https://ritis.org (Accessed: 27 May 2026).
48. Rosenbloom, S. (1978) 'Peak-period traffic congestion: A state-of-the-art analysis and evaluation of effective solutions', Transportation, 7(2), pp. 167-191.

49. Rothenberg, M.J. (1985) 'Urban congestion in the United States: What does the future hold?', ITE Journal, 55(7), pp. 22-39.
50. Schonhof, M. and Helbing, D. (2007) 'Empirical features of congested traffic states and their implications for traffic modeling', Transportation Science, 41(2), pp. 135-166.
51. Schonhof, M. and Helbing, D. (2008) 'Criticism of three-phase traffic theory', Transportation Research Part B: Methodological, 43(7), pp. 784-797.
52. Schrank, D. and Lomax, T. (2019) 2019 Urban Mobility Report. College Station: Texas A&M Transportation Institute.
53. Sen Luan, R., Ke, R., Huang, Z. and Ma, X. (2021) 'Traffic congestion propagation inference using dynamic Bayesian graph convolution network', Transportation Research Part C: Emerging Technologies, 128, 103193.
54. Smit, R., Brown, A.L. and Chan, Y.C. (2008) 'Do air pollution emissions and fuel consumption models for roadways include the effects of congestion in the roadway traffic flow?', Environmental Modelling and Software, 23(10-11), pp. 1262-1270.
55. Smith, R.A. and Jones, T. (2021) 'The role of centrality in urban traffic management', International Journal of Traffic and Transportation Engineering, 9(2), pp. 75-88.
56. Sole-Ribalta, A., Gomez, S. and Arenas, A. (2016) 'A model to identify urban traffic congestion hotspots in complex networks', Royal Society Open Science, 3(10), 160098.
57. Stipancic, J., Miranda-Moreno, L. and Saunier, N. (2016) 'Estimating congestion with a smartphone trip-based approach', Transportation Research Procedia, 19, pp. 65-78.
58. Tang, J. and Heinimann, H.R. (2022) 'A resilience-oriented approach for quantitatively assessing recurrent spatial-temporal congestion on urban roads', PLOS ONE, 13(1), e0190616.
59. Vickrey, W.S. (1969) 'Congestion theory and transport investment', American Economic Review, 59(2), pp. 251-260.
60. Victoria Transport Policy Institute (VTPI) (2005) Transportation Cost and Benefit Analysis. Victoria: VTPI.
61. Vuchic, V.R. and Kikuchi, S. (1994) 'The bus transit system: Its underutilized potential', Transportation Quarterly, 48(1), pp. 75-89.
62. Wang, M. and Debbage, N. (2021) 'Urban morphology and traffic congestion: Longitudinal evidence from US cities', Computers, Environment and Urban Systems, 89, 101676.
63. Wang, T., Cao, J. and Hussain, A. (2022) 'Adaptive traffic signal control for large-scale scenario with cooperative group-based multi-agent reinforcement learning', Transportation Research Part C: Emerging Technologies, 125, 103046.
64. Wang, Y. and Kockelman, K. (2013) 'Analyzing traffic congestion patterns: Use of spatially aggregated, time-varying variables', Transportation Research Record, 2387, pp. 19-29.
65. Weisbrod, G., Vary, D. and Treyz, G. (2001) Economic Implications of Congestion. NCHRP Report 463. Washington, D.C.: Transportation Research Board.
66. World Bank (2014) Identification of Causes, Types and Locations of Traffic Congestion. Washington, D.C.: World Bank.
67. Zhang, J., Liu, P. and Cao, S. (2023) 'Network-wide traffic signal control optimization using multi-agent deep reinforcement learning', IEEE Transactions on Intelligent Transportation Systems, 24(5), pp. 4856-4867.

68. Zhang, L. and Levinson, D. (2010) 'Ramp metering and freeway bottleneck capacity', Transportation Research Part A: Policy and Practice, 44(4), pp. 218-235.
69. Zhang, L. and Wang, H. (2024) 'Traffic signal coordination control along oversaturated two-way arterials', Journal of Transportation Engineering, Part A: Systems, 150(4), 04024012.
70. Zhang, L. and Wang, J. (2021) 'Resilience assessment of urban road networks under recurrent congestion', Transportation Research Part D: Transport and Environment, 95, 102855.
71. Zhang, L. and Wang, J. (2023) 'Analyzing urban traffic flow with centrality measures', Transportation Research Part A: Policy and Practice, 158, pp. 123-134.
72. Zhang, L. and Zhao, Y. (2019) Traffic Congestion: Causes and Solutions. Beijing: Science Press.
73. Zhou, Z. (2021) 'A second-order Lagrangian macroscopic traffic flow model for freeways', Transportation Research Part B: Methodological, 145, pp. 124-149.