Resilience and efficiency in transportation networks

Alexander A. Ganin,1,2 Maksim Kitsak,3 Dayton Marchese,2 Jeffrey M. Keisler,4 Thomas Seager,5 Igor Linkov2*

Urban transportation systems are vulnerable to congestion, accidents, weather, special events, and other costly delays. Whereas typical policy responses prioritize reduction of delays under normal conditions to improve the efficiency of urban road systems, analytic support for investments that improve resilience (defined as system recovery from additional disruptions) is still scarce. In this effort, we represent paved roads as a transportation network by mapping intersections to nodes and road segments between the intersections to links. We built road networks for 40 of the urban areas defined by the U.S. Census Bureau. We developed and calibrated a model to evaluate traffic delays using link loads. The loads may be regarded as traffic-based centrality measures, estimating the number of individuals using corresponding road segments. Efficiency was estimated as the average annual delay per peak-period auto commuter, and modeled results were found to be close to observed data, with the notable exception of New York City. Resilience was estimated as the change in efficiency resulting from roadway disruptions and was found to vary between cities, with increased delays due to a 5% random loss of road linkages ranging from 9.5% in Los Angeles to 56.0% in San Francisco. The results demonstrate that many urban road systems that operate inefficiently under normal conditions are nevertheless resilient to disruption, whereas some more efficient cities are more fragile. The implication is that resilience, not just efficiency, should be considered explicitly in roadway project selection and justify investment opportunities related to disaster and other disruptions.

INTRODUCTION

Existing roadway design standards emphasize the efficient movement of vehicles through a transportation network (1–4). Efficiency in this context may include identification of the shortest or fastest route (1, 5–7), or the route that minimizes congestion (8). It is the primary criterion on which road networks are modeled and design alternatives are considered (6, 7, 9, 10). The Texas A&M Transportation Institute defines and reports traffic delay in urban areas as the annual delay per auto commuter (11). Other studies define efficiency as delay for the individual driver in terms of time spent moving or stopped (7), or mean travel time between all origin-destination pairs in the network (9). However, as the experience of any motorist in large American cities can attest, conditions beyond the scope of the roadway design, including congestion, accidents, bad weather, construction, and special events (for example, a marathon race), can cause costly delays and frustrating inefficiencies that result in fuel waste, infrastructure deterioration, and increased pollution (12, 13). Evaluating road networks based only on efficiency under normal operating conditions results in little to no information about how the system performs under suboptimal or disrupted conditions.

Infrastructure systems that exhibit adaptive response to stress are typically characterized as resilient (14–21). Given the essential role of transportation in emergency response, provision of essential services, and economic well-being, the resilience of roadway networks has received increasing policy attention. Nonetheless, scholars have yet to converge on a shared understanding of resilience suitable to guide design, operation, and reconstruction of roadway networks. Although resilience in infrastructure systems is characterized as a multidimensional concept (22, 23), in many engineering and civil infrastructure implementations, resilience is defined as the ability of a system to prepare for, absorb, recover from, and adapt to disturbances (16). Specific to transportation, resilience has been defined as “the ability of the system to maintain its demonstrated level of service or to restore itself to that level of service in a specified timeframe” (24). Others describe transportation resilience as simply the ability of a system to minimize operational loss (25) or use the term synonymously with robustness, redundancy, reliability, or vulnerability (26–28).

Current efforts in transportation resilience research have focused on framework development and quantification methods. These efforts include the specification of resilience indicators, such as total traffic delay (24), economic loss (29), post-disaster maximum flow (30), and autonomous system components (31). Practical concerns with this type of resilience evaluation are that it relies on uncertain performance data and often omits indicators that are unquantifiable (19). Other resilience approaches apply traffic network modeling to identify locations for critical buildings (for example, hospitals and fire stations) (32), minimize trip distance for individual passengers (33), and minimize travel time across the system (12). One drawback of existing network resilience methods is that they are data-intensive, often requiring limited information about resources for unusual road system repair (26, 28) or network behavior following a disruptive event (34). Moreover, existing resilience quantification approaches lack calibration and testing across a range of transportation systems. Because many disruptive events, and their associated consequences, are difficult to predict, resilient road systems must be characterized and evaluated by the capacity to adapt to a variety of different stress scenarios. Partly because of these obstacles, joint consideration of efficiency and resilience has yet to be implemented for transportation networks.

Here, we study the interconnections between resilience and efficiency (20) among road transportation networks in 40 major U.S. cities. We develop an urban roadway efficiency model, calibrate it on the basis of the observed data (11) of annual delay per peak-period auto commuter, and apply the model to calculate efficiency in 40 cities.
Then, we model traffic response to random roadway disruptions and recalculate expected delays to determine the sensitivity of each city to loss of roadway linkages. The results may reveal important considerations for assessing proposals for improvement of roadway infrastructure that maintain efficiency under stress conditions.

METHODS
The Methods section appears here to help clarify the subsequent sections. To develop the urban roadway efficiency model, we defined the urban area boundaries, constructed the road networks, and evaluated the population density within cities using the Census Bureau data sets (35, 36) and OpenStreetMap (OSM) data sets (37). We relied on these data to assess commuter patterns, which we used to measure efficiency and resilience of road networks.

Alternative approaches to transportation have been offered and include those based on percolation theory and cascading failures (38–40), human mobility pattern studies (41–43), queueing (44, 45), and the use of historical data to predict traffic. We review these approaches in the Supplementary Materials and note that the main benefit of our model is that it relies solely on readily available public data, rather than on particular data sets that may or may not be practical to obtain for any particular region. The model’s algorithmic simplicity allows us to consider spatial topologies of cities in high resolution including tens of thousands of nodes and links. We did not create a more accurate transportation model than the existing ones, but we were able to obtain measurable characteristics of transportation systems (average delays) using our model.

Geospatial boundaries and population density
To define geospatial boundaries for the transportation infrastructure networks, we used the U.S. Census Bureau geospatial data set (35) for urban areas—densely developed residential, commercial, and other nonresidential areas (46). We approximated the exact urban area polygon with a simplified manually drawn one (Fig. 1A) and included all roadways within 40 km (25 miles) of it in the network. For each of the links, we calculated its length on the basis of the polyline defining the link and assigned a number of lanes $m$ and the FFSs (see the Supplementary Materials).

We next estimated population in vicinity of each intersection $i$ using the Census Tract data (36). To this end, we split the map into Voronoi cells centered at intersections and then evaluated the population of each cell $N_i$ as

$$N_i = \sum_t N_t \frac{\text{Area}(P_t \cap P_i)}{\text{Area}(P_t)}$$

Above, $N_t$ is the population of Census Tract $t$, and $P_i$ and $P_t$ are the polygons of the cell and the tract, respectively (Fig. 1B and table S2).

Transportation model
We built on the gravity model to generate commuting patterns. The gravity model (47) is a classical model for trip distribution assignment and is extensively adopted in most metropolitan planning and statewide travel demand models in the United States (48–51). Other trip
distribution models include, for example, destination choice models (52, 53). However, these models are not as widely used in large scale, because the detailed data required by these models are frequently unavailable (48).

We assumed that (i) the flow of commuters from origin region $o$ to destination region $d$ is proportional to the population at the destination $N_d$ and that (ii) the flow of commuters depends on the distance $x_{od}$ between the origin and destination and is given by a distance factor, $P(x_{od})$. Using these assumptions, we assessed the fraction of individuals commuting from region $o$ to destination region $d$, $f_{od}$ as

$$f_{od} = \frac{N_d P(x_{od})}{\sum_k N_k P(x_{ok})} \quad (2)$$

Then, the commuter flow from origin region $o$ to destination region $d$ is

$$F_{od} = N_d f_{od} \quad (3)$$

Although individual driving habits may vary (54), we assumed that all drivers tended to optimize their commute paths such that their travel time was minimized. This assumption allowed us to calculate commute paths for every origin-destination pair using inferred FFSs. To calculate commuter flows between all pairs of intersections, we estimated distances $x_{od}$ as the distance of the shortest time path from $o$ to $d$. Furthermore, in place of the distance factor $P(x_{od})$, we used the distribution of trip lengths from the U.S. Federal Highway Administration National Household Travel Survey (55, 56), which we approximated with the exponential function (Fig. 2A and table S3).

Next, we defined the commuter load on each road segment as

$$L_{ij} = \sum_{o,d} F_{od} \theta_{od}(ij) \quad (4)$$

where $\theta_{od}(ij)$ is a binary variable equal to 0 when the link $ij$ is not on the shortest path connecting nodes $o$ and $d$, and 1 otherwise. Note that in Eq. 4, we only considered origins that were not farther than 30 km from the urban area boundary polygon. The nodes farther than 30 km from the boundary were only used as destinations to evaluate the fraction of commuters not going toward the urban area (Eq. 2).

Because most commuters travel during peak periods, commuter loads $L_{ij}$ can be regarded as traffic-based centrality measures estimating the number of individuals using corresponding road segments. Then, the cumulative time lost by all commuters is

$$\Delta T = \beta \sum_{ij} L_{ij} \left( \frac{\left( l_j + l_0 \right)}{v_{ij}} - \frac{\left( l_j + l_0 \right)}{V_j} \right) \quad (5)$$

where $V_j$ and $v_{ij}$ are, respectively, the FFS and the actual traffic speed along the $ij$ road segment, $l_j$ is its length, $l_0$ is the length correction due to traffic signals, and $\beta$ is the proportionality coefficient same for all urban areas. The summation in Eq. 5 includes only links, whose origins and destinations are within the boundary polygon. A similar equation was obtained for the moving delay in the study of Jiang and Adeli (45), where the authors looked at the delay induced from road repairs.

The actual traffic speed $v_{ij}$ depends on many factors including the speed limit, the number of drivers on the road, and road conditions. Although there exist a number of approaches to estimate actual traffic speed (57, 58), we chose to use the Daganzo model (59) to derive the traffic speed, as shown in the Supplementary Materials.

$$v_{ij} = \alpha \frac{l_ij m_{ij}}{L_{ij}} - v_{veh}, \text{ subject to } v_{ij} \in [v_{min}, V_j] \quad (6)$$

where $v_{min}$ is the minimum speed in the traffic, $v_{veh}$ is the correction for the finite size of the car, and $\alpha$ is the proportionality coefficient (Fig. 2B).

**Efficiency and resilience metrics**

We measured efficiency as the average annual delay per peak-period auto commuter. In practice, lower delay means higher efficiency. There are multiple ways to map from delays to efficiency, such as taking the inverse values of delays, taking negative values of delays, etc. To avoid ambiguity and facilitate the interpretation of results, we used the delays themselves to quantify the transportation efficiency of urban areas.

We operationalized resilience through the change in traffic delays relative to stress, which is modeled as loss or impairment of roadway...
linkages. Looking at resilience from the network science perspective, we focused on topological features of cities, rather than on recovery resources available. Sterbenz et al. (60) evaluated a network’s resilience as a range of operational conditions for which it stays in the acceptable service region and highlighted that remediation mechanisms drive the operational state toward improvement. We are studying how availability of alternate routes helps remediate the consequences of the initial disruption to the network. In the traffic context, the immediate impact of a given physical disruption (and the time for it to unfold) in terms of closing lanes or reducing speed limits on affected roads will not vary much from network to network, although the number and type of these disruptions will. Likewise, the speed of restoring full functionality (through action in the physical domain) is not so much dependent on the road network as it is on the nature of the disruption (snow versus earthquake versus flood) and the resources that the city allocates to such repair. The level of functionality that these repairs achieve ought to be the full predisruption functionality, that is, eventually all roads can be fully cleared or restored. However, the immediate loss of function for a given traffic flow can very quickly be partially recovered after a disruption by action in the information domain, namely, rerouting of traffic. From the new steady state at that level of functionality, full functionality is gradually restored. Thus, our model proxies for resilience and is calibrated against the data that proxy for efficiency. At the same time, we note that to fully capture resilience characteristics of a transportation system, it is required to analyze recovery resources available and the effectiveness of coordination between the relevant authorities. Lower additional delay corresponds to higher resilience, but using the same reasoning that we had for efficiency, we quantified resilience through additional delays.

RESULTS

Efficiency

Together, our traffic model has three parameters (proportionality coefficient $a$, minimum speed $v_{\text{min}}$, and finite vehicle size correction $v_{\text{veh}}$) and is summarized in Eqs. 5 and 6. Given parameter values of the model, one can estimate the total delay incurred by all commuters in any given suburban area or, equivalently, the average delay per commuter. We take $v_{\text{veh}} = 9 \text{ km/h}$ and $v_{\text{min}} = 5 \text{ km/hour}$ and calibrate the model to determine the value of $a$ to match the real data on the annual average delay per peak-period auto commuter provided by the Urban Mobility Scorecard (11).

We divide the 40 urban areas into two equally sized groups for model calibration and validation, respectively. We have found that for the 20 urban areas used for calibration, the $R$-squared coefficient took values in the range ($-0.01$ to $0.83$) (Fig. 3 and Supplementary Materials). This allows us to set model parameters $a$ and $b$ (see Methods) as follows: $a = 4.30 \times 10^4 \text{ hour}^{-1}$ and $b = 10.59$. These values correspond to the Pearson coefficient of 0.91 ($P = 2.17 \times 10^{-8}$).

To validate the model, we estimate travel delays in 20 different urban areas. As seen from Fig. 3, the estimated travel delays are

![Fig. 3. Modeled and observed delays in 40 urban areas. Pearson correlation coefficients and $P$ values between observed and modeled delays are (0.91, 2.17 $\times 10^{-8}$) for the 20 cities used to calibrate the model and (0.63, 3.00 $\times 10^{-3}$) for the 20 cities used to validate the model. Observed delays were taken from the Texas A&M Transportation Institute Urban Mobility Scorecard (11).](http://advances.sciencemag.org/content/3/10/e1701079/F3)

significantly correlated ($R = 0.63, P = 3.00 \times 10^{-3}$) with actual delay times ($11$), validating the transportation model. Figure 4 is a Google Maps representation of real and modeled results for Los Angeles and San Francisco. Road conditions under real, average traffic patterns at 8 a.m. provided by Google Maps are in Fig. 4 (A and D). Modeled conditions are given for comparison in Fig. 4 (B and E). Finally, Fig. 4 (C and F) shows the new, modeled traffic patterns that result from redistribution of travel in response to a disruption of 5% of the links.

Resilience

Our approach to model stress is inspired by percolation theory. For every independent simulation of stress, we select a finite fraction of affected road segments $r$ at random, with the probability of failure proportional to segment length. We collect statistics for 20 realizations of the percolation. On failed segments, free-flow speeds (FFSs) are reduced to 1 km/hour (representing near-total loss), and loads $L$ and traffic delays are then recalculated using the updated FFSs. Low-stress scenarios ($r < 0.1$) might be caused by accidents or construction. Larger disruptions might occur during power failures that disrupt traffic signals or severe flooding that makes many roadways nearly impassable. Finally, widespread stress might be caused by snow, ice, or dust storms that affect nearly the entire roadway system. Figure 5 displays the analysis of delay times in six representative urban areas for the full spectrum of adverse event severities, $r \in [0; 1]$. In addition, fig. S5 shows the results for all urban areas. Some routes within a single urban area experience longer delays than others. The inset of Fig. 5 shows the delay distribution for both Los Angeles, which is narrowly clustered, and Boston, where greater variability between roadways is evident.

Traffic delay times grow rapidly as $r$ increases and reach saturation (all routes moving at 1 km/hour) as $r$ approaches 1. We determine the most resilient urban transportation network to be Salt Lake City, UT, whereas the least resilient among the 40 metropolitans is shown to be Washington, DC.

Figure 6 shows both the efficiency (in blue) and resilience response (additional delays due to 5% link disruption, in orange) for the 40 urban areas modeled. Some cities with high efficiency under normal operating conditions (that is, low delays) nevertheless exhibit low resilience (that is, a sharp increase in traffic delays) under stress. Virginia Beach, VA; Providence, RI; and Jacksonville, FL all fall into this category of urban areas in which traffic operates well under ordinary circumstances but rapidly become snarled under mild stress. On the other hand, Los Angeles is notorious for traffic delays under all conditions—and minor stress levels result in little degradation of efficiency. By contrast, normal traffic delays in San Francisco are comparable to Los Angeles, but mild stress in San Francisco results in large increases in additional delays. These examples indicate that resilience (that is, additional delay response to stress) is independent of normal operating efficiency.

DISCUSSION

The disturbances affecting the road infrastructure are often complex, and their impact on the structure and function of roadway systems may be unknown ($28, 31$). These disturbances might be natural and irregular, such as distributed road closures caused by an earthquake or a snowstorm. The disturbances might also be anthropogenic and intentional, such as a...
street fair or marathon race. Whatever the disturbance, the results of this analysis allow several meaningful inferences to be made that may have important implications for highway transportation policy. The first is that resilience and efficiency represent different aspects related to the nature of transportation systems; they are not correlated and should be considered jointly as complementary characteristics of roadway networks.

Second, there are characteristic differences in the resilience of different urban areas, and these differences are persistent at mild, medium, or widespread levels of stress (Fig. 5). Except for San Francisco,
CA, which is the most fragile of all cities represented in Fig. 5 at stress levels $r < 20\%$ but then surpassed by Boston, MA and Washington, DC, the rank ordering of urban area resilience is insensitive to stress levels. That is, cities that exhibit relatively low resilience under mild stress are the same cities that exhibit low levels of resilience (relative to peers) under widespread roadway impairment. This suggests that the characteristics that impart resilience (such as availability or alternate routes through redundancy of links) are protective against both the intermittent outages caused by occasional car crashes and those caused by snow and ice storms. For cities without resilience, a widespread hazard such as snow may lead to a cascade of conditions (for example, crashes) that rapidly deteriorate into gridlock. This was exactly the case for Washington, DC 20 January 2016 under only $2.5 \times 10^{-2} \text{ m}$ or $2.5 \text{ cm}$ of snow (61), and for Atlanta, GA 2 years earlier, which experienced $5.1 \times 10^{-2} \text{ m}$ or $5.1 \text{ cm}$ of snow in the middle of the day that resulted in traffic jams that took days to disentangle (62). Whereas popular explanations of these traffic catastrophes focus on the failure of roadway managers to prepare plows and emergency response equipment, Fig. 5 suggests that cities with similar climates (Memphis, TN and Richmond, VA) are less likely to be affected, regardless of the availability of plow or sand trucks.

The third inference follows from Fig. 6, which suggests that urban areas that make capital investments to reduce traffic delays under normal operating conditions may nevertheless be vulnerable to traffic delays under mild stress conditions. Because these stressors are inevitable, whether from crashes, construction, special events, extreme weather, equipment malfunctions, or even deliberate attack, investment strategies that prioritize reduction of normal operating delays may have the unintended consequence of exacerbating tail risks—that is, the risk of worse catastrophe under unlikely but possible conditions.

Finally, the exceptional position of New York City in Fig. 3 calls attention to the fact that substitutes for roadway transportation are available in many cities and have an important role to play in relieving traffic congestion. According to the Texas A&M Institute (63, 64), public transit reduces delays per peak-period auto commuter in the New York urban area by 63 hours, in Chicago by 23 hours, and by less than 20 hours in other urban areas. Because our model considers only roadway transit, and New York City contains a myriad of nonroad-based options to avoid roadway congestion, it is unlikely that our model can provide informative results for the New York urban area.

Although interest has increased in policies that enhance roadway resilience, few analytic tools are available to guide new investments in achieving resilience goals. It is widely understood that roadway infrastructure is expensive, both in acquiring land for rights-of-way and in construction of improvements, and thus, decisions regarding alignment, crossing, and access made over a period of decades may have long-lasting consequences that are observable in traffic data today. Consequently, urban areas exhibit different unintentional traffic characteristics, including delays under normal and random stress conditions. Investments motivated exclusively by expected efficiencies under normal operating conditions are unreliable safeguards against loss of efficiency under stress conditions. Therefore, new analytic tools are required that allow designers to assess the adaptive capacity of roadway infrastructure and assess the potential of new investments to provide enhanced resilience. The adaptive network-based model described herein is one such approach.

**SUPPLEMENTARY MATERIALS**

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/3/12/e1701079/DC1

**ALTERNATIVE APPROACHES TO MODEL TRANSPORTATION**

Mapping from OSM Foundation shapefiles to network nodes and links

**POPULATION ASSIGNMENT ALGORITHM**

Distance factor of the likelihood of travel between nodes

**ESTIMATION OF THE TRAFFIC SPEED FROM THE DENSITY OF VEHICLES**

Model calibration procedure

**SENSITIVITY OF THE MODEL TO RAMP SPEEDS**

Additional delay as a function of the severity of link disruption

**Density-flow relationship in the Daganzo traffic model**

**Model calibration**

**Modeled delays for ramp speed coefficients of $1/3$ and $1/2$.**

**DEPENDENCY OF THE ADDITIONAL DELAY ON THE SEVERITY OF THE LINK DISRUPTION FOR ALL 40 URBAN AREAS**

**REFERENCES AND NOTES**


2. K. Beverly, Efficient Use of Highway Capacity (FHWA-HOP-10-023, Texas Transportation Institute, 2010), p. 100.


34. S. Simonovic, From risk management to quantitative disaster resilience. 2010 Census Data.
35. Planet dump (OpenStreetMap contributors, 2017); https://planet.openstreetmap.org.

Acknowledgments: We would like to thank S. Buldyrev (Yeshiva University) and J. Palma-Oliveira (University of Lisbon) for their insightful comments. Funding: This study was supported by the U.S. Army Engineer Research and Development Center and by the Defense Threat Reduction Agency, Basic Research Program (P. Tandy, program manager). A.A.G. was additionally supported by the Virginia Transportation Research Council and Virginia Department of Transportation. T.S. was supported by the NSF under grant no. 1411352. Author contributions: A.A.G., M.K., and IL conceived the model and designed the simulations. A.A.G. developed software and performed data retrieval and simulations. A.A.G. and M.K. analyzed results. IL provided senior guidance. A.A.G., MK, JMK, T.S., and IL wrote the paper and contributed to the interpretation of the results.

Competing interests: The authors declare that they have no competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors. Map data were copyrighted by OSM contributors and are available at www.openstreetmap.org.