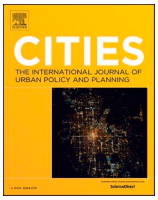
Cities 162 (2025) 105980

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities



Causal discovery from city data, where urban scaling meets information theory

Tian Gan [a,b], Rayan Succar [a,b], Simone Macrì [c], Manuel Ruiz Marín [d], Maurizio Porfiri [a,b,e,*]

[a] Department of Mechanical and Aerospace Engineering, Tandon School of Engineering, New York University, Brooklyn 11201, New York, USA
[b] Center for Urban Science and Progress, Tandon School of Engineering, New York University, Brooklyn 11201, New York, USA
[c] Centre for Behavioural Sciences and Mental Health, Istituto Superiore di Sanità, Rome 00161, Italy
[d] Department of Quantitative Methods, Law and Modern Languages, Universidad Politécnica de Cartagena, Murcia 30201, Spain
[e] Department of Biomedical Engineering, Tandon School of Engineering, New York University, Brooklyn 11201, New York, USA

ARTICLE INFO

Keywords:
Causal inference
Information theory
Open urban data
Urban science

ABSTRACT

Causal discovery from urban data offers an unprecedented opportunity for research and practice in urban science. Whether they implement traditional regression analyses or more sophisticated tools for time series analysis, existing approaches to causal discovery consider cities to be statistically equivalent – an assumption that is seldom met in real urban systems, where cities dramatically differ in population size, land area, etc. In this study, we embrace the heterogeneity between cities within a novel approach to causal discovery that integrates urban scaling and information theory, towards the discovery of associations between urban processes. Our approach takes as input time series of salient urban variables for causal discovery, along with a set of static features that capture the heterogeneities within the system. Using a Cobb-Douglas function, we extract features-adjusted metropolitan indicators (FsAMIs) that mitigate spurious dependencies among variables due to the underlying variations in spatial features. Through the application of transfer entropy on FsAMIs, we ultimately perform model-free causal inference between urban processes. We validate our approach on synthetic datasets representative of a pair of time-varying urban mock variables and on two real-world open urban datasets related to climate change and infectious diseases. Results demonstrate that our framework outperforms the state of the art across diverse scenarios, minimizing false (positive and negative) inferences. Our methodology offers a powerful tool that advances data-driven urban policy and planning, leveraging open urban datasets to gain robust insight into relationships among urban variables.

1. Introduction

"Openness of data" is universally regarded as a critical requirement for scientific progress, allowing multiple scholars to synergistically build on each other's work (Nosek et al., 2015; Wilkinson et al., 2016). Recently, the meaning of this idiom has transitioned to a more comprehensive view, beyond academia, where citizens and stakeholders can easily access local government data that matter to them (Robinson et al., 2008). Consequently, several initiatives have been launched globally, including the Public Sector Information Directive in Europe, the United States (US) President Obama Open Data Initiative, the Open Government Partnership, and the G8 Open Data Charter (Attard et al., 2015). Governmental open data projects typically take the form of web portals, wherein data are accessible via the internet in a machine-readable format. These initiatives offer several advantages to citizens and stakeholders, such as fostering transparency, shared governance, and innovation, as well as enhancing the efficiency of operations through data-driven decisions (Attard et al., 2015; Mergel et al., 2018).

The growth of open government initiatives has greatly benefited urban science research, the interdisciplinary field that pursues scientific and community-based research, to establish new insights on how cities function and engineer solutions to the challenges posed by urbanization (Batty, 2013; Bettencourt, 2021). During the last two decades, urban scientists have used open urban data that capture several aspects of the fabric and function of cities, from infrastructure and transportation networks to public health and the environment (Barbosa et al., 2014;

* Corresponding author at: Department of Mechanical and Aerospace Engineering, Tandon School of Engineering, New York University, Brooklyn 11201, New York, USA.
E-mail addresses: tian.gan@nyu.edu (T. Gan), rayan.succar@nyu.edu (R. Succar), simone.macri@iss.it (S. Macrì), manuel.ruiz@upct.es (M.R. Marín), mporfiri@nyu.edu (M. Porfiri).

https://doi.org/10.1016/j.cities.2025.105980
Received 16 February 2025; Accepted 1 April 2025
Available online 15 April 2025

Lim et al., 2018). Most of these data are in the form of spatiotemporal series, in which salient variables are collected for several cities within a region or country and sampled in time over a given observation period. For example, city-level data from US official sources (Federal Bureau of Investigation and National Immigration Law Center) helped debunk the proposition that sanctuary policies promote city-level violence in a systematic way (Martínez-Schuldt & Martínez, 2019). Likewise, open city-level data from the National Aeronautics and Space Administration allowed researchers to pinpoint an association between urbanization and warming in Indian cities (Sethi & Vinoj, 2024).

Linear regression models have been the preferred approach for examining spatiotemporal series related to urban processes (Dalziel et al., 2018; Hsu et al., 2020; Hsu et al., 2022; Martínez-Schuldt & Martínez, 2019; Tanaka & Okamoto, 2021; Uto et al., 2023). While some causal associations can be inferred from this class of models, dynamical systems research in the area of causal discovery warns prudence against the overreliance on linearity as a working assumption (Schreiber, 2000; Sugihara et al., 2012). In this vein, a few studies have considered the application of state-of-the-art causal discovery tools for the study of urban data (Kissler et al., 2020; Wang et al., 2024). Wang et al. (2024) applied convergent cross mapping – a method based on nonlinear state space reconstruction (Sugihara et al., 2012) – to study causal associations between ambient temperature and other meteorological variables in US cities. Kissler et al. (2020) applied transfer entropy – an information-theoretic metric that generalizes Granger causality to nonlinear systems (Schreiber, 2000) – to infer the role of age on influenza transmission. Irrespective of their degree of sophistication, the main hurdle in all these approaches resides in the way time series from multiple cities are aggregated in the statistical analysis. Whether they score the causal effect as a correlation coefficient in a linear model (Dalziel et al., 2018; Hsu et al., 2020, 2022; Martínez-Schuldt & Martínez, 2019; Tanaka & Okamoto, 2021; Uto et al., 2023), a correlation coefficient in a phase space embedding (Wang et al., 2024), or an uncertainty reduction (Kissler et al., 2020), all these studies average the observed causal associations across multiple cities, thereby assuming all cities to be statistically equivalent.

To fully harness the potential of open urban data, we require mathematical tools capable of integrating spatiotemporal series generated from heterogeneous cities with vastly differing characteristics. Consider, for instance, the simplest, yet most fundamental, feature of a city: its population size. In the US, population size can vary from 10,000 to over 10 million (Arshad et al., 2018), challenging the tenet of a statistical equivalence among cities. A person living in a large metropolis is likely to experience a radically different milieu compared to a person living in a small city. How can we then aggregate urban data from different cities to identify causal associations about fundamental processes? For example, the number of social interactions that a person will experience in a metropolis is certainly higher on average than those experienced in a small city, making it difficult to draw causal inferences between variables that are themselves affected by social behavior. Were we to study the microbiological determinants of or the effects of vaccination on an airborne disease, our analysis would be complicated by the variations in the number of social interactions that a person experiences in cities of different sizes.

The theory of urban scaling provides an elegant framework to compare cities of widely different population sizes (Bettencourt, 2021). Through urban scaling, one can identify how chosen variables of interest vary with the population size of the city and then debias the role of the population size when comparing one city to another through scale-adjusted metropolitan indicators (SAMIs) (Bettencourt et al., 2010). Given a specific variable of interest, the SAMI of a city quantifies the extent to which that particular city differs from the theoretical norm, predicted from its population size. For example, while New York City (NYC) has a high number of homicides per resident compared to other US cities, recent work (Succar & Porfiri, 2024) indicates that the city has one of the lowest SAMIs in the US with respect to the number of homicides. Population size is an important feature, but it is unlikely to be the only one moderator of urban processes that has to be controlled for during causal discovery. Returning to the example of the impact of vaccination on airborne diseases, the number of social contacts will not only be related to population size, but also to the specific lifestyle of its dwellers and the health policies in place (Adolph et al., 2021). In this vein, one can extend the notion of SAMIs to include multivariate factors, honing the concept of features-adjusted metropolitan indicators (FsAMIs). FsAMIs replace the simple allometry used in urban scaling theory with a Cobb-Douglas function (Zellner et al., 1966), thereby allowing one to control for multiple urban features at once.

Our methodology utilizes FsAMIs to integrate data from multiple places and transfer entropy to perform causal discovery. We start from: i) time series for a set of urban variables of different places, among which we want to study causal associations (for example, vaccination and infection for an airborne disease), and ii) an array of urban features that are expected to moderate these causal associations (for example, population size, lifestyle, and health policies), see Fig. 1, left panel. Given the set of urban features, we generate time series of FsAMIs (see Fig. 1, top right panel) and apply transfer entropy to detect causal associations among them (see Fig. 1, bottom right panel). To demonstrate the viability of the proposed approach, we explored both synthetic and real-world open urban datasets, related to climate change and airborne diseases in the US. Synthetic data are representative of a pair of noisy time-varying urban variables, whose one-directional interaction is moderated by two static urban features. In the case study of climate change, we investigated the causal relationship between CO_2 emissions and temperature, controlling for population size, land area size, and unemployment rates (Reckien et al., 2015). With respect to infectious disease, we explored the causal structure among increases in COVID-19 cases, deaths, and vaccinations, controlling for population size, number of available intensive care unit beds, number of older city residents, and political leaning (Adolph et al., 2021; Pijls et al., 2021).

2. Results

2.1. Inference of causal structure from urban data

Urban scaling helps us investigate how urban variables scale with respect to population size (Bettencourt, 2021). Given an urban variable X_t, urban scaling takes the form of the following power law:

$$X_t \sim N^{\beta_t} \tag{1}$$

where N is the population size and β_t is the scaling exponent at time t. In general, we argue that the scaling at time t is superlinear if $\beta_t > 1$, sublinear if $\beta_t < 1$, and linear if $\beta_t = 1$. Scale-adjusted metropolitan indicators (SAMIs) (Bettencourt et al., 2010) allow one to compare places with respect to their nominal behavior given by power law (1). The SAMI of place i (typically, a city) at time t, corresponds to the specific residuals of the power law in logarithmic scale,

$$\xi_{i,t} = ln\frac{X_{i,t}}{C_t N_i^{\beta_t}} \tag{2}$$

where $X_{i,t}$ is the value of the chosen urban variable and C_t is a function of time that, together with β_t, is typically estimated using ordinary least squares (OLS) regression.

SAMIs offer a population-independent metric for the chosen urban variable, allowing us to compare places of different populations. However, population size may not be the sole moderator of urban dynamics, so that SAMIs may still contain spurious dependencies that hinder causal discovery. To overcome this issue, we introduce a multivariate version of the power law in the form of a Cobb-Douglas function (Zellner et al., 1966) that controls for other static spatial features besides the population size. To illustrate the idea, we consider one other static urban feature, denoted as R; the multivariate scaling becomes

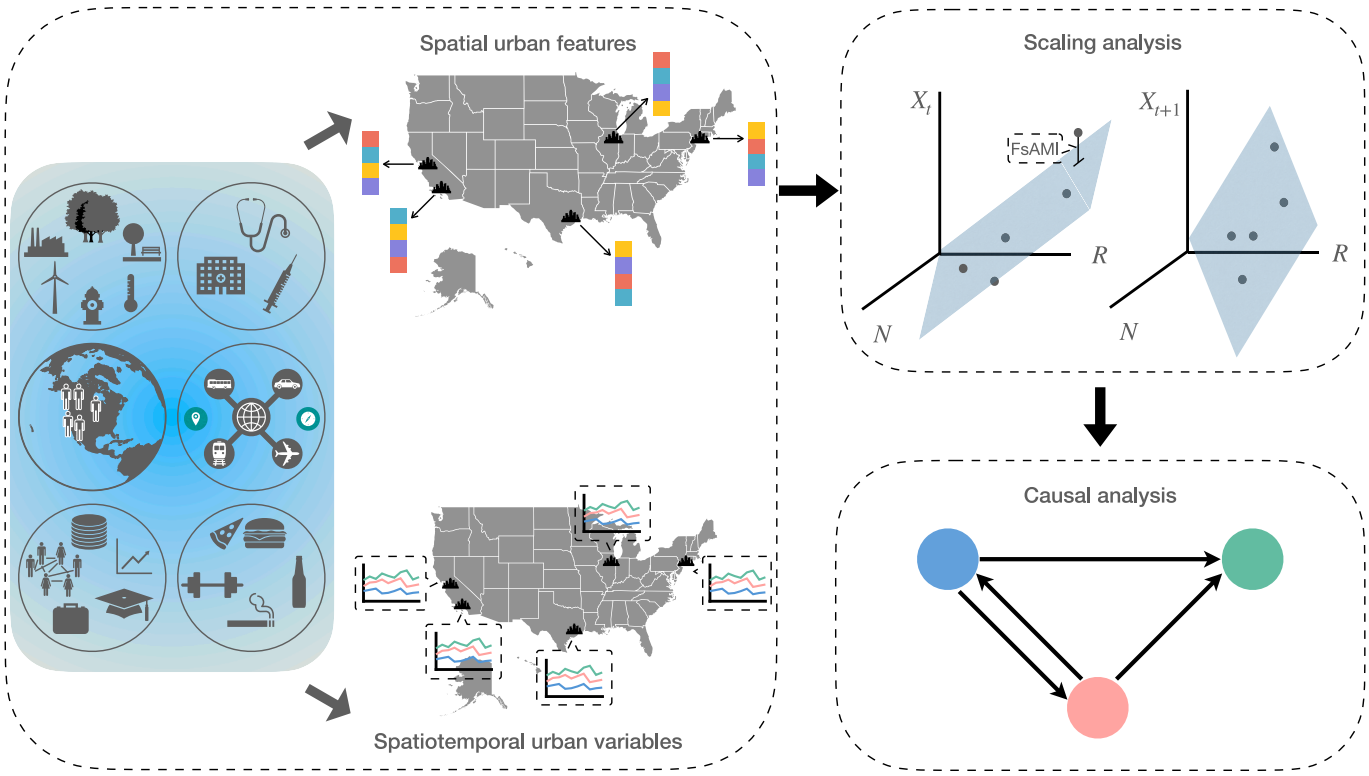


Fig. 1. Sketch of the proposed framework for causal discovery from urban spatiotemporal data. The framework takes data of mixed nature from cities. Some data remain stable over time (N and R), while others vary in time (X_t and Y_t), as shown in the left panel. The first step of the framework is to compute features-adjusted metropolitan indicators (FsAMIs) for the time-varying variables, by controlling for the effect of the static urban features via multivariate urban scaling, as depicted in the top right panel. Subsequently, these FsAMIs, in the form of city-level time series, are fed into an information-theoretic causal discovery algorithm, which yields the corresponding causal structure of the spatiotemporal urban variables, as illustrated in the bottom right panel.

$$X_t \sim N^{\beta_{N,t}} R^{\beta_{R,t}} \tag{3}$$

where $\beta_{N,t}$ and $\beta_{R,t}$ are scaling exponents for the population size and the other confounding urban feature, respectively. Based on the scale-invariant elasticity of the Cobb-Douglas model (Ribeiro et al., 2019), $\beta_{N,t} + \beta_{R,t} > 1$, $\beta_{N,t} + \beta_{R,t} < 1$, and $\beta_{N,t} + \beta_{R,t} = 1$ imply, respectively, superlinear, sublinear, and linear scaling in the two urban features. Within this multivariate scaling, one can define features-adjusted metropolitan indicators (FsAMIs), generalizing Eq. (2). Specifically, the FsAMI of place i at time t is computed as

$$\zeta_{i,t} = ln \frac{X_{i,t}}{C_t N_i^{\beta_{N,t}} R_i^{\beta_{R,t}}} \tag{4}$$

where C_t is a function of time that is computed from OLS estimation of the logarithmically transformed Cobb-Douglas function; such a function is generally different from the one obtained from traditional scaling.

FsAMIs allow one to eliminate spurious dependencies from multiple urban features, facilitating the recovery of true associations between urban processes. An effective approach to infer these associations in a causal sense is via transfer entropy – an extension of Granger causality for nonlinear dependencies (Bossomaier & Green, 2000). Given two urban variables X_t (source) and Y_t (target), we compute their respective FsAMIs with respect to a chosen set of urban features. Let us denote these FsAMIs as $\zeta^X_{i,t}$ and $\zeta^Y_{i,t}$, where, once again, i is used to specify the place and these computations are carried out over the entire system for all the available times.

We treat $\zeta^X_{i,t}$ and $\zeta^Y_{i,t}$ as the realizations of two spatiotemporal stochastic processes Z^X_t and Z^Y_t, which we assume to have the following two properties: i) stationarity in time and ii) independent and identically distributed (i.i.d.) in space. The first property allows one to focus on the time lag between the source and the target variable, without bookkeeping their individual time steps – note that the assumption can be released upon detrending the time series, so that the key requirement is for the relationship between the variables to be stable in time (see Methods). The second property allows one to assimilate all cities, without the need to retain their individual identities in any computation. In this way, we virtually aggregate space and time to increase the statistical power of our inferences, which could be otherwise unfeasible with typical open datasets (see Methods). The accuracy of these assumptions depends on the selection of the static urban features, so that the more these features are representative of potential moderators of the causal links under study.

Hence, we compute the features-adjusted transfer entropy (FsATE) from X_t to Y_t as (Bossomaier et al., 2016)

$$\mathrm{FsATE}^{(\tau_1,\tau_2)}_{X\to Y} = H\left(Z^Y_t | Z^Y_{t-1}, \ldots, Z^Y_{t-\tau_2}\right) - H\left(Z^Y_t | Z^X_{t-\tau_1}, Z^Y_{t-1}, \ldots, Z^Y_{t-\tau_2}\right), \tag{5}$$

where $H(\cdot)$ is the entropy of a random variable, $\tau_1 \geq 1$ identifies the time lag of the source variable with respect to the target variable, and $\tau_2 \geq 1$ represents the duration of the time history of the target variable. Eq. (5) could be interpreted as the reduction in the uncertainty about the knowledge of Z^Y_t from its past up to τ_2 previous times (target variable) due to additional knowledge about Z^X_t at τ_1 time steps before (source variable). To estimate transfer entropy from FsAMIs, we binned the variables into equal quantiles (Bossomaier et al., 2016) (see Methods). To test for statistical significance with respect to chance, we performed a permutation test by shuffling the time series (Bossomaier et al., 2016) with a significance level set at 0.050 (see Methods). Finally, partial correlations (noted as ρ) between variables were used to identify the signs of associations (see Methods). When working with more than two variables, we implemented features-adjusted conditional transfer

entropy (FsACTE, see Methods) that mitigates false inferences due to chain effects and multiple drivers (Bossomaier et al., 2016).

2.2. An example of bivariate linear system

To demonstrate the validity of the proposed approach and contrast its performance with the state-of-the-art, we consider a synthetic example consisting of the inference of the relationship between two mock urban variables (X_t and Y_t) from time series of length T over M cities. Our ground truth is that the variables are one-directionally coupled (Y_t influences X_t, but not vice versa) and that the FsAMIs are stationary in time and i.i.d. across cities.

We generate the data starting from the following dynamics for the FsAMIs:

$$\begin{aligned} Z^X_{t+1} &= 0.5Z^X_t + 2Z^Y_t + \varepsilon^X \\ Z^Y_{t+1} &= 0.6Z^Y_t + \varepsilon^Y \end{aligned} \tag{6}$$

where ε^X and ε^Y are independent standard Gaussian variables. By running Eq. (6) over windows of length $T = 100$, we generate our synthetic FsAMIs, which, in turn, are used to form $M = 100$ city-level time series X_t and Y_t by considering two urban mock feature (population size, N, and an additional mock urban features, R), that is,

$$\begin{aligned} X_t &= N^{\beta^X_N} R^{\beta^X_R} e^{Z^X_t}, \\ Y_t &= N^{\beta^Y_N} R^{\beta^Y_R} e^{Z^Y_t}. \end{aligned} \tag{7}$$

where $\beta^X_N = 1.2, \beta^Y_N = 0.8, \beta^X_R = 0.3$, and $\beta^Y_R = 0.1$.

This synthetic dataset is input to our algorithm, which applies the OLS regression on the logarithmically transformed data to estimate the FsAMIs for X_t and Y_t in Eq. (4) and then FsATE in Eq. (5) with $\tau_1 = \tau_2 = 1$ to detect the presence of a directional relationship between them (see Methods). Over 100 independent replications, our algorithm does not produce false negative results and has a false positive rate of 10 %. Thus, while our approach never missed the relationship from Y_t to X_t, only in 10 of the 100 repetitions did it lead to the prediction of a directional association from X_t to Y_t.

We compare our framework with a traditional linear regression between the per-capita variables, using the moderators as exposure variables. Specifically, we estimated the per-capita values of X_t and Y_t by dividing the spatiotemporal series X_t and Y_t by the corresponding N. In this case, we register a false positive rate of 70 %, in stark contrast to the findings through our approach. The estimated causal graphs constructed using our approach and traditional regression are, respectively, shown in Fig. 2. The improvement offered by our approach over traditional regression is not limited to a specific set of model parameters. To substantiate this statement, we repeated the comparison by systematically varying the scaling exponents for the two urban mock features ($\beta_{N,t}$ and $\beta_{R,t}$) in Eq. (7), as well as the number of cities (M). As shown in Fig. 3, our approach consistently yields lower false positive rates than traditional regression, especially when integrating observations from a large number of cities. In almost all cases, traditional regression underperforms compared to our approach, exhibiting higher false positive rates. In the Supplementary Information S1, we present an extended comparison against variations of traditional regression, in which we consider autoregression or undertake the analysis using transfer entropy (akin to the proposed approach, but without scaling). Across a wide range of parameter combinations, our method demonstrates superior performance in minimizing false rates.

2.3. Real-world case studies

We explored two open urban datasets in the real world to demonstrate the inference of causal associations between urban variables in pairs and triplets, moderated by population size and other urban features. Values of FsATE, Eq. (5) and FsACTE, Eq. (10), were computed

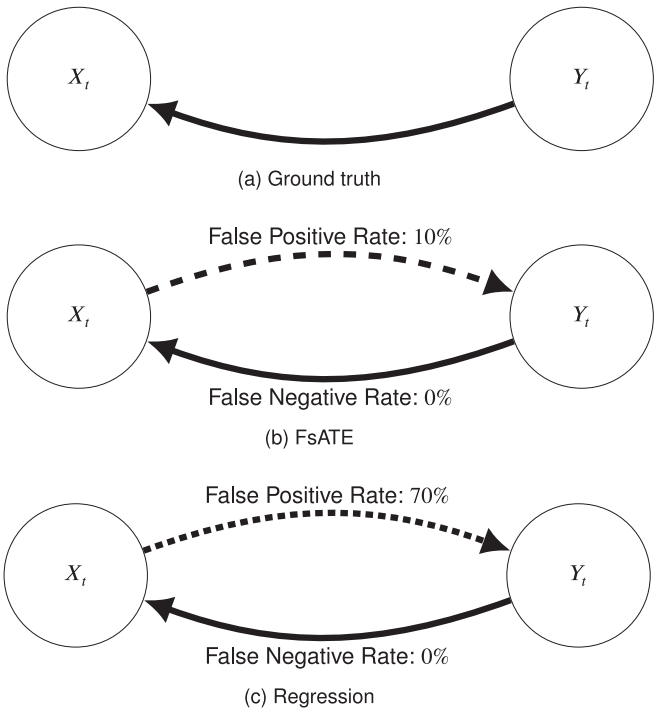


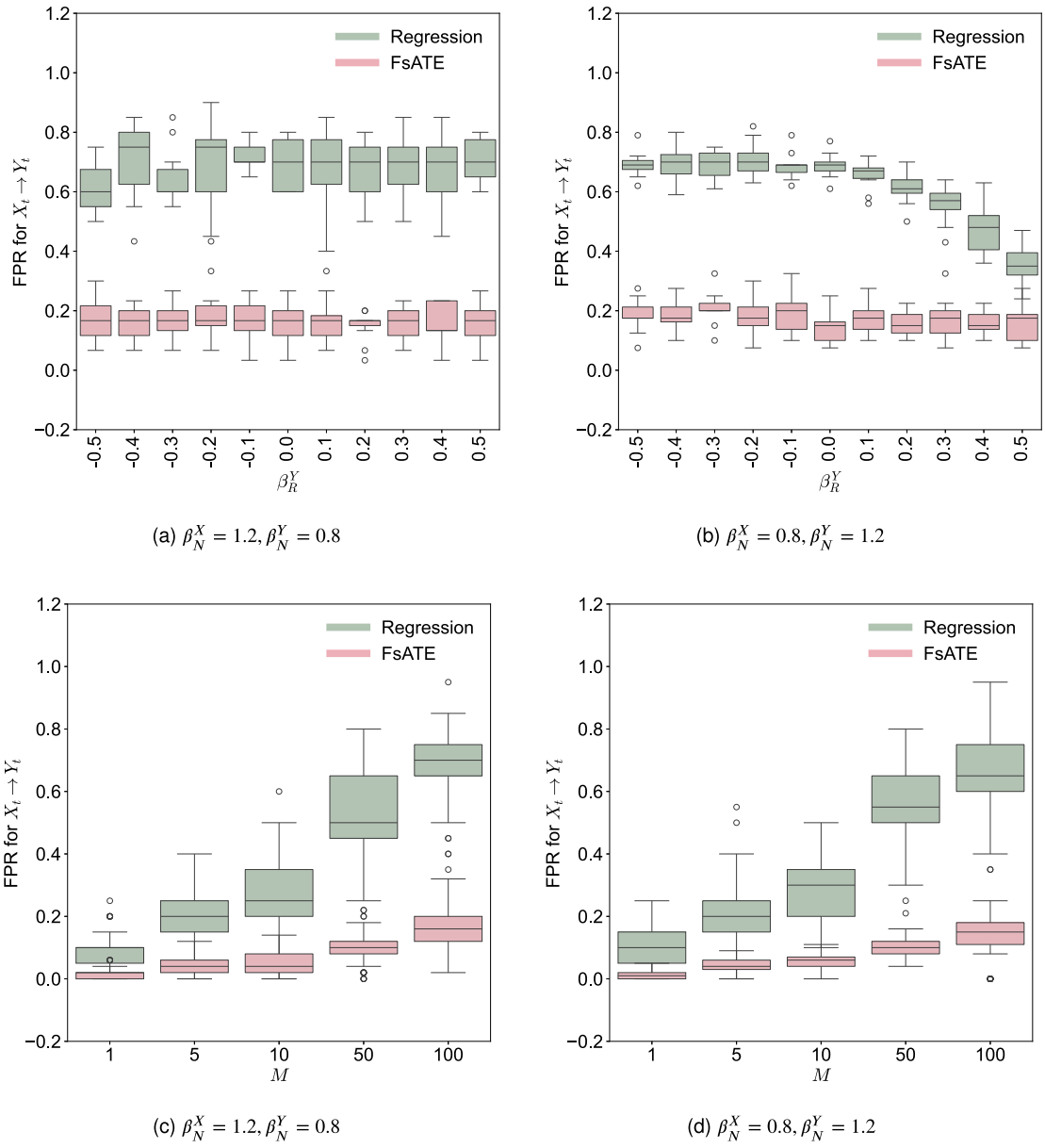
Fig. 2. Causal graphs derived from synthetic data generated by a bivariate linear system. (a) Ground truth, pertaining to model data from Eqs. (6) and (7). (b) Estimated causal graph using features-adjusted transfer entropy (FsATE), computed based on Eqs. (4) and (5). (c) Estimated causal graph using regression on per-capita variables, while controlling for *logN* and *logR* as exposures (see Eq. (11) in Methods). False positive (X_t influences Y_t) and negative (Y_t does not influence X_t) rates are computed over 100 repetitions.

using four bins (see Methods) and the length of the time histories was adjusted accordingly to balance between the need to properly describe the dynamics of the variables and accurately reconstructing the probability mass functions for the computation of transfer entropy measures (see Methods). All multivariate scaling analyses based on Eq. (3) are reported in the Supplementary Information S2.

2.3.1. Case study I: Climate change

As a case study with real data that would mirror the bivariate system examined above, we investigated the relationship between CO_2 emissions (in kilograms) and average temperature (in Kelvin) in the US from 1995 to 2015 (see Methods). Given the yearly resolution of the data, we considered the smallest lag of $\tau_1 = 1$ for the source variable and we performed the analysis at the county-level (see Methods). We chose population size, land area, and prevalence of unemployment as the urban features in the analysis based on prior work suggesting that these factors significantly affect emissions and climate across regions (Reckien et al., 2015) (see Methods). Upon controlling for these factors, FsAMIs were found to be spatially decoupled over the entire observation window (Moran's I of CO_2 emission across counties range from -0.0008 to 0.0010; Moran's I for average temperature ranges between 0.0402 and 0.0540, see Methods). Counties were also found to be nearly identical in both CO_2 emissions (100.0 % of counties are indistinguishable through Kolmogorov-Smirnov test, see Supplementary Information S3) and average temperature (69.84 % of counties are indistinguishable through Kolmogorov-Smirnov test, see Supplementary Information S3).

Through the application of our approach, we verified the prediction (Madden & Ramanathan, 1980) that CO_2 emission influences average temperature (FsATE $= 0.0264$; $p < 0.001$ and 95 % quantile from permutation tests $= 0.0239$), but not vice versa (FsATE $= 0.0201$; $p = 0.291$ and 95 % quantile from permutation tests $= 0.0208$), see Fig. 4. From partial correlation ($\rho = 0.0224$), we further confirmed (Madden &



(a) $\beta_N^X = 1.2, \beta_N^Y = 0.8$ (b) $\beta_N^X = 0.8, \beta_N^Y = 1.2$

(c) $\beta_N^X = 1.2, \beta_N^Y = 0.8$ (d) $\beta_N^X = 0.8, \beta_N^Y = 1.2$

Fig. 3. Sensitivity analyses for synthetic data generated by a bivariate linear system. Boxplots of false positive rates (FPR) computed from 100 trials of a bivariate linear system over $T = 100$ time steps for different scaling exponents (β_N^X, β_N^Y, β_R^X, and β_R^Y) and number of cities (M): (a) $M = 100$, $\beta_N^X = 1.2$, and $\beta_N^Y = 0.8$, β_R^Y varying from -0.5 to 0.5 in steps of 0.1 (on the horizontal axis), and β_R^X varying from -0.5 to 0.5 in steps of 0.1 (condensed in the box plots); (b) $M = 100$, $\beta_N^X = 0.8$, and $\beta_N^Y = 1.2$, β_R^Y varying from -0.5 to 0.5 in steps of 0.1 (on the horizontal axis), and β_R^X varying from -0.5 to 0.5 in steps of 0.1 (condensed in the box plots); (c) $M \in \{1, 5, 10, 50, 100\}$ (on the horizontal axis) $\beta_N^X = 1.2$, $\beta_N^Y = 0.8$, and both β_R^X and β_R^Y varying from -0.5 to 0.5 in steps of 0.1 (condensed in the box plots); and (d) $M \in \{1, 5, 10, 50, 100\}$ (on the horizontal axis) $\beta_N^X = 0.8$, $\beta_N^Y = 1.2$, and both β_R^X and β_R^Y varying from -0.5 to 0.5 in steps of 0.1 (condensed in the box plots). Each box plot in (a) and (b) comprises 1100 points (100 repetitions and 11 values of β_R^X), while each box plot in (c) and (d) comprises 12,100 points (100 repetitions and 121 combinations of β_R^X and β_R^Y). The green boxes show FPRs from traditional regression, while the pink boxes show FPRs from FsATE. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Ramanathan, 1980) that higher CO_2 emissions are associated with rising temperature.

2.3.2. *Case study II: Infectious diseases*

As a case study of a problem in which three urban variables interact with each other, we considered the COVID-19 pandemic in the US. Specifically, we tested our framework against the following predictions: i) a higher incidence of cases is conducive to more deaths (Stokes et al., 2021; Woolf et al., 2021), and ii) higher levels of vaccination in the population can reduce infections (Tregoning et al., 2021). We used weekly data at the level of metropolitan and micropolitan statistical areas (MSAs) in the US from 01/27/2021 to 11/23/2022 for new infection cases, deaths, and increases in the number of persons who completed the full vaccination series. Given prior evidence that 20 days is the median time duration for viral shedding (Zhou et al., 2020), we chose $\tau_1 = 3$. Based on prior work identifying social contacts, access to care, age, and political leaning as candidate moderators (Adolph et al., 2021; Pijls et al., 2021), we chose population size, number of ICU beds, prevalence of people aged more than 60-year-old, and number of persons who voted Democrat in the 2020 election as urban features. Upon controlling for these factors, the FsAMIs were found to be spatially decoupled over the entire observation window (Moran's I for the

(a)

X	Y	τ_1	τ_2	FsÂTE	p-value	ρ
CO_2	TEMP	1	4	0.0264 (0.0239)	< 0.001	0.0224
TEMP	CO_2	1	4	0.0201 (0.0208)	0.291	−0.0116

(b)

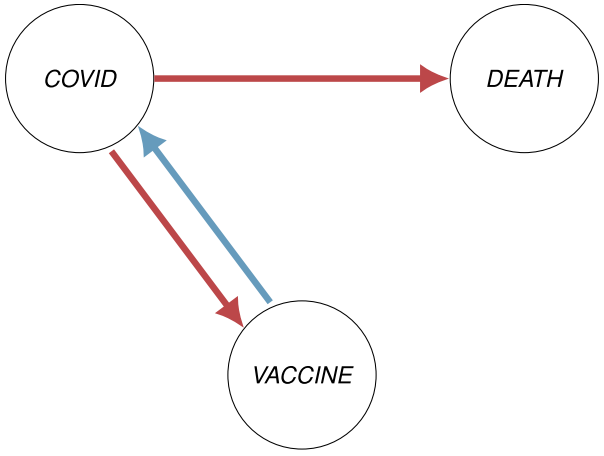
Fig. 4. Causal results established from scaled climate change data. (a) Causal graphs between FsAMIs of spatio-temporal urban variables, measured by FsATE: total emissions in kilograms CO_2 per year (CO_2), and average temperature (in Kelvin) per year (*TEMP*). (b) Synoptic presentation of statistical results; columns from left to right respectively contain: source variables, target variables, time lag between source and target variables, time history of target variable, estimated FsATE, p-values, and partial correlations. The numbers in parentheses denote the 95 % quantile obtained from a permutation test, in which each observed FsATE was compared with the random FsATE estimated from the FsAMIs of 10,000 surrogate time series.

number of weekly new infections ranges between −0.0140 and 0.1442 with only five of the total 96 weeks registering values above 0.1; Moran's I for deaths ranges between −0.0147 and 0.0496; and Moran's I for the increase in the number of people who completed full vaccination series ranges between −0.0191 and 0.0793). MSAs were also found to be nearly identical in number of weekly new infections (94.31 % of MSAs are indistinguishable through Kolmogorov-Smirnov test, see Supplementary Information S3), weekly new deaths (60.14 % of MSAs are indistinguishable through Kolmogorov-Smirnov test, see Supplementary Information S3), and the increase in the number of people who completed full vaccination series (80.83 % of MSAs are indistinguishable through Kolmogorov-Smirnov test, see Supplementary Information S3).

Using our approach, we verified that the incidence of COVID-19 affects the death toll (FsACTE = 0.0310; $p < 0.001$ and 95 % quantile from permutation tests = 0.0220), but not vice versa (FsACTE = 0.0186; $p = 0.506$ and 95 % quantile from permutation tests = 0.0195), see Fig. 5. We further verified that higher number of new COVID-19 infection cases elevates the number of new COVID-19 death cases ($\rho = 0.1116$). Besides, we confirmed that higher number of people who completed full vaccination series reduced the risk of COVID-19 infections (FsACTE = 0.0194; $p = 0.005$ and 95 % quantile from permutation tests = 0.0189; and partial correlation $\rho = -0.0227$). Likewise, we observed that higher levels of infections were conducive to higher vaccination rates (FsACTE = 0.0208; $p < 0.001$ and 95 % quantile from permutation tests = 0.0191; and partial correlation $\rho = 0.0033$).

3. Discussion

Causal discovery from observational data is a challenging problem that is central to urban policy-making (Kandt & Batty, 2021). The availability of open data to support these analyses has fueled several research directions, spanning from the application of well-established linear regression models to state-of-the-art causal discovery tools from



(a)

X	Y	Z	τ_1	τ_2	FsÂCTE	p-value	ρ
COVID	DEATH	VACCINE	3	3	0.0310 (0.0220)	< 0.001	0.1116
COVID	VACCINE	DEATH	3	3	0.0208 (0.0191)	< 0.001	0.0033
DEATH	COVID	VACCINE	3	3	0.0186 (0.0195)	0.506	−0.0160
DEATH	VACCINE	COVID	3	3	0.0174 (0.0217)	1.0	0.0043
VACCINE	COVID	DEATH	3	3	0.0194 (0.0189)	0.005	−0.0227
VACCINE	DEATH	COVID	3	3	0.0180 (0.0288)	1.0	−0.0214

(b)

Fig. 5. Causal results established from scaled infectious disease data. (a) Causal graph and detailed causal results among FsAMIs of spatio-temporal urban variables, measured by FsATE: new COVID-19 infection cases (*COVID*), deaths (*DEATH*), and increases in the number of people who completed the full vaccination series (*VACCINE*). Red arrows indicate positive association. (b) Synoptic presentation of statistical results; columns from left to right respectively contain: source variables, target variables, driver variables, time lag between source and target variables, time history of target variable, estimated FsACTE, p-values, and partial correlations. The numbers in parentheses denote the 95 % quantile obtained from a permutation test, in which each observed FsACTE was compared with the random FsACTE estimated from the FsAMIs of 10,000 surrogate time series. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

dynamical systems theory (Dalziel et al., 2018; Hsu et al., 2020, 2022; Kissler et al., 2020; Martínez-Schuldt & Martínez, 2019; Tanaka & Okamoto, 2021; Uto et al., 2023; Wang et al., 2024). Common to these approaches is the premise that cities are statistically equivalent. In this paper, we challenge this proposition and demonstrate that overlooking differences among cities with respect to their population and other salient features could hinder reliable inference of causal associations from urban datasets. We introduce a new urban metric – features-adjusted metropolitan indicators (FsAMIs) – as an extension of classic urban scaling law to integrate a mixture of urban factors. Such a metric allows controlling for the effect of multivariate urban moderators and mitigating spurious dependencies across urban variables. Our approach to causal discovery is based on the use of transfer entropy on FsAMIs, within the premise of what we call features-adjusted transfer entropy (FsATE).

To demonstrate the viability of our approach, we conducted numerical experiments on multiple synthetic datasets associated with a simple bivariate linear system. Our numerical results indicate that FsATE consistently outperforms traditional regression models across a wide range of scenarios, minimizing false positive results. Regression models are particularly unreliable in the presence of urban features with datasets comprising many cities. In these cases, the tails in the distributions of the urban features tend to magnify the role of heterogeneities and skew inferences based on per-capita values. The quality of the inferences does not improve when accounting for autoregression or undertaking the inference using transfer entropy, pointing at the critical need to debiasing urban data through scaling before attempting causal discovery.

Working with synthetic data is an elegant way to test the viability of an approach against ground truth, yet, synthetic data fail to capture the intricacies of real-world datasets for which our methodology was ultimately designed. To demonstrate our approach on real-world data, we considered two examples that are very well studied in the literature: the role of CO_2 emissions on temperature and the role of vaccines on COVID-19 spreading and death in the US. Working with county-level data at a yearly resolution and controlling for population size, land area, and prevalence of unemployment, we confirmed established evidence (Madden & Ramanathan, 1980) of a causal role of CO_2 emission on temperature. Working with MSA-level data at a weekly resolution and controlling for population size, number of ICU beds, prevalence of people aged more than 60-year-old, and number of persons who voted Democrat in the 2020 election, we found evidence supporting the beneficial role of vaccination on infections and the consequent effect on deaths (Tregoning et al., 2021). Our analysis also unveiled a positive relationship between number of infections and vaccination rates, so that the more the disease would spread in a city, the more would be the vaccine uptake in the US. While not part of our original set of expectations, this effect could have been anticipated based on prior work (Karlsson et al., 2021), suggesting that vaccine uptake may rise during infection waves out of an increased fear of being infected.

In this work, we adopted the concept of causality known as Granger (or Wiener-Granger) causality (Granger, 1969), originally developed to analyze temporal data in economics. The core principle of this notion is that the predictability of a future "target" variable improves when information about the present state of a "source" variable is taken into account. In the original framework by Granger, predictability is assessed by measuring the covariance between the autoregressive processes (source and target). Transfer entropy (Schreiber, 2000) generalizes this notion to nonlinear systems, scoring predictability in an information-theoretic sense, rather than through a covariance, and it leads to equivalent predictions to Granger causality for linear systems (Barnett et al., 2009). Whether it is traditional Granger causality or transfer entropy, the notion of causality does not rely on interventions with randomization or controlled conditions in laboratory settings. Other notions of causality, such as those pioneered by Pearl (2009) that focus on structural causal models, or machine-learning-based approaches (Banerjee et al., 2023) offer alternative frameworks to transfer entropy-based causal inference. While distinct in their assumptions and methodologies, these methods could potentially be integrated as part of our framework; for example, we envision the use of the PCMCI algorithm developed by Runge et al. (2019) on FsAMIs to account for broader definitions of causal inference.

Our approach is not free of limitations. First, the analysis implemented in the multivariate urban scaling assumes the absence of hidden urban variables or spatial features. The resulting FsAMIs from multivariate urban scaling are assumed to have no spurious relationship upon controlling spatial confounding features and be self-contained with respect to any other process within the urban system. A potential way to mitigate the presence of hidden variables or at least to alert if they have some role in the observed causal structures could leverage recent advancements in the PCMCI algorithm, such as the latent-PCMCI which does not assume the absence of hidden variables (Gerhardus & Runge, 2020). Alternatively, one might attempt at expanding the notion of a detection matrix (Haehne et al., 2019) to a spatiotemporal process to infer the number of hidden variables from the rank of ancillary matrices. Second, we assume that spatiotemporal variables do not cause the static features, thereby avoiding spurious causality arising from conditioning on colliders (Pearl, 2009). In most cases, this assumption poses no significant challenge, as the static features remain invariant over time, allowing us to argue that temporal variables cannot influence them within the studied timescales. However, for relatively slow-changing variables, this assumption may be limiting. Third, all the spatiotemporal variables taken into account for Granger causality are assumed to be i.i.d. in space and stationary in time. While the requirement of stationarity in time is not as stringent, given the possibility to detrend the time series, being i.i.d. in space is a rather strict requirement. Should that condition not be fulfilled, one may attempt to expand the Cobb-Douglas function to include spatial interactions, following the recent work by Xiao and Gong (2022).

Open data fosters a culture of transparency and collaboration, enabling researchers, policymakers, and community stakeholders to work together in addressing urban challenges, such as housing affordability, transportation efficiency, environmental sustainability, and social equity. We call for local governments, organizations, and other data custodians to expand their sharing of updated data in machine-readable formats, supplemented by relevant metadata to facilitate accurate interpretation. By fostering an ecosystem of open data, cities can create the foundations for mathematical tools like ours to be successfully implemented, towards new discoveries in urban science and informed decision. The application of these tools on large datasets can help unveil relationships between urban processes that are presently elusive and empower communities with the knowledge needed to address everyday urban challenges and plan for a more sustainable and inclusive future.

4. Methods

4.1. Datasets

4.1.1. Climate change

We obtained the yearly CO_2 emission data between 1995 and 2015 and land area data for 3071 US counties from the Oak Ridge National Laboratory Distributed Active Archive Center (DACC) (Environmental Sciences Division, Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC), 2024)– an archive center focused on Earth science data, particularly in areas related to biogeochemical and ecological processes. County-level temperature data in 1995–2015 were obtained from the Parameter-elevation Regressions on the Independent Slopes Model (PRISM) dataset (PRISM Climate Group, 2024). The employment data were retrieved from the 2020 County Health Rankings National Findings Report (University of Wisconsin Population Health Institute, 2020).

4.1.2. Infectious disease

To evaluate how the COVID-19 pandemic, deaths, and vaccinations vary based on population size and relevant urban characteristics in the US, and to assess whether and how they are associated, we utilized freely available county-level data sources. Specifically, we used county-level data between 01/27/2021 and 11/23/2022 for new infection cases, deaths, and increases in the number of persons who completed the full vaccination series, collected and provided by the Johns Hopkins Coronavirus Resource Center (Dong et al., 2020). Additionally, we collected the capacity of intensive care unit (ICU) beds which identify the number of ICU beds each hospital reported most recently, and the number of people aged 60 and older released by *Kaiser Health News* (Schulte et al., 2020) (these data were available between fiscal year 2018 and fiscal year 2019 at the time of this study), as well as the county-level presidential election returns in 2020, published by the MIT Election Data and Science Lab (Data & Lab, 2018). We grouped the counties in 721 MSAs using CBSA to FIPS County Crosswalk files (of Economic Research, 2023).

4.2. Fitting of urban scaling laws

To determine the scaling exponents of population size and other urban moderators ($\beta_{N,t}$ and $\beta_{R,t}$), and the FsAMIs of all spatiotemporal variables of interest per place at each time step, we applied OLS regression to the logarithmically transformed data. Heteroskedasticity was considered by using a heteroskedasticity-consistent covariance estimator (so-called HC1) (MacKinnon & White, 1985) in OLS regression when calculating standard errors for R^2 and the 95 % confidence interval for $\beta_{N,t}$ and $\beta_{R,t}$. The linear regression analyses were conducted using the Python library *Statsmodels* (Seabold & Perktold, 2010).

4.3. Estimating features-adjusted transfer entropy

The entropy of a discrete random variable X, measured in nats, is defined as (Bossomaier & Green, 2000)

$$H(X) = -\sum \mathrm{P}(X=x) ln \mathrm{P}(X=x), \tag{8}$$

where $\mathrm{P}(\cdot)$ is the probability, and the sum is over the sample space. The other measure needed to compute transfer entropy is the conditional entropy between two random variables X and Y which reads (Bossomaier & Green, 2000)

$$H(Y|X) = -\sum \mathrm{P}(X=x, Y=y) ln \frac{\mathrm{P}(X=x, Y=y)}{\mathrm{P}(X=x)}. \tag{9}$$

Eqs. (8) and (9) can be replaced in Eq. (5) to express FsATE in terms of probability mass functions, which should be estimated from data. More specifically, under the assumption that the FsAMIs are independently and identically distributed in space, we combined all places for the estimation of the probability mass functions through binning. Such an approach begets a dataset of MT samples (T time steps and M cities). With such a dataset, one can tackle the b^{τ_2+2} probabilities needed for the probability mass functions in Eq. (5), with b being the number of bins. Given that at least five-to-ten samples are needed for estimating any given probability values (Rohatgi, 1976), the values of b and τ_2 should satisfy $MT > 10b^{\tau_2+2}$. We used four bins ($b = 4$), each containing 25 % of the data, for all the computations. Using $\tau_2 = 4$, this implies that $MT > 40,960$, which, for the US, we satisfy by working with county-level data ($M = 3,071$) when only yearly observations ($T \in [10, 100]$) are available and with MSAs ($M = 721$) when we have access to weekly data ($T \in [100, 1000]$). Note that a transfer entropy analysis within each place would be impossible as one would be asked to estimate the same amount of probability values with only T data.

To assess the statistical significance of the estimated FsATE, we relied on bootstrapping with a significance level set at 0.050. We generated a surrogate distribution under the null hypothesis of independence. This distribution was created by shuffling the source variable randomly within each place 10,000 times and calculating FsATE on the resulting shuffled data. Upon generating the surrogate distribution, the p-values were then computed as the quantile of the 10,000 surrogate FsATE that is greater than the observed FsATE of the original data. Once a causal link is deemed statistically significant, we computed $\rho\left(Z^X_{t-\tau_1}; Z^Y_t | Z^Y_{t-1}, \ldots, Z^Y_{t-\tau_2}\right)$ to detail the sign of the association.

In systems with multiple spatiotemporal urban variables, one should utilize features-adjusted conditional transfer entropy (FsACTE) to control for temporal confounders. FsACTE from X_t to Y_t, given W_t, is defined as

$$\mathrm{FsACTE}^{(\tau_1,\tau_2)}_{X\to Y|W} = H\left(Z^Y_t | Z^W_{t-\tau_1}, Z^Y_{t-1}, \ldots, Z^Y_{t-\tau_2}\right) - H\left(Z^Y_t | Z^X_{t-\tau_1}, Z^W_{t-\tau_1}, Z^Y_{t-1}, \ldots, Z^Y_{t-\tau_2}\right) \tag{10}$$

where τ_1 is the time lag between the source X_t (or control variable W_t) and the target variable Y_t, and τ_2 represents the time history of the target variable. To identify the sign of the causal associations, we computed $\rho\left(Z^X_{t-\tau_1}; Z^Y_t | Z^W_{t-\tau_1}, Z^Y_{t-1}, \ldots, Z^Y_{t-\tau_2}\right)$ for trivariate FsAMIs. For the estimation of probability mass functions, the value of τ_2 shall be reduced compared to what is done in standard transfer entropy computation to account for the larger dimensionality of the probability mass functions. To assess the statistical significance of the estimated FsACTE, we compared the observed FsACTE with the random FsACTE estimated from the empirical distribution over 10,000 surrogate computations with a significance level set at 0.050. We first identified $\left(Z^Y_t, Z^W_t\right)$ with 16 possible pairs and allocated them to 16 different symbols. Secondly, we generated 10,000 surrogate symbolic time series by shuffling the source variable Z^X_t randomly within each of these 16 symbols for each MSA while preserving the association between $\left(Z^Y_t, Z^W_t\right)$ and Z^X_t; such an approach mirrors the strategy proposed by Runge (2018).

4.4. Traditional regression model

To compare with traditional methods (Dalziel et al., 2018; Hsu et al., 2020, 2022; Martínez-Schuldt & Martínez, 2019; Tanaka & Okamoto, 2021; Uto et al., 2023), we implemented the regression model using the OLS estimator on per-capita values, while controlling for the moderators ($logN$ and $logR$) as exposure variables, that is,

$$\frac{Y_t}{N} = \alpha_N logN + \alpha_R logR + \alpha_X \frac{X_{t-1}}{N} + \epsilon_t, \tag{11}$$

where $\frac{Y_t}{N}$ is the per-capita measures of Y_t, $\frac{X_t}{N}$ is the per-capita measures of X_t, α_X denotes the effect of X_t on Y_t, with one time lag, and ϵ_t is the error term. To identify the causal relationship from X_t to Y_t, we checked if the p-value of coefficient α_X is less than 0.050. We note that the specific form of this model might slightly change depending on the problem at hand.

4.5. Synthetic data through bivariate linear systems

To generate the synthetic dataset, we modeled $M = 100$ cities with two urban variables X_t and Y_t, and two spatial urban moderators, population size N, and another spatial urban feature R. We first generated the FsAMIs (Z^X_t and Z^Y_t) of the spatiotemporal variables X_t and Y_t over $T = 100$ time steps following a bivariate linear system in which Y_t causes X_t, but X_t does not cause Y_t, as presented in Eq. (6). Population sizes were generated using ranks based on Zipf's law (Gabaix, 1999), where higher-ranked cities have smaller populations. These were computed using the formula

$$N_i = \frac{i^{-\gamma}}{max\{1^{-\gamma}, \ldots, M^{-\gamma}\}}(Max - Min) + Min \tag{12}$$

with $\gamma = 1$, $i = [1, \cdots, M]$ represents a city, so that the resulting popula-

tion values fall between $Min = 10,000$ and $Max = 10,000,000$. Similarly, the other spatial feature vector R was generated using $\gamma = 1.5$, $Min = 1,000$, and $Max = 100,000$. The spatiotemporal series X_t and Y_t were then generated using multivariate urban scaling according to Eq. (4). Specifically, given a city i, we write

$$\begin{aligned} X_{i,t} &= N_i^{\beta_N^X} R_i^{\beta_R^X} e^{Z_t^X}, \\ Y_{i,t} &= N_i^{\beta_N^Y} R_i^{\beta_R^Y} e^{Z_t^Y}. \end{aligned} \tag{13}$$

where $\beta_N^X = 1.2, \beta_N^Y = 0.8, \beta_R^X = 0.3$, and $\beta_R^Y = 0.1$. We then applied OLS regression on the logarithmic transformation of the synthetic data to estimate the FsAMIs for X_t and Y_t, and estimated the per-capita values of X_t and Y_t by dividing the spatiotemporal series X_t and Y_t by the corresponding N.

To compare our approach with traditional regression across a range of scenarios, we generated multiple synthetic datasets through the bivariate linear system. We first modeled cities with two spatiotemporal variables X_t and Y_t, in which the FsAMI time series were generated over 100 time steps following the bivariate linear system in Eq. (6), and the two urban mock moderators N and R were generated using Eq. (12) by systematically varying parameters. Specifically, we tested four sets of parameters: i) $M = 100, \beta_N^X = 1.2, \beta_N^Y = 0.8$; ii) $M = 100, \beta_N^X = 0.8, \beta_N^Y = 1.2$; iii) $M \in \{1, 5, 10, 50, 100\}$, $\beta_N^X = 1.2$, $\beta_N^Y = 0.8$; and iiii) $M \in \{1, 5, 10, 50, 100\}, \beta_N^X = 0.8, \beta_N^Y = 1.2$. In all these cases, we had β_R^X and β_R^Y varied incrementally from -0.5 to 0.5 in steps of 0.1.

4.6. Testing the assumptions on the urban variables

Each FsAMI time series was tested for stationarity using the augmented Dickey-Fuller test using the Python library *Statsmodels* (Seabold & Perktold, 2010). Time series failing to meet the stationarity assumption at a confidence level of 0.05 were linearly detrended using the Python library *SciPy* (Virtanen et al., 2020).

We then tested if these stationary time series were spatially dependent at each time step using Moran's I (Li et al., 2007),

$$I_t = \frac{M}{W} \frac{\sum_{i=1}^{M} \sum_{j=1}^{M} w_{ij} \left(Z_{i,t} - \overline{Z}_t\right) \left(Z_{j,t} - \overline{Z}_t\right)}{\sum_{i=1}^{M} \left(Z_{i,t} - \overline{Z}_t\right)^2}, \tag{14}$$

where $W = \sum_{i=1}^{M} \sum_{j=1}^{M} w_{ij}$, $\overline{Z}_t$ represents the spatial average, and w_{ij} is the spatial weight between places i and j, which we took to be the distance between them. We computed Moran's I using Python library *PySAL* (Rey & Anselin, 2009). Based on Li et al. (2007) we used $-0.1 < I < 0.1$ as a threshold for arguing in favor of independence, at least to a first degree of an approximation.

Finally, we tested if the distributions of the FsAMI time series between every two places are statistically identical using the Kolmogorov-Smirnov test from *SciPy*. The null hypothesis of the test is that the distributions are identical, so that p-values above $\frac{0.05}{M-1}$ (where we employed the Bonferroni correction (Hochberg, 1987)) were used to advocate in favor of FsAMIs from different cities being identical. The test was performed on the binned time series, which were used to compute the FsATE and FsACTE.

CRediT authorship contribution statement

Tian Gan: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Rayan Succar:** Conceptualization, Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Simone Macrì:** Conceptualization, Investigation, Supervision, Validation, Writing – original draft, Writing – review & editing. **Manuel Ruiz Marín:** Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – review & editing. **Maurizio Porfiri:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

We declare that we have no competing interests

Acknowledgments

T.G., R.S., and M.P. were supported by National Science Foundation grant numbers ECCS-1928614, DUE-2129076, and CMMI-1953135 and by the Pilot Project Program grant number A23-0539 from New York University's Collaborative Center in Children's Environmental Health Research and Translation. M.R.M. was supported by the Spanish Ministry of Science and Innovation (MCIN/AEI/10.13039/ 501100011033) through grant number (PID2022-136252NB-I00).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cities.2025.105980.

Data availability

Scripts and datasets that are allowed to be shared are available on Dynamical Systems Laboratory's Github.

References

Adolph, C., Amano, K., Bang-Jensen, B., Fullman, N., & Wilkerson, J. (2021). Pandemic politics: Timing state-level social distancing responses to COVID-19. *Journal of Health Politics, Policy and Law, 46*, 211–233.

Arshad, S., Hu, S., & Ashraf, B. N. (2018). Zipf's law and city size distribution: A survey of the literature and future research agenda. *Physica A: Statistical Mechanics and its Applications, 492*, 75–92.

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly, 32*, 399–418.

Banerjee, A., Chandra, S., Ott, E., 2023. Network inference from short, noisy, low time-resolution, partial measurements: Application to C. *Elegans* neuronal calcium dynamics. Proceedings of the National Academy of Sciences 120, e2216030120.

Barbosa, L., Pham, K., Silva, C., Vieira, M. R., & Freire, J. (2014). Structured open urban data: Understanding the landscape. *Big Data, 2*, 144–154.

Barnett, L., Barrett, A. B., & Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for Gaussian variables. *Physical Review Letters, 103*, Article 238701.

Batty, M. (2013). *The new science of cities*. Cambridge, Massachusetts: MIT Press.

Bettencourt, L. M. (2021). *Introduction to urban science: Evidence and theory of cities as complex systems*. Cambridge, Massachusetts: MIT Press.

Bettencourt, L. M., Lobo, J., Strumsky, D., & West, G. B. (2010). Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PLoS One, 5*, Article e13541.

Bossomaier, T., Barnett, L., Harré, M., Lizier, J. T., Bossomaier, T., Barnett, L., Harré, M., & Lizier, J. T. (2016). *Transfer entropy*. New York City, NY: Springer.

Bossomaier, T. R., & Green, D. G. (2000). *Complex systems*. United Kingdom: Cambridge University Press.

Dalziel, B. D., Kissler, S., Gog, J. R., Viboud, C., Bjørnstad, O. N., Metcalf, C. J. E., & Grenfell, B. T. (2018). Urbanization and humidity shape the intensity of influenza epidemics in US cities. *Science, 362*, 75–79.

Data, M. E., & Lab, S. (2018). County presidential election returns 2000-2020. *URL*. https://doi.org/10.7910/DVN/VOQCHQ

Dong, E., Du, H., Gardner, L., 2020. An interactive web-based dashboard to track COVID-19 in real time. The lancet infectious diseases published online Feb 19.

Environmental Sciences Division, Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC), 2024. DARTE Annual On-road CO_2 Emissions (Version 2.0). URL: https://daac.ornl.gov/CMS/guides/CMS_DARTE_V2.html. accessed: 2024-08-22.

Gabaix, X. (1999). Zipf's law for cities: An explanation. *The Quarterly Journal of Economics, 114*, 739–767.

Gerhardus, A., & Runge, J. (2020). High-recall causal discovery for autocorrelated time series with latent confounders. *Advances in Neural Information Processing Systems, 33*, 12615–12625.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424–438.

Haehne, H., Casadiego, J., Peinke, J., & Timme, M. (2019). Detecting hidden units and network size from perceptible dynamics. *Physical Review Letters, 122*, Article 158301.

Hochberg, Y. (1987). *Multiple comparison procedures*. New York, US: Wiley.

Hsu, A., Tan, J., Ng, Y. M., Toh, W., Vanda, R., & Goyal, N. (2020). Performance determinants show European cities are delivering on climate mitigation. *Nature Climate Change, 10*, 1015–1022.
Hsu, A., Wang, X., Tan, J., Toh, W., & Goyal, N. (2022). Predicting European cities' climate mitigation performance using machine learning. *Nature Communications, 13*, 7487.
Kandt, J., & Batty, M. (2021). Smart cities, big data and urban policy: Towards urban analytics for the long run. *Cities, 109*, Article 102992.
Karlsson, L. C., Soveri, A., Lewandowsky, S., Karlsson, L., Karlsson, H., Nolvi, S., Karukivi, M., Lindfelt, M., & Antfolk, J. (2021). Fearing the disease or the vaccine: The case of COVID-19. *Personality and Individual Differences, 172*, Article 110590.
Kissler, S. M., Viboud, C., Grenfell, B. T., & Gog, J. R. (2020). Symbolic transfer entropy reveals the age structure of pandemic influenza transmission from high-volume influenza-like illness data. *Journal of the Royal Society Interface, 17*, 20190628.
Li, H., Calder, C. A., & Cressie, N. (2007). Beyond Moran's i: Testing for spatial dependence based on the spatial autoregressive model. *Geographical Analysis, 39*, 357–375.
Lim, C., Kim, K. J., & Maglio, P. P. (2018). Smart cities with big data: Reference models, challenges, and considerations. *Cities, 82*, 86–99.
MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics, 29*, 305–325.
Madden, R. A., & Ramanathan, V. (1980). Detecting climate change due to increasing carbon dioxide. *Science, 209*, 763–768.
Martínez-Schuldt, R. D., & Martínez, D. E. (2019). Sanctuary policies and city-level incidents of violence, 1990 to 2010. *Justice Quarterly, 36*, 567–593.
Mergel, I., Kleibrink, A., & Sörvik, J. (2018). Open data outcomes: US cities between product and process innovation. *Government Information Quarterly, 35*, 622–632.
Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., et al. (2015). Promoting an open research culture. *Science, 348*, 1422–1425.
of Economic Research, N.B., 2023. Hud usps zip code crosswalk files. https://www.huduser.gov/portal/datasets/usps_crosswalk.html.
Pearl, J. (2009). *Causality*. United Kingdom: Cambridge University Press.
Pijls, B. G., Jolani, S., Atherley, A., Derckx, R. T., Dijkstra, J. I., Franssen, G. H., , … Zalpuri, S., et al. (2021). Demographic risk factors for COVID-19 infection, severity, ICU admission and death: A meta-analysis of 59 studies. *BMJ Open, 11*, Article e044640.
PRISM Climate Group, 2024. PRISM weather Data. URL: https://asmith.ucdavis.edu/data/prism-weather. accessed: 2024-08-22.
Reckien, D., Flacke, J., Olazabal, M., & Heidrich, O. (2015). The influence of drivers and barriers on urban adaptation and mitigation plans—An empirical analysis of European cities. *PLoS One, 10*, Article e0135597.
Rey, S. J., & Anselin, L. (2009). *Pysal: A python library of spatial analytical methods*.
Ribeiro, H. V., Rybski, D., & Kropp, J. P. (2019). Effects of changing population or density on urban carbon dioxide emissions. *Nature Communications, 10*, 3204.
Robinson, D., Yu, H., Zeller, W. P., & Felten, E. W. (2008). Government data and the invisible hand. *Yale Journal of Law & Technology, 11*, 159.
Rohatgi, V. K. (1976). *An introduction to probability theory and mathematical statistics*. New York: Wiley.
Runge, J. (2018). Causal network reconstruction from time series: From theoretical assumptions to practical estimation. Chaos: An interdisciplinary. *Journal of Nonlinear Science, 28*.
Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., & Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. Science. *Advances, 5*, eaau4996.
Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters, 85*, 461.
Schulte, F., Lucas, E., Rau, J., Szabo, L., & Hancock, J. (2020). Millions of older Americans live in counties with no ICU beds as pandemic intensifies. Kaiser health news. https://kffhealthnews.org/news/as-coronavirus-spreads-widely-millions-of-older-americans-live-in-counties-with-no-icu-beds/.
Seabold, S., Perktold, J., 2010. Statsmodels: Econometric and statistical modeling with python.
Sethi, S. S., & Vinoj, V. (2024). Urbanization and regional climate change-linked warming of Indian cities. *Nature Cities*, 1–4.
Stokes, A. C., Lundberg, D. J., Elo, I. T., Hempstead, K., Bor, J., & Preston, S. H. (2021). COVID-19 and excess mortality in the United States: A county-level analysis. *PLoS Medicine, 18*, Article e1003571.
Succar, R., & Porfiri, M. (2024). Urban scaling of firearm violence, ownership and accessibility in the United States. *Nature Cities, 1*, 216–224.
Sugihara, G., May, R., Ye, H., Hsieh, C.h., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science, 338*, 496–500.
Tanaka, T., & Okamoto, S. (2021). Increase in suicide following an initial decline during the COVID-19 pandemic in Japan. *Nature Human Behaviour, 5*, 229–238.
Tregoning, J. S., Flight, K. E., Higham, S. L., Wang, Z., & Pierce, B. F. (2021). Progress of the COVID-19 vaccine effort: Viruses, vaccines and variants versus efficacy, effectiveness and escape. *Nature Reviews Immunology, 21*, 626–636.
University of Wisconsin Population Health Institute, 2020. County Health Rankings & Roadmaps. www.countyhealthrankings.org. Accessed: 2023-10-20.
Uto, M., Nakagawa, M., & Buhnik, S. (2023). Effects of housing asset deflation on shrinking cities: A case of the Tokyo metropolitan area. *Cities, 132*, Article 104062.
Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods, 17*, 261–272.
Wang, Y., Yang, X., & Wang, Z. H. (2024). Causal mediation of urban temperature by geopotential height in US cities. *Sustainable Cities and Society, 100*, Article 105010.
Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., et al. (2016). The fair guiding principles for scientific data management and stewardship. *Scientific Data, 3*, 1–9.
Woolf, S. H., Chapman, D. A., Sabo, R. T., & Zimmerman, E. B. (2021). Excess deaths from COVID-19 and other causes in the US, march 1, 2020, to January 2, 2021. *JAMA, 325*, 1786–1789.
Xiao, Y., & Gong, P. (2022). Removing spatial autocorrelation in urban scaling analysis. *Cities, 124*, Article 103600.
Zellner, A., Kmenta, J., & Dreze, J. (1966). Specification and estimation of cobb-Douglas production function models. *Econometrica: Journal of the Econometric Society*, 784–795.
Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., , … Gu, X., et al. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *The Lancet, 395*, 1054–1062.