PNAS NEXUS

PNAS Nexus, 2025, 4, pgaf075
https://doi.org/10.1093/pnasnexus/pgaf075
Advance access publication 3 March 2025
Physical Sciences and Engineering

Hidden high-risk states identification from routine urban traffic

Shiyan Liu[a], Mingyang Bai[a], Shengmin Guo[b], Jianxi Gao[c], Huijun Sun[d], Zi-You Gao[d,*] and Daqing Li[a,*]

[a]School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China
[b]State Key Laboratory of Software Development Environment, Beihang University, Beijing 100191, China
[c]Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
[d]School of Systems Science, Beijing Jiaotong University, No. 3 Shangyuancun Haidian District, Beijing 100044, China
*To whom correspondence should be addressed: Email: zygao@bjtu.edu.cn (Z.-Y.G.); Email: daqingl@buaa.edu.cn (D.L.)
Edited By Attila Szolnoki

Abstract

One of the core risk management tasks is to identify hidden high-risk states that may lead to system breakdown, which can provide valuable early warning knowledge. However, due to the high dimensionality and nonlinear interactions embedded in large-scale complex systems like urban traffic, it remains challenging to identify hidden high-risk states from huge system state space where over 99% of possible system states are not yet visited in empirical data. Based on the maximum entropy model, we infer the underlying interaction network from complicated dynamical processes of urban traffic and construct the system energy landscape. In this way, we can locate hidden high-risk states that may have never been observed from real data. These states can serve as risk signals with a high probability of entering hazardous minima in the energy landscape, which lead to huge recovery cost. Our findings might provide insights for complex system risk management.

Keywords: system risk, maximum entropy model, early warning signals, risk management, urban traffic

Significance Statement

Hidden high-risk states are lethal system states with catastrophic consequences for large-scale complex systems. Existing methods for identifying hidden high-risk states face challenges in routine urban traffic, where over 99% of possible system states are not yet visited in empirical data. We utilize large-scale traffic speed data to learn a least structured network entropy model that can capture the intricate traffic interactions. This model can generate the entire state space and help the identification of hidden high-risk states that may lead to critical system disruptions. Our results suggest a new way to provide reliable early warning signals for system breakdown based on critical risk scenarios.

Introduction

Hidden high-risk states are system states that have not occurred yet but are likely to happen in the future, with catastrophic consequences. Once these states occur, they may lead to system breakdowns (1) in natural, technical, and social systems, such as earthquakes (2), tsunamis (3), financial crises (4), blackout (5), extreme climate (6), and neuronal avalanches (7). In urban traffic, system breakdown is generally understood as conditions in which traffic is completely or largely paralyzed, leading to enormous damage. For example, on 2012 July 21, Beijing was hit by a violent rainstorm that paralyzed the backbone of its road network (8). According to an empirical analysis of 2017 data, during peak hours, most roads in the San Francisco Road network operated at low levels of service, resulting in a significant breakdown in traffic flow (9). In future traffic management scenarios, with more connected vehicles (10) on the road, cyberattacks against smart traffic management systems may lead to widespread traffic paralysis (11). To prevent these system breakdowns from occurring, the identification of hidden high-risk states is needed. However, the state space of an urban traffic system is so huge that traditional methods can hardly recognize the small number of states that lead to breakdown. One of the essential questions is how to effectively identify these hidden high-risk states that serve as precursors to system breakdown in urban traffic.

To effectively identify hidden high-risk states that may lead to system breakdown, researchers have dedicated extensive efforts. There are several traditional methods for identifying system breakdown which mainly focus on statistics and indicators, such as extreme value theory (12) and large deviation theory (13). To further address this challenge, a wide range of advanced

Competing Interest: The authors declare no competing interests.
Received: September 25, 2024. **Accepted:** February 19, 2025

methods have been developed, which can be broadly classified into data-driven methods or model-driven methods. In the first approach, historical spatiotemporal observations are used to determine whether and when system breakdown will occur without knowing the system dynamic models. Examples include forecasting extreme floods using an event synchronization measure (14), global seasonal forecasts of marine heatwaves based on 30 years of historical observational data (15), system breakdown prediction via active learning (16), and deep learning-based methods (17, 18). The second approach is based upon identified dynamic models of systems to uncover breakdown mechanisms, such as bursting phenomena of multiple timescales in deterministic systems (19, 20), and noise-induced transitions to extreme state attractor in nondeterministic systems (21–23). In the above approaches, data-driven methods often require extensive sample data, while the highly cyclical and repetitive nature of daily traffic patterns (24) results in an exceptionally low occurrence of system breakdowns. Meanwhile, current model-driven approaches typically face challenges in complex systems like urban traffic, which are controlled by a large family of hidden correlated parameters. To overcome these issues, we propose a method for identifying hidden high-risk states in urban traffic using a pairwise maximum entropy model, which captures intricate traffic interactions and reconstructs the entire state space.

Pairwise maximum entropy models (25, 26), originally developed in statistical mechanics and information theory, have been applied to complex systems to uncover internal interaction relationships, such as finding weak interactions between neurons behind strongly collective behavior (27), deducing the network structure of brain functional areas (28), and analyzing gene–gene interaction networks underlying cell metabolism (29). In traffic systems, researchers have also used pairwise maximum entropy model to predict local traffic flow (30), but they did not consider how interactions between regions could lead to system breakdown of urban traffic. In our work, we use spatiotemporal data of urban road network to learn the pairwise maximum entropy model, inferring the underlying interaction network between local regions. Based on this model, we define hidden high-risk states and identify them by constructing system energy landscape and analyzing their failure processes. Our proposed method can identify hidden high-risk states from routine urban traffic, providing a list of potential management targets often overlooked in traffic management practices. The identification of extreme outliers mentioned here is based on the relatively stable correlation patterns observed in daily data. In addition, the model is fitted by considering interactions between local regions rather than individual road segments (reducing the system state space from $2^{33,000}$ to 2^{20}), which avoids the issue of dimensionality explosion. This approach makes it possible to identify hidden high-risk states more effectively and analyze their dynamic properties across the entire state space. Our results suggest a new way to provide reliable early warning signals for system breakdowns and reconstruct critical risk scenarios for traffic management practices.

Results

Maximum entropy model for urban traffic

Here, we consider a specific local road network area as a hexagonal region (see Supplementary Note S1), as shown in Fig. 1a. The overall road network states are represented by the binary states of 20 regions (i.e. free region or jammed region, see Materials and methods), resulting in a total of 2^{20} possible system states. As a large-scale and high-dimensional complex system, there are a huge number of possible states within the state space of urban traffic. Despite advances in traffic big data technology that enable real-time monitoring of urban traffic states, the highly repetitive nature (24) of daily traffic means that only a small portion of the state space can be observed. In real data, we find that the observed states (Fig. 1b and c) during rush-hour periods (i.e. 7:30–8:30 AM) from 1 month of working days are fewer than 10^3 (see Materials and methods), which is only 0.1% of the entire state space. This implies that over 99% of possible states have not yet been visited in empirical data. Therefore, relying solely on real-time data makes it nearly impossible to detect all states of the entire road network (32), especially when it comes to identifying hidden high-risk states. To effectively identify these hidden high-risk states of urban traffic, a comprehensive understanding of the entire state space is required, the core of which is to infer a dynamical model capable of capturing complex system interactions.

To address this issue, we use a pairwise maximum entropy model to capture the interaction network (Fig. 1d) between different local road network regions based on real data. The model parameters are determined by maximizing the (Gibbs) entropy H of the joint probability distribution, subject to the normalization condition and constraints of the first and second moments (see Materials and methods for more details):

$$\text{Max} \quad H = -\sum_{1\le k\le n} p(s_k)\log p(s_k), \tag{1}$$

where s_k represents a global road network state, n is the number of road network states in the entire state space, and $p(s_k)$ is the probability of road network state s_k. This probability distribution can be obtained by solving the Lagrange function (33) (Supplementary Note S2) as:

$$p(s_k) = (1/Z)e^{\sum_{1\le i\le m} h_i s_k^i + \sum_{1\le i<j\le m} J_{ij} s_k^i s_k^j}, \tag{2}$$

where m is the number of regions and s_k^i is the state of region i with a value of +1 (i.e. jammed region) or −1 (i.e. free region) when the system is in state s_k, and Z is the partition function, which equals $\sum_{1\le k\le n} e^{\sum_{1\le i\le m} h_i s_k^i + \sum_{1\le i<j\le m} J_{ij} s_k^i s_k^j}$. The pairwise maximum entropy model is the least structured model incorporating second-order interaction terms, which can be mapped onto the Ising model (27). The model parameter h_i represents the congestion tendency of local road network region i, while J_{ij} characterizes the interaction between region i and j. Specifically, a positive J_{ij} value indicates a positive interaction (27), implying that congestion in one region is likely to trigger congestion in the other. Conversely, a negative J_{ij} value represents a negative interaction (27), suggesting that congestion in one region can alleviate congestion in the other. These model parameters (i.e. h_i and J_{ij}) can be estimated from real data (see Materials and methods and Fig. S1). The input data for the model consist of global road network states, aggregated from raw road segment-level speed data during a specific time period across 17 workdays in October 2015 (see Materials and methods). Here, we focus on the rush-hour period (i.e. 7:30–8:30 AM) and nonrush-hour period (i.e. 6:00–7:00 AM). We then use J_{ij} to infer the underlying interaction network between local regions (Fig. 1d), which exhibits high heterogeneity during rush-hour periods. For example, region 20 tends to have more positive interactions with other regions, while region 15 tends to have more negative interactions. The heterogeneity embedded in underlying interaction network of urban traffic can lead to highly complicated dynamic behavior (34).

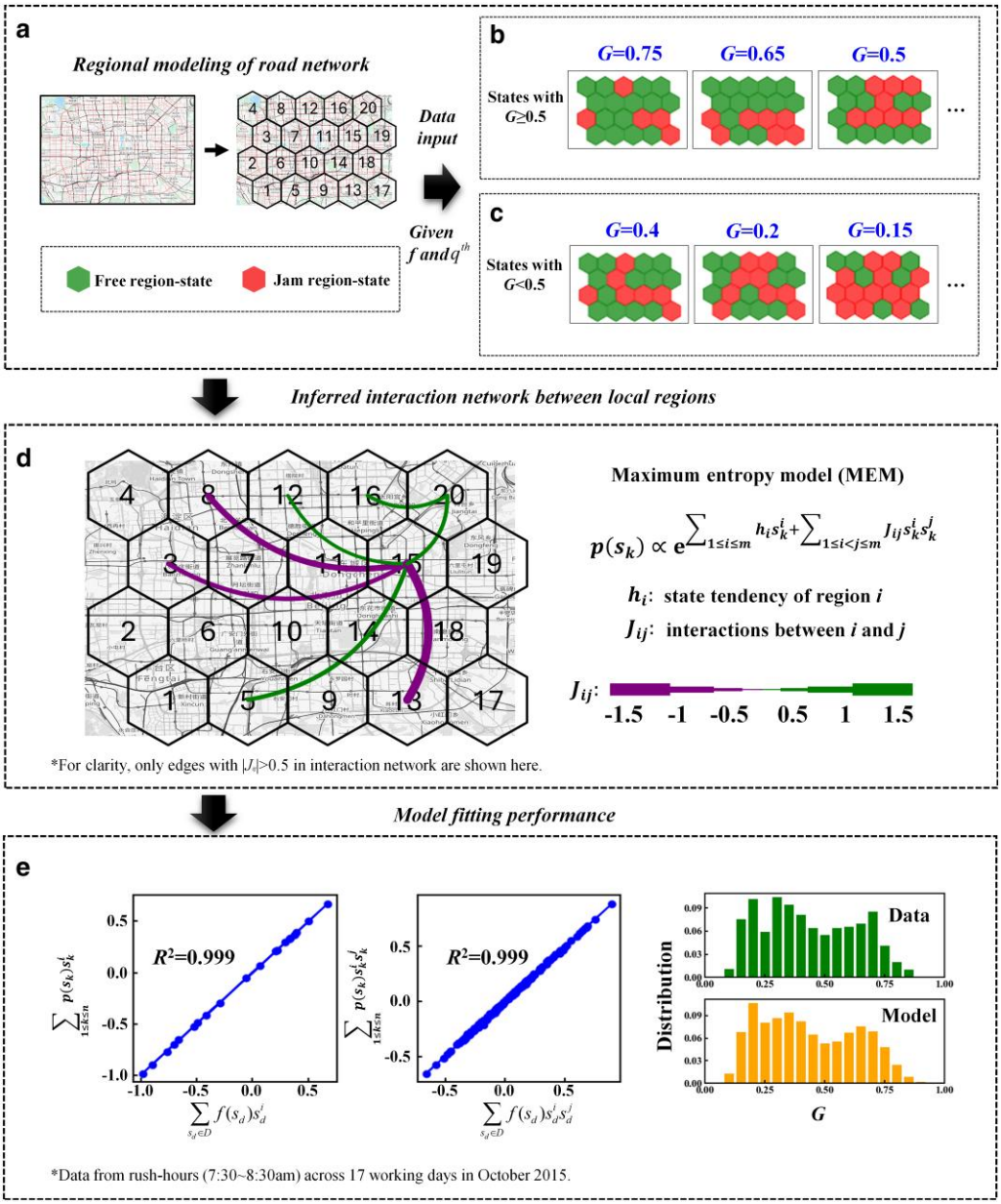

Fig. 1. Pairwise maximum entropy model for urban traffic. a) Regional modeling of the road network within the Fourth Ring Road of Beijing (© OpenStreetMap contributors, https://www.openstreetmap.org/copyright). The entire road network is divided into 20 equal-sized, regular hexagonal local regions. The state of each local region can be either free or jammed based on the corresponding local road network (see Materials and methods). b) Typical road network states with $G \geq 0.5$ observed in real data. Traffic functional performance (G) of road network states is defined as the ratio of the largest functional cluster (31) formed by connected free regions to the total number of regions. States with $G \geq 0.5$ represent good traffic operational connectivity. c) Typical road network states with $G < 0.5$ observed in real data. These states represent poor traffic operational connectivity. d) Inferred interaction network between local regions based on the pairwise maximum entropy model. Model parameters J_{ij} inferred from real data represent interactions between local regions. e) Model goodness of fit. The fitting performance of the model to rush-hour data (7:30–8:30 AM) is good (see main text for details).

The pairwise maximum entropy model constructed for urban traffic has good performance. At the microscopic level, the two-moment constraints derived from the model match well with actual data during rush-hour periods (Fig. 1e, with $R^2 > 0.99$) and nonrush-hour periods (Fig. S2a, with $R^2 > 0.99$). At the macroscopic level, given that the number of observed states is significantly smaller than the total state space, directly validating the model by comparing the probability of each system state between the data and the model is not feasible (26, 27). Instead, previous studies (35) have validated maxent models by examining the consistency of key macroscopic property distributions between the model and observed data. Here, we select three key macroscopic properties of traffic network—traffic functional performance ($G^{\text{free}}_{\text{region}}$, hereafter referred to as G), congestion concentration degree ($G^{\text{jam}}_{\text{region}}$), and the proportion of jammed region ($F^{\text{jam}}_{\text{region}}$), with their calculation methods detailed in the Materials and methods section. As shown in Fig. S3, the distributions of three macroscopic properties demonstrate a high degree of consistency between the data and the model, with R^2 values of 0.949, 0.977, and 0.982 during rush-hour periods and 0.997, 0.999, and 0.997 during nonrush-hour periods. Notably, the model successfully reproduces the multimodal characteristic observed in the distribution of traffic

functional performance (i.e. G) during rush-hour periods (Fig. 1e), which indicates the presence of multistable states in the road network during rush hours (36). When we change the parameters that determine local region states (Fig. S2c), coarse-graining methods (Figs. S4–S6), and alternative urban road networks (Figs. S10 and S11), the model is still able to capture different distribution patterns of traffic functional performance. Additionally, we observe a relatively consistent distribution of traffic functional performance for different number of regions (Supplementary Note S3 and Fig. S8), indicating the feasibility of the coarse-graining approach applied to urban traffic.

Although the pairwise maximum entropy model constrains only the first moments (i.e. the average state of each region) and second moments (i.e. the average correlation between pairs of regions), it successfully reproduces several key macroscopic features of the system. This indicates that the pairwise maximum entropy model can effectively capture the local interactions underlying macroscopic behaviors of urban traffic system. Unlike traditional correlation methods including Pearson's correlation coefficient, our pairwise maximum entropy model considers global system behavior as a result of the inferred interaction network embedded in urban traffic, rather than just linear relationships between pairs of regions. This interaction network ensures consistency of global system properties (e.g. the distribution of traffic functional performance) between the model and actual data, thereby better capturing the intrinsic properties of urban traffic. Moreover, the interaction pattern between local regions (i.e. J_{ij} distributions) inferred by the pairwise maximum entropy model is robust across different coarse-graining methods (Fig. S7), numbers of regions (Fig. S9), and different samples (Figs. S15 and S23).

Hidden high-risk states defined in system state transition network

To clarify the definition of hidden high-risk states, we first construct entire system state space of urban traffic, which can be represented by a state transition network. As shown in Fig. 2a, a node in the state transition network represents a road network state s_k ($k = 1, 2, 3, \ldots, 2^{20}$), with edges representing state transitions (i.e. one-region-state-flip). According to the traffic functional performance (i.e. G, defined as the ratio of the largest functional cluster (31) formed by connected free regions to the total number of regions), road network states can be classified as normal states with $G \geq 0.5$ (e.g. s_1, s_4, and s_5 in Fig. 2a) or hazardous states with $G < 0.5$ (e.g. s_2, s_3, and s_6 in Fig. 2a). Among them, some states (e.g. s_3, s_4, and s_5 in Fig. 2a) have never been observed in empirical data (i.e. hidden states), making them potential risky states that traffic managers may overlook. Here, the observed states are aggregated from rush-hour road segment-level data during 17 workdays in October 2015, with the congestion ratio $f = 0.25$ and performance threshold $q^{\text{th}} = 0.09$ (see Materials and methods for more details). When urban traffic is in normal states (e.g. s_5 in Fig. 2a) with also many neighboring normal states, there is a high probability that it will remain normal, representing stable healthy system operation. In contrast, normal states (e.g. s_4 in Fig. 2a) surrounded by many hazardous states are highly likely to fall into hazardous states, signaling that the system is at risk boundary of hazardous states. Regarding hazardous states, those like s_6 in Fig. 2a with normal neighbors, may spontaneously recover to normal, reflecting system resilience. Hazardous states (e.g. s_3 in Fig. 2a) with neighboring hazardous states are likely to remain hazardous states, suggesting relatively stable hazardous condition that could lead to significant recovery cost for the system. Here, we aim to identify hidden high-risk states that currently appear normal but have a high probability of transitioning into relatively stable hazardous states. This dynamic characteristic of hidden high-risk states requires early intervention by traffic managers or smart control centers to prevent the system from persisting in poor traffic operational connectivity. The key question is how to explore the dynamical transition processes of hidden high-risk states.

Based on the maximum entropy model of urban traffic, the energy of all road network states can be calculated, which is a crucial parameter for describing the system dynamical behavior. Similar to the definition of energy in the Ising model, the energy value $E(s_k)$ of each global road network state s_k can be calculated as:

$$E(s_k) = -\sum_{1\leq i\leq m} h_i s_k^i - \sum_{1\leq i<j\leq m} J_{ij} s_k^i s_k^j, \tag{3}$$

where h_i and J_{ij} are parameters of the pairwise maximum entropy model, fitted using data from rush hours (i.e. 7:30~8:30 AM) of 17 workdays in October 2015. According to Eq. 2, the energy value $E(s_k)$ of road network state s_k represents its probability of occurrence $p(s_k)$:

$$p(s_k) \propto e^{-E(s_k)}, \tag{4}$$

indicating that a road network state with a lower energy value has a higher probability of occurrence. It can be seen from Eq. 4 that the occurrence probability distribution of road network states follows the same form as the Boltzmann distribution, meaning that energy of each road network state can be determined by counting its frequency. However, the observed states that account for <0.1% are quite sparse, making it impractical to calculate the energy of all road network states directly from the actual data. Hence, the fitted pairwise maximum entropy model is necessary to estimate the energy of states in the entire state space. More importantly, the energy values can be used to describe system transition dynamics, as the system tends to move from higher-energy states to lower-energy states, with the energy difference between adjacent states determining the transition direction.

To explore the energy characteristics of road network states, we analyze the energy distribution across all states within the urban traffic state space and find that there are a huge number of hidden states with high energy. This explains the observed rarity of these states in real data. As shown in Fig. 2b, the energy of all road network states is distributed over a wide range between −20 and 30. While a few observed states are concentrated around −15, there is still a fraction of hidden states with low energy that have never been observed in real data. For example, it is shown in Fig. 2b that there are 8,877 road network states with energy values below −14, of which only 570 are observed states, indicating that over 93% of these low-energy road network states remain unobserved in realistic operation. Compared with hidden states with higher energy, hidden states with lower energy have a greater occurrence probability according to Eq. 4. As a result, we can filter hidden high-risk states from low-energy states (i.e. high-p states), with energy lower than a certain threshold (E^{th}).

By combining above the dynamic characteristic with occurrence probability constraints, hidden high-risk states are defined as unobserved normal states with a high likelihood of occurring and falling into hazardous states (see Materials and methods). It is shown in Fig. 2c that high-p states (i.e. states with a high likelihood of occurring) consist of 17,121 road network states, about 1.6% of the total states, distributed across different values of G. Since states tend to move from high to low energy, hidden high-risk states will transition into lower-energy states, thus becoming

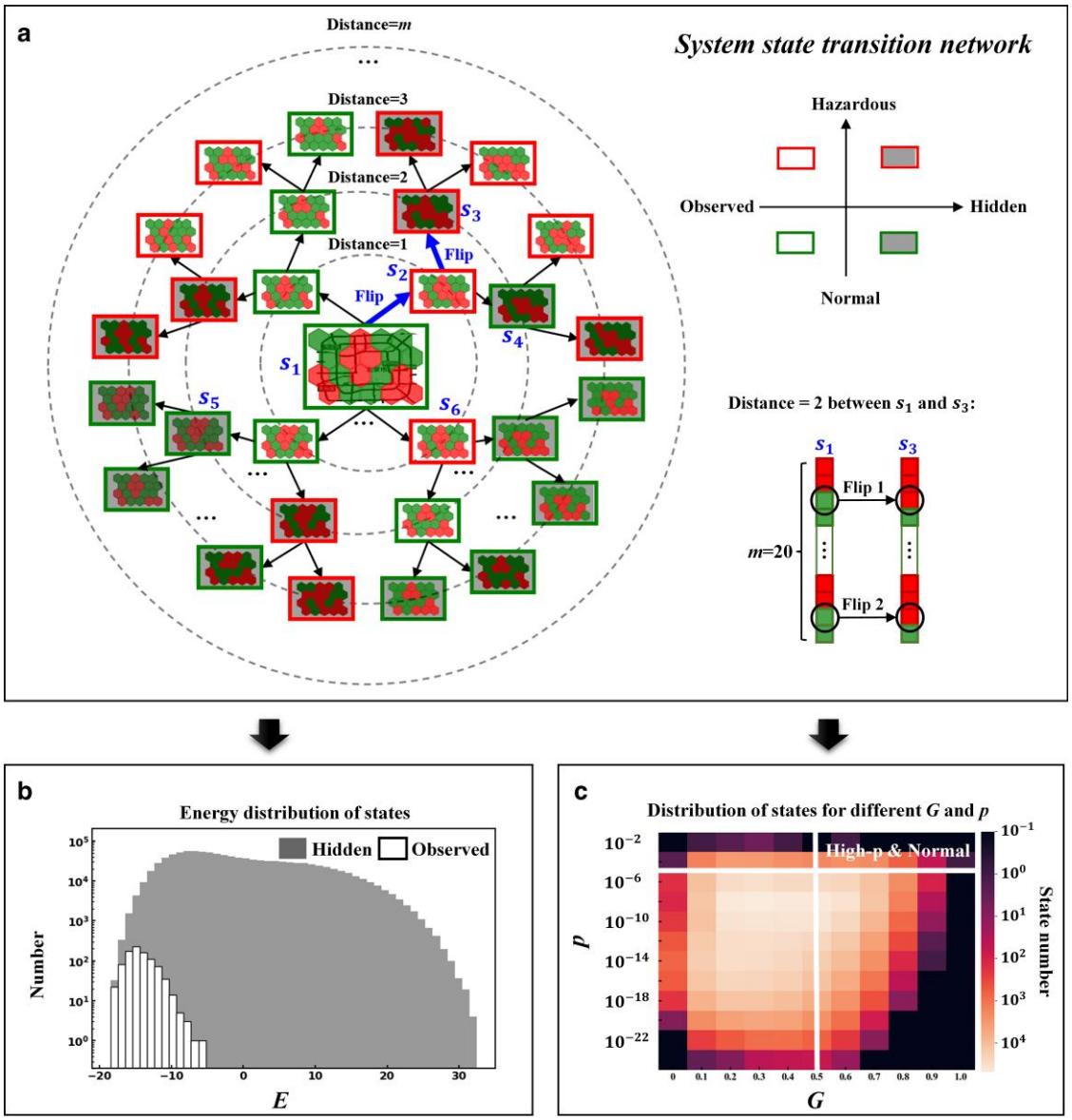

Fig. 2. State space analysis based on pairwise maximum entropy model. a) System state transition network. Nodes in this network represent the possible states of the entire road network, categorized into four classes: hidden hazardous states (e.g. s_3), hidden normal states (e.g. s_4), observed hazardous states (e.g. s_2), and observed normal states (e.g. s_1). The distance between two states is defined as the number of region-state changes required, e.g. the distance between s_3 and s_1 is 2. b) Energy distribution of states. The energy values of all road network states are distributed over a wide range, from −20 and 30. While a few observed states are concentrated around −15, there are still a fraction of hidden states with low energy that have never been observed in real data. c) Distribution of states for different G and p. High-p states are defined as those with $E < -13.2365$ (i.e. $p > 1 \times 10^{-5}$), distributed over different values of G. Hidden high-risk states are confined within these high-p and normal states (i.e. $G \geq 0.5$).

stable in the low-energy zone. Therefore, for simplification, we will focus on the transition dynamics of low-energy states (i.e. high-p states) rather than all states to further identify hidden high-risk states.

Vulnerability origin from system energy landscape

To quantify the transition dynamics of road network states, we construct the energy landscape of urban traffic, characterized by the system state transition network and the energy values of road network states (see Materials and methods). The energy landscape approach has been successfully used to understand the dynamic processes of complex system, such as specific folding pathways of proteins (37) and dynamic brain activities during bistable perception (38). In our constructed energy landscape, states move along energy-decrease paths until they reach local minima, corresponding to local lower-energy states in Fig. 3a. Such local energy minima are critical in various fields, such as differentiated cells (39), native state of a protein (37), and stable brain states (38), due to their stability during dynamical processes. This dynamical stability may be linked to stable hazardous states, which generates high risk. For example, if a hazardous state (where G is small) in the system state space has strong stability and a large basin of attraction, the system is more likely to be trapped into dangerous states. To further analyze these local energy minima, we construct a disconnectivity graph (40) (Fig. 3b) that illustrates the energy levels of the minima and the energy barriers between them (see Materials and methods). We find seven local energy minima, representing metastable states of urban

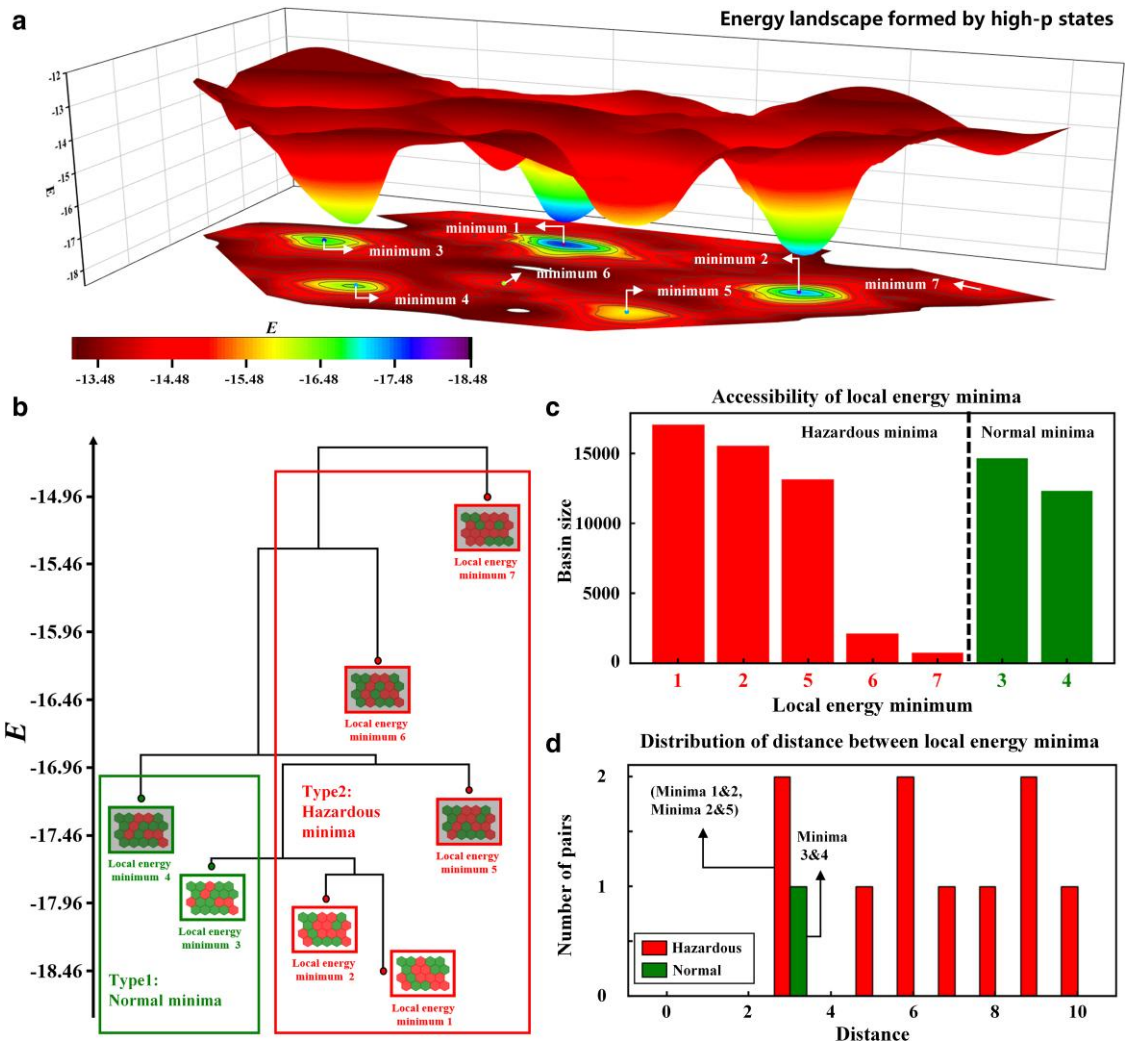

Fig. 3. System energy landscape. a) Energy landscape formed by high-p states. To highlight local energy minima, the visualization of the energy landscape simplifies the dynamical processes of states (see Materials and methods), with the relative positions inside each minimum basin provided only for illustration purposes. b) Disconnectivity graph of energy landscape. This graph illustrates the energy levels of local energy minima and the energy barriers between them. There are seven energy minima, with energy values ranging between −14 and −19, classified into two types: normal ($G \geq 0.5$) and hazardous ($G < 0.5$). c) Accessibility of local energy minima. Compared with normal minima, the basin sizes of hazardous minima are relatively larger. d) Distribution of distances between local minima. Compared with distances between other hazardous minima, the distances between minima 1 and 2, as well as between minima 2 and 5, are shorter.

traffic (36). This may be attributed to the frustration (41) arising from the complex positive and negative interactions between local regions.

Among these seven local minima, there are two normal minima and five hazardous minima, with three observed and four hidden. The accessibility of each minimum is analyzed based on its basin size, which is the number of states that can reach it along energy-decrease paths. Compared with normal minima, hazardous minima have relatively higher accessibility. As shown in Fig. 3c, the basin sizes of normal minima 3 and 4 are smaller than those of hazardous minima 1 and 2, and the basin size of normal minimum 4 is also smaller than that of hazardous minimum 5. The higher accessibility of hazardous minima indicates great vulnerability in the urban traffic. It is worth noting that hazardous minimum 5 in Fig. 3c is unobserved but highly accessible, which represents a hidden hazardous state that the system is likely to fall into, revealing the potential vulnerability of urban traffic. To further explore the system intrinsic vulnerability behavior, we analyze the transitions between local minima. It is shown in Fig. 3d that the distances between hazardous minima 1 and 2, as well as between hazardous minima 2 and 5, are relatively shorter compared with other hazardous minima. Specifically, transitions between them can be achieved by changing the states of only three regions. Due to these shorter distances, hazardous minima 1, 2, and 5 in Fig. 3b are likely to form a larger vulnerability valley. Once the system enters this vulnerable valley, returning to normal requires a high cost, highlighting the significant risk of urban traffic (during rush hours).

Hidden high-risk states identification

Hidden high-risk states are unobserved normal states with a high likelihood of occurring and transitioning into hazardous states. Given that hidden normal states (Fig. 2a) and high-p states (Fig. 2c) have been clarified, the dynamic characteristic is the key aspect in identifying hidden high-risk states. Compared with states that only reach normal minima (i.e. class 2 in Fig. 4a), states that can reach hazardous minima (i.e. class 1 and class 3 in Fig. 4a)

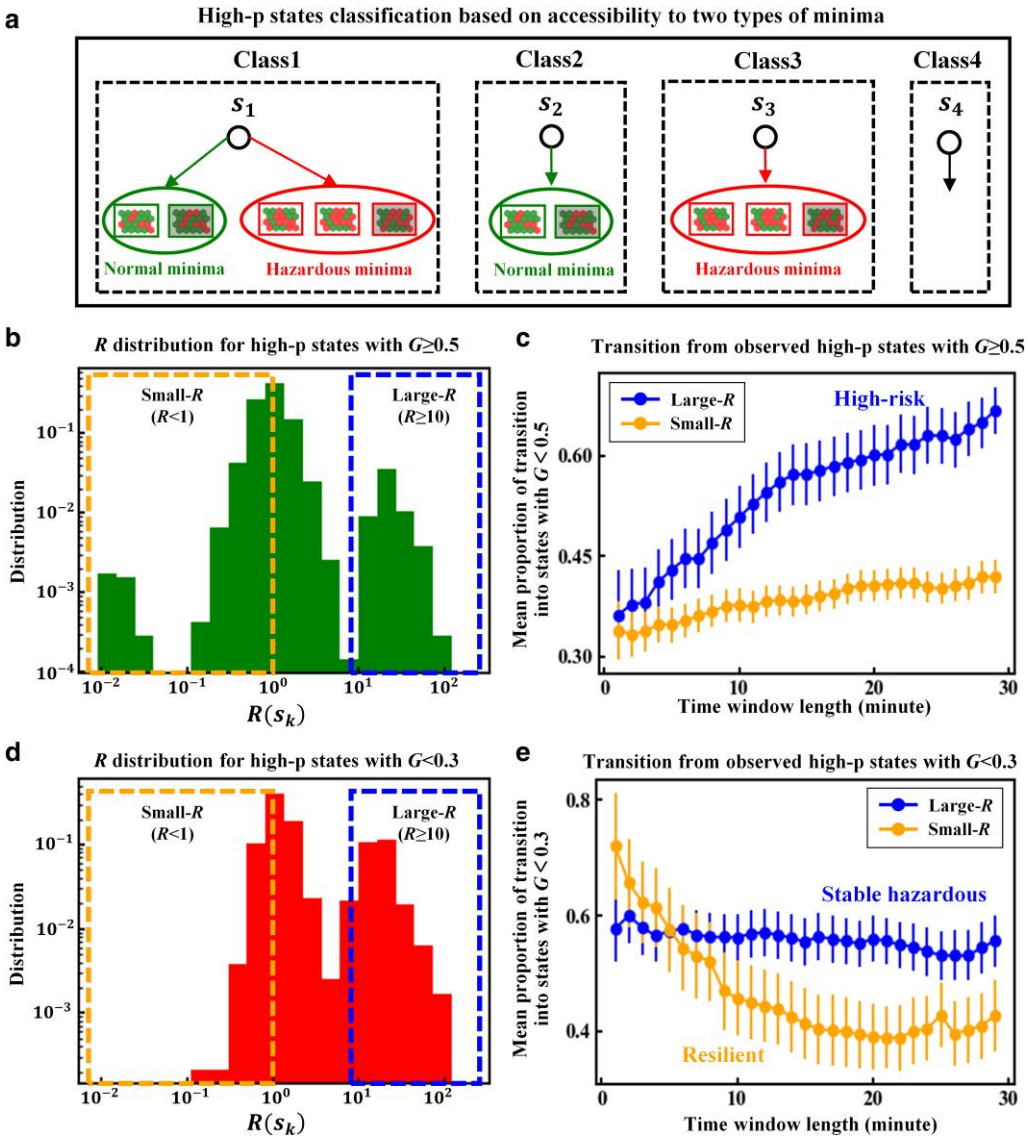

Fig. 4. Hidden high-risk states identification. a) High-*p* states classification based on accessibility to two types of minima. High-*p* states are classified into four categories based on their accessibility to normal minima and hazardous minima: (i) states that can reach both types of minima (class 1), (ii) states that can only reach normal minima (class 2), (iii) states that can only reach hazardous minima (class 3), and (iv) states that are inaccessible to either type of minima (class 4). b) *R* distribution for high-*p* states with $G \geq 0.5$. There is a broad distribution of *R* values ranging from 0.01 to 100 for high-*p* states with $G \geq 0.5$. c) Actual dynamical transitions from observed high-*p* states with $G \geq 0.5$. Compared with observed normal (i.e. $G \geq 0.5$) high-*p* states with small-*R*, those with large-*R* are more likely to transition into hazardous states (i.e. $G < 0.5$). Error bars represent the SEM proportion of transition into states with $G < 0.5$. d) *R* distribution for high-*p* states with $G < 0.3$. There is a broad distribution of *R* values ranging from 0.1 to 100 for high-*p* states with $G < 0.3$. e) Actual dynamical transitions from observed high-*p* states with $G < 0.3$. High-*p* states with $G < 0.3$ and small-*R* possess an inherent ability to spontaneously escape such states, highlighting system resilience. Conversely, high-*p* states with $G < 0.3$ and large-*R* tend to trap the system in long-time poor traffic operational connectivity, which can be regarded as stable hazardous states. Error bars represent the SEM proportion of transition into states with $G < 0.3$.

are more likely to fall into hazardous states. Therefore, the dynamic characteristic of hidden high-risk states is quantified by introducing a risk-level indicator (*R*; see Materials and methods), where a high *R* value indicates a short path length toward hazardous minima, suggesting great risk.

It is shown in Fig. 4b that there is a broad distribution of *R* value ranging from 0.01 to 100 for all normal high-*p* states, indicating significant differences in their risk levels. To verify that this proposed risk-level indicator can measure the dynamic characteristic of system falling into hazardous states, we analyze the actual evolution processes of observed high-*p* states at different *R* values. Compared with small-*R* (i.e. $R < 1$), normal high-*p* states with large-*R* (i.e. $R \geq 10$) are more likely to transition into hazardous states during real operations. As shown in Fig. 4c, as the time window length increases, the probability of the system transitioning into hazardous states from observed normal high-*p* states with large-*R* increases rapidly, rising from 35% to nearly 60% within 15 min and reaching 70% within 30 min. Conversely, the likelihood of observed normal high-*p* states with small-*R* falling into hazardous states remains below 45%. This result indicates that even though normal high-*p* states with large-*R* may appear to have good traffic functional connectivity, they tend to fall into hazardous states during dynamical processes, thus becoming high-risk states. Among these high-risk states, the unobserved states are the hidden high-risk states we aim to detect. Traffic managers or smart control centers may easily ignore hidden

high-risk states because they show good traffic operational connectivity and have not been observed in the data yet. However, without timely intervention in these high-risk states, the system will fall into hazardous states with high probability.

To further clarify the consequence of the system falling into hazardous states, we use the risk-level indicator (R) to analyze the dynamic characteristic of hazardous high-p states. As shown in Fig. 4d, the R values of high-p states with $G < 0.3$ vary significantly, ranging from 0.1 to 100. High-p states with $G < 0.3$ and large-R (i.e. $R \geq 10$) persist in hazardous states due to their proximity to hazardous minima (Fig. 4e), demonstrating stable hazardous dynamics. Conversely, high-p states with $G < 0.3$ and small-R (i.e. $R < 1$) are closer to normal minima, which can escape from such hazardous states spontaneously (Fig. 4e), highlighting system resilience. Among high-p states with $G < 0.3$, only 30.4% have small-R value (i.e. $R < 1$), indicating that most high-p states with $G < 0.3$ are unable to recover spontaneously. As a result, once the system falls into such hazardous states, it is likely to remain hazardous, leading to system breakdowns. Therefore, our identified hidden high-risk states can offer reliable early warning signals for the system entering potential stable hazardous states and enhance the resilience of urban traffic management. We further analyze the spatial characteristics of the top 10 hidden high-risk states with the largest R values (Fig. S14a) and find that jammed regions in these states form several larger jam clusters, the aggregation of which may trigger global-level failures (42, 43). In contrast, the jammed regions of the top 10 hidden normal high-p states with the smallest R values (Fig. S14b) are relatively dispersed, making it less likely to form a large congestion cluster. Therefore, avoiding the concentration of congested areas may be a primary goal in managing hidden high-risk states.

Discussion

Risk is commonly characterized by the severity of its consequences and associated uncertainties (44). In the transportation domain, studies on traffic risk can be broadly categorized into local perspectives and global perspectives. From the local perspective, traffic risk is often defined as the probability of accidents occurring due to specific factors related to drivers, vehicles, or road conditions. Examples include risks from drunk driving (45), lane-changing behavior (46), and road geometric design (47). These studies focus on identifying and mitigating risks at localized points or road segments. From the global perspective, traffic risk is closely related to transportation system vulnerability (48), which describes the loss of functionality in response to external disruptions such as extreme weather (49) or natural disasters (50). These studies examine how external shocks impact the overall system performance. While these existing studies provide valuable insights, they often overlook the dynamic propagation of risks within the road network—how localized risks interact and spread across local regions, potentially leading to global congestion risks. Addressing this gap is critical for mitigating secondary impacts of local risks and preventing system breakdowns. In our study, hidden high-risk states are defined as system states that appear normal under current conditions but have a high probability of transitioning into dangerous states due to interregional traffic interactions. The pairwise maximum entropy model used in our study captures these interregional interactions and reveals how risks may dynamically propagate across the network. This approach aligns with existing studies on traffic percolation (31, 51) and the spatiotemporal propagation of congestion (42, 52), which emphasize the importance of understanding the dynamic traffic evolution to prevent large-scale breakdowns. By integrating dynamic propagation into the risk definition, our study moves beyond traditional traffic risk analysis and provides a new perspective on predicting large-scale congestion risks.

Our identified dynamical processes of high-risk states transitioning from normal to hazardous states are actually phase transitions (53), which are widely found in complex systems. It is challenging to discover early warning signals for phase transitions. There are efforts including critical slowing down near tipping points (6) and changing stability landscapes in stochastic systems (54), with relevant metrics such as recovery rate (55) and autocorrelation (56) proposed to quantify these dynamics. In addition to these time-series-based indicators, researchers have also proposed phase transition metrics from a spatial perspective, such as spatial variance and skewness (57) and spatial permutation entropy (PE) (58). Different from these methods that consider the spatial and temporal perspectives separately, we examine evolutionary paths toward hazardous states to identify high-risk states, which could serve as good spatiotemporal signals for phase transition in urban traffic. On this basis, resilience regulation of traffic congestion at different transition stages is possible. When the traffic system enters states that may fall into hazardous attractors, timely intervention is required to avoid greater losses. Together with smart control agents for urban traffic (59), it is possible to recover system with high resilience.

Our results can also serve for reliability testing scenarios of various smart control agents. With the rapid progress of new generation information technologies (e.g. cloud computing, big data, and artificial intelligence), smart traffic control agents (e.g. Intelligent Transportation Systems (60), City Traffic Brain, and Smart Traffic Signal Systems) have been widely applied, aiming at efficient, intelligent, and sustainable control over the entire road network state. A key challenge for these smart control agents is identifying critical scenarios to accelerate large-scale reliability testing. Based on our constructed energy landscape, high-risk states can be sampled as critical scenarios, which may help accelerate reliability assessing and testing of various smart control agents in urban traffic.

While the model in our study remains relatively stable under routine traffic conditions (Figs. S15 and S23), exceptional events (e.g. large public events or unusual disruptions) may significantly alter the dynamics of traffic flow and the correlation patterns between regions. In real-world complex systems, the interactions between system components may evolve over time in response to various factors, a phenomenon that has been explored using modified kinetic Ising models (61, 62). Therefore, adaptive modeling techniques should be incorporated in future work. Specifically, machine learning methods such as generative models can be employed to generate dynamics of interactions. These methods will allow us to simulate how the system behaves under varying traffic conditions and identify the most dangerous states that emerge in extreme scenarios. Additionally, due to the limitations of the available dataset, we are unable to fully test the model on unseen critical states under substantial variations in time, space, and entity, especially considering the sparsity of such critical states. To address this, we plan to explore large-scale city simulations, which can generate diverse traffic states under varying conditions (63). By integrating such simulations with our model, we can test the model under a wider range of dynamic conditions, enabling more comprehensive validation and optimization.

Materials and methods

Datasets description

Raw data

Our raw dataset includes real-time velocity records for each road segment in the Beijing Road network, covering 00:00 to 24:00 each day from 2015 October 1 to 30 with the exception of October 19 due to data quality issues. The spatial scope consists of the road network within the Fourth Ring Road of Beijing (Fig. S16a and b), comprising over 18,000 nodes and more than 33,000 edges. The road speed data are derived from the global positioning system recorded by floating cars at 1-min intervals. The spatial distribution of the top 20 fastest and bottom 20 slowest road segments at a given time is shown in Fig. S17. The dataset includes speed data from 17 workdays (i.e. 20151008, 20151009, 20151010, 20151012, 20151013, 20151014, 20151015, 20151016, 20151020, 20151021, 20151022, 20151023, 20151026, 20151027, 20151028, 20151029, and 20151030) and 12 nonworkdays. Due to the distinct traffic patterns observed on workdays and nonworkdays (Figs. S18 and S19), the pairwise maximum entropy model is trained using data from a specific time period within the 17 workdays of October 2015. Two time periods are selected: the rush-hour period (7:30–8:30 AM) and the nonrush-hour period (6:00–7:00 AM), as the road speed distributions differ significantly between these two periods (Fig. S16c and d). Furthermore, considering that the road network includes different levels of roads, such as national highways, provincial roads, and local roads, each with varying speed limits, we use the relative velocity (31) to measure the operational condition of a road segment at a given time. The relative velocity of a road segment is defined as the ratio of its actual speed to its standard maximum velocity, which is the 95th percentile speed for that segment throughout the day.

Data aggregation

Each global road network state is formed by a 20-dimensional vector, consisting of the region-level binary states at a time point. It is shown in Fig. S20 that the global road network state is aggregated from road segment-level speeds, as described below. (i) *Identify congested road segments*: At time t, congested road segments are determined by the congestion ratio f. The congestion ratio f is an adjustable parameter used to classify road segments as congested or uncongested based on their relative velocities (Figs. S21 and S22), with road segments ranked from highest to lowest relative velocity at time t, and the lowest f-percent is classified as congested. (ii) *Determine region-level binary states*: The operational performance $q^i(t)$ ($i = 1, 2, 3, \ldots, 20$) for region i at time t is computed (Fig. S20b), which is defined as the maximum size of the connected cluster formed by congested road segments within that region. A higher value of $q^i(t)$ indicates worse operational performance of region i, meaning that vehicles in this region are traveling at relatively lower speeds. The binary state $s^i(t)$ ($i = 1, 2, 3, \ldots, 20$) for region i at time t is assigned based on a performance threshold q^{th}. If $q^i(t)$ exceeds q^{th}, the state of region i at time t is classified as jammed state (i.e. $s^i(t) = 1$); otherwise, it is classified as free state (i.e. $s^i(t) = -1$). (iii) *Determine global road network state*: Collectively, the global road network state at time t is represented as a 20-dimensional vector of region-level binary states $s(t) = \{s^1(t), s^2(t), \ldots, s^i(t), \ldots, s^{20}(t)\}$, where $s^i(t) \in \{1, -1\}$, as shown in Fig. S20c. Therefore, the global road network state formed by the region-level binary states at a given time point is determined by the road segment-level speed data at that time point, along with the values of congestion ratio f and the region performance threshold q^{th}. Specifically, we set f to 0.25 and q^{th} to 0.09 (our results hold across a broad range of f and q^{th}; see Fig. S2c) and record road network states at 1-min intervals over 17 working days in October 2015. According to the count, the number of road network states observed in the data during rush hours (7:30–8:30 AM) is only 910 and during nonrush hours is 362 (6:00–7:00 AM). Compared with the entire system space, the number of observed road network states in the data is very small.

Pairwise maximum entropy model of urban traffic

Model introduction

The maximum entropy model is a modeling approach that uses the maximum entropy principle (25) to find a probability distribution based on given information. If our knowledge of a system is limited, we should choose a probability distribution that maximizes system uncertainty under those constraints. This is the least structured or most unbiased model among all probability distributions satisfying certain given constraints. In this context, pairwise maximum entropy model maximizes the entropy of the probability distribution under two-moment constraints (i.e. attributes of both individual variables and pairwise correlations). Pairwise maximum entropy model has been widely applied in complex systems to addresses inverse statistical problems (33), such as finding weak interactions between neurons behind strongly collective behavior of neurons (27), deducing the network structure of brain functional areas (28), and analyzing the gene–gene interaction network that underlies cell metabolism (29). The core of these problems is to infer interaction relationships between local entities from system behavior data while preserving some macroscopic behavioral characteristics of the system.

Model definition

A pairwise maximum entropy model is used to estimate the interactions between local regions in the road network for a given time period. The pairwise maximum entropy model aims to find a probability distribution that maximizes (Gibbs) entropy under the constraints of the first and second moments.

$$\begin{aligned} \text{Max} \quad & H = -\sum_{1\le k\le n} p(s_k)\log p(s_k) \\ \text{s.t.} \quad & \sum_{1\le k\le n} p(s_k) = 1 \\ & \sum_{1\le k\le n} p(s_k)s_k^i = \sum_{s_d\in D} f(s_d)s_d^i, \quad 1 \le i \le m \\ & \sum_{1\le k\le n} p(s_k)s_k^i s_k^j = \sum_{s_d\in D} f(s_d)s_d^i s_d^j, \quad 1 \le i < j \le m \end{aligned}, \tag{5}$$

where s_k represents a global road network state, n is the number of global road network states in the entire state space, m is the number of regions, $p(s_k)$ is the probability of global road network state s_k, and s_k^i is the state of region i with a value of +1 (i.e. jammed region) or −1 (i.e. free region) when the system is in state s_k. $\sum_{1\le k\le n} p(s_k)s_k^i$ is the expected state of region i and $\sum_{1\le k\le n} p(s_k)s_k^i s_k^j$ is the expected pairwise joint state of region i and region j, which can be calculated by the estimated probability distribution. The sample dataset D consists of global road network states observed at all time points within the selected time period across 17 workdays, $f(s_d)$ is the occurrence frequency of global road network state s_d in D, and s_d^i is the state of region i with a value of +1 (i.e. jammed region) or −1 (i.e. free region) when the system is in observed global road network state s_d. $\sum_{s_d\in D} f(s_d)s_d^i$ is the average state of region i in the empirical data, where a larger value represents that the region is more congested. $\sum_{s_d\in D} f(s_d)s_d^i s_d^j$ is the average

pairwise joint state of region i and region j in the empirical data, where a larger value represents a higher correlation between two regions.

The form of this probability distribution can be obtained by solving the Lagrange function as:

$$p(s_k) = (1/Z)e^{\sum_{1\le i\le m} h_i s_k^i + \sum_{1\le i<j\le m} J_{ij} s_k^i s_k^j}, \tag{6}$$

where h_i represents the congestion tendency of local road network region i, J_{ij} characterizes the interaction between region i and j, and Z is the partition function ensuring normalization, which equals $\sum_{1\le k\le n} e^{\sum_{1\le i\le m} h_i s_k^i + \sum_{1\le i<j\le m} J_{ij} s_k^i s_k^j}$.

Data input

The data feeding into the model consist of global road network states during a specific time period across 17 workdays in October 2015. In our study, two time periods are selected: rush hour (7:30–8:30 AM) and nonrush hour (6:00–7:00 AM). Each global road network state in the data is represented by a 20-dimensional vector $s(t) = \{s^1(t), s^2(t), \ldots, s^i(t), \ldots, s^{20}(t)\}$, where $s^i(t)$ is the binary state of region i at a certain time point t during selected time periods, with $s^i(t) \in \{1, -1\}$ (1 for jammed, −1 for free). From these global road network states, the empirical first moments $\sum_{s_d\in D} f(s_d) s_d^i$ (average state of region i) and second moments $\sum_{s_d\in D} f(s_d) s_d^i s_d^j$ (average pairwise joint state of region i and region j) are derived. These moments serve as constraints in Eq. 5 for fitting the maximum entropy model.

Model algorithm

Model parameters can be inferred by ensuring that Eq. 6 satisfies the constraints of the first and second moments obtained from data, which is equivalent to minimizing the KL divergence between the model distribution (i.e. Eq. 6) and the empirical data distribution. However, given the large state space of our road network (with 20 regions, the total state space is 2^{20}), directly computing the partition function Z in Eq. 6 is computationally intractable due to its exponential growth with the number of regions.

To address this, we employ the minimum probability flow (MPF) algorithm (64) to estimate the parameters h_i and J_{ij} efficiently (i.e. coniii (65)). In the MPF algorithm, parameter estimation problem is regarded as a dynamical process, where the observed data distribution $\boldsymbol{p}^{(0)}$ is treated as the initial state and the model distribution $\boldsymbol{p}^{(\infty)}(\theta)$ is treated as the steady-state equilibrium. Here, θ represents model parameters and θ=$\{h_i, J_{ij}\}$ in the pairwise maximum entropy model. The evolution from $\boldsymbol{p}^{(0)}$ to $\boldsymbol{p}^{(\infty)}(\theta)$ is described as a probability flow, driven by the transition rates Γ_{lk}, which depend on the difference in energy between states. The transition rate Γ_{lk} between neighboring states s_l and s_k (differing by a single region-state flip) is defined as:

$$\Gamma_{lk} = g_{lk} \exp\left(\frac{1}{2}(E(s_l) - E(s_k))\right), \tag{7}$$

where $g_{lk} = 1$ when s_l and s_k are neighboring states; otherwise, $g_{lk} = 0$. $E(s_k)$ represents the energy of state s_k, which is calculated using the model parameters θ. In the pairwise maximum entropy model,

$$E(s_k) = \sum_{1\le i\le m} -h_i s_k^i - \sum_{1\le i<j\le m} J_{ij} s_k^i s_k^j. \tag{8}$$

The MPF algorithm aims to minimize the probability flow from observed states to unobserved states. This is equivalent to minimizing the KL divergence between the data distribution and the model distribution after an infinitesimal time ϵ. The objective function is given by:

$$K(\theta) = \frac{\varepsilon}{M} \sum_{s_l\in D} \sum_{s_k\notin D} \Gamma_{lk}, \tag{9}$$

where M is the total number of observed states in the dataset, ϵ is a small time step representing the initial dynamics, and D denotes the set of observed data states. By minimizing $K(\theta)$, MPF can estimate the model parameters θ effectively, avoiding the need to compute the partition function Z.

Model fitting performance evaluation

We evaluate the fitting performance of model through two ways: (i) by assessing whether the model satisfies the two-moment constraints at the microscopic level and (ii) by checking whether the distribution of macroscopic properties obtained by the model is consistent with the data at the macroscopic level. Here, we select three macroscopic properties of traffic network: traffic functional performance ($G^{\text{free}}_{\text{region}}$), congestion concentration degree ($G^{\text{jam}}_{\text{region}}$), and the proportion of jammed region ($F^{\text{jam}}_{\text{region}}$). The formulas for calculating these macroscopic properties are as follows:

$$G^{\text{free}}_{\text{region}} = \frac{C^{\text{free}}_{\text{region}}}{N_{\text{region}}}, \tag{10}$$

$$G^{\text{jam}}_{\text{region}} = \frac{C^{\text{jam}}_{\text{region}}}{N_{\text{region}}}, \tag{11}$$

$$F^{\text{jam}}_{\text{region}} = \frac{N^{\text{jam}}_{\text{region}}}{N_{\text{region}}}, \tag{12}$$

where $C^{\text{free}}_{\text{region}}$ is the size of the largest functional cluster (31) formed by connected free regions, $C^{\text{jam}}_{\text{region}}$ is the size of the largest jam cluster (52) formed by connected jam regions, $N^{\text{jam}}_{\text{region}}$ is the number of jam regions, and N_{region} is the number of regions.

Definition of hidden high-risk states

We define hidden high-risk states as unobserved normal states with high likelihood of occurring and falling into hazardous states. According to this definition, hidden high-risk states, S^{HHR}, can be represented as:

$$S^{\text{HHR}} = S^{\text{U}} \cap S^{\text{N}} \cap S^{\text{HP}} \cap S^{\rightarrow \text{SH}}, \tag{13}$$

where S^{U} represents hidden states that have never been observed, S^{N} represents normal states (i.e. $G \ge 0.5$), and S^{HP} represents low-energy states (i.e. high-p states), with energy lower than a certain threshold (E^{th}). Here, we set the energy threshold (E^{th}) to −13.2365, corresponding to low-energy states with a probability of occurrence exceeding 1×10^{-5}. $S^{\rightarrow \text{SH}}$ represents states that are likely to fall into hazardous states.

Dynamical processes of road network states

Energy landscape

The energy landscape of urban traffic is characterized by the system state transition network and the energy values of high-p states (Fig. 2c). The dynamics of this energy landscape are defined such that states can only transition to neighbors (i.e. those differing by only one region state) with lower energy. Consequently, states move along energy-decrease paths in the energy landscape until they reach a local minimum. The basin size of each local

minimum is determined by the number of states that can reach it along energy-decrease paths, thereby representing its accessibility. To highlight these local energy minima, the visualization of the energy landscape (Fig. 3a) simplifies the system state dynamics, allowing each state to transition only to its lowest energy neighbor. This ensures that each state reaches a single local minimum along the steepest energy-decreasing path. The relative positions within each basin are only for illustration.

Disconnectivity graph

The disconnectivity graph (40) is commonly used to analyze the energy landscape of high-dimension complex systems. In such a graph, each endpoint on a branch represents a local energy minimum. The height of these endpoints reflects their energy levels, with lower endpoints indicating lower energy. The energy barrier is defined as the energy difference that must be overcome for the system to transition from one minimum to another, which equals the vertical distance from the starting minimum up to the connecting saddle point (i.e. point where branches intersect between these two minima).

Risk-level indicator of system states

To quantify the risk level of system states, we propose an indicator, $R(s_k)$, based on the relative path length of states transitioning toward two types of high-accessibility minima. High-accessibility minima are defined as those with large basin sizes, as illustrated by minima 1, 2, 3, 4, and 5 in Fig. 3b and c. Among these, minima 1, 2, and 5 are hazardous high-accessibility minima, while minima 3 and 4 are normal high-accessibility minima. The indicator $R(s_k)$ can be denoted as:

$$R(s_k) = l^{\to \text{Normal}}(s_k)/l^{\to \text{Hazardous}}(s_k), \tag{14}$$

where $l^{\to \text{Normal}}(s_k)$ represents the shortest path length from state s_k to the nearest of normal high-accessibility minima and $l^{\to \text{Hazardous}}(s_k)$ represents the shortest path length from state s_k to the nearest of hazardous high-accessibility minima. Considering that some states cannot reach either hazardous or normal minima (i.e. class 4 in Fig. S12), we set the corresponding value of $l^{\to \text{Normal}}(s_k)$ or $l^{\to \text{Hazardous}}(s_k)$ to 100, which is significantly greater than any shortest path length toward a specific type of minima (Fig. S13). According to this definition, a higher R value indicates a shorter path length toward hazardous minima, signifying greater risk.

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Funding

This work was supported by the National Natural Science Foundation of China (grant nos. 72225012, 72288101, and 71822101).

Author Contributions

S.L. contributed the conceptualization, methodology, investigation, validation, formal analysis, visualization, writing—original draft, writing—review & editing. M.B. contributed the conceptualization, methodology, investigation, validation, formal analysis, visualization, writing—review & editing. S.G. contributed the data curation, formal analysis, writing—review & editing. J.G. contributed the methodology, formal analysis, writing—review & editing. H.S. contributed the methodology, formal analysis, writing—review & editing. Z.-Y.G. contributed the methodology, funding acquisition, formal analysis, writing—review & editing. D.L. contributed the conceptualization, methodology, funding acquisition, project administration, supervision, formal analysis, writing—original draft, writing—review & editing.

Preprints

This manuscript was posted on a preprint: https://arxiv.org/abs/2407.20478.

Data Availability

The data and code, including all scripts and necessary files for replicating the results, are openly available in GitHub at https://github.com/lsy2220124130/HiddenHighRiskStates/. The raw velocity data are not publicly available due to privacy restrictions. More details and figures are presented in Supplementary Information.

References

1 Farazmand M, Sapsis TP. 2019. Extreme events: mechanisms and prediction. *Appl Mech Rev*. 71(5):050801.
2 Knopoff L, Kagan Y. 1977. Analysis of the theory of extremes as applied to earthquake problems. *J Geophys Res*. 82(36):5647–5657.
3 Danielsen F, *et al.* 2005. The Asian tsunami: a protective role for coastal vegetation. *Science*. 310(5748):643–643.
4 Poon S-H, Rockinger M, Tawn J. 2004. Extreme value dependence in financial markets: diagnostics, models, and financial implications. *Rev Financ Stud*. 17(2):581–610.
5 Haes Alhelou H, Hamedani-Golshan ME, Njenda TC, Siano P. 2019. A survey on power system blackout and cascading events: research motivations and challenges. *Energies (Basel)*. 12(4):682.
6 Dakos V, *et al.* 2008. Slowing down as an early warning signal for abrupt climate change. *Proc Natl Acad Sci U S A*. 105(38):14308–14312.
7 Lombardi F, Pepić S, Shriki O, Tkačik G, De Martino D. 2023. Statistical modeling of adaptive neural networks explains co-existence of avalanches and oscillations in resting human brain. *Nat Comput Sci*. 3(3):254–263.
8 Wang K, Wang L, Wei Y-M, Ye M. 2013. Beijing storm of July 21, 2012: observations and reflections. *Nat Hazards*. 67:969–974.
9 Erhardt GD, *et al.* 2019. Do transportation network companies decrease or increase congestion? *Sci Adv*. 5(5):eaau2670.
10 Chu W, *et al.* 2021. Cloud control system architectures, technologies and applications on intelligent and connected vehicles: a review. *Chin J Mech Eng*. 34(1):139.
11 Petit J, Shladover SE. 2014. Potential cyberattacks on automated vehicles. *IEEE Trans Intell Transp Syst*. 16(2):546–556.
12 Castillo E. *Extreme value theory in engineering*. Academic Press, Boston, 1988.
13 Ragone F, Wouters J, Bouchet F. 2018. Computation of extreme heat waves in climate models using a large deviation algorithm. *Proc Natl Acad Sci U S A*. 115(1):24–29.
14 Boers N, *et al.* 2014. Prediction of extreme floods in the eastern Central Andes based on a complex networks approach. *Nat Commun*. 5(1):5199.
15 Jacox MG, *et al.* 2022. Global seasonal forecasts of marine heatwaves. *Nature*. 604(7906):486–490.
16 Pickering E, Guth S, Karniadakis GE, Sapsis TP. 2022. Discovering and forecasting extreme events via active learning in neural operators. *Nat Comput Sci*. 2(12):823–833.

17 Qi D, Majda AJ. 2020. Using machine learning to predict extreme events in complex systems. *Proc Natl Acad Sci U S A*. 117(1):52–59.
18 Chattopadhyay A, Hassanzadeh P, Subramanian D. 2020. Data-driven predictions of a multiscale Lorenz 96 chaotic system using machine-learning methods: reservoir computing, artificial neural network, and long short-term memory network. *Nonlinear Process Geophys*. 27(3):373–389.
19 Farazmand M, Sapsis TP. 2016. Dynamical indicators for the prediction of bursting phenomena in high-dimensional systems. *Phy Rev E*. 94(3-1):032212.
20 Latif M, Keenlyside NS. 2009. El Niño/Southern Oscillation response to global warming. *Proc Natl Acad Sci U S A*. 106(49): 20578–20583.
21 Van den Broeck C, Parrondo JMR, Toral R. 1994. Noise-induced non-equilibrium phase transition. *Phys Rev Lett*. 73(25):3395.
22 D'Odorico P, Laio F, Ridolfi L. 2005. Noise-induced stability in dryland plant ecosystems. *Proc Natl Acad Sci U S A*. 102(31): 10819–10822.
23 Xu L, Patterson D, Levin SA, Wang J. 2023. Non-equilibrium early-warning signals for critical transitions in ecological systems. *Proc Natl Acad Sci U S A*. 120(5):e2218663120.
24 Gonzalez MC, Hidalgo CA, Barabasi AL. 2008. Understanding individual human mobility patterns. *Nature*. 453(7196):779–782.
25 Jaynes ET. 1957. Information theory and statistical mechanics. *Phys Rev*. 106(4):620–630.
26 Yeh FC, *et al.* 2010. Maximum entropy approaches to living neural networks. *Entropy*. 12(1):89–106.
27 Schneidman E, Berry MJ, Segev R, Bialek W. 2006. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*. 440(7087):1007–1012.
28 Watanabe T, *et al.* 2013. A pairwise maximum entropy model accurately describes resting-state human brain networks. *Nat Commun*. 4(1):1370.
29 De Martino A, De Martino D. 2018. An introduction to the maximum entropy approach and its application to inference problems in biology. *Heliyon*. 4(4):e00596.
30 Daniotti S, Monechi B, Ubaldi E. 2023. A maximum entropy approach for the modelling of car-sharing parking dynamics. *Sci Rep*. 13(1):2993.
31 Li D, *et al.* 2015. Percolation transition in dynamical traffic network with evolving critical bottlenecks. *Proc Natl Acad Sci U S A*. 112(3):669–672.
32 Yildirimoglu M, Ramezani M, Geroliminis N. 2015. Equilibrium analysis and route guidance in large-scale networks with MFD dynamics. *Transp Res Proc*. 9:185–204.
33 Nguyen HC, Zecchina R, Berg J. 2017. Inverse statistical problems: from the inverse Ising problem to data science. *Adv Phys*. 66(3): 197–261.
34 Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D. 2006. Complex networks: structure and dynamics. *Phys Rep*. 424(4-5): 175–308.
35 Weistuch C, *et al.* 2021. Metabolism modulates network synchrony in the aging brain. *Proc Natl Acad Sci U S A*. 118(40): e2025727118.
36 Zeng G, *et al.* 2020. Multiple metastable network states in urban traffic. *Proc Natl Acad Sci U S A*. 117(30):17528–17534.
37 Bryngelson JD, *et al.* 1995. Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins*. 21(3): 167–195.
38 Watanabe T, Masuda N, Megumi F, Kanai R, Rees G. 2014. Energy landscape and dynamics of brain activity during human bistable perception. *Nat Commun*. 5(1):4765.
39 Li C, Wang J. 2013. Quantifying cell fate decisions for differentiation and reprogramming of a human stem cell network: landscape and biological paths. *PLoS Comput Biol*. 9(8):e1003165.
40 Becker OM, Karplus M. 1997. The topology of multidimensional potential energy surfaces: theory and application to peptide structure and kinetics. *J Chem Phys*. 106(4):1495–1517.
41 Ferreiro DU, *et al.* 2011. On the role of frustration in the energy landscapes of allosteric proteins. *Proc Natl Acad Sci U S A*. 108(9):3499–3503.
42 Zhao J, Li D, Sanhedrai H, Cohen R, Havlin S. 2016. Spatio-temporal propagation of cascading overload failures in spatially embedded networks. *Nat Commun*. 7(1):10094.
43 Duan J, Li D, Huang HJ. 2023. Reliability of the traffic network against cascading failures with individuals acting independently or collectively. *Transp Res Part C Emerg Technol*. 147:104017.
44 Aven T. 2016. Risk assessment and risk management: review of recent advances on their foundation. *Eur J Oper Res*. 253(1): 1–13.
45 Lin HA, *et al.* 2022. Evaluating the effect of drunk driving on fatal injuries among vulnerable road users in Taiwan: a population-based study. *BMC Public Health*. 22(1):2059.
46 Mahajan V, Katrakazas C, Antoniou C. 2020. Crash risk estimation due to lane changing: a data-driven approach using naturalistic data. *IEEE Trans Intell Transp Syst*. 23(4):3756–3765.
47 Álvarez P, Fernández MA, Gordaliza A, Mansilla A, Molinero A. 2020. Geometric road design factors affecting the risk of urban run-off crashes. A case-control study. *PLoS One*. 15(6): e0234564.
48 Berdica K. 2002. An introduction to road vulnerability: what has been done, is done and should be done. *Trans Policy*. 9(2): 117–127.
49 Molarius R, *et al.* 2014. The extreme weather risk indicators (EWRI) for the European transport system. *Nat Hazards*. 72: 189–210.
50 Toma-Danila D. 2018. A GIS framework for evaluating the implications of urban road network failure due to earthquakes: Bucharest (Romania) case study. *Nat Hazards*. 93(Suppl 1): 97–111.
51 Zeng G, *et al.* 2019. Switch between critical percolation modes in city traffic dynamics. *Proc Natl Acad Sci U S A*. 116(1):23–28.
52 Zhang L, *et al.* 2019. Scale-free resilience of real traffic jams. *Proc Natl Acad Sci U S A*. 116(18):8673–8678.
53 Scheffer M, *et al.* 2012. Anticipating critical transitions. *Science*. 338(6105):344–348.
54 Hirota M, Holmgren M, Van Nes EH, Scheffer M. 2011. Global resilience of tropical forest and savanna to critical transitions. *Science*. 334(6053):232–235.
55 Veraart AJ, *et al.* 2012. Recovery rates reflect distance to a tipping point in a living system. *Nature*. 481(7381):357–359.
56 Dai L, Vorselen D, Korolev KS, Gore J. 2012. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science*. 336(6085):1175–1177.
57 Guttal V, Jayaprakash C. 2009. Spatial variance and spatial skewness: leading indicators of regime shifts in spatial ecological systems. *Theor Ecol*. 2:3–12.
58 Tirabassi G, Masoller C. 2023. Entropy-based early detection of critical transitions in spatial vegetation fields. *Proc Natl Acad Sci U S A*. 120(1):e2215667120.
59 Liu Y, Liu L, Chen WP. 2017. Intelligent traffic light control using distributed multi-agent Q learning. In: *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Yokohama, Japan. p. 1–8.

60 Zhang J, *et al.* 2011. Data-driven intelligent transportation systems: a survey. *IEEE Trans Intell Transp Syst*. 12(4): 1624–1639.
61 Aguilera M, Moosavi SA, Shimazaki H. 2021. A unifying framework for mean-field theories of asymmetric kinetic Ising systems. *Nat Commun*. 12(1):1197.
62 Campajola C, Gangi DD, Lillo F, Tantari D. 2022. Modelling time-varying interactions in complex systems: the score driven kinetic Ising model. *Sci Rep*. 12(1):19339.
63 Jiang S, *et al.* 2016. The TimeGeo modeling framework for urban mobility without travel surveys. *Proc Natl Acad Sci U S A*. 113(37):E5370–E5378.
64 Sohl-Dickstein J, Battaglino PB, DeWeese MR. 2011. New method for parameter estimation in probabilistic models: minimum probability flow. *Phys Rev Lett*. 107(22):220601.
65 Lee ED, Daniels BC. 2019. Convenient interface to inverse Ising (ConIII): a Python 3 package for solving Ising-type maximum entropy models. *J Open Res Soft*. 7(1):3.