Scientific Reports

https://doi.org/10.1038/s41598-026-58442-w

Article in Press

A multiscale interpretability framework for identifying actionable road network features to mitigate congestion in highly congested cities

Received: 11 February 2026

Accepted: 15 June 2026

Published online: 26 June 2026

Cite this article as: Kumar N., Zhang Y., Wiedemann N. *et al.* A multiscale interpretability framework for identifying actionable road network features to mitigate congestion in highly congested cities. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-58442-w

Nishant Kumar, Yatao Zhang, Nina Wiedemann, Jimi Oke & Martin Raubal

A multiscale interpretability framework for identifying actionable road network features to mitigate congestion in highly congested cities

Nishant Kumar[1], **Yatao Zhang**[2], **Nina Wiedemann**[3], **Jimi Oke**[4], **and Martin Raubal**[3,*]

[1]National University of Singapore (NUS), Singapore
[2]University College London (UCL), London, United Kingdom
[3]Institute of Cartography and Geoinformation, ETH Zurich, Zurich, Switzerland
[4]University of Massachusetts Amherst, Amherst, MA, USA
[*]mraubal@ethz.ch

ABSTRACT

Transferring congestion mitigation strategies across cities remains challenging due to two compounding issues: road network features affect congestion differently in different cities, and their influence may vary across spatial scales. We propose a multiscale framework to identify "actionable" road network features, defined as those maintaining a consistent directional association with traffic congestion across varying spatial resolutions. Through systematic investigation of seven highly congested cities worldwide, our findings reveal strong city-specific signatures in feature importance, indicating supply-side planning must be tailored individually. However, the most consistently exacerbating features for demand-side consideration show notable similarities across cities. We emphasise that the directional associations identified reflect patterns in the training data and should not be interpreted as causal effects; the identified features are policy-relevant candidates subject to validation through counterfactual modelling. Feature attribution is model-dependent: results are derived from RF-based TreeSHAP and should be interpreted accordingly. In the short term, our results directly inform ongoing measures like congestion pricing and personalised route choice applications, making them more robust and publicly acceptable. Long-term, our insights encourage re-evaluating scepticism toward infrastructure planning and highlight opportunities for designing congestion-resistant future cities. Given the diverse city selection, common observations identified can likely be extrapolated to new cities when detailed analysis is constrained by data limitations. We provide key recommendations for researchers. First, data-driven congestion studies must explicitly interrogate microscopic spatial scales, because coarse-scale models can conceal scale-dependent reversals in feature influence that directly undermine policy robustness. Second, simulation-based counterfactual approaches should test road network features beyond known value ranges.

The relationship between road networks and urban traffic congestion is complex and dynamic, influenced by scale-dependence[1], socioeconomic factors[1, 2], and is not fully understood[3]. Moreover, cities worldwide vary significantly in driving behaviour, infrastructure constraints, and the degree of heterogeneity in traffic streams[4], resulting in significant variations in the road-network-congestion-relationship[5]. The complexity of the relationship and its variation across cities presents transportation authorities and urban planners worldwide with difficult choices in policymaking with respect to urban road networks. Policy transfer, though often considered an extremely efficient approach to policymaking[6, 7], has limited effectiveness for urban congestion mitigation policies and designing policies tailored to individual cities is required[8]. The explicit focus of this paper on seven highly congested cities[9] across several continents aims to determine the cross-city similarities and differences in the road-network-congestion relationships.

Congestion arises in the short term when road network capacity fails to meet travel demands. Over the long term, these networks shape regional demand by influencing the development of Points of Interest (POIs) within accessibility constraints[10]. Efforts to enhance road infrastructure are often seen as minimally effective in reducing congestion due to the high costs and the induced demand, whereby capacity increases ironically lead to more traffic[2, 4, 11]. While polycentric urban models, like the 15-minute city concept, which aims to reduce travel times, are being researched, these research directions also face challenges such as the risk of spatial segregation and amenity quality disparities[12]. As visualised in Figure 1, here we take a fundamentally different approach and argue that since less congested areas exist even within highly congested cities, a thorough city-specific, multiscale investigation that considers both recurrent congestion (RC) and non-recurrent congestion (NRC) can provide insights that are based on road-network features and have the potential to resonate well with policymakers and at the same time increase trust in existing solutions.

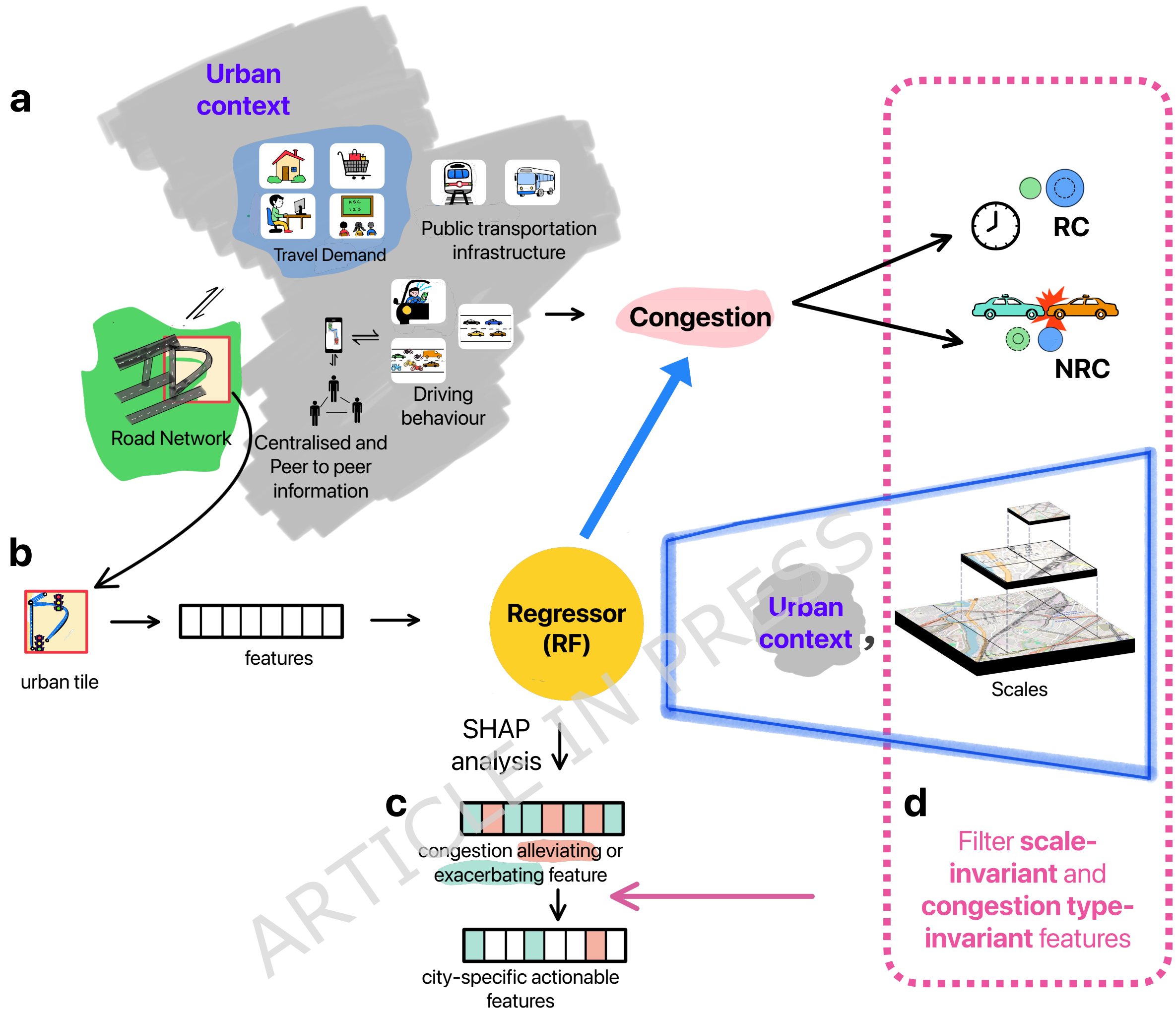


Figure 1. Illustration of our approach to extract the key actionable road network features for each city. **(a)** shows the several factors that influence the traffic in a city and result in the formation of peak hour recurrent congestion (RC) and crash-induced non-recurrent congestion (NRC). **(b)** shows a single urban tile from which the road network features are extracted. These features are used to predict the RC and NRC for each city using Random Forests (RF). The city-specific models account for the contextual differences. A SHAP-based feature importance analysis is performed, resulting in **(c)** where features are identified as associatively linked to higher or lower congestion levels. These features are candidate targets for further policy investigation. **(d)** illustrates the multiscale filtering process: features are assessed for sign consistency across all evaluated spatial resolutions (500 m to 4000 m) and both congestion types, and only those that are scale-consistent and congestion-type-consistent are retained as actionable. **Image sources**: The figure was created by the first author on an iPad using Freeform drawing app and includes only built-in or hand-sketched icons. The basemaps used in the three scales (panel **(d)**) were sourced from © OpenStreetMap contributors and plotted using contextily.

A multiscale problem warrants a multiscale solution

While spatio-temporal insights from geographical perspectives, including human geography, have been long known to be scale-dependent[13, 14], the scale-dependencies in the urban network spaces are particularly pronounced[15–17].

For instance, a traffic jam can almost always be tracked from being sourced at the microscopic scale (often a single event) to spreading to nearby locations and finally manifesting at macroscopic scales[4]. This multiscale nature of traffic jams calls for multi-resolution modelling techniques[18, 19]. It has resulted in the spatiotemporal dynamics of congestion propagation being modelled as epidemic spreading[20], percolation[21], and diffusion[22] processes and in the development of highly capable models that leverage the spatial dependencies at various scales to improve the model performances[4, 23]. From a policymaking perspective, the spatial scale of the analyses has profound implications for congestion-mitigation. On the one hand, if a road network feature is associated with high traffic congestion across various scales, it is the most likely candidate for targeted congestion alleviation. On the other hand, if a feature's direction of influence (on traffic congestion) varies at different scales, the choice becomes complex, and trade-offs are involved.

Targeting peak hour and incident-induced congestion simultaneously

RC typically occurs during peak hours due to increased demand and NRC forms due to reduced supply capacity, which might be caused by various *unusual* events such as inclement weather[24, 25], planned events and roadworks[26], with accidents being the most significant contributor[27]. Since RC and NRC have different formation mechanisms, the congestion mitigation approaches are also different. For instance, reducing RC might involve redirecting traffic, while reducing NRC might involve imposing lower speed limits. This paper identifies city-specific road network features that are robustly associated with both RC and NRC across all spatial scales, representing candidate targets for further policy investigation. The findings are intended to support data-driven discussions on supply-side strategies for urban traffic congestion management and to provide a transparent, reproducible evidence base that policymakers can draw upon alongside other considerations.

Key contributions and definitions

Actionable features and their implications: We define an 'actionable feature' as a road network feature that exhibits the same directional relationship (positive or negative) with both types of congestion (RC and NRC) across all spatial scales. We refrain from referring to these features as scale-invariant because even when the directional relationship is the same, the varying strength of influence across scales implies these features are not scale-invariant in the true sense. Among the actionable features, the top three are called 'highly actionable'. Features marked as 'highly actionable' imply that these features represent robust predictive associations and strong candidate features for congestion-mitigating consideration through supply or demand measures. Different actionable features present distinct opportunities for interventions while being subject to various constraints related to space and funding allocation. For example, exploring interventions such as reducing the number of lanes can be strategically achieved by reallocating lanes for bicycles, thereby addressing multiple urban mobility goals at once[28]. Similarly, modifying intersection typologies (e.g., traffic lights) represents a relatively low-cost intervention consideration, whereas altering road curvature (circuity) is nearly impossible once construction is complete. Our findings are also important for rapidly expanding suburbs like those around Mumbai that could benefit from supply-side planning by incorporating robust indicators from this study into early phases[29]. When modifying features through supply-side interventions is not feasible, demand-side measures can be considered. For instance, congestion tolling, currently under discussion in several cities studied here, can be tailored based on actionable features to efficiently manage traffic flow into specific urban areas. The identification of actionable features enables policy recommendations to be conveyed to policymakers in straightforward terms. For example, a simple statement such as "reducing the number of traffic lights is strongly associated with lower congestion in City B" can be derived based on our results.

Feature names: Throughout the manuscript, we use the shorthand # to represent the number and *avg* to represent the *average*. The feature names are summarised in Table 1.

Results

Highly congested cities show scale-dependent congestion complexity

Across 10 spatial scales and 7 cities, a total of 140 Random Forest (RF) models were trained—70 for RC and 70 for NRC. The goodness of fit (GoF) of our models, as measured by R^2 , varied across cities and showed a systematic drop at finer spatial scales and for NRC.

First, Figure 2**c** shows that the GoF varies significantly across cities and two types of congestion, thereby suggesting inherent differences in the complexities of road-network-congestion relationships. These variations strongly support city-specific models or an ensemble of models for multicity studies targeting congestion prediction to ensure effective dataset utilisation[30].

Table 1. 14 chosen road network features representing five feature groups

Feature Category	Details
Classic Graph Features	Number of nodes (#*nodes*), Number of edges (#*edges*)
Intersection-Count-Related	Total crossings (#*total crossings*), Traffic lights (#*traffic lights*), Free turns (#*free turns*)
Intersection-Type-Related	Average streets per node (*SPN-avg*), *SPN-n* refers to the number of intersections with n streets. Possible values are #*SPN–2*, #*SPN–3*, #*SPN–4*, and #*SPN–5*
Network Structure-Related	Circuity (*circuity*), Average number of lanes (#*avg lanes*)
Demand Attraction Related	Average degree (*average degree*), Local centrality (*local centrality*)

Table 2. Goodness-of-Fit (GoF) for Recurrent Congestion (RC) and Non-Recurrent Congestion (NRC) across cities.

City	GoF (R^2) RC	GoF (R^2) NRC
Auckland	0.43 ± 0.07	0.51 ± 0.05
Bogota	0.38 ± 0.05	0.50 ± 0.06
Capetown	0.39 ± 0.08	0.12 ± 0.05
Istanbul	0.32 ± 0.07	0.32 ± 0.06
Mexico City	0.49 ± 0.06	0.28 ± 0.05
Mumbai	0.03 ± 0.16	0.06 ± 0.30
New York City	0.52 ± 0.07	0.33 ± 0.04
Mean across all cities	0.36 ± 0.17	0.30 ± 0.20

Second, Figure 2**c** shows that the overall GoF for RC is better than that of NRC, thereby suggesting that RC is more amenable to road-network-based interventions. This observation can be attributed to the inherent unpredictable nature of non-recurrent congestion. Specifically, traffic accidents, which are the cause behind most instances of non-recurrent congestion, are often due to the human factor and, hence, inherently difficult to model. City-specific values show a similar trend, with some cities experiencing a significant drop in R^2 between RC and NRC. Specifically, for Cape Town, Mexico City, and New York City, the mean R^2 across all scales drops to almost half of that for RC. Cities that exhibit a significant drop in R^2 between RC and NRC are less likely to be amenable to road-network-based interventions to reduce NRC, and policymaking should be targeted at reducing RC instead. Several data-driven studies have shown that when RC is reduced to the levels of maximum traffic flow, it also reduces the rate of accidents, and hence NRC[4].

Third, Figure 2**c** and Table 2 show that the GoF decreases for finer scales for both types of congestion, implying an increase in the complexity of the road-network-congestion relationship at these finer scales. Such a drop in model fitness at finer scales suggests that extreme care is needed when inferring congestion-related insights from machine learning models at these finer spatial scales, especially below $1\mathrm{km}^2$. This trend has considerable implications for urban studies due to the widespread adoption of the $1\mathrm{km}^2$scale in urban studies, which can be primarily attributed to the resolution of geospatial data available through satellite imagery[31–33]. However, this popularity of the $1\mathrm{km}^2$scale presents a significant challenge to understanding the mechanisms of urban traffic congestion at microscopic levels. Hence, enhancing the policymakers' confidence in actionable insights arising from congestion-predicting models necessitates more research in understanding congestion mechanisms at these microscopic spatial scales.

City-specific policymaking is the key to effective congestion mitigation

Highly congested cities exhibit unique feature importance (FI) signatures, transcending typical socioeconomic groupings such as the Global North and South. Feature importances are computed with SHAP values (cf. Methods Figure). The fingerprint-like city-specific FI value distribution in Figure 2**a** and **b** raise concerns about the conventional approach where contextual similarities within SES groups have been leveraged to recommend similar policy interventions[34] and indicates that policy transfer between cities is more complex than previously understood. The FI values visualised in Figure 2**a** and **b** shed light on the underlying mechanisms of congestion formation in these highly congested cities.

First, we observe significant differences in FI values for RC and NRC. For RC (Figure 2**a**), the features with the top three total FI across all scales and cities are: the number of traffic lights (*#traffic lights*), the total number of intersections (*#total crossings*), and the number of nodes with four streets (*#SPN-4*, where SPN stands for Streets Per Node). The high FI of these features corroborates the recognised role of intersections, especially traffic lights, as sources of bottlenecks in recurrent congestion (RC)[17]. The findings underscore the importance of continued research into traffic control strategies aimed at reducing stop-time at intersections through adaptive Traffic Signal Controls (TSC). Additionally, some cities show a bimodal grouping of FI values, with high FI values found for intersection-related features and less FI for others. The cities which exhibit a bimodal distribution of FI, such as Mexico City, New York City, Bogota, and Mumbai, might show increased amenability to interventions on intersection-related features. For NRC (Figure 2**b**), while *#total crossings* remains significant in some cities, the FI for *#traffic lights* and *#SPN-4* diminishes, shifting the importance to features such as *avg lanes*, *local centrality*, and *avg degree*. The heightened FI for centrality-related and degree-related features, alongside *avg lanes*, underscores the dominance of travel demand factors over the supply constraints typical of recurrent congestion (RC). Specifically, since *avg degree* and *centrality*, known to be typically associated with higher attractions for travel demand, suggest that increased demand leads to unexpectedly high traffic volumes in certain areas, thereby increasing the probability of potential conflicts. This implies that mitigation measures should prioritise demand-side solutions for NRC. Similarly, the high FI of *avg lanes* in cities such as Bogota, Istanbul, and New York City suggest that effective lane reduction measures, such as allocating lanes to other modes[28] could effectively mitigate NRC for these cities.

Second, Figure 2 shows that the overall FI values are higher for RC than NRC, as is evident by the different ranges of values in the colour bars of Figure 2**a** and Figure 2**b**; maximum FI of 0.314 for RC versus 0.038 for NRC. The high FI values of RC compared to NRC suggest a stronger predictive association between the road network and RC. Specifically, since the intersection-related features (*e.g. #total crossings*, *#traffic lights*, and *#SPN-4*) have high FI values across most cities and scales, these robustly identify supply configurations as primary candidate features for recurrent congestion policies. While the high penetration of suggestive route choice applications is extremely effective, these findings suggest that by incorporating the intersection counts into their route choice sets, such applications can reduce the overall congestion in the city. Moreover, these findings lend support to the ongoing research in traffic signal control (TSC), wherein the goal is to search for intelligent ways to make the traffic light adaptive to the congestion levels and, thus, manage the traffic flow.

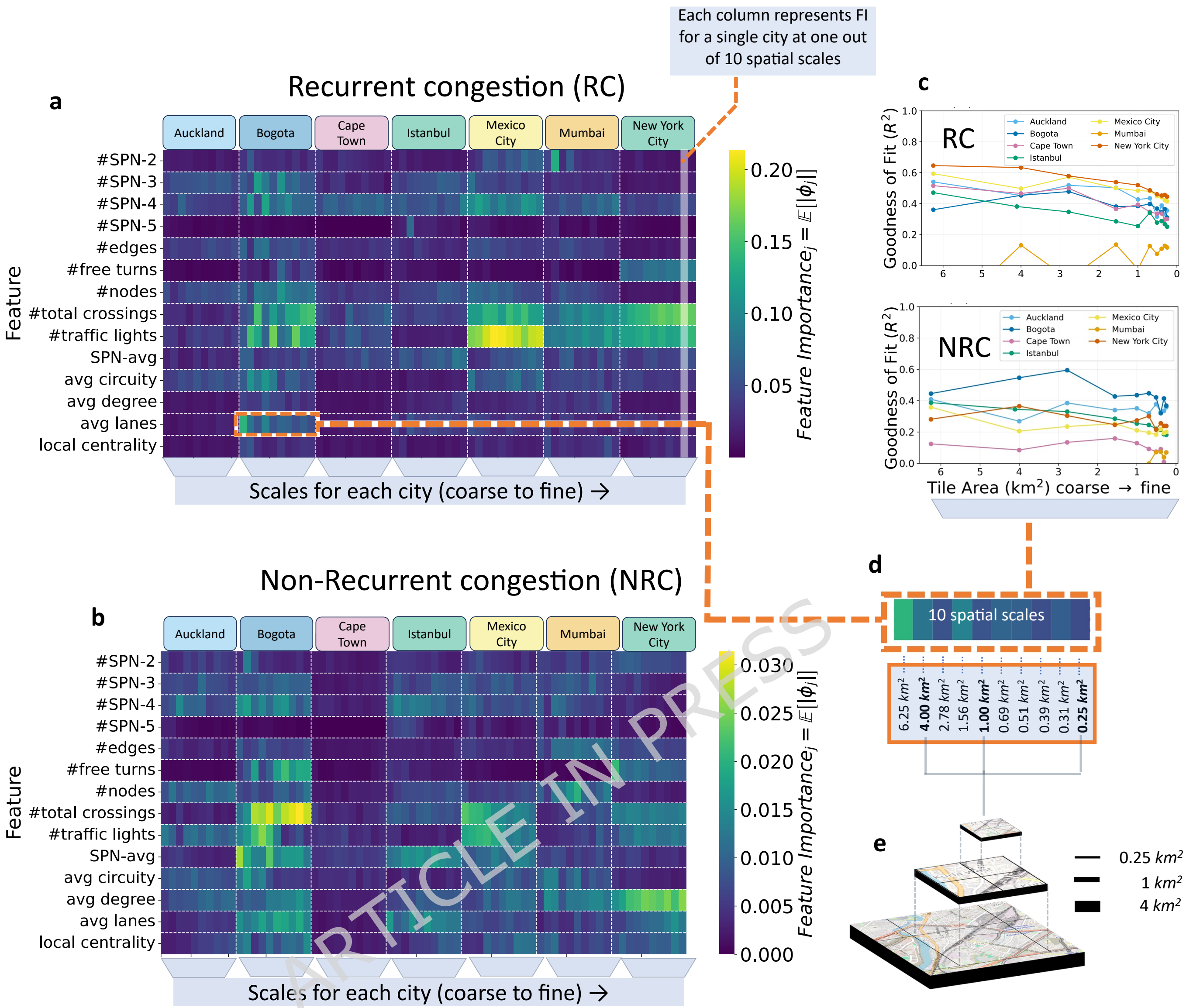


Figure 2. Heatmap showing the feature importance (FI) for each of the 14 features across 7 cities and 10 spatial scales for **(a)** RC and **(b)** NRC. Each column of this Heatmap is the FI value using one RF model. **(c)** GoF across all scales and cities. The low GoF for Mumbai is due to the small number of tiles. **(d)** The scales defined using tile areas as varying from *fine* (0.25km^2) to *coarse* (6.25km^2). **(e)** Urban tiles corresponding to three scales that fit perfectly- 4km^2, 1km^2, and 0.25km^{2}provide a computationally cheap way to evaluate a multiscale feature vector for any part of the city. (The basemaps used in the three scales (panel **(e)**) were sourced from © OpenStreetMap contributors and plotted using contextily).

A domain-specific application of the robust FI map, depicted in Figure 2**a** and **b**, is its use in enhancing congestion prediction models by integrating, as an additional layer, the multiscale city-specific road network features weighted by the corresponding FI values. Particularly for scale combinations such as 4 km^2, 1 km^2, and 0.25 km^2—formed by subdividing urban tiles into equal-sized sub-tiles—our approach facilitates the efficient integration of multiscale features without the computational overhead typically associated with tile matching, as shown in Figure 2**d**.

City-specific highly actionable features

While the FI values help determine the features that best explain the intra-city variations in congestion levels, from a policymaking perspective, the direction of influence of these features is crucial for designing targeted interventions. Typically, the direction of influence is captured using beeswarm plots as shown in Figure 3**c** and **d**. However, while studying multiple cities and scales, beeswarms are not practical for observing the overall trends across cities and scales. Hence, we introduce a simplified metric called the Sensitivity Ratio (SR) that relates the SHAP values to the feature values and produces a single scalar that summarises the aggregated influence of the feature (cf. Methods Equation 1). Features with a positive SR value are referred to as 'exacerbating features' since the increase in feature values is likely to increase the SHAP values and, hence, congestion. Similarly, the features with a negative SR value are referred to as 'alleviating features'. Recall that increasing a feature to achieve policy intervention value does not necessarily require changing the road infrastructure since a similar effect is possible by nudging traffic through (or away from) specific urban tiles.

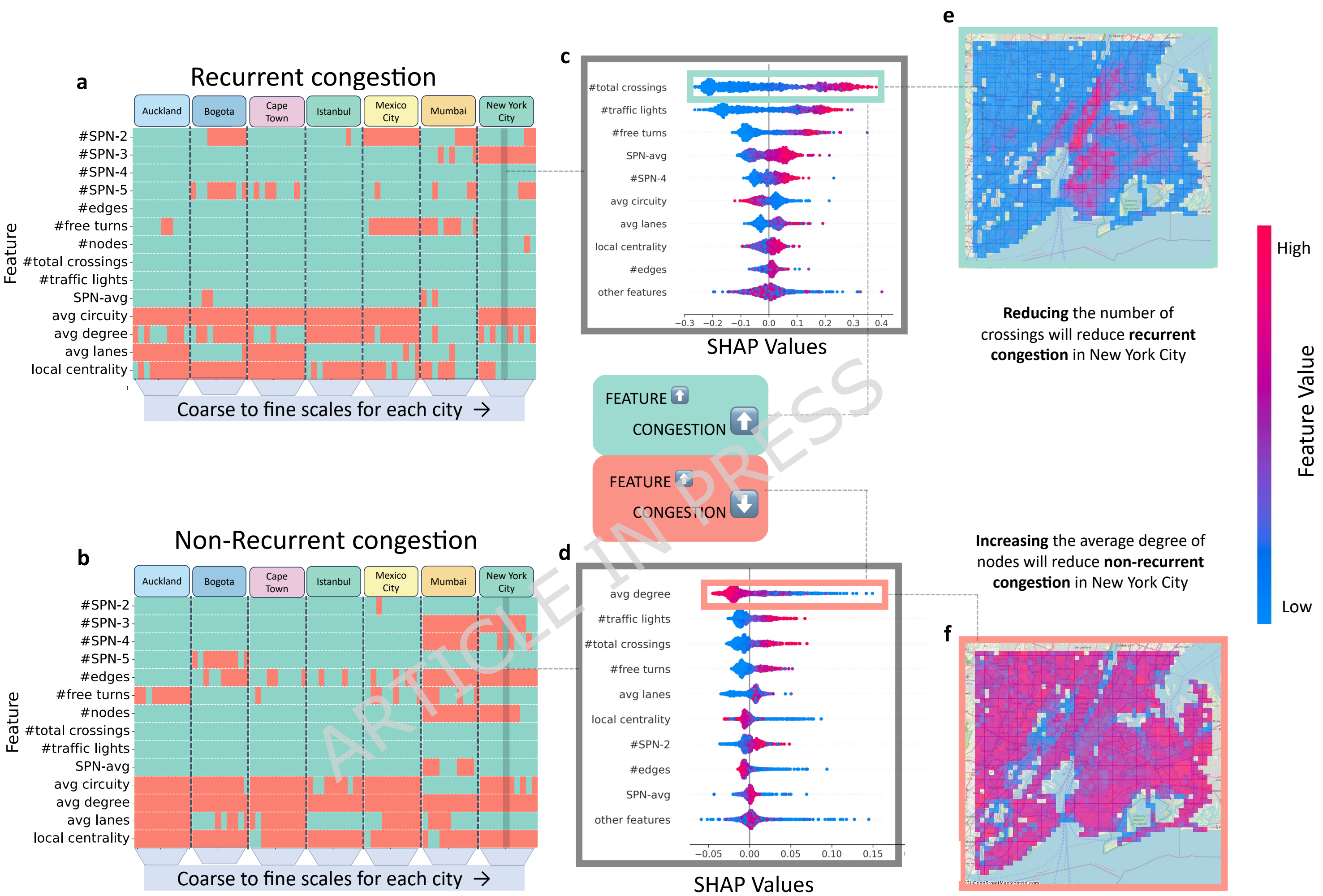


Figure 3. **(a)** and **(b)** visualise the direction of influence for RC and NRC, respectively. The features with a positive direction of influence are exacerbating and are marked with a soft green colour. These features, if increased, result in higher SHAP values and, hence, higher congestion. The soft red features are congestion-alleviating features. **(c)** and **(d)** visualise the beeswarm plots for selected models to demonstrate how the colour captures the direction of influence. **(e)** and **(f)** visualise the physical location of tiles with low and high feature values for a single feature. (The basemaps used in the panels **(e)** and **(f)** were sourced from © OpenStreetMap contributors and plotted using contextily).

Figure 3**a** shows some features exhibiting a consistent sign for SR across all scales. These features are excellent candidates for targeting policy interventions. On the contrary, we find some features that had high FI values in the preceding analysis exhibit a switching of the sign of SR between scales (*e.g.#avg lanes* for Bogota NRC). These features pose a challenge to policy interventions despite their high FI values since any mitigation of congestion obtained through an intervention on these features will likely offset the exacerbation caused at a different scale. Along similar lines, we find some features to be changing the SR sign at a specific scale but remain consistent after that. These features are still helpful in mitigating congestion, depending on the problem at hand. For instance, for New York, if the goal is to mitigate RC in larger parts of the city (at scales greater than 1km^2), reducing the lanes #*local centrality* is likely to be effective in reducing RC, albeit at the cost of small congested pockets within otherwise less-congested areas.

The literature on congestion mitigation has argued that from a policy perspective, the most effective interventions for congestion target both RC and NRC together[4]. Hence, when considering a single city, the most effective congestion mitigation insights should be based on features with the same sign across both RC and NRC. For each city, we filtered the features with the same sign across all scales, sorted them in decreasing order of magnitude of SR, selected the top 3, filtered the features that have the same sign for both RC and NRC and marked these features as the highly actionable features for each city as shown in Figure 4**a**. (cf. SI **??** for the unfiltered table of features). Intra-city heterogeneity profiling confirmed substantial cross-city distributional variation in the independent variables (SI **??**). The directional stability of the actionable features reported in Figure 4**a** was evaluated against this background of heterogeneity; its persistence within all GoF-passing model configurations supports the interpretation that these features reflect robust structural associations and are not distributional artefacts.

Another interesting observation from Figure 3 is that among the actionable features, some features show a reversal of signs between cities. For instance, for RC, Figure 3**a** shows that *avg circuity* is an alleviating feature for all cities except Mumbai. Similarly, for NRC, Figure 3**b** shows that the *avg lanes* is an exacerbating feature for Istanbul but an alleviating feature for Auckland. Such reversals might be attributed to the interplay between the alternate hypotheses regarding travel demand attractions and incident rates. Let us consider the case of *avg degree* for RC. On the one hand, parts of the city with a higher *avg degree* typically have higher levels of economic activities and hence more travel demand to them, thus resulting in higher RC. On the other hand, higher *avg degree* is also associated with providing more access and egress points to the traffic, thus resulting in lower RC. Similarly, let us consider *avg lanes* for NRC. On the one hand, since roads with more lanes are typically associated with higher shock absorption as described in the Highway Capacity Manual[35] by a lower percentage of traffic flow drop compared to the percentage of lanes blocked, the #*avg lanes* is likely to reduce NRC. On the other hand, since the #*avg lanes* is associated with higher speed limits, it is likely to result in a higher number and severity of accidents[4], thereby increasing NRC. Which of the two sets of forces emerges as the winner for any city might be moderated by some other city-dependent contextual factors—thereby suggesting the need to enrich the feature variables by quantifying such contextual factors.

Finally, Figure 3 shows that #*total crossings* and #*traffic lights* are two features which are exacerbating features across all scales, cities, and both congestion types. To assess the sensitivity of the actionable feature identification to methodological choices, we stress-tested the SR cutoff ($\text{SR} \in \{0, 10^{-4}, 10^{-3}, 10^{-2}\}$) and the directional consistency criterion (strict: all scales; majority: $\geq$75% of scales). The top-3 actionable feature sets remained fully stable across all combinations in all GoF-passing city groups, confirming that our reported results are not an artefact of the specific threshold or consistency rule employed. Hence, in the absence of detailed analyses for a new city, these features are likely to be actionable features, specifically because the studied cities varied in socioeconomic groupings, precisely, the Global North and Global South, had differing population sizes, densities, and levels of public transportation reliance (cf. SI **??**).

Figure 4**a** identifies city-specific actionable road network features. Features marked with an asterisk (*) represent those with the largest influence, indicating they are likely to be the most effective at reducing congestion. Such features are ideal for supply-side interventions, primarily infrastructural investments at any urban tile and for directing route-suggestion-based strategies, such as preferred routing, to mitigate congestion. However, solutions that involve directing traffic towards areas with alleviating features—or away from areas with exacerbating features—such as congestion pricing often favour contiguous regions. Thus, we specifically show features that demonstrate strong spatial clustering within the features marked as highly actionable for these cities in Figure 4**b**. Coincidentally, all of these features exacerbate congestion, suggesting that traffic redirection away from these areas will be effective. More importantly, these features are amenable to demand-side interventions and are remarkably similar across cities, suggesting the high transferability of area-based, demand-side solutions such as congestion pricing. (cf. SI Fig. **13-19** for the spatial distribution of all 14 features to choose alternate actionable features if the most actionable suffers additional challenges in selected cities).

Discussion

We systematically analysed the relationship between road network features and urban traffic congestion across seven of the world's most congested cities[9]. The multiscale analysis employed SHAP values to assess feature importance. A novel metric called the Sensitivity Ratio (SR) was proposed to quantify the direction and strength of each feature's

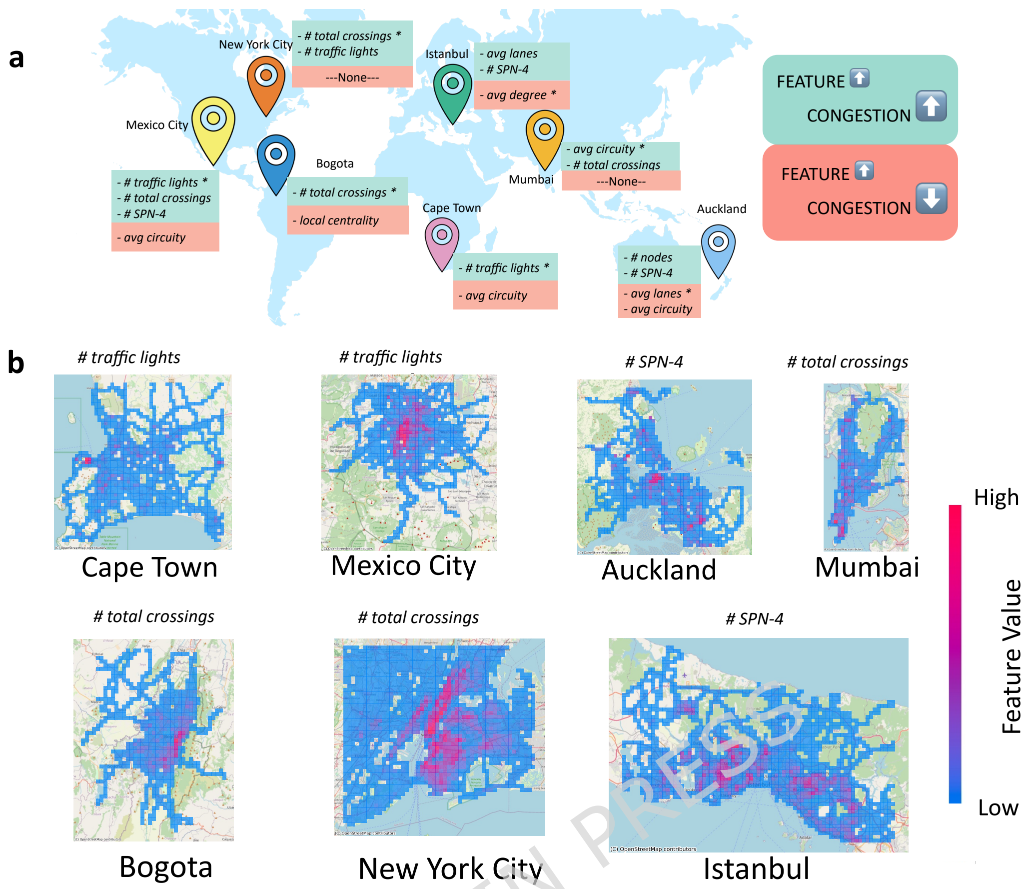


Figure 4. **(a)** City-specific highly actionable features for targeting our congestion-mitigation efforts. **(b)** Features preferred for demand-side interventions due to strong spatial autocorrelation. Each pixel in **b** represents 1km^2 urban tile. We note that Mumbai failed the $\mathrm{R}^2 \geq 0.20$ GoF threshold across all spatial configurations. Its entries are retained here for completeness but should not be interpreted as validated actionable features; they represent an open modelling challenge for future research. (The basemaps used in the subfigures in panel **(b)**) were sourced from © OpenStreetMap contributors and plotted using contextily).

influence on congestion (cf. Equation 1). The features with the same influence direction across all spatial scales were called actionable features. Based on findings from similar efforts to establish the link between road network features and traffic congestion[2], we expected to observe commonalities in FI across these highly congested cities. However, to our surprise, we found minimal similarities, with each city demonstrating a unique pattern of feature importance, underscoring the necessity for tailored, city-specific policymaking.

We emphasise that the directional associations identified in this study reflect patterns observed within the training data and should not be interpreted as causal effects. The identified actionable features represent robust associative signals that warrant consideration as policy-relevant candidates, subject to validation through counterfactual simulation and demand-response modelling prior to implementation.

When a feature is identified as alleviating in Figure 4**a**, it implies a structural association where higher feature values correlate with lower expected congestion. Alternatively, manipulating the travel demand to direct more traffic through areas with higher values of alleviating features could also mitigate congestion levels. The choice between supply and demand measures for any given city depends on various factors, including political, economic, and social considerations.

Supply side implications

On the supply side, our findings challenge the prevalent scepticism about the effectiveness of supply-side interventions—often due to concerns about induced demand. City-specific road network features influencing congestion across various spatial scales prove that congestion mitigation is possible, provided robust scale-invariance and context-specificity are acknowledged. Another aspect of supply-side interventions, mainly when physical upgrades are impractical, involves Traffic Signal Control (TSC). TSC can regulate traffic flow into specific urban tiles located within road segments controlled by traffic signals. This control alters the supply-to-demand ratio and effectively alters the road network features experienced by the passing traffic. Furthermore, consistent with the widely established spatial autocorrelation of urban spaces, as demonstrated in Figure 4**b**, similar values of selected road network features are found in nearby tiles. This autocorrelation suggests that supply-side interventions like TSC can be strategically targeted at specific zones within a city based on the features shown in Figure 4**b**, thereby reducing the complexity of implementing such measures.

We present two strategic recommendations in light of the substantial costs to the exchequer while modifying the road networks. First, for established, densely populated cities where constructing new infrastructure such as flyovers or overpasses is challenging, our findings are particularly insightful for developing new urban areas or satellite cities often developed near highly congested cities. Such areas are likely to share the contextual factors, thus making the actionable features identified in our study relevant and potentially effective for these new developments. Second, we advocate for changes to road network features such that the new values are within the existing range of values associated with parts of the city that have lower congestion. A thorough evaluation is needed if changes are to extend beyond current parameter ranges. Being a purely data-driven approach, our paper did not explore expansions beyond observed values. Future counterfactual studies using simulation-based approaches could provide valuable insights into the impact of changing the feature values beyond the city-specific ranges of existing feature values.

Demand side implications

On the demand side, our research has profound implications for congestion pricing, which has been demonstrated to reduce travel demand while improving air quality effectively, minimising road wear, a source of revenue, and decreasing accident rates[36]. Innovations such as GPS-based tolling have dramatically reduced operational costs compared to older number plate recognition systems[36]. The successful implementation and public acceptance of congestion pricing depend on effective communication. The cities in this study are at various stages of discussing or implementing congestion pricing. It has been previously implemented in Mumbai[37], Bogotá[38], and Mexico City[39]. Discussions are ongoing in Istanbul[40], New York City[41], and Auckland[42]. In some cities, substantial deviations from expected revenues have been reported, alongside significant confusion in driving behaviour[39, 43]. The road network features identified in our research as explaining the congestion levels present an opportunity to enhance public understanding and support for these measures. Interestingly, most features that were found to be actionable across different cities have alternate names that can resonate with the members of the public, such as road curvature in place of *avg circuity*, 4-way intersections in place of *#SPN-4*, and the number of crossings in place of # *total crossings*. Hence, using these road network features to explain why certain areas are tolled can help demystify the pricing structures and align public perception with the overarching goals of congestion management strategies.

Moreover, when complex pricing schemes are inevitable, navigation aids help simplify route-choice decision-making in dynamically priced environments[36]. The insights from our study can help enhance public trust in such route-choice suggestions by explaining the rationale behind suggested routes using road network features. Once drivers believe their actions are for the common good, there is an increased likelihood of compliance[44]. The increase of trust is particularly crucial as profit-driven entities operate the most popular route choice applications, and their net impact on urban congestion remains uncertain[4]. Along similar lines, another use case for our findings is during disaster scenarios where real-time information might be unavailable due to disasters or IT infrastructure malfunctions. In such a scenario, a shortest-path algorithm based on road network features could be a reliable fallback option for route-choice algorithms. Specifically, two approaches are possible: one could utilise the most significant actionable feature (indicated with an asterisk in the Figure 4) for each city to design disaster-scenario route choices, or multiple actionable features could be employed in a weighted manner based on their FI values.

Limitations and future research

Despite the multiscale-multicity nature of our study, several key limitations bound the application of this framework and highlight avenues for future research.

First, our static analytical window is limited to a one-month temporal period (September 2022). While this captured a stable baseline of recurrent congestion, it cannot fully account for long-term seasonal variations, school calendar effects, or the full range of weather conditions influencing non-recurrent congestion. A temporal ablation across AM peak (07:00–09:00) and PM peak (17:00–19:00) windows confirmed GoF stability within the observation period (cf. SI **??**); however, the generalisability of the findings across seasons and years remains an avenue for future research.

Second, the current framework identifies predictive associations but lacks dynamic feedback loops; it does not model the phenomenon of 'induced demand,' where increased supply capacity dynamically generates new travel demand over time. While our framework incorporates demand-attraction network features (average degree and local centrality) as structural representations of spatial demand distribution, it does not dynamically control for exogenous demand volume variance across cities. Supply-side candidate features identified as alleviating should therefore be validated through simulation-based counterfactual studies. Consistent with our earlier recommendation that such studies should test feature values beyond observed ranges, we additionally argue that they should model day-to-day demand response scenarios to assess the potential for induced demand to erode projected service gains.

Third, our API-constrained sampling strategy prioritised global geographic and socioeconomic variance by selecting the single most congested city per country based on the TomTom Congestion Index. While this maximises cross-continental coverage, it necessarily limits the representation of distinct congested hubs within the same nation, such as the morphologically distinct cases of New York City and Los Angeles within the United States (*e.g.*, a transit-oriented dense grid versus a largely auto-dependent polycentric metropolis).

Fourth, our use of fixed square bounding boxes rather than Functional Urban Area (FUA) boundaries was a design choice motivated primarily by our focus on the dense urban road network structures. Migrating to official FUA geometries (as defined by the OECD Functional Urban Areas database and regional metropolitan authorities), could have introduced a fundamental 'dilution effect' that conflicted with the research focus of this paper. The official FUAs for the studied cities extend deep into rural and low-density commuter hinterlands: for example, moving to true functional boundaries would expand our study areas substantially (*e.g.* to $\approx$ 11,000 km^{2}for New York City (OECD FUA database[45])). Flooding the dataset with such a large number of low-density, uncongested peripheral tiles would have mathematically obscured the structural bottleneck signals we seek to isolate in highly congested urban cores. The bounding box design was therefore a deliberate choice to protect signal concentration. The distributional variation within those boxes was characterised through intra-city heterogeneity profiling (cf. SI **??**), and the persistence of directional SR stability within GoF-passing models provides strong evidence that the main conclusions are not artefacts of the specific bounding box placement. As a partial sensitivity check, we additionally tested the stability of actionable features under −10% and +10% bounding-box extent variants for Istanbul and New York City. We found that the sign consistency of overlapping features was found to be perfect across all GoF-passing comparisons (cf. **??**). However, evaluating whether these conclusions hold under consistently defined FUA boundaries remains an interesting direction for future research.

Fifth, our reliance on the Goodness-of-Fit gating mechanism explicitly defines the boundaries of the framework. As demonstrated by the consistently poor model fit for Mumbai, not all complex urban topologies yield predictable congestion outcomes based solely on these specific road network features. The framework structurally limits interpretation to highly predictable regimes, meaning these insights cannot be universally applied to cities that fail to meet minimum predictive thresholds. To confirm the stability of GoF-based exclusion decisions across random initialisations, the core setup was re-run with five independent random seeds for both RC and NRC; GoF pass rates were found to be fully consistent with single-seed results across all cities (cf. SI **??**). Finally, due to the small number of cities in this study, the selection of features that show spatial clustering was performed manually. An automated spatial autocorrelation index, such as Moranś I[46], is recommended for studies considering a large number of cities to ensure scalability.

Methods

Choice of cities

The total number of cities included in this study (7 cities) was limited by the API quota of the HERE traffic API. The chosen 7 cities comprise a list of highly congested cities worldwide in 2021. The cities are Auckland, Bogota, Capetown, Istanbul, New York City, Mexico City, and Mumbai. First, we created a list of the most congested cities according to the Tom-Tom congestion index from 2021[9]. From this list, we removed the cities for which either the traffic data (speed reduction index) from the HERE API or the OSM API was unavailable during the period of data collection. To ensure a reasonable sample across different parts of the world, for countries that had multiple cities in the top congested cities, only the most congested were chosen. Similarly, in the case of a tie where two cities had similar scores according to the Tom-Tom index, cities that belonged to different continents were given preference over a continent that was already represented by another city. We queried the traffic speed data every 10 minutes over a period of 1 month (September 2022).

Choice of features

Each city is segmented into uniform square-shaped urban tiles of varying sizes and the road network corresponding to each tile is represented by a tile-subgraph g. Table 3 visualises the extracted urban tiles corresponding to the maximum and the minimum value of each feature for scale $1\text{km}^2$for Auckland. The extracted subgraphs from each of those tiles are also visualised. Our choice of features is targeted at selecting features that have known predictive associations with both types of traffic congestion, sometimes affecting both types of congestion simultaneously. Road network features that attract or repel traffic demand are more likely to show a relationship with recurrent congestion. In contrast, those that provide avenues for conflict between traffic streams are more likely to cause non-recurrent congestion. However, often, the pathways are mixed due to the human factor involved- drivers often use their experience to deviate from the computed most optimal routes in order to avoid congestion and traffic lights[47]. To account for the discrepancies in the representation of road network as a graph (particularly, nodes versus intersections, edges versus streets, and *average degree* versus *streets per node (SPN)*), the chosen features are a mix of classic graph features and those that capture the structure of the road network. The graph-based features are particularly interesting for the data-driven traffic prediction community, whereas the streets-based features are interesting to the policymakers. Five feature group features are identified, each supported by the current understanding of congestion mechanisms in the transportation literature (cf. Table 1).

Finally, to ensure that our framework is transferable to other cities, we removed the features with more than 10% of missing data in the OSM. For instance, the *#SPN-1* (which counts the number of dead ends) and average speed were removed due to a high percentage of missing data. While the classic graph features have universal definitions, the intersection-related, network structure-related and demand-attraction-related are explained below to avoid any ambiguity.

Urban tiles at various scales

In a multi-city analysis like the one presented in this paper, defining city boundaries poses a significant challenge due to cities' diverse spatial extents and orientations. To address this challenge, we used large square bounding boxes for each city to ensure that most of the city's area is contained within the box. The selected bounding box was centred approximately at the city centre for each city. The total size of the square area was $75\times75\text{km}^2$for Istanbul and $50\times50\text{km}^2$for other cities. The various spatial scales are defined by dividing the square into uniformly sized $n\times n$ urban tiles, where $n \in \{20, 25, 30, 40, 50, 60, 70, 80, 90, 100\}$ for all cities except Istanbul and $\{30, 37, 45, 60, 75, 90, 105, 120, 135, 150\}$ for Istanbul. The study areas are shown in Figure 5 using the scale corresponding to 1km^2 of tile areas.

For all cities except Istanbul, the scales defined with n=25, n=50, and n=100 correspond to 4 km^2, 1km^2, and 0.25km^2. For Istanbul, the scales defined with n=37, n=75, and n =150 correspond to 4.1km^2, 1km^2, and 0.25km^2.

Independent and target variables

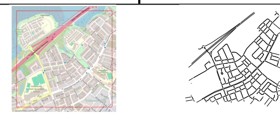
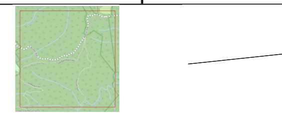
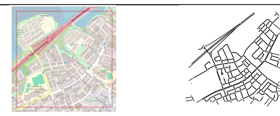
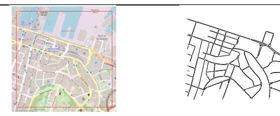
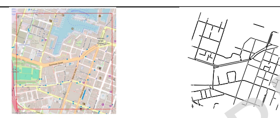
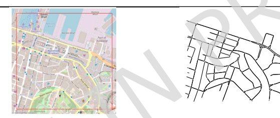
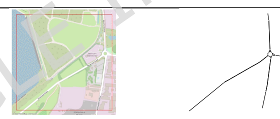
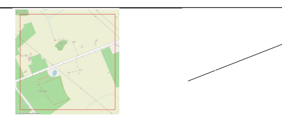
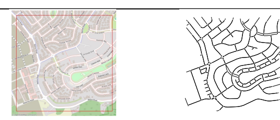
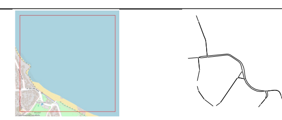
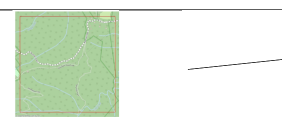
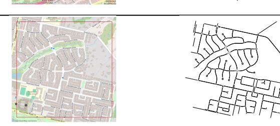
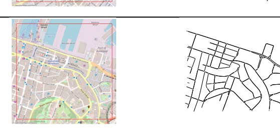
For each scale under consideration, the independent variable (X) is a vector of selected road network features, and the dependent variable (Y) is a scalar quantity that quantifies the mean value of congestion in the tile for the entire month.

Independent variable

Each city was then divided into square tiles of varying sizes. For the majority of this study, we consider three sizes of tiles, corresponding to coarse (tile area 4km^2), medium (1km^2), and fine (0.25km^2) scales. For any given scale, the road network graph corresponding to each tile is extracted from OSM, and the 14 features are extracted using 'osmnx'[48]. The selected features have been widely studied in the literature and have established structural associations with urban traffic congestion. To extract the tile subgraph, first, we compute the tile boundaries represented by a smaller bounding box. Corresponding to the smaller bounding box, we filter out the nodes from the entire city graph that lie inside the tile boundaries. Additionally, to avoid missing out on edges that are not fully contained within the tile boundaries, the edges that intersect with the tile bounding box are cut into two parts by the bounding box. Hence, the tile subgraph includes all the edges due to nodes that are neighbours of the filtered nodes.

Dependent variables (RCI and NRCI)

The data on urban traffic congestion was collected using a commercially available API from HERE technologies[49, 50]. This dataset quantifies traffic congestion using a variable called Jam Factor. As per the API documentation[51], the Jam Factor (JF) is computed as the ratio of the current space mean speed to the free-flow speed. The free flow speed corresponds to the speed when very few vehicles are on the road so that the movement of one vehicle is not

Extracted feature for $g(V,E)$	Highest feature value: Urban tile, Subgraph (g)	Lowest feature value: Urban tile, Subgraph (g)
$\#nodes{=}\lvert V\rvert$		
$\#edges{=}\lvert E\rvert$		
$avg\ degree{=}2\frac{\lvert E\rvert}{\lvert V\rvert}$		
$avg\ circuity{=}\frac{\sum_{e\in E} d_{\text{network}}(e)}{\sum_{e\in E} d_{\text{euclidean}}(e)}$		
$\#total\ crossings{=}\sum_{v\in V}\mathbb{1}_{\lvert E(v)_{\text{undirected}}\rvert\geq 2}$		
#traffic lights: #metered crossings		
*#free turns: #**un**metered crossings*		
$local\ centrality{=}\sum_{v\in V}\frac{C_{\text{bet}}(v,g)}{\lvert V\rvert}$		
$avg\ lanes{=}\frac{\sum_{e\in E}\#\text{lanes}(e)}{\lvert E\rvert}$		
$\#SPN\text{-}2{=}\sum_{v\in V}\mathbb{1}_{\lvert E(v)_{\text{undirected}}=2\rvert}$		
$\#SPN\text{-}3{=}\sum_{v\in V}\mathbb{1}_{\lvert E(v)_{\text{undirected}}=3\rvert}$		
$\#SPN\text{-}4{=}\sum_{v\in V}\mathbb{1}_{\lvert E(v)_{\text{undirected}}=4\rvert}$		
$\#SPN\text{-}5{=}\sum_{v\in V}\mathbb{1}_{\lvert E(v)_{\text{undirected}}=5\rvert}$		
$\#SPN\text{-}avg{=}\sum_{v\in V}\frac{\lvert E(v)_{\text{undirected}}\rvert}{\lvert V\rvert}$		

SI

Table 3. Visualization of subgraphs for urban tiles of 1 km^2 for Auckland corresponding to the tiles that had the maximum and minimum values for each feature. # represents the count for any feature and $|x|$ represents the cardinality of set x. C_{bet} refers to betweenness centrality. (Image basemap source: ©OpenStreetMap contributors; plotted using contextily)

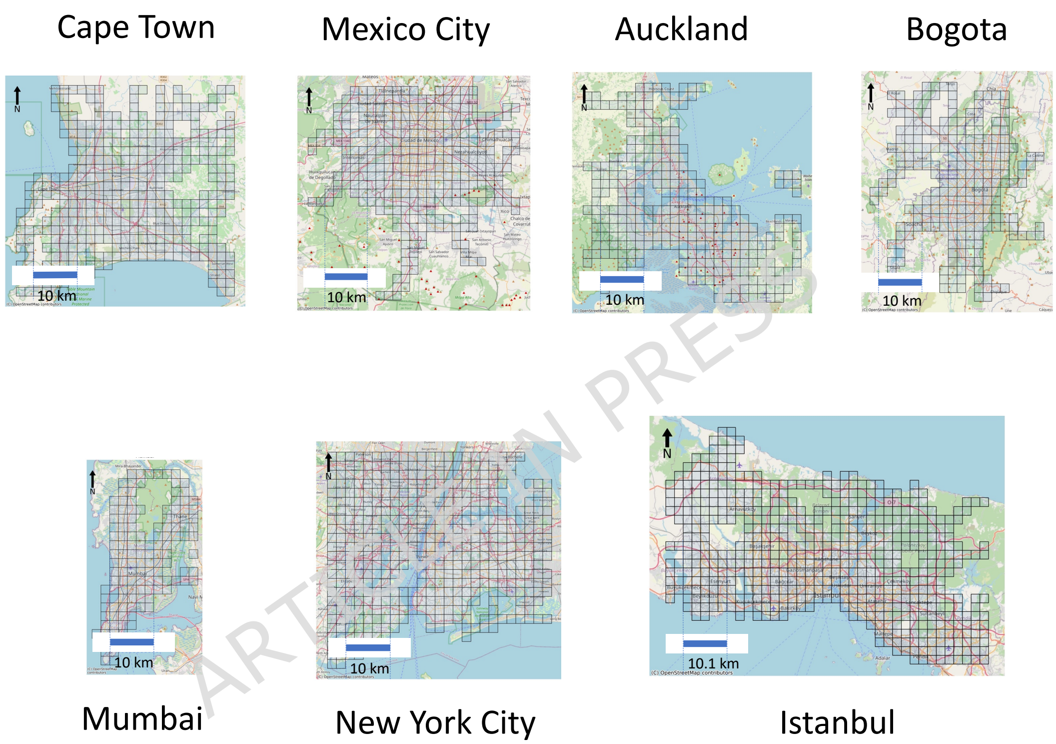


Figure 5. Study areas with square tiles of area ≈ 4 km^2. The tiles at this scale are formed by dividing the chosen $50{\times}50$ km^{2}study area into $25{\times}25 = 625$ tiles and then filtering the tiles for which OSM had street network and traffic data was collected from HERE API. For Istanbul, the tiles are ≈ 4.1 km^{2}since the chosen study area was $75{\times}75$ km^{2}and the same was divided into $37{\times}37 = 1369$ tiles. (Image basemap sources: © OpenStreetMap contributors; plotted using contextily)

affected by the other. It is generally lower than the speed limit due to poor weather or road conditions. Conceptually, JF is the same as the Speed reduction index, SRI, which is an extensively studied variable of interest for traffic congestion research[2, 52]. Proposed in Lomax et al. (1997) [53], the SRI varies continuously from 0 to 10, with 0 implying the speed corresponding to very few vehicles on the road and 10 implying a complete traffic jam when vehicles barely move. Since SRI is normalised by the free flow of each road segment, it ensures that the variation in speed limits and flow capacities of different roads does not affect the congestion metrics. Typically, SRI values below 4 are not perceptible to drivers[53], nevertheless useful for transportation researchers. For our dataset, a JF of 0 indicates free-flow traffic conditions. JF values up to 3 suggest light congestion, values between 3 and 7 indicate moderate congestion, values from 7 to 9 imply severe congestion, and finally, values from 9 to 10 signify that the road is completely blocked, indicating a complete traffic jam. The SRI was collected every 10 minutes for each road segment, and the data was collected for one month.

The SRI, being a spatiotemporally varying variable, must be aggregated temporally and spatially to assign a congestion value to each urban tile. Several temporal and spatial aggregations have been commonly used in the transportation literature to formulate unified indices to assign a single value to a temporal bin (hour, day, or month) and spatial bins (urban tile, administrative regions, special economic zones and so on)[54]. We approach the temporal and spatial aggregations sequentially. First, the temporal aggregations are used to assign a single value of congestion level to each road segment, followed by spatial aggregations that assign a single value of congestion level to each urban tile (cf. SI Section 2 for details of the two aggregations).

After pre-processing the independent and dependent variables, we used Random Forest models to predict the RCI and NRCI congestion indices for each urban tile. (cf. SI Section 3 for details on model selection, hyperparameter tuning, spatial cross validation, and spatial perturbation tests).

Model Validation and Goodness-of-Fit Gating

To ensure interpretability is derived exclusively from mathematically robust models, we instituted a strict Goodness-of-Fit (GoF) gating mechanism. Prior to SHAP value and Sensitivity Ratio (SR) extraction, each city-scale-time slice was evaluated for predictive quality. A minimum cross-validated threshold of $R^2 \geq 0.2$ was required. Models failing to meet this threshold (as consistently observed in the highly complex topologies such as Mumbai) were systematically excluded from the interpretability pipeline. The full GoF summary under this gating rule is reported in SI **??**. A temporal ablation stratifying the data into AM peak (07:00–09:00) and PM peak (17:00–19:00) windows confirmed that the GoF pass/fail structure is stable across both temporal strata: 12 of 14 city–window combinations passed the threshold, with Mumbai remaining the only consistent outlier in both windows (cf. SI **??**).

Multiscale Spatial and Heterogeneity Profiling

Spatial perturbation tests were expanded beyond the baseline 1 km^2 scale to include representative fine (0.25km^2), medium (1km^2), and coarse (4km^2) spatial resolutions. Temporal parameters were held strictly fixed during spatial iterations to isolate scale-dependent variance. Furthermore, intra-city heterogeneity was quantified by calculating the spatial variance and skewness of all independent variables across the constituent urban tiles for each city, establishing the distributional diversity underlying the extracted SHAP values. The full per-city heterogeneity profile is reported in SI **??**.

Feature importance through SHAP values

In recent years, applying SHAP (Shapley Additive exPlanations) values has significantly advanced the investigation into model explainability, contributing to the development of explainable AI[55]. SHAP values address the attribution problem in machine learning, which aims to assign a feature importance value to each raw data feature[56]. Originally conceptualized in[57], the relevance of SHAP values in explainable machine learning was first demonstrated in[58], which also introduced certain approximations for efficient computation. Additionally, these values align well with human intuition regarding feature importance. Subsequently, exact methods for computing SHAP values for tree-based models (such as RFs) were developed[59] and released as an open-sourced software, forming the basis of the SHAP-based analyses presented in this paper.

Recently, SHapley Additive exPlanations (shortened to SHAP) have been shown to explain the performance of black box models. Specifically, the SHAP values can be computed efficiently for tree-based models such as Random Forests due to Lundberg et al.[59]. A single SHAP value is assigned to each feature for each data point (urban tile) in the model. For any data point, a large positive SHAP value for a feature j indicates that the feature j contributes positively to the model output for that data point and vice versa. The target variables in this paper (RCI and NRCI, both derived from the Speed Reduction Index) increase for higher congestion levels, so a higher SHAP value for a feature j implies that more congestion can be attributed to the feature j. Aggregating SHAP values is necessary to ensure inferences are based on easily interpretable plots. Here, we are interested in determining the feature importance (FI), so we compute the mean FI using the mean of absolute SHAP values across all data points, as recommended in Lundberg et al. (2020) [59]. The heatmaps in Figure 2 show the variation of FI across scales and cities.

SHAP values assign a single value to each raw feature for each model prediction. A large positive SHAP value indicates that an increase in the feature value will significantly shift the prediction away from the mean across all data points. However, visualizing SHAP values can be challenging, as they are computed for each feature and data point. Therefore, aggregation is necessary to ensure inferences are based on easily interpretable plots. In this study, we compute the mean feature importance using the average of the absolute SHAP values across all data points, as recommended in[59]. As a result, the feature importance reported in this study is always ≥ 0. Hence, further analysis of the SHAP value distribution is needed to ascertain the direction of the relationship between a feature and its SHAP values.

Given a set of SHAP values $\{SHAP_{ij}\}$ for feature j across all data points i, we can divide the data points into two subsets: pos_j representing all data points which have positive SHAP and pos_j as all data points having negative SHAP. Formally,

$$pos_j = \{i \mid SHAP_{ij} > 0\}$$

$$neg_j = \{i \mid SHAP_{ij} < 0\}$$

Let MeanSHAP_{pos_j} and MeanSHAP_{neg_j} represent the mean SHAP value for feature j calculated over the subset pos_j and neg_j respectively. Formally,

$$\text{MeanSHAP}_{pos_j} = \frac{1}{|pos_j|} \sum_{i \in pos_j} SHAP_{ij}$$

$$\text{MeanSHAP}_{neg_j} = \frac{1}{|neg_j|} \sum_{i \in neg_j} SHAP_{ij}$$

Similarly, let MeanX_{pos_j} and MeanX_{neg_j} represent the average value of feature j, but as per the data subset pos_j and neg_j respectively. Formally,

$$\text{MeanX}_{pos_j} = \frac{1}{|os_j|} \sum_{i \in os_j} x_{ij}$$

$$\text{MeanX}_{neg_j} = \frac{1}{|neg_j|} \sum_{i \in neg_j} x_{ij}$$

where x_{ij} is the value of feature j at data point i.

Finally, we define the Sensitivity Ratio (SR) value for feature j as:

$$\text{Sensitivity Ratio}_j = \frac{\text{MeanSHAP}_{pos_j} - \text{MeanSHAP}_{neg_j}}{\text{MeanX}_{pos_j} - \text{MeanX}_{neg_j}} \tag{1}$$

The ratio provides a dimensionless, standardized measure of how the model's output is sensitive to changes in feature j. The *absolute value* of SR indicates the strength of influence, while the *sign* of SR shows the direction of this influence (positive or negative); positive implies an increase in feature increases the SHAP value, implying more congestion. For linear models, the SR approximates the model coefficients while for non-linear models, which is the focus of this study, the SR recovers the overall impact of the features on the SHAP values (cf. SI Fig. 8 for the demonstration of the usefulness of SR values using simulated data).

As a model-agnostic sensitivity check, RF-based TreeSHAP attribution was compared against an equivalent XGBoost model at the 1 km^2 baseline scale; results and theoretical justification are reported in SI Section 4 and **??**.

Declaration of competing interest

The authors declare no known competing interests.

Spatial Splitting method	Recurrent ($1km^2$)	Non-Recurrent @ ($1km^2$)
Vertical-Spatial	New York City, Mexico City, Bogota, Cape Town, Istanbul	New York City, Mexico City, Cape Town Bogota, Istanbul
Horizontal-Spatial	Auckland, New York City, Capetown, Istanbul	Auckland, Cape Town
Grid-Spatial	New York City, Istanbul, Capetown	(No cities passed)

Table 4. Summary of cities passing the Spatial Cross-Validation test for Recurrent and Non-Recurrent cases at a tile size of 1 km^2.

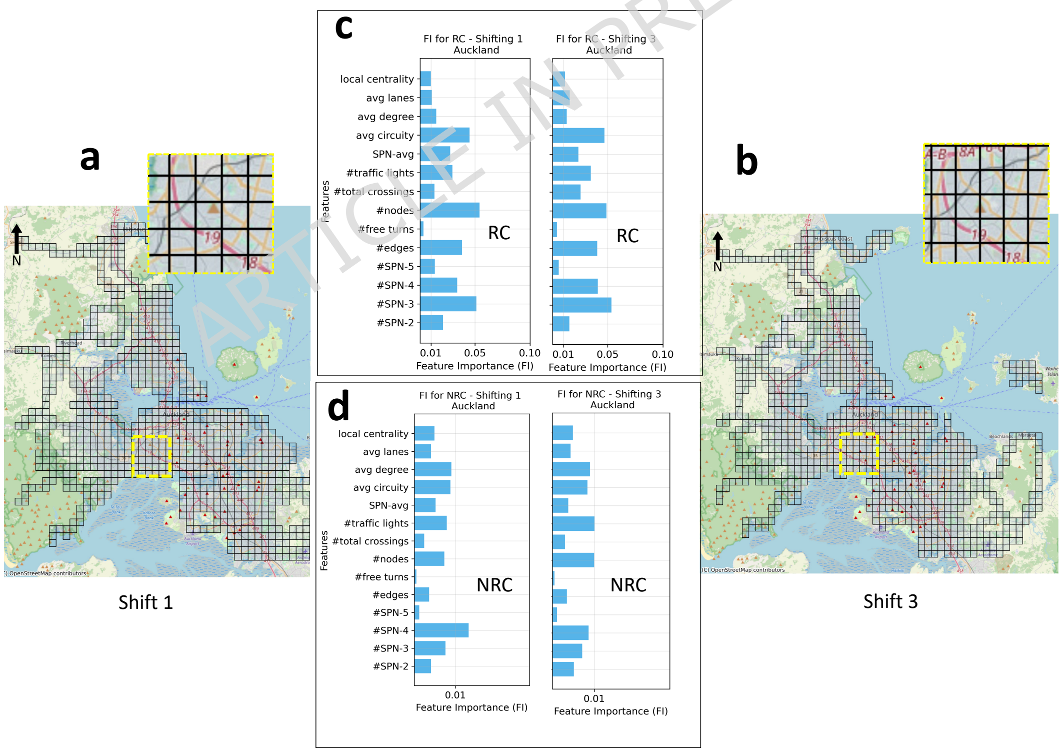


Figure 6. Spatial perturbations by 2/3 km to the **(a)** North and **(b)** East for Auckland at 1 km^2 tiles. The Feature importance plots in **(c)** and **(d)** show comparable values for both RC and NRC, thereby validating our results. (Image basemap source: ©OpenStreetMap contributors; plotted using contextily)

Authorship contributions

N.K., Y.Z., and M.R. conceptualised the study. N.K. programmed the framework and wrote the first draft. Y.Z. processed the raw data files into segment-level variables and contributed to programming an early version of the framework. N.K., Y.Z., and N.W. developed the methods. N.W. and J.O. critically reviewed the methods. J.O. and M.R. helped improve the manuscript organization. M.R. secured funding for the study and supervised it. All authors reviewed the manuscript and approved the final version.

Acknowledgments

We thank HERE Technologies for providing the Traffic4cast dataset. The authors thank Dr. Kiran Bacsa and Dr. Wei Liu from Future Resilient Systems, Singapore, for their insightful discussions and valuable feedback. Finally, the authors thank the two anonymous reviewers for their detailed comments and valuable suggestions in improving this manuscript.

Data and codes availability

All codes to reproduce the results from this paper are released under an MIT license and are hosted at our lab's GitHub repository: https://github.com/mie-lab/Urbanscales. The repository also contains instructions on how to transfer the study to a new city. The raw traffic data can be downloaded from the HERE API (https://www.here.com/developer). Image basemap sources throughout the manuscript: ©OpenStreetMap contributors; plotted using contextily (https://github.com/geopandas/contextily))

Funding declaration

This work is an outcome of the Future Resilient Systems project at the Singapore-ETH Centre (SEC) supported by the National Research Foundation, Prime Minister's Office, Singapore, under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

References

1. Li, R. *et al.* Simple spatial scaling rules behind complex cities. *Nat. communications* **8**, 1841 (2017).

2. Çolak, S., Lima, A. & González, M. C. Understanding congested travel in urban areas. *Nat. communications* **7**, 10793 (2016).

3. Taillanter, E. & Barthelemy, M. Empirical evidence for a jamming transition in urban traffic. *J. Royal Soc. Interface* **18**, 20210391 (2021).

4. Kumar, N. & Raubal, M. Applications of deep learning in congestion detection, prediction and alleviation: A survey. *Transp. Res. Part C: Emerg. Technol.* **133**, 103432 (2021).

5. Tsekeris, T. & Geroliminis, N. City size, network structure and traffic congestion. *J. Urban Econ.* **76**, 1–14 (2013).

6. Dolowitz, D. P. & Marsh, D. Learning from abroad: The role of policy transfer in contemporary policy-making. *Governance* **13**, 5–23 (2000).

7. Pojani, D. *Theoretical Approaches to Studying Policy Transfer*, 9–16 (Springer International Publishing, Cham, 2020).

8. Pojani, D. *Policy Transfer as an Emerging Field of Study*, 1–8 (Springer International Publishing, Cham, 2020).

9. TomTom. TomTom Congestion Index 2021. https://nonews.co/wp-content/uploads/2022/02/TomTom2021.pdf (2021). Accessed: 2024-03-20.

10. Wu, J., Zou, Q., Claramunt, C., Cheng, P. & Gu, H. A poi-constrained space-time accessibility model. *Geogr. Analysis* **54**, 104–123 (2022).

11. Bansal, P., Graham, D. J. *et al.* Congestion in cities: Can road capacity expansions provide a solution? *Transp. Res. Part A: Policy Pract.* **174**, 103726 (2023).

12. Bruno, M., Monteiro Melo, H. P., Campanelli, B. & Loreto, V. A universal framework for inclusive 15-minute cities. *Nat. Cities* 1–9 (2024).

13. Montello, D. Scale, in geography. In Smelser, N. & Baltes, P. (eds.) *International Encyclopedia of the Social & Behavioral Sciences*, 13501–13504 (Pergamon Press, United Kingdom, 2001).

14. O'sullivan, D. & Unwin, D. *Geographic information analysis* (John Wiley & Sons, USA, 2003).

15. Unwin, D. J. Gis, spatial analysis and spatial statistics. *Prog. Hum. Geogr.* **20**, 540–551 (1996).

16. Batty, M. *The New Science of Cities* (The MIT Press, Cambridge, MA, 2013).

17. Serok, N., Havlin, S. & Blumenfeld Lieberthal, E. Identification, cost evaluation, and prioritization of urban traffic congestions and their origin. *Sci. Reports* **12**, 13026 (2022).

18. Zhou, X., Hadi, M., Hale, D. K. *et al.* Multiresolution modeling for traffic analysis: State-of-practice and gap analysis report. Tech. Rep., United States. Federal Highway Administration (2021).

19. Khoudari, N. & Seibold, B. *Multiscale Properties of Traffic Flow: The Macroscopic Impact of Traffic Waves*, 267–296 (Springer International Publishing, Cham, 2024).

20. Saberi, M. *et al.* A simple contagion process describes spreading of traffic jams in urban networks. *Nat. communications* **11**, 1616 (2020).

21. Ambühl, L., Menendez, M. & González, M. C. Understanding congestion propagation by combining percolation theory with the macroscopic fundamental diagram. *Commun. Phys.* **6**, 26 (2023).

22. Bellocchi, L. & Geroliminis, N. Unraveling reaction-diffusion-like dynamics in urban congestion propagation: Insights from a large-scale road network. *Sci. reports* **10**, 4876 (2020).

23. Song, C., Lin, Y., Guo, S. & Wan, H. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, 914–921 (2020).

24. Böcker, L., Dijst, M. & Prillwitz, J. Impact of everyday weather on individual daily travel behaviours in perspective: a literature review. *Transp. reviews* **33**, 71–91 (2013).

25. Theofilatos, A. & Yannis, G. A review of the effect of traffic and weather characteristics on road safety. *Accid. Analysis & Prev.* **72**, 244–256 (2014).

26. Kwon, J., Mauch, M. & Varaiya, P. Components of congestion: Delay from incidents, special events, lane closures, weather, potential ramp metering gain, and excess demand. *Transp. Res. Rec.* **1959**, 84–91 (2006).

27. Hallenbeck, M. E., Ishimaru, J., Nee, J. *et al.* Measurement of recurring versus non-recurring congestion. Tech. Rep., Washington (State). Dept. of Transportation (2003).

28. Ballo, L., de Freitas, L. M., Meister, A. & Axhausen, K. W. The e-bike city as a radical shift toward zero-emission transport: Sustainable? equitable? desirable? *J. Transp. Geogr.* **111**, 103663 (2023).

29. Yu, S., Zhang, Z., Liu, F., Wang, X. & Hu, S. Urban expansion in the megacity since 1970s: A case study in mumbai. *Geocarto Int.* **36**, 603–621 (2021).

30. Kumar, N., Martin, H. & Raubal, M. Enhancing deep learning-based city-wide traffic prediction pipelines through complexity analysis. *Data Sci. for Transp.* **6**, 24 (2024).

31. Balenzano, A. *et al.* Sentinel-1 soil moisture at 1 km resolution: a validation study. *Remote. Sens. Environ.* **263**, 112554, DOI: https://doi.org/10.1016/j.rse.2021.112554 (2021).

32. Martino, P. & Sergio, F. Ghs-smod r2016a-ghs settlement grid following the regio model 2014 in application to ghsl landsat and ciesin gpw v4-multitemporal (1975-1990-2000-2015). *Eur. Comm. Jt. Res. Centre (JRC)* (2016).

33. National Geophysical Data Center. Report MGG-6: The TerrainBase global digital terrain model. Retrieved from https://www.ngdc.noaa.gov/mgg/topo/report/s6/s6A.html (1988). Accessed on 2024-03-13.

34. Venter, C., Mahendra, A. & Hidalgo, D. From mobility to access for all: Expanding urban transportation choices in the global south. *World Resour. Institute, Washington, DC* 1–48 (2019).

35. "TRB", T. R. B., of Sciences Engineering, N. A. & Medicine". *Highway Capacity Manual 7th Edition: A Guide for Multimodal Mobility Analysis* (The National Academies Press, Washington, DC, 2022).

36. De Palma, A. & Lindsey, R. Traffic congestion pricing methodologies and technologies. *Transp. Res. Part C: Emerg. Technol.* **19**, 1377–1399 (2011).

37. Hindustan Times. Day 1 confusion, traffic congestion prevail at the nakas. *Hindustan Times* (2023).

38. Torrico, A., Boonsiriphatthanajaroen, N., Garg, N., Lodi, A. & Mainguy, H. Equitable congestion pricing under the markovian traffic model: An application to bogota. *arXiv preprint arXiv:2407.05035* (2024).

39. Guerra, E. & Reyes, A. Examining behavioral responses to mexico city's driving restriction: A mixed methods approach. *Transp. research part D: transport environment* **104**, 103191 (2022).

40. Hurriyet Daily News. Municipality plans road fee to alleviate traffic congestion (2023).

41. Ghassabian, A. *et al.* Beyond traffic jam alleviation: evaluating the health and health equity impacts of new york city's congestion pricing plan. *J Epidemiol Community Heal.* **78**, 273–276 (2024).

42. New Zealand Transport Agency. The congestion question in auckland (2023).

43. Unknown, A. Day 1 confusion, traffic congestion prevail at the nakas. *Hindustan Times* (2024).

44. Ringhand, M. & Vollrath, M. Make this detour and be unselfish! influencing urban route choice by explaining traffic management. *Transp. research part F: traffic psychology behaviour* **53**, 99–116 (2018).

45. OECD. OECD Definition of Cities and Functional Urban Areas. https://www.oecd.org/en/data/datasets/oecd-definition-of-cities-and-functional-urban-areas.html (n.d.). Accessed 2026-04-21.

46. Moran, P. A. Notes on continuous stochastic phenomena. *Biometrika* **37**, 17–23 (1950).

47. Papinski, D., Scott, D. M. & Doherty, S. T. Exploring the route choice decision-making process: A comparison of planned and observed routes obtained using person-based gps. *Transp. research part F: traffic psychology behaviour* **12**, 347–358 (2009).

48. Boeing, G. Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Comput. Environ. Urban Syst.* **65**, 126–139 (2017).

49. HERE. Jam Factor Documentation.

50. Skanda, V. Visualizing real-time traffic patterns using HERE traffic api | by Skanda Vivek | Towards Data Science (2020).

51. HERE. Jam Factor Documentation.

52. Afrin, T. & Yodo, N. A survey of road traffic congestion measures towards a sustainable and resilient transportation system. *Sustainability* **12**, 4660 (2020).

53. Lomax, T. J. Quantifying congestion. *Transp. Res. Board* **398** (1997).

54. Seong, J., Kim, Y., Goh, H., Kim, H. & Stanescu, A. Measuring traffic congestion with novel metrics: A case study of six us metropolitan areas. *ISPRS Int. J. Geo-Information* **12**, 130 (2023).

55. Fryer, D., Strümke, I. & Nguyen, H. Shapley values for feature selection: The good, the bad, and the axioms. *Ieee Access* **9**, 144352–144360 (2021).

56. Sundararajan, M. & Najmi, A. The many shapley values for model explanation. In *International conference on machine learning*, 9269–9278 (PMLR, 2020).

57. Shapley, L. S. *Notes on the N-Person Game; I: Characteristic-Point Solutions of the Four-Person Game* (RAND Corporation, Santa Monica, CA, 1951).

58. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. *Adv. neural information processing systems* **30** (2017).

59. Lundberg, S. M. *et al.* From local explanations to global understanding with explainable ai for trees. *Nat. machine intelligence* **2**, 56–67 (2020).