Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Transportation Research Part A 172 (2023) 103669

Contents lists available at ScienceDirect

Transportation Research Part A

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

A social media Data-Driven analysis for transport policy response to the COVID-19 pandemic outbreak in Wuhan, China

Wenbin Zha[a], Qian Ye[b], Jian Li[c,*], Kaan Ozbay[d]

[a] *Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai 201804, China*
[b] *Transport Planning and Research Institute of Ministry of Transport P.R. China, Beijing 100028, China, Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai 201804, China*
[c] *Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, College of Transportation Engineering, Tongji University, 4800 Cao'an Road, Shanghai 201804, China*
[d] *C2SMART Center, Department of Civil and Urban Engineering & Center for Urban Science and Progress (CUSP), Tandon School of Engineering, New York University, 15 MetroTech Center, 6th Floor, Brooklyn, NY 11201, USA*

ARTICLE INFO

Keywords:
Pandemics
Urban Transport Policy
Social Media Data
Latent Dirichlet Allocation
Jaccard Distance

ABSTRACT

Non-pharmacological interventions (NPI) such as social distancing and lockdown are essential in preventing and controlling emerging pandemic outbreaks. Many countries worldwide implemented lockdowns during the COVID-19 outbreaks. However, due to the lack of prior experience and knowledge about the pandemic, it is challenging to deal with short-term polices decision-making due to the highly stochastic and dynamic nature of the COVID-19. Thus, there is a need for the exploration of policy decision analysis to help agencies to adjust their current policies and adopt quickly. In this study, an analytical methodology is developed to analysis urban transport policy response for pandemic control based on social media data. Compared to traditional surveys or interviews, social media can provide timely data based on the feedback from public in terms of public demands, opinions, and acceptance of policy implementations. In particular, a sentiment-aware pre-trained language model is fine-tuned for sentiment analysis of policy. The Latent Dirichlet Allocation (LDA) model is used to classify documents, e.g., posts collected from social media, into specific topics in an unsupervised manner. Then, entropy weights method (EWM) is used to extract public policy demands based on the classified topics. Meanwhile, a Jaccard distance-based approach is proposed to conduct the response analysis of policy adjustments. A retrospective analysis of transport policies during the COVID-19 pandemic in Wuhan, China is presented using the developed methodology. The results show that the developed policymaking support methodology can be an effective tool to evaluate the acceptance of anti-pandemic policies from the public's perspective, to assess the balance between policies and people's demands, and to further perform the response analysis of a series of policy adjustments based on online feedback.

* Corresponding author.
E-mail addresses: wenbinzha@tongji.edu.cn (W. Zha), yeqian122@163.com (Q. Ye), jianli@tongji.edu.cn (J. Li), kaan.ozbay@nyu.edu (K. Ozbay).

https://doi.org/10.1016/j.tra.2023.103669

Available online 29 March 2023
0965-8564/© 2023 Elsevier Ltd. All rights reserved.

1. Introduction

In recent years, the emergence of new infectious diseases in the world has become a grim reality that human beings must face, e.g., the 21st-century pandemics including SARS, H1N1 Flu, MERS, Ebola, and the ongoing COVID-19. The substantial human disaster caused by infectious diseases has returned to public view. With the acceleration of globalization and urbanization, the highly concentrated population may lead to the rapid spread of infectious diseases and result in severe economic and social impacts. Therefore, how to minimize the impact of the pandemic and reduce losses has become a major issue facing the international community.

Pandemics have had significant influences on human activities and mobility, and transport policy on emergency preparedness and response are effective ways to mitigate viral spread (An et al., 2021). For example, Fang et al. (2020) concluded that the COVID-19 lockdown in Wuhan, the capital of Hubei Province, China, reduced inflows to Wuhan by 76.64 %, outflows from Wuhan by 56.35 %, and movement within the city by 54.15 %. If no travel restrictions had been imposed, the number of COVID-19 cases would have increased by 64.81 % in 347 Chinese cities outside Hubei and 52.64 % in the other 16 cities within Hubei. However, due to the unpredictability and unknown nature of emerging pandemics (Benini et al., 2020), policymakers face significant transport management and emergency response issues. For instance, the major issue during the COVID-19 pandemic has been determining how to provide adapted travel services to meet dynamic activity demands and maintain the balance between safety and mobility.

Compared to regular situations, transportation policy decisions on responses to pandemics are challenging in terms of the inevitable trade-off between safety and mobility. Due to the uncertainty during pandemics, decision-makers are required to formulate time-sensitive transport policies and adjust these based on the real-world feedback. In general, surveys and interviews are widely used to analyze attitudes and behavior of community members towards regular transport policies. Quantitative methods based on these and other relevant data include regressions, structural equation modeling, cost-benefit analysis, etc. For example, a cost-benefit analysis can be performed to assess value of the effects of a policy intended or unintended for the target population and society as a whole (Batarce et al., 2016). However, these analysis patterns may not be applicable to conducting policy response analysis and policymaking in the emergency contexts, where policy-relevant issues need to be captured dynamically and responded to in a timely manner since surveys and interviews are often time-consuming and significantly delayed. As digital technology advances, public opinion information on social media platforms provide an opportunity to make the process of policy formulation, implementation and evaluation more citizen-centric (Rogers, 2019). It enables policy makers to identify policy feedback in a timely manner, e.g., hot topic discussions and debates, to improve the subsequent implementation.

This study aims to develop a social media-based approach to support urban transportation policymaking in the context of a pandemic, especially from the policy response perspective. Social media data has been widely recognized as an emerging data source to support crisis communication and disaster management in natural or man-made hazards (Kumar and Ukkusuri, 2020; Ogie et al., 2019). It also allows governments to understand the impact of policy on society (Leavey, 2013), while its rising popularity has made it a necessary part of the evidence-based policymaking (Belkahla Driss et al., 2019; Nguyen, 2021). . In this regard, an analytical methodology which uses policy-relevant tweets and comments from the Chinese social media platform Weibo (commonly referred to as "Chinese Twitter") is developed. The analytical methodology works on the monitoring phase of the public policy process, providing a complementary tool for policy response analysis that can be applied to future similar emergency situations. The methodology includes policy sentiment analysis, extraction of policy demands, and the analysis of policy adjustments. In particular, a sentiment-aware pre-trained language model is fine-tuned to classify public sentiments on the implementation of policies as positive, neutral, and negative, with positive public sentiment meaning higher policy acceptance while negative sentiment meaning lower policy acceptance. Next, policy demands representing the values, interests, and benefits of the public are further retrieved from the negative opinions using the developed EWM-based Latent Dirichlet Allocation model for understanding the issues of policies on society. Then, the Jaccard distance, a measure of how dissimilar two sets are, is used to conduct the response analysis of the policy adjustments by measuring the similarity of the Latent Dirichlet Allocation (LDA) topics. A retrospective analysis of transport policies during the COVID-19 pandemic in Wuhan, China is also conducted using the developed methodology.

The rest of paper is organized as follows. Section 2 provides a brief literature review. Section 3 introduces methodological methodology. Section 4 and 5 describes the data and results & discussion in the case study. Section 6 provides conclusions.

2. Literature REVIEW

Previous studies have explored the role of social media in supporting transportation planning, management, and operations (Gal-Tzur et al., 2014; Khan et al., 2017; Rashidi et al., 2017). The main research topics include travel demand modeling, aggregate/individual travel behavior analysis, transport service assessment, and traffic condition prediction. For example, Zhan and Ukkusuri (2014) provided evidence that social media can support the process of land use and transport policymaking. Hasan and Ukkusuri (2014) developed a topic-based activity behavior modeling approach to infer individuals' multi-day activity patterns from social media check-in services. Rodrigues et al. (2018) used the data for taxi demand analysis. Li et al. (2019) evaluated the service quality of the metro system in Nanjing, China, by mining the feedback comments. Lucini et al. (2020) identified twenty-seven topics related to passengers' experiences by applying the topic model to the collected feedback data from airline passengers and developed a logistic regression modeling to measure passengers' satisfaction.

From the perspective of emergency management, social media data has proven to be effective in the application of natural disaster detection (Li et al., 2018; Singh et al., 2019) (as well as traffic accident or emergency events extraction (Lian et al., 2020; Ma et al., 2019) . In the case of disaster detection, Singh et al. (2019) proposed a disaster location detection system based on a stochastic Markov chain model. Further, they developed a Twitter classification model to filter disaster-related blogs. Zhang et al. (2018) conducted traffic accident detection using individual social media posts. Lian et al. (2020) provided an online information management framework that integrates spatial and temporal information extraction, sentiment identification, and opinion classification. Ma et al. (2019) proposed an emergency natural disaster finding and analysis model combining graph analysis and keyword filtering. Kumar and Ukkusuri (2020) explored the factors influencing the evacuation behavior of various groups of coastal residents based on geo-tagged Twitter posts. Such analyses can help monitor the progress of interventions or relief efforts or improve situational awareness. A detailed review of emergency management using social media data can be seen in Alexander et al. (2014).

Social media also plays an increasingly important role in shaping public policy (Rogers, 2019). In most cases, policy decisions are guided by community attitudes and behaviors, which are usually measured through quantitative studies involving labor-intensive surveys or interviews (Nguyen, 2021). For example, a cost-benefit analysis can be performed to assess whether the effects of a policy are intended or unintended for the target population and society as a whole (Batarce et al., 2016). Social media provides policymakers with the ability to access broad and near real-time information, such as opinions, demands, and other raw materials related to the policy process (Belkahla Driss et al., 2019) , offering the chance for the policy development, implementation, and evaluation to be more citizen-centered in a faster and more cost-effective way (Nguyen, 2021).

For the application of social media in policy response analysis, studies have been conducted by extracting the debated topics and sentiments of citizens' discussions during the policy implementation process. Sentiment analysis can quantify the public's satisfaction level, with positive public sentiment means higher policy acceptance, and vice versa for negative (Chen et al., 2020b; Dandannavar et al., 2019). For example, Dandannavar et al. (2019) proposed a framework for assessing the acceptance of government welfare policy based on tweets sentiment analysis. Recently, academia has also explored the potential of social media data to support the COVID-19 response, Wang et al. (2022) analyzed the sentiment tendencies (positive, neutral, and negative) of the public towards the lockdown policies in different countries during the COVID-19 pandemic; Monmousseau et al. (2020) analyzed the impact of different airline policies on passengers under COVID-19 by assessing the sentiment expressed in tweets, and gave each airline its own "Twitter profile" based on the keyword analysis, and stated that discussions between federal agencies, airlines and passengers should be held via Twitter to manage customer satisfaction.

In summary, existing studies have illustrated that social media can be used as a strategic tool in the transportation sector to improve management and operational efficiency. In the context of pandemic emergencies, transport policy decision-makers need to make a series of timely responses to help citizens and travelers adjust their activities accordingly. However, to our best knowledge, from the perspective of policy evaluation, that is, policy implementation monitoring and policy performance assessment, quantitative analytical solutions that assist urban transport decision-making have not been fully explored. Therefore, there is a need to develop an integrated data-driven methodology by utilizing social media data to effectively evaluate the performance of anti-epidemic policies and understand the specific demands of the public.

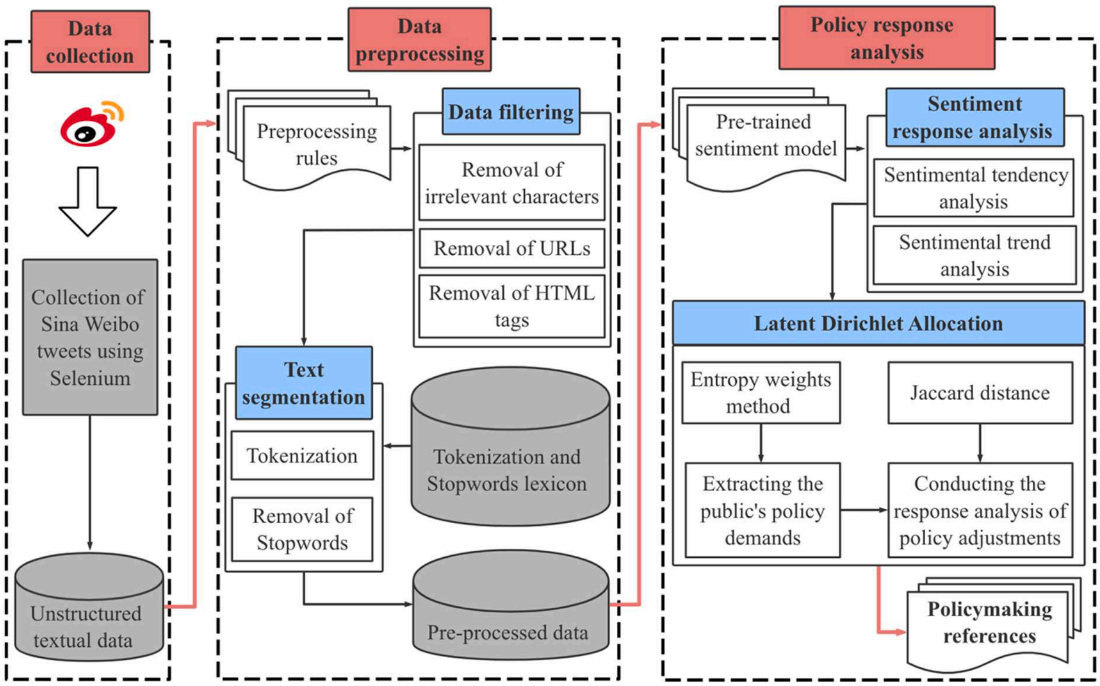

Fig. 1. Analytical methodology.

3. Methodology

This section provides a detailed introduction of the methodology and associated methods. It aims at transport anti-epidemic policy evaluation during implementation, which will help policymakers understand the current situation and identify emerging policy issues. Fig. 1 presents the logic of the proposed methodology. The general analytical process can be seen as follows:

(1) First, textual tweets, retweets, and comments related to the transport anti-epidemic policy are extracted based on the web scraping tool Selenium. The next step is data cleaning, which includes stripping out whitespace, repetitive contents, HTML tags, and nonsensical special characters.
(2) Second, fine-tuning the sentiment-aware pre-trained language model (SKEP) for sentiment analysis to evaluate the level of public acceptance or satisfaction towards the policy.
(3) Third, inferring the appropriate number of topics for the classical LDA model based on the entropy weights method. And topics retrieved from the expressed negative opinions are considered public policy demands.
(4) Fourth, the Jaccard distance is adopted to conduct the response analysis of policy adjustments by quantifying the similarity or overlap share of topics between policies.

3.1. Data collection and preprocessing

Using the policy executive summary as the search string, policy-related textual posts are scratched from Sina Weibo with web scraping. With over 500 million monthly active users, Sina Weibo gives the government the opportunity to understand the impact of policy on society. Specifically, with the executive summary of the released administrative policy as the search string, the web scraping was performed by Selenium, which is a powerful web browser automation tool that provides extensions to emulate the user's interaction with the browser. Furthermore, related raw data such as username, user profiles, date of posting, number of likes, tweets' content, and comments were extracted. Further, data filtering involves stripping out whitespace, repetitive contents, HTML tags, and nonsensical special characters. The filtered data is then used for sentiment analysis, which is further processed by text tokenization and removal of stop words for LDA topic modeling. In addition, there is a consensus that very few posts in social media are geo-targeted, so spatial heterogeneity is not considered in this study.

3.2. Policy sentiment analysis

The pre-trained sentiment-aware language model (SKEP) is fine-tuned for sentiment analysis, and the growth rate of sentiment based on daily collected data is used to evaluate the level of public acceptance or satisfaction of policy over time. In recent years, the natural language processing (NLP) community has made many breakthroughs, especially the shift from traditional statistical and machine learning approaches to pre-training-based transfer learning. Early results of NLP on transfer learning are largely based on recurrent neural networks, but recently the use of models based on the "Transformer" architecture has become more common (Raffel et al., 2020). And transformer-based models have emerged as a new benchmark for sentiment analysis tasks, as these models can extract more contextual semantic information due to the self-attention mechanism (Liu et al., 2020b) . In particular, the recent proposed transformer-based Sentiment Knowledge Enhanced Pre-training (SKEP) model (Tian et al., 2020), in which sentiment knowledge about words, polarity, and aspect-sentiment pairs are included to guide the process of pre-training, has achieved state-of-the-art results on the sentiment analysis benchmark dataset SST-2 (Socher et al., 2013). The SKEP model compensates for the fact that sentiment-prone words are not sufficiently considered during the pre-training process in versatile pre-trained language models, and is therefore suitable for sentence-level sentiment analysis. In this study, we will fine-tune the SKEP model to classify policy-related opinions into positive, neutral, and negative.

The process of fine-tuning includes domain data preparation and fine-tuning strategies configuration. The input data for fine-tuning is a text set (policy-responsive posts) and a corresponding target set with manually annotated sentiment labels (positive, neutral, and negative). It should be noted that tweets and retweets refer specifically to individual accounts since government accounts typically do not have additional private opinions, and comments refer to individual comments on government or individual tweets and retweets. As for the fine-tuning options, the task in this study is a single sentence-level sentiment classification, which is more straightforward than other complex tasks such as machine translation. Therefore, instead of invoking other new state-of-the-art strategies such as multi-task fine-tuning and prompt-based tuning (Qiu et al., 2020), we consider fine-tuning the whole model structure along the lines of the original paper (Tian et al., 2020), that is, a classification layer is added on top of the pre-trained model to compute the sentiment probability based on the learned sentence sentiment representation. The detailed fine-tuning process and results are described in the Results section.

3.3. Policy demands extraction

Latent Dirichlet Allocation (LDA) is one of the most popular methods for performing topic modeling, and topics extracted from negative policy opinions are considered as public policy demands. LDA uncovers underlying "topics" in a given collection of documents based on the Bayesian inference approach. It considers documents as the result of a probabilistic generation process, with topic distributions in each document and word distributions in each topic based on the Dirichlet-distributed Multinomial distribution (Blei et al., 2003). Fig. 2 illustrates the LDA modeling process, in which the plate notation (Bayesian network graph) is used to reveal the dependencies and derivative relationships between the variables of the LDA model. M denotes the number of documents, K denotes the number of topics,N_m is number of words in a given document,$\vec{\alpha}$ and $\vec{\beta}$ are the parameters of the Dirichlet prior for the document-topic and topic-word distributions, $\vec{\theta}_m$ is the distribution of topics in document m, $\vec{\varphi}_k$ is the distribution of words in topic k, $z_{m,n}$ is the topic for the n-th word in document m and $w_{m,n}$ is the specific n-th word in document m. The detailed LDA theory and derivation process are presented in Appendix A.

Despite its success and popularity in retrieving information from mammoth text corpus in the form of topic structure, LDA remains a challenging task for selecting the appropriate number of topics. As an unsupervised method, LDA must pre-specify the number of topics to be implemented. An under-specified model is too coarse to help find the underlying topic structure, whereas an over-specified model could instead generate non-informative and potentially redundant topics. So far, several ad-hoc evaluation strategies exist that guide how to identify the appropriate number of topics. Some studies suggest setting to an arbitrary value as judged by human experts (Sun and Yin, 2017); some suggest Perplexity, with lower Perplexity scores indicating better generalization performance (Blei et al., 2003; Lucini et al., 2020); some are based on Rate of Perplexity Change (RPC) (Zhao et al., 2015), they tried also proved that RPC works better than Perplexity; some suggest Topic Coherence (Röder et al., 2015) . The study proposed an iterative approach based on simulated annealing (SA) algorithm named SA-LDA to find the appropriate (Pathik and Shukla, 2020). As such, finding the appropriate number of topics remains challenging, as different configurations will likely result in various interpretations of the corpus. Instead of extending the existing studies lines that can be characterized as intrinsic methods combining the distribution of corpus words and topics, this study proposes a comprehensive judgment index based on sophisticated LDA model evaluation metrics to direct the selection of topic numbers. Specifically, two model evaluation metrics known as topic ***Perplexity*** (Blei et al., 2003) and ***Topic coherence*** $\boldsymbol{C_V}$ (Röder et al., 2015) are integrated into a comprehensive index based on the entropy weights method (EWM). A more detailed analysis is presented in Appendix B.

From a theoretical perspective, the basic idea of EWM is rooted in weighting and fusing information from the multi-level evaluation metrics of the LDA model. The evaluation metrics ***Perplexity*** and ***Topic coherence*** $\boldsymbol{C_V}$ represent the uncertainty about the goodness-of-fit of the LDA model from the document level and the topic level. Therefore, a comprehensive weighted index of ***Perplexity*** and ***Topic coherence*** $\boldsymbol{C_V}$ was constructed to judge the appropriate number of topics for the LDA topic model. If a metric tends to exhibit central tendency rather than statistical dispersion under variation of different topic model parameters, which means larger information entropy, we believe that it provides relatively little information to the modeler about model selection and interpretation, and vice versa for the one with greater dispersion. Moreover, the two metrics involve time-consuming human expert-assisted operations to get the appropriate number of topics. At the same time, the lowest perplexity or highest coherence also does not always account for the human interpretability of the topics generated (Blair et al., 2020). If one considers a comprehensive approach that would include a number of metrics, this may actually involve multi-criteria decision making (MCDM), which uses computational methods to incorporate several criteria and preference orders based on the desired outcome when evaluating and selecting the best of many alternatives. Since EWM can avoid human factors on the index weights thereby improving the objectivity of the comprehensive evaluation results, it is widely used to solve MCDM problems compared to various subjective weighting methods such as Analytic Hierarchy Process (AHP) (El-Araby et al., 2022).

The EWM-based approach can be seen as extrinsic evaluation utilizing additional tasks that employ the outputs of the LDA model (i. e., topic-word distribution), and the Jaccard distance is employed to verify the effectiveness of the EWM-based approach. For LDA

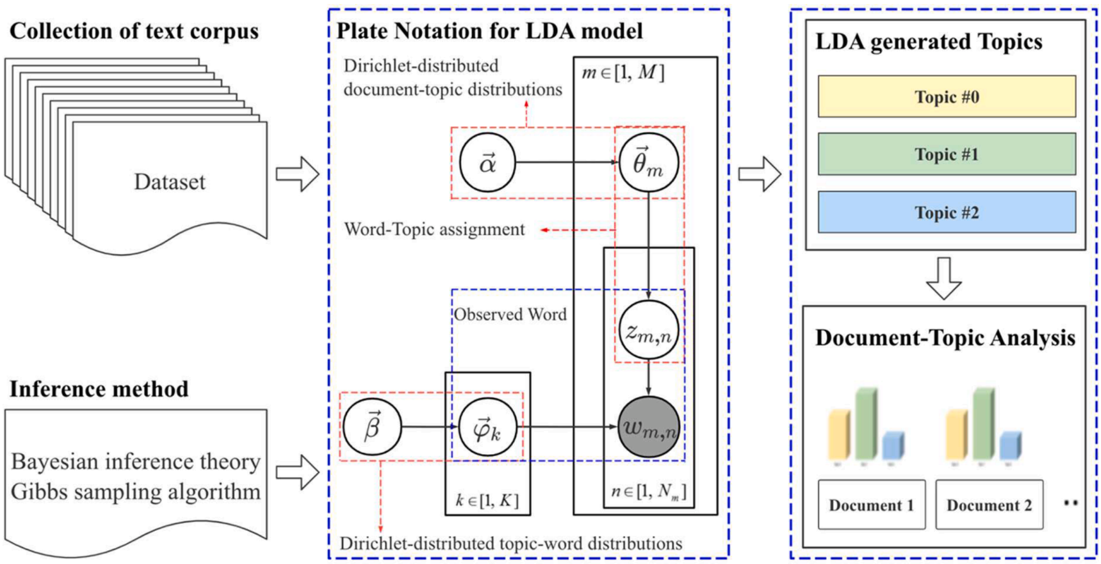

Fig. 2. LDA modeling process.

topic modeling, the generated topics need to be interpretable, and the similarity between topics should be small to avoid duplication (Deng et al., 2020; Zhao et al., 2015). Also, the distinctiveness of individual topics based on similarity measures has been used as an evaluation criterion for the quality of LDA derived topics (Vega et al., 2020). Therefore, the validity of the proposed EWM-based LDA is measured by the Jaccard distance between the topics of the results. The Jaccard coefficient illustrates the number of identical words between topics, i.e., it reveals another point of view of the similarity/reduplicate of topics, with lower implying better model performance. In addition, the EWM-based approach is compared with baseline models configured by coherence $\boldsymbol{C_V}$ and ***Perplexity*** metrics, namely the Coherence Model (CM) and the Perplexity Model (PM). The analysis results are described in Section 5.2.

3.4. Response analysis of policy adjustment

The Jaccard distance is also taken to conduct the response analysis of policy adjustments. Suppose policy P_1 and P_2 are consecutive policies that address the same subject event, and policy P_2 is an improvement and enhancement to the policy P_1. The set of topic-specific words/terms for policy P_1 is described as $Termset_1$ with size K_1, and the set of topic-specific terms for policy P_2 is described as $Termset_2$ with size K_2. To measure the similarity between topics, the difference matrix $\mathrm{Diff\_Matrix}_{P1P2}$ is defined as follows:

$$\mathrm{Diff\_Matrix}_{P1P2} = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,K_2-1} & P_{1,K_2} \\ P_{2,1} & P_{2,2} & & P_{2,K_2-1} & P_{2,K_2} \\ \vdots & & \ddots & \vdots & \\ P_{K_1-1,1} & P_{K_1-1,1} & \cdots & P_{K_1-1,K_2-1} & P_{K_1-1,K_2} \\ P_{K_1,1} & P_{K_1,1} & & P_{K_1,K_2-1} & P_{K_1,K_2} \end{bmatrix}_{K_1 \times K_2} \tag{1}$$

$$P_{i,j} = \frac{|Termset_i \cup Termset_j| - |Termset_i \bigcap Termset_j|}{|Termset_i \cup Termset_j|} \tag{2}$$

where $T\ ermset = [\{term_1, term_2, \cdots, term_n\}_1, \cdots, \{term_1, term_2, \cdots, term_n\}_K]$.

$Termset_i = \{term_1, term_2, \cdots, term_n\}_i$ represents the i-th topic,$1 \le i \le K_1$, $1 \le j \le K_2$, $0 \le P_{i,j} \le 1$. Focus can be placed on elements in $\mathrm{Diff\_Matrix}_{P1P2}$ with values less than a certain threshold value, such as 0.9; the smaller the value, the more relevant the topics from the two policies, which means that the public's demands from the former policy are not effectively addressed by the latter.

Timely policy feedback is essential in emergency management. Social media access to public opinion related to policy adjustments can help first responders and emergency management teams gain situational awareness. On the one hand, techniques based on Jaccard analysis can be implemented in real-time news feeds during crisis events to capture aggregated reactions and dynamically track the conversation about the event. On the other hand, we believe that a term-based Jaccard distance analysis is more feasible to perform the response analysis of the policy adjustments than the time-consuming survey-based cost-benefit analysis. In essence, according to Zipf's

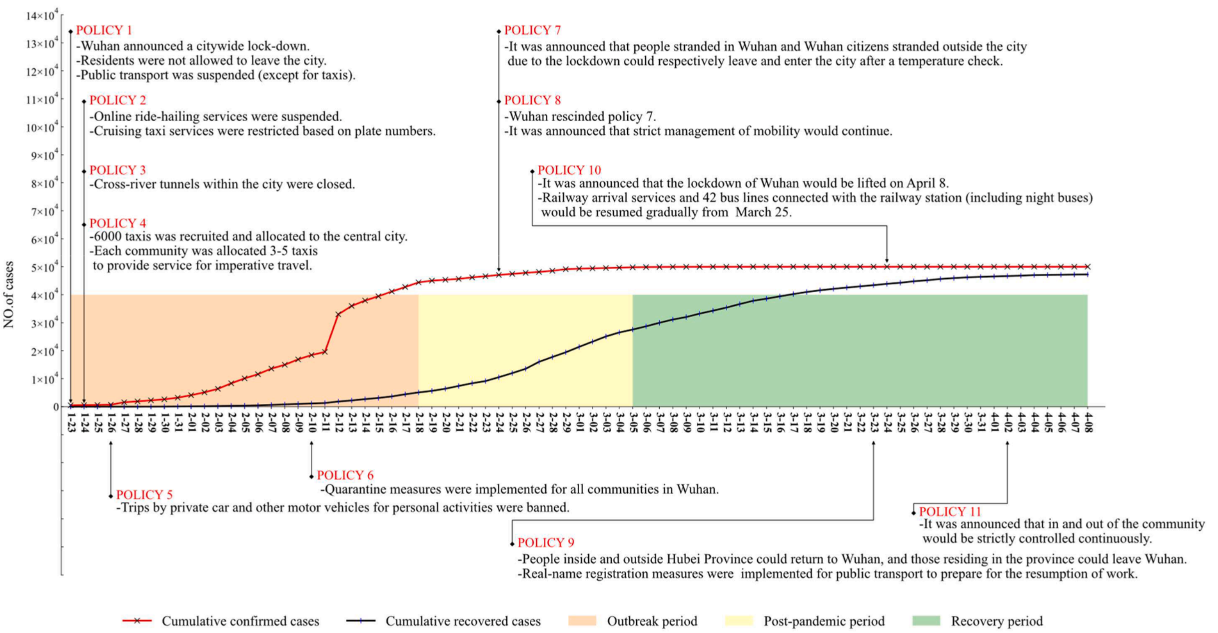

Fig. 3. Wuhan COVID-19 transport policies from January 23 to April 8, 2020 (lockdown period). Data source: http://en.hubei.gov.cn/special/coronavirus_2019/index.shtml (COVID-19 cases). http://www.gov.cn/xinwen/2020-01/23/content_5471751.htm (COVID-19 policies).

law (Zipf, 1949), the most basic fact of human language is that few very high-frequency words account for most of the tokens in the text and many low-frequency words. And Zipf's law is applicable in social media data, such as Zipf's law on Twitter (Bokányi et al., 2021). Moreover, it is relatively simple to collect hundreds or even thousands of words related to emergency policies from multi-channel social media platforms.

4. Data

4.1. Study area and background

The study was conducted in Wuhan, where the COVID-19 pandemic first occurred in China. As a nationally important industrial, scientific and educational base, and the capital of Hubei Province, Wuhan has a well-developed infrastructure of transportation systems and functions as one of the comprehensive transportation hubs in central China. According to the Statistical Yearbook 2020 published by the Wuhan Municipal Government (Wuhan Bureau of Statistics, 2020), Wuhan Airport covers 72 cities nationwide, and has convenient high-speed train and highway networks to connect with external traffic. In addition, the rail transit in the city has reached 410 km of operation, and the bus stop coverage within the 500-meter radius is 100 %. As such, the outbreak of the pandemic in such an important transportation hub was an unfortunate situation.

To control the spread of COVID-19 and protect people's lives, Wuhan implemented the world's first Covid lockdown after the virus was confirmed in early 2020. On January 23, 2020, all public transport, from buses to subway services, were suspended, and residents were urged not to leave the district unless necessary. Retrospectively, it was the first time in decades that China had imposed a lockdown on a major city with more than 10 million people. As shown in Fig. 3, after the lockdown, the government released a series of transport policies related to cab services, driving bans, road closures, and support for essential travel demands from the city's functioning and communities. Although some mishandling can be seen in the released policies, such as Policy 7 was abolished and replaced with Policy 8 after its release, probably due to the complex decision-making environment and lack of response experience. These policies effectively contained the pandemic, and recovery was achieved three months after the lockdown.

4.2. Summary of social media data

To conduct the policy response analysis, public opinions corresponding to the policy were collected from Sina Weibo with the geographic boundary set to Wuhan. Descriptive statistical analysis found that the primary role of public agency accounts on Weibo is to disseminate policy content, so only the individual posts with subjective judgments were chosen for the case study. Specifically, 211,731 data rows were obtained, including individual tweets, retweets, and comments during the lockdown period. It should be noted that Policies 7 and 8 were combined in the analysis, given that Policy 7 was announced to be abolished a few hours after its release and was replaced by Policy 8. Furthermore, posts that contained less than ten tokens were removed after tokenization and stop words removal. As a result, 154,474 data rows were obtained, with a maximum number of 48,963 (Policy 1) and a minimum of 351 (Policy 9). Table 1 provides more detailed descriptive statistics for the data used in the subsequent analysis.

Compared with the results of the seventh national census of Wuhan City in 2020, this study has limitations in terms of population sample. To be precise, the policy target population on Weibo during the lockdown period is skewed towards the labor force group of highly educated people aged 15–59 years old, with a female majority. In detail, a total of 61,684 Weibo users were identified, 82.3 % of users were labeled with university degree, higher than the 33.9 % disclosed in the census, and 8.1 % with a secondary school degree lower than the 45.1 % disclosed. For gender composition, the proportion of female users was 40.4 % close to the 48.1 % in the census; the proportion of male users was 29.5 %, lower than the 51.9 % disclosed. Moreover, the age of Weibo users was concentrated in the 15–59 age group, accounting for 92.4 % higher than the 69.7 % disclosed in the census; while users over 60 and under 15 years old accounted for 0.9 % and 6.7 %, both lower than the announced 17.2 % and 13.1 %. Table 2 provides more detailed demographics of the

Table 1
The summary of data.

	Policy 1	**Policy 2**	**Policy 3**	**Policy 4**	**Policy 5**
(A) Corpus Size					
Data rows	48,963	4208	10,315	26,401	20,527
Tokens	21,555	5167	5357	7692	12,323
(B) Post Length Statistics (Tokens)					
Min	10	10	10	10	10
Max	258	151	157	155	188
Mean	40	40	41	40	41
	Policy 6	**Policy 7–8**	**Policy 9**	**Policy 10**	**Policy 11**
(A) Corpus Size					
Data rows	10,206	30,834	351	2249	420
Tokens	8683	9307	434	4268	907
(B) Post Length Statistics (Tokens)					
Min	10	10	10	10	10
Max	177	162	139	149	153
Mean	43	33	30	34	39

Table 2
Demographics of the Weibo users.

	Variables	Weibo data	Census data
Gender			
	Male	29.5 %	51.9 %
	Female	40.4 %	48.1 %
	Unknown	30.1 %	–
Age			
	less than15 years	6.7 %	13.1 %
	15–59 years	92.4 %	69.7 %
	greater than60 years	0.9 %	17.2 %
Education level			
	University (college and above)	82.3 %	33.9 %
	Secondary education (high and junior high school)	8.1 %	45.1 %
	Primary school	2.4 %	13.6 %
	Unknown	7.2 %	7.4 %

Note: the variable "Unknown" in Weibo data refers to the portion of users who cannot be identified due to the lack of their personal information.

Weibo users.

5. Results And Discussion

5.1. Policy sentiment analysis based on pre-trained model SKEP

Social media-based policy sentiment analysis is used to understand the level of acceptance of policies. The sentiment knowledge enhanced language pre-training model SKEP was employed to conduct the sentiment analysis of the transport policies in Wuhan. To begin with, the Python Senta library wraps the Chinese version of the SKEP pre-trained model, and it achieves accuracy of 67.8 % when used directly on 24,000 manually labeled samples of our Weibo data (8000 randomly sampled posts for each of the three labels: positive, negative and neutral). Therefore, the SKEP model must be fine-tuned before it can be used for policy sentiment analysis tasks. For fine-tuning options, a classification layer was added on top of the pre-trained model to compute the sentiment probability based on the learned sentence sentiment representation. The size of the classification layer equals the number of labels. SoftMax activation with cross-entropy loss was added on top of the fine-tuning model to predict the likelihood of labels. In addition, for the training parameters,

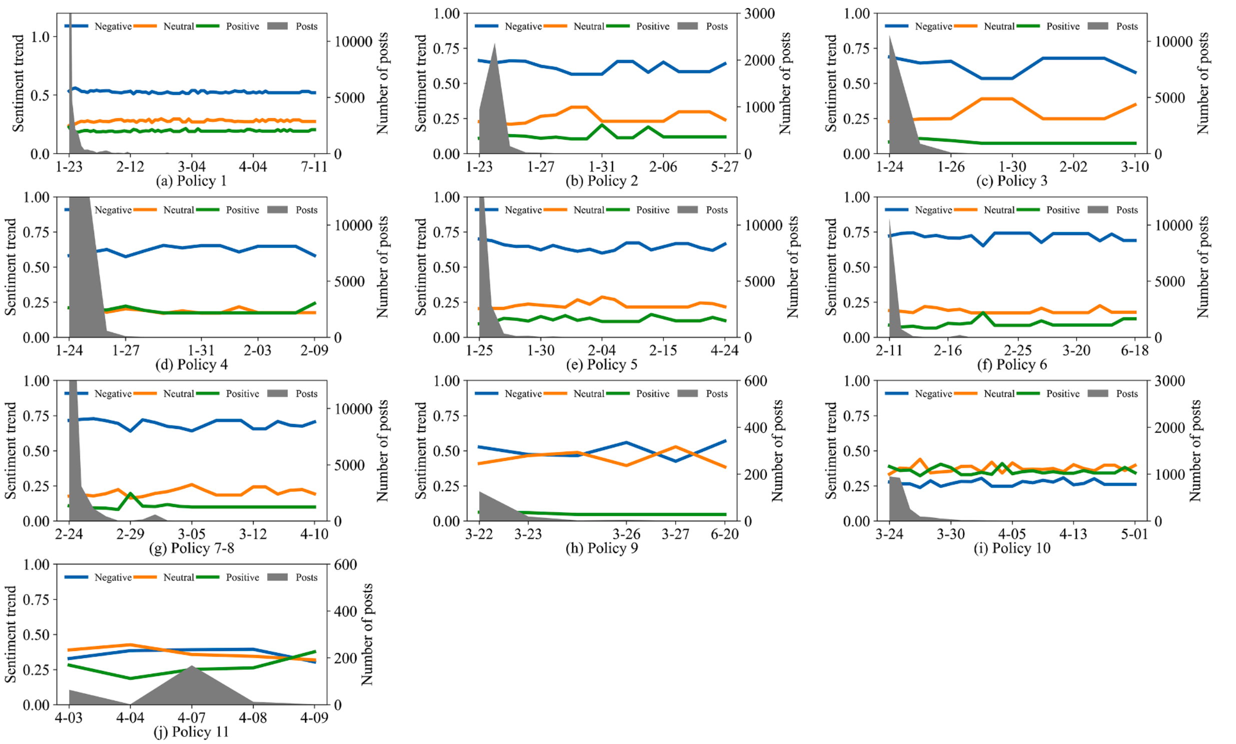

Fig. 4. Sentiment analysis of Wuhan COVID-19 transport policies.

the maximum length of the text sequence and the batch size were set as 160 and 32; the maximum learning rate during training was set as 4e-5, while the learning rate scheduler adopted a linear decay mechanism; the weight decay coefficient to avoid model overfitting and the round training epochs were set to 0.9 and 50. The training dataset consists of 24,000 manually labeled and calibrated data, which were further divided into the training set, validation set, and test set according to the ratio of 6:2:2. The final accuracy reached 98.7 % on the training set, 92.1 % and 91.6 % on the validation set and test set, respectively.

We use the growth rate of sentiment in the daily cumulative data to reflect the fluctuations of policy acceptance. The basic idea is that positive public sentiment implies a high level of acceptance, which promotes policy implementation and effectiveness, and vice versa. Precisely, the sentiment rate of the day was calculated by dividing the sum of the sentiment rate of the current day and the previous day by the total number of days. As shown in Fig. 4, such an approach can be seen as smoothing because some days have no posts data, and we are actually more interested in the upward part of the negative growth rate curve or the downward part of the positive growth rate curve because it means that the current day's sentiment rate is higher or lower than the average sentiment rate of all the days before it, which may represent the presence of a potential policy disruption or failure. By tracking the trends of people's attitudes toward policies, the potential unintended risks that need to be emphasized can be detected by identifying sentiment shifts (Wang et al., 2019). Further, the data points at the end of the growth rate curve represent the corresponding average sentiment over the policy life cycle. In general, the average percentage of people's negative attitudes toward Wuhan's COVID-19 policies was 50.1 %, while the neutral attitudes fluctuated at around 33.6 % and the positive attitudes were 16.3 %. In which the highest average negative sentiment rate was 75.2 % and the lowest was 28.7 %; the highest and lowest average neutral sentiment rates were 47.6 % and 10.7 %; while the average positive sentiment rates were 37.8 % and 2 %, respectively.

Notably, the negative sentiment shows an "up-down" trend for the transport policies issued in Wuhan during the lockdown period, while the positive and neutral sentiments show a "down-up" trend accordingly. In detail, as depicted in Fig. 4, Policy 4 shows the highest percentage of negative sentiment among all policies, averaging 75.2 % over its implementation period. It is noteworthy that while Policy 1 on lockdown attracts the highest public attention, i.e., had the highest number of posts, it corresponds to a relatively low percentage of negative sentiment at 34.5 %, followed by an upward trend in percentage from Policy 1 to Policy 4, and then down to 36.8 % at Policy 11. The results suggest that public sentiment has a process of accumulation and bursts out at a certain point if the policy does not sufficiently consider the public interest. Since public transport was unavailable after the lockdown, Policy 2 and Policy 3 on taxi cabs and road closures reduced the supply of intra-city travel, leading to the tension between supply and travel demand highlighted in Policy 4. It also shows that in an emergency, hot topics may not imply more negative sentiments. Next, Policy 6 had the second highest proportion of negative sentiment, averaging 66.1 % over the policy life cycle. This is likely because Policy 6 corresponded to the announcement of home quarantine and the strict access management of communities, thus leading to massive complaints about the lack of convenience for residents' daily necessities and activities. Indeed, although home and community quarantines with strengthened enforcement have proven to be effective pandemic control measures (Zhu and Tan, 2021), it may magnify social inequalities, with populations of higher socioeconomic status being less affected than those of lower socioeconomic status, as they tend to be able to work remotely under quarantine and often have better access to daily necessities and protective supplies, employment status, and community environments; and vice versa for those of lower socioeconomic status (Pullano et al., 2020).

In addition, Fig. 4 also reveals that the public's negative sentiment toward Policies 7 and 8 remained relatively high at 57.7 %. This may be due to the public's questioning of inconsistent policy decisions. On February 24, the Wuhan COVID-19 Working Group released Policy 7, which announced that people stranded in Wuhan due to the lockdown (long-distance commercial and recreational travelers, etc.) could leave the city after passing a temperature test. However, this announcement was rescinded three hours after its release and replaced by Policy 8, which announced that trips in and out of Wuhan city would be continuously strictly controlled. As Zhang et al. (2020) stated, policymakers should avoid making unstable policy decisions and providing inconsistent information in emergency situations. Further, unstable policy decisions may worsen the emergency, especially in a public health crisis. Policy changes regarding mobility management in outbreak-hit areas may increase the risk of virus transmission and lead to multiple waves of outbreaks. However, dynamic policy changes seem inevitable because the bounded rationality of policymakers (Simon, 2014), which stems from the complex policy-making environment that includes multiple stakeholder groups and the lack of extensive historical knowledge and response mechanisms for new outbreaks. For this, social media-based policy sentiment analysis and key themes from policy debate provide practical solutions to make early warnings about possible failures in policy implementation (Rogers, 2019), and contribute to the shaping of a new wave of policy making in a timely and relatively low-cost manner. Also, there is a need to strengthen communication and cooperation mechanisms between central and local governments, as well as between local government agencies (Malandrino and Demichelis, 2020). Then a month later, Wuhan issued Policy 10, which included the date of lifting the lockdown, the laws, regulations, and rules for resuming social production. For this policy, there were approximately 15 % more positive attitudes than negative ones, likely due to the recovery of people's expectations for the return to normal life. All in all, despite the paralyzing lockdown and anti-epidemic transport policies that took a huge personal cost on residents and led to an overall negative public sentiment, the virus was effectively contained.

5.2. Policy demands of Wuhan COVID-19 transport policies

The LDA model provides alternative sources of knowledge on policy demands by providing insight into people's discussions of policies that attract negative publicity and complaints. In this study, policy demands were retrieved in the form of topics using the developed entropy weight method (EWM) based LDA model. Meanwhile, the effectiveness of the EWM-based LDA model for topic extraction is compared with the LDA baseline model by measuring the similarity between the generated topics, with lower implying better performance. In the experiments, the number of topics K, α, and β was taken as the tuning hyperparameters; topics K were selected in the range [1,2,3,4,5,6,7,8,9,10] as the number of negative posts corresponding to the policies was relatively small averaging 7,734; The hyperparameters α and β were chosen in the range [0.01,0.11,0.21,0.31,0.41,0.51,0.61,0.71,0.81,0.91]. Through the grid search, each K matched 100 combinations of α and β, and each combination has two traditional metric scores derived from the dataset. For the baseline model, each K has an optimal combination (K, α, β) for model selection based on its evaluation metrics. Fig. 5 presents the coherence and perplexity scores for the candidate number of topics in the baseline model. In the coherence model (CM), a value of K that marks the end of the rapid growth of topic coherence (see red lines) was chosen, as it usually provides meaningful and interpretable topics; choosing a higher value can sometimes provide more fine-grained sub-topics (Islam, 2019). For the perplexity model (PM), the rate of perplexity change (RPC) based heuristic approach implemented by Zhao et al. (2015) was followed (see green lines), in which for the increasing number of candidate topic numbers $T_1, T_2, \cdots T_K$, the first T_i satisfying RPC (T_i) < RPC (T_{i+1}) is considered as the appropriate topic number.

According to the analysis, the two baseline models tend to select a relatively large number of topics, while the EWM-based model shows the opposite and achieves better inter-topic distinctiveness. As shown in Fig. 5, in the coherence model, the number of selected topics corresponding to policies 1 to 11 are 4,5,3,8,6,5,4,3,6,4, respectively; the perplexity model derives the results of 5,4,4,6,6,6,3,2,8,3. Correspondingly, the number of topics selected in the EWM-based model are 3,2,2,3,2,2,2,2,9,10,6. For further insight, Table 3 presents the selected hyperparameter combinations (K, α, β) after modeling different policies with EWM-based and baseline LDA models, while the validation of the derived topics was examined using the Jaccard distance, Perplexity and Coherence metrics. In detail, from Policy 1 to Policies 7 & 8, the EWM-based model tended to select a relatively small number of topics compared to the baseline model. As measured by the Jaccard distance, the EWM-based model outperformed the baseline model in terms of topic distinctiveness, for which the average Jaccard distance between topics is equal to or near 1, with the number of words ranging from 50 to 2050. Whereas the small number of topics resulted in relatively high perplexity values and lower coherence values on average for the EWM-based model. However, unexpected results emerged from policies 9, 10 and 11, where the validation of the EWM-based model shows the opposite results, i.e., the Jaccard distance measure is lower than the baseline model, while the perplexity and coherence values are higher as compared to the baseline model. The results suggest that the appropriate number of topics for EWM-based LDA topic modeling is sensitive to the corresponding dataset size. Since COVID-19 had been effectively controlled by the time of the release of Policies 9, 10, and 11, little attention was attracted to them.

To further clarify the reasons for the anomalous results of Policies 9, 10 and 11, the distance between topics corresponding to each of these policies were illustrated. As shown in Fig. 6, with the number of topic words ranging from 50 to 2050 in steps of 100, the left panel shows the average Jaccard distance between topics derived from policies 9, 10, and 11, while the right one reveals the distance under the specified number of words. Notably, the average Jaccard distance between the topics of policies 9 and 11 becomes 0 when

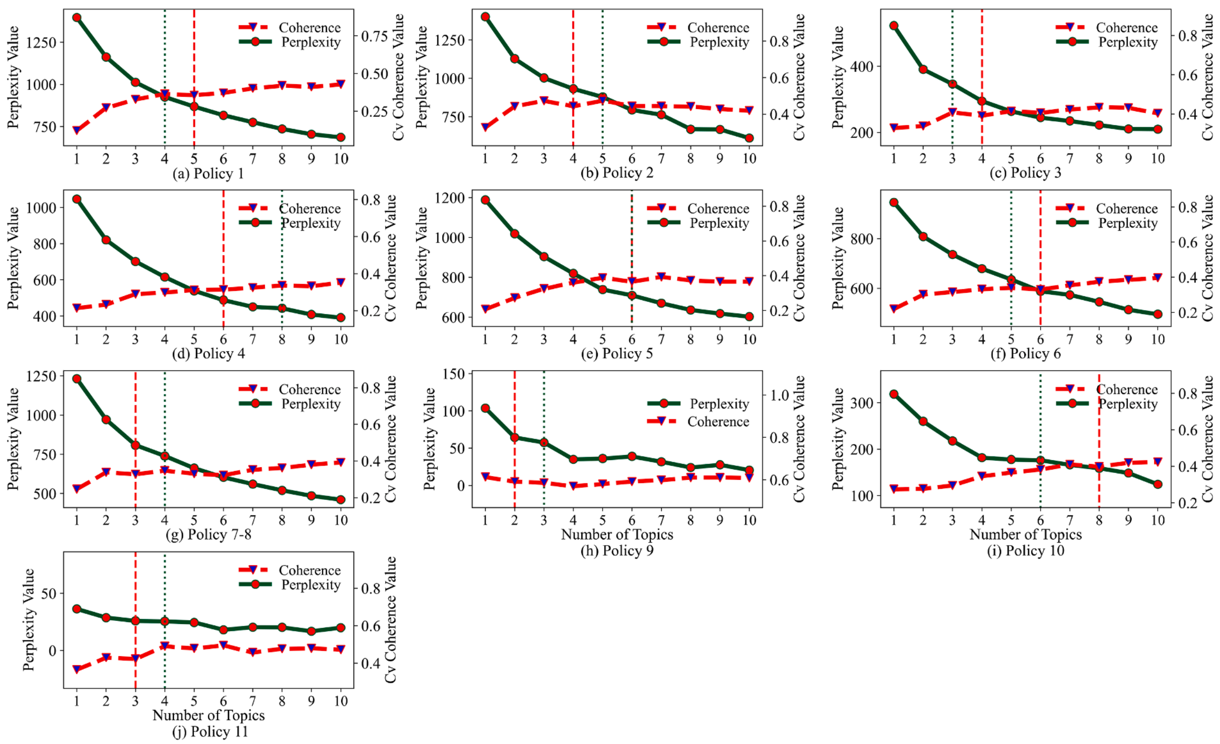

Fig. 5. LDA baseline models fit for each candidate topic number. The red and green vertical dashed lines indicate the number of topics selected by the baseline model CM and PM. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
The optimal number of topics and hyper-parameters selected by the developed EWM-based and baseline models.

	Parameter (K,α,β)			Jaccard Distance			Perplexity			Coherence		
	CM	PM	EWM	CM	PM	EWM	CM	PM	EWM	CM	PM	EWM
Policy 1	(4,0.11,0.11)	(5,0.01,0.21)	**(3,0.91,0.01)**	0.97	0.96	**1.0**	989.3	868.9	**2447.1**	0.36	0.31	**0.29**
Policy 2	(5,0.31,0.91)	(4,0.01,0.11)	**(2,0.91,0.01)**	0.92	0.96	**0.98**	936.6	932.1	**1450.4**	0.47	0.41	**0.47**
Policy 3	(3,0.41,0.01)	(4,0.01,0.11)	**(2,0.41,0.01)**	0.97	0.97	**0.98**	417.2	294.7	**536.2**	0.41	0.38	**0.40**
Policy 4	(8,0.51,0.51)	(6,0.01,0.21)	**(3,0.51,0.01)**	0.96	0.97	**1.0**	517.1	488.4	**1217.6**	0.34	0.29	**0.24**
Policy 5	(6,0.31,0.21)	(6,0.01,0.31)	**(2,0.31,0.01)**	0.95	0.94	**1.0**	729.8	709.5	**1667.8**	0.37	0.33	**0.26**
Policy 6	(5,0.51,0.71)	(6,0.01,0.21)	**(2,0.31,0.01)**	0.96	0.94	**1.0**	701.1	587.7	**1199.1**	0.34	0.33	**0.29**
Policies 7 & 8	(4,0.51,0.41)	(3,0.01,0.21)	**(2,0.31,0.01)**	0.97	0.98	**1.0**	800.1	807.6	**1276.3**	0.34	0.31	**0.33**
Policy 9	(3,0.91,0.01)	(2,0.01,0.91)	**(9,0.91,0.01)**	0.17	0.23	**0.15**	102.3	64.3	**106.7**	0.58	0.55	**0.61**
Policy 10	(6,0.21,0.01)	(8,0.01,0.01)	**(10,0.11,0.91)**	0.76	0.80	**0.65**	179.6	158.8	**331.6**	0.38	0.37	**0.42**
Policy 11	(4,0.01,0.01)	(3,0.01,0.01)	**(6,0.01,0.91)**	0.30	0.32	**0.28**	25.8	25.9	**45.9**	0.49	0.45	**0.49**

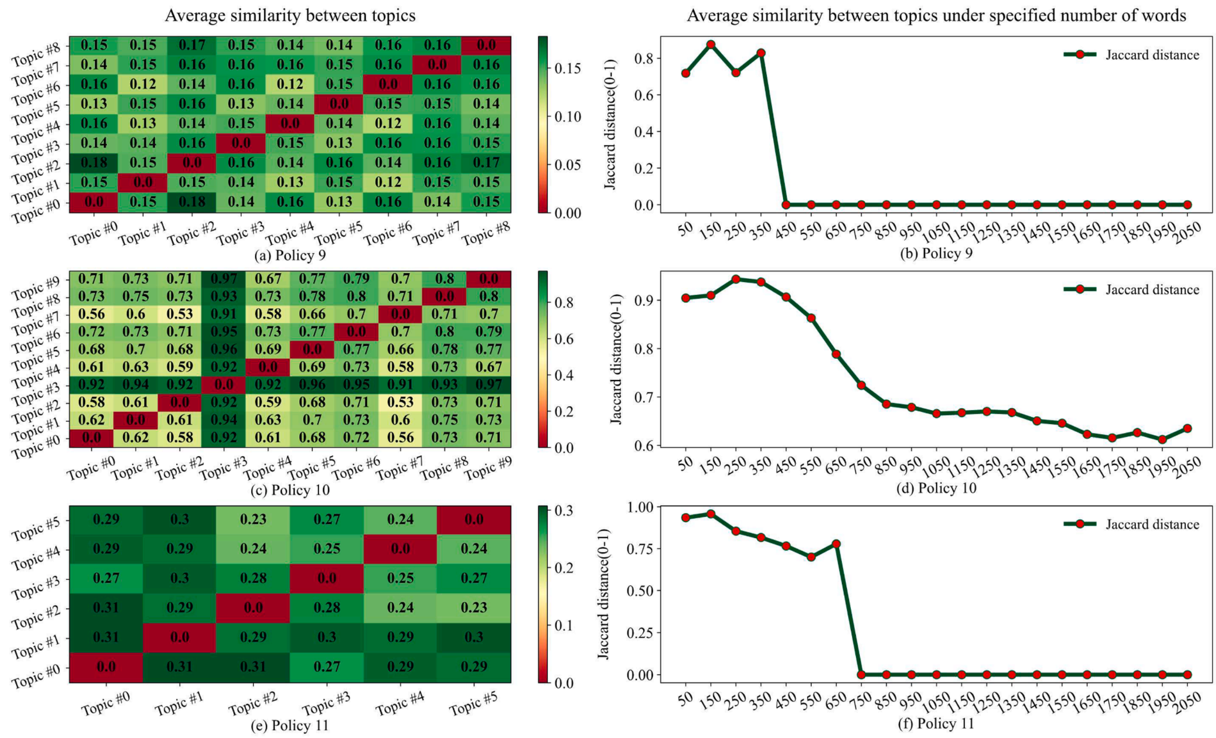

Fig. 6. The average Jaccard distances between topics in Policies 9,10, and 11.

the number of words increases to 450 and 750 onwards. Therefore, the anomalous experimental results may be caused by the small size of the modeled data set. The results suggest that EWM-based LDA topic modeling may not be applicable to small datasets. At the same time, small datasets have also been proven to be problematic in traditional LDA modeling (Qiang et al., 2020) .

For this, the topics corresponding to policies 9, 10, and 11 were further post-processed to obtain more interpretable outcomes, with specific Jaccard similar topics being merged. Using Policy 9 as an example, Fig. 6**(a)** reveals that the Jaccard distance between Topics #0 and #1 was 0.15. This implies a small difference between these two topics and poor results. Accordingly, Topics #0 and #1 were then merged into one topic. This synthesized topic was then merged into the next topic closest to it. In this study, the human-in-the-loop process was repeated until the distance between topics was not less than 0.90. Finally, the number of topics for policies 9, 10, and 11 was set to 1, 2, and 1 subjectively combined with the Jaccard distance metric. Typically, the topics derived from the LDA model need to be interpreted manually with labels assigned to the topics. Fig. 7 gives word cloud representations of the Wuhan COVID-19 transport policies discussed topics. In each topic, the top 20 words with high probability are displayed, and the font size is proportional to the probability value, indicating that the high probability words have more capability to explain the topic. Therefore, considering the policy context, the label for each topic was intuitively defined based on the connotations of the words.

The results of the topic analysis show that public concerns about COVID-19 transport policies focus on commuting of anti-epidemic related staff, necessary travel for vulnerable groups in the community, information and communication mechanisms, disinfection of available travel tools, logistics and courier services, the safety of resuming production, unstable policy decisions. This implies that the

Fig. 7. The word cloud representations of topics.

policies issued did not adequately consider the roles and interests of various segments of the public in Wuhan early before the effective vaccine and medical aid was distributed. Nevertheless, the transport policies in Wuhan played a prominent role in the response, preparedness, mitigation, and recovery from COVID-19 (Kim, 2022). Moreover, the COVID-19 pandemic also provides an opportunity to improve the resilience of urban transportation systems in response to future similar crises, and some of the lessons learned and policy implications are summarized below.

1. **Transport organization and services for essential trips:** The results of the topic analysis revealed an insufficient number of commuting vehicles for COVID-19 affairs-related staff, and taxi service regulations could not address pandemic conditions. The difficulties in commuting faced by medical staff persisted since the implementation of the lockdown from Policy 1 to Policy 5. Tweets like “My brother was notified that he needed to return to work at the hospital, but after public transportation was suspended, he had to commute to work by riding a shared bike, which took about 30 min each way” and “When I asked the driver to take me to the hospital, the driver refused to provide service”; these tweets indicate that the commuting of staff was not adequately considered in the early stages of the outbreak.
2. **Travel equity and accessibility:** Policies 2 and 4 both address the issue of the supply of taxis. The results of the topic analysis indicate that the 6,000 government-owned recruited cabs are insufficient to meet the community’s necessary travel demands. The high rigidity of travel demand from vulnerable groups persisted since the implementation of the lockdown, as the population of each community in Wuhan is approximately 8,000 to 10,000, and the city had 1159 residential communities as of 2019 (Chen et al., 2020a) .
3. **Controlling fake news and updating transportation information during the pandemic:** Topic #0 in Policy 1 and Topic #1 in Policy 10 reveal insufficient openness, lack of transparency, and a limited number of open information channels. For example, people were not adequately informed about highway operations after the lockdown, and social media users had divergent views; some said that highways were not completely blocked, while others had the opposite perspective. In such cases, the public often does not know how to distinguish between rumor and truth, and the panic caused by this is akin to “adding fuel to the fire” in emergency management. Simultaneously, the risk of negative health effects (including death) from misguided news in a pandemic has been verified (Cuello-Garcia et al., 2020). Therefore, information filtering, verification, and communication mechanisms that integrate multiple sources of social media channels should be explored and developed, thus enabling travelers to make safer and more orderly choices.
4. **Public health (the need to contain the spread of the virus and protect travelers and worker):** Topic #1 in Policy 3 reveals the insufficient epidemic prevention and management measures implemented for public transport facilities. According to the analysis results, many citizens expressed concern about the management of bicycle sharing, such as disinfection; the government had not launched a management policy in response to this, which may have increased the risk of the spread of the virus, as most viruses can remain on a physical surface for several hours (Teixeira and Lopes, 2020). The same problem also arose in response to Policy 4, as the recruited taxis were only equipped with simple disinfection materials, thereby increasing the risk of the further transmission of

the virus due to the cramped and unventilated environment to which passengers and drivers are exposed (Chen et al., 2020a) . In addition, some scholars have suggested that active transport systems should be improved, which can reduce the risk of infection while promoting sustainable development and healthy lifestyles (Zhang et al., 2021; Mogaji et al., 2022).

5. **Logistics and courier delivery services:** The COVID-19 outbreak also harmed logistics and courier services in Wuhan. After the lockdown, some large shopping centers were closed, especially when communities were quarantined, the residents' online shopping behavior increased the demand for logistics transportation. However, cumbersome procedures in and out of Wuhan via highways due to traffic control measures and labor shortages caused by the Chinese New Year holiday led to a significant negative impact on express logistics (as indicated by Topic #0 in Policy 5). For this, express logistics that consider factors such as labor shortages, traffic restrictions, and service costs should be explored and developed. Also, disinfection programs and automation aids in the logistics chain should be developed to prevent infection among residents and employees (Yang et al., 2021; Zhang et al., 2021). For long-term perspective, the planning of logistics hubs around the city should be promoted so that emergency supplies can be delivered quickly with shortened transportation times.
6. **Asymptomatic infected people in the return-to-work population flow:** Topic #0 in Policy 10 expresses concern about asymptomatic infected people in population flows. Essentially, the most effective way of pandemic control is tracking individuals who have had close contact with infected patients, combined with antigen testing of potential carriers (Kim, 2022). Meanwhile, mobility data from traffic cards, mobile phone signals and app-based location records offer opportunities for pandemic contact tracing. In response, screening methods for infected and at-risk populations based on travel and activity information derived from mobility data should be developed.
7. **Avoiding unstable policy decisions:** Topic #0 derived from the negative opinions of policies 7&8 implies a low public acceptance of inconsistent policy decisions. Indeed, in contrast to one-off natural or manufactured hazards, pandemics may have multiple waves of outbreaks lasting months or longer, and are often accompanied by unknown and more severe social impacts (Fakhruddin et al., 2020). Simultaneously, decision makers must make a reasonable equilibrium between travel demand and supply contexts under abnormal conditions, given the bounded rationality resulting from the lack of historical experience and knowledge (Simon, 2014), trial-and-error policymaking to adapt to changing situations seems inevitable. For this, an evidence-based decision-making paradigm using models and simulations should be promoted and adopted, and candidate transport policies should be pre-evaluated, while the implemented policies should be monitored and post-evaluated to identify potential policy disruption and failure in a timely manner.

5.3. Response analysis of policy adjustment

This section provides an example to conduct the response analysis of policy adjustments based on the Jaccard distances. Policies 2 and 4 are taken as the case study. The Wuhan COVID-19 Working Group released Policy 2 to reduce the supply of taxis in the early days of the outbreak, followed by Policy 4, which announced the recruitment of an additional 6,000 taxis, namely 3–5 per community to offer travel services for essential travel demands. Thus, Policy 4 can be considered as a complement and improvement to Policy 2. It is assumed that, ideally, the shortage of taxi supply that occurred after the release of Policy 2 would no longer recur during the implementation of Policy 4, i.e., the release of Policy 4 effectively guarantees essential trips from the community, and vice versa, while Policy 4 needs further improvement.

The average Jaccard distance was used to measure the difference in topic distribution between Policies 2 and 4 (see Fig. 8). The results demonstrate that Topic #0 in Policy 2 had a high correlation with Topic #2 in Policy 4, as the Jaccard distance was 0.83, i.e., below the threshold of 0.90. Both topics shared the same concern about commuting travel for medical staff who did not own private

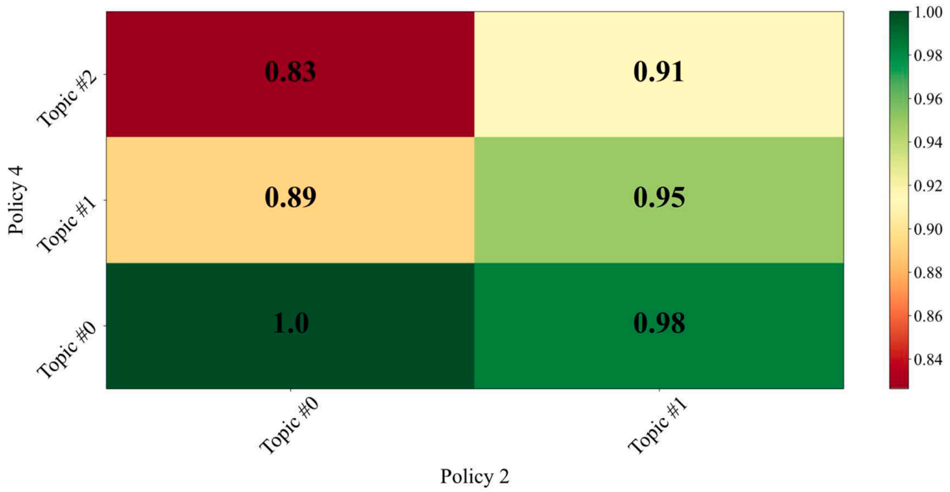

Fig. 8. The response analysis of policy adjustment between Policies 2 and 4.

cars. This is likely because Policy 2 provided taxi service for all community members, including vulnerable groups and community workers, medical staff, and volunteers, and the Policy 4 still does not address the commuting difficulties of medical staff. In the early stages of the outbreak, a certain percentage of cruising taxis was still in operation based on license plate numbers after Policy 2. Topic #0 in Policy 2 and Topic #1 in Policy 4 also exhibited a high correlation with a Jaccard distance of 0.89. Similar tweets on these two topics, such as " The taxi driver refused to pick me up when I asked him/her to take me to the hospital," show that people paid attention to the service quality of cruising taxis. In summary, the analysis indicates that Policy 4 of the adjustments to Policy 2 requires further improvement.

There is a wealth of document guidance and technical tools for the evaluation of the policy implementation, such as survey-based before-and-after comparisons (Burchell et al., 2019), economic evaluations, e.g., cost-benefit and cost-effectiveness analyses (Batarce et al., 2016; Basso et al., 2021), and case study evaluations (Liu et al., 2018, 2020a) . However, constructing these experiments or scenarios is often time-consuming and labor-intensive, especially in complex environments with dynamic pandemic impacts and public awareness and demand. The proposed social media data-based bottom-up methodology provides a feasible solution to this problem, as it can measure the response of policy adjustments in a timely and citizen-centered manner.

6. Conclusion

This study aims to provide an integrated analytical methodology for quantitatively evaluating urban transport policies during a pandemic. An analytical methodology and corresponding methods based on social media data were developed. The analytical methodology integrates data acquisition and processing, sentiment analysis, and topic modeling. Specifically, the sentiment-aware pre-trained language model was fine-tuned to measure the policy's public acceptance level. Next, policy demands representing the values, interests, and benefits of the public were further retrieved from the negative opinions using the developed EWM-based LDA model for understanding the issues of policies on society. Then, the Jaccard distance was used to conduct the response analysis of the policy adjustment by measuring the similarity of the LDA topics. An empirical investigation was conducted using the transport policies of Wuhan, China, during the COVID-19 pandemic. The major contributions of this study can be summarized as follows.

1. An analytical methodology based on social media data is developed for pandemic transport policy response analysis. The methodology enables two-way communication with stakeholders and the broader public that makes it possible to gain key insights into the opinions and attitudes of individual citizens and inform the policy making process.
2. An entropy weight-based method is developed to judge the appropriate number of topics for the LDA model. The developed method constructs a comprehensive metric to judge the number of topics for the LDA model, making the topics more interpretable in representing policy demands.
3. A Jaccard distance-based method is proposed to conduct the response analysis of policy adjustments. The similarity of LDA topics between policies before and after the adjustment is measured by the Jaccard distance as an evaluation criterion for how well the policy adjustment works.

The findings showed that the proposed policy-making support approach is effective to evaluate the acceptance of anti-pandemic policies from the public's perspective, to assess the balance between policies and people's demands, and to conduct the response analysis of a series of policy adjustments based on people's feedback. The case study suggests that the Wuhan COVID-19 transport anti-epidemic policies did not adequately consider the roles and interests of various segments of the public early before the effective vaccine and medical aid was distributed. And public concerns mainly focus on commuting of anti-epidemic related staff, necessary travel for vulnerable groups in the community, information and communication mechanisms, disinfection of available travel tools, logistics and courier services, the safety of resuming production, unstable policy decisions.

The social media-based analytical methodology proposed in this study works on the monitoring phase of the policy making and implementation process, providing a complementary tool for policy response analysis that can be applied to future similar emergency situations. However, the proposed analytical methodology does have some limitations and challenges. First, the analysis procedure for policy interpretation is not fully automated and the analysis results rely on subjective human interpretation. In the near future, we will consider the application of more sophisticated text analysis or natural language processing techniques such as Transformer-based text summarization to enable more automated and reliable policy analysis. Second, the methodology currently involves only a single data source due to limited data access. Data obtained from different social media platforms may show different opinions towards the same event due to the variability of user groups. In the future, we will integrate data sources from multiple social media platforms in China (e.g., WeChat, public transport forums) into the developed methodology to increase the robustness of analysis results.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Key R&D Program of China (2018YFB1601100). We also thank the anonymous reviewers for their insightful comments.

Appendix A. Mathematical Derivations For LDA Inference

In this appendix, we provide the implementation of Latent Dirichlet Allocation (LDA) model for parameter estimation and inference. The variable names are defined as follows:

- M denotes the number of documents. V denotes the number of words in the vocabulary. K denotes the number of topics.
- N_m is number of words in a given document (document m has N_m words).
- $\overrightarrow{\alpha}$ is the parameter of the Dirichlet prior on the per-document topic distributions. K-dimensional vector of positive reals, and α_k is the prior weight of topic k in a document.
- $\overrightarrow{\beta}$ is the parameter of the Dirichlet prior on the per-topic word distributions. V-dimensional vector of positive reals, and β_t is the prior weight of word t in a topic.
- $\overrightarrow{\theta}_m$ is the distribution of topics in document m. K-dimensional multinomial parameters vector of probabilities, which sums to 1.
- $\Theta = \left(\overrightarrow{\theta}_1, \cdots, \overrightarrow{\theta}_m\right)^T$ consists of rows defined by documents and columns defined by topics, $M \times K$ matrix.
- $\overrightarrow{\varphi}_k$ is the distribution of words in topic k. V-dimensional multinomial parameters vector of probabilities, which sums to 1.
- $\Phi = \left(\overrightarrow{\varphi}_1, \cdots, \overrightarrow{\varphi}_k\right)^T$ consists of rows defined by topics and columns defined by words, $K \times V$ matrix.
- $\overrightarrow{z}$ is the identity topic of all words in all documents.N-dimensional vector of integers between 1 and K. N is the total number of words in all documents. $z_{m,n}$ is the topic for the n-th word in document m.
- $\overrightarrow{w}$ is the identity of all words in all documents.N-dimensional vector of integers between 1 and V. $w_{m,n}$ is the specific n-th word in document m.

In the modeling, the LDA parameters need to be inferred given a set of documents, namely the topic of each word $\overrightarrow{z}$, document-topic distribution Θ, and topic-word distribution Φ. As depicted in Fig. 2, LDA document generation consists of two processes, $\overrightarrow{\alpha}\rightarrow\overrightarrow{\theta}_m\rightarrow z_{m,n}$ and $\overrightarrow{\beta}\rightarrow\overrightarrow{\varphi}_k\rightarrow w_{m,n}\,|\,k = z_{m,n}$,and in general $p(\overrightarrow{z})$ with $p(\overrightarrow{w}\,|\,\overrightarrow{z})$. The primary goal of LDA inference is to determine the posterior topic distribution of each word $p(\overrightarrow{z}\,|\overrightarrow{w})$, i.e., topic-word co-occurrence frequency matrix, which is directly proportional to the joint distribution $p(\overrightarrow{w}, \overrightarrow{z})$ (Blei et al., 2003):

$$p(\overrightarrow{z}\,|\overrightarrow{w}) = \frac{p(\overrightarrow{w}|\overrightarrow{z})p(\overrightarrow{z})}{p(\overrightarrow{w})} = \frac{\prod_{i=1}^{W} p(z_i, w_i)}{\prod_{i=1}^{W}\sum_{k=1}^{K} p(z_i = k, w_i)} \tag{A.1}$$

W represents the number of corpus words, which is smaller than the vocabulary size V. However, this distribution cannot be solved analytically due to the intractability of solving for the marginal likelihood $p(\overrightarrow{w})$ (the denominator would be computationally expensive, as it covers a large space of discrete random variables and represents a summation over of K^W terms) (Heinrich, 2009). At this point, Gibbs sampling based on the Markov chain Monte Carlo (MCMC) algorithm is widely implemented to approximate the posterior distribution $p(\overrightarrow{z}\,|\overrightarrow{w})$ under the observed corpus, due to its sufficient detailed balance condition for converging to the sampled target distribution (Heinrich, 2009).

The Gibbs sampling procedure involves the definition of a Gibbs sampler, which requires to formulate the joint distribution of $\overrightarrow{z}$ and $\overrightarrow{w}$. Considering the Dirichlet prior parameters, the joint distribution can be factorized:

$$p\left(\overrightarrow{w}, \overrightarrow{z}\,|\overrightarrow{\alpha}, \overrightarrow{\beta}\right) = p\left(\overrightarrow{w}|\overrightarrow{z}, \overrightarrow{\beta}\right)p(\overrightarrow{z}\,|\overrightarrow{\alpha}) \tag{A.2}$$

In equation (A.2), the term $\overrightarrow{w}$ is independent of $\overrightarrow{\alpha}$, while the second term $\overrightarrow{z}$ is independent of $\overrightarrow{\beta}$. Thus, the two elements of the joint distribution can now be addressed separately. The target distribution $p\left(\overrightarrow{w}|\overrightarrow{z}, \overrightarrow{\beta}\right)$ on the right side of the equation is obtained by integrating over the parameter Φ:

$$p\left(\overrightarrow{w}|\overrightarrow{z}, \overrightarrow{\beta}\right) = \int p(\overrightarrow{w}|\overrightarrow{z}, \Phi)p\left(\Phi|\overrightarrow{\beta}\right)\mathrm{d}\Phi \tag{A.3}$$

$$= \int \prod_{z=1}^{K}\prod_{t=1}^{V}\varphi_{z,t}^{n_k^{(t)}} \bullet \prod_{z=1}^{K}\frac{1}{\Delta\left(\overrightarrow{\beta}\right)}\prod_{t=1}^{V}\varphi_{z,t}^{\beta_t-1}\mathrm{d}\overrightarrow{\varphi}_z \tag{A.4}$$

$$= \int \prod_{z=1}^{K}\frac{1}{\Delta\left(\overrightarrow{\beta}\right)} \bullet \prod_{t=1}^{V}\varphi_{z,t}^{n_k^{(t)}+\beta_t-1}\mathrm{d}\overrightarrow{\varphi}_z \tag{A.5}$$

$$= \prod_{z=1}^{K} \frac{\Delta\left(\vec{n}_z + \vec{\beta}\right)}{\Delta\left(\vec{\beta}\right)} \tag{A.6}$$

where $n_k^{(t)}$ is the number of times that term t is observed in topic k, $\vec{n}_z = \left\{n_z^{(t)}\right\}_{t=1}^{V}$. $\Delta(\bullet)$ is a multivariate beta function that represents the normalized constant of the Dirichlet distribution. Eq. (A.6) can be explained as the product of K Dirichlet-multinomial models, representing the corpus with K independent "topic texts" (Heinrich, 2009). Likewise, the topic distribution $p(\vec{z}\,|\vec{\alpha})$ can be derived by integrating over the parameter Θ, starting with rewriting it as the product of the multinomial distribution and the Dirichlet prior distribution:

$$p(\vec{z}\,|\vec{\alpha}) = \int p(\vec{z}|\Theta)p(\Theta|\vec{\alpha})\mathrm{d}\Theta \tag{A.7}$$

$$= \int \prod_{m=1}^{M} \prod_{k=1}^{K} \theta_{m,k}^{n_m^{(k)}} \bullet \prod_{m=1}^{M} \frac{1}{\Delta(\vec{\alpha})} \prod_{k=1}^{K} \theta_{m,k}^{\alpha_k - 1} \mathrm{d}\vec{\theta}_m \tag{A.8}$$

$$= \int \prod_{m=1}^{M} \frac{1}{\Delta(\vec{\alpha})} \bullet \prod_{k=1}^{K} \theta_{m,k}^{n_m^{(k)} + \alpha_k - 1} \mathrm{d}\vec{\theta}_m \tag{A.9}$$

$$= \prod_{m=1}^{M} \frac{\Delta\left(\vec{n}_m + \vec{\alpha}\right)}{\Delta(\vec{\alpha})} \tag{A.10}$$

where $n_m^{(k)}$ refers to the number of times that topic k has been observed with a word of document m, $\vec{n}_m = \left\{n_m^{(k)}\right\}_{k=1}^{K}$. Then the joint distribution therefore becomes:

$$p\left(\vec{w}, \vec{z}\,|\vec{\alpha}, \vec{\beta}\right) = \prod_{z=1}^{K} \frac{\Delta\left(\vec{n}_z + \vec{\beta}\right)}{\Delta\left(\vec{\beta}\right)} \bullet \prod_{m=1}^{M} \frac{\Delta\left(\vec{n}_m + \vec{\alpha}\right)}{\Delta(\vec{\alpha})} \tag{A.11}$$

Further, from the joint distribution, the sampling formula for the topic of the n-th word with index $i = (m, n)$ in the m-th document can be derived:

$$p\left(z_i = k|\vec{z}_{\neg i}, \vec{w}\right) = \frac{p(\vec{w}, \vec{z})}{p\left(\vec{w}, \vec{z}_{\neg i}\right)} \propto \frac{n_{k,\neg i}^{(t)} + \beta_t}{\sum_{t=1}^{V}\left(n_{k,\neg i}^{(t)} + \beta_t\right)} \bullet \frac{n_{m,\neg i}^{(k)} + \alpha_k}{\sum_{k=1}^{K}\left(n_{m,\neg i}^{(k)} + \alpha_k\right)} \tag{A.12}$$

Where the counts $n_{\bullet,\neg i}^{(\bullet)}$ indicate that the word i is excluded from the corresponding document or topic. The Gibbs sampler runs in cycles through an iterative Monte Carlo scheme, a topic is randomly assigned for each word within each document from the multinomial distribution during the initialization phase, and then the corpus is iteratively scanned, and for each word its topic is resampled and updated according to the Gibbs sampling formula (Heinrich, 2009), until the Gibbs sampling converges. Finally, the co-occurrence frequency matrix of the topic-word is obtained, which is the basis for the interpretation of the LDA model results, while the values of the parameters Θ and Φ can then also be derived.

Appendix B. EWM-BASED COMPREHENSIVE EVALUATION INDEX

To begin with, the number of topics K and the Dirichlet prior distribution parameters α, β are taken as the tuning hyperparameters. Then the grid search method is applied to fit the LDA model under the parameter combination (K, α, β), resulting in multiple LDA models over the same documents, and the coherence and perplexity scores for each parameter combination (K, α, β) are recorded. Further, the EWM is adopted to derive the comprehensive evaluation index score, and its highest value is considered to correspond to the optimal parameters. Based on the algorithm described in Yue et al. (2017), the following steps is carried out:

Suppose that there are X options for the parameter combinations (K, α, β), with Y refers to the evaluation metric types. The x-th option with the y-th metric has a value of q_{xy}, the original matrix $\mathbf{Q}_{X\times Y}$ can be constructed.
The matrix $\mathbf{Q}$ is standardized and marked as $\mathbf{D}$, $\mathbf{D} = (d_{xy})_{X\times Y}$, where d_{xy} is calculated as follows:

$$d_{xy} = \frac{q_{xy} - \min(\{q_{1Y}, q_{2Y}, ..., q_{XY}\})}{\max(\{q_{1Y}, q_{2Y}, ..., q_{XY}\}) - \min(\{q_{1Y}, q_{2Y}, ..., q_{XY}\})} \tag{B.1}$$

The matrix $\mathbf{D}$ is normalized and marked as $\mathbf{G} = (g_{xy})_{X\times Y}$, $g_{xy} = d_{xy}/\sum_{x=1}^{X} d_{xy}$, $\sum_{x=1}^{X} g_{xy} = 1, x = 1, 2, ..., X$, and $y = 1, 2, ..., Y$. The information entropy of the y-th metric is:

$$E_y = \frac{\sum_{x=1}^{X} g_{xy} \ln(g_{xy})}{-\ln(X)} \tag{B.2}$$

The weight of each metric is calculated with the information entropy. The weight of the y-th metric is:

$$w_y = \frac{1 - E_y}{Y - \sum_{y=1}^{Y} E_y} \tag{B.3}$$

The comprehensive index score of each parameter combination (K, α, β) is then calculated as follows:

$$score_x = \sum_{y=1}^{Y} w_y g_{xy} \tag{B.4}$$

References

Alexander, D.E., 2014. Social media in disaster risk reduction and crisis management. Sci. Eng. Ethics 20, 717–733. https://doi.org/10.1007/s11948-013-9502-z.

An, Y., Lin, X., Li, M., He, F., 2021. Dynamic governance decisions on multi-modal inter-city travel during a large-scale epidemic spreading. Transp. Policy 104, 29–42. https://doi.org/10.1016/j.tranpol.2021.01.008.

Basso, L.J., Montero, J.P., Sepúlveda, F., 2021. A practical approach for curbing congestion and air pollution: driving restrictions with toll and vintage exemptions. Transp. Res. Part A Policy Pract. 148, 330–352. https://doi.org/10.1016/j.tra.2021.02.011.

Batarce, M., Muñoz, J.C., Ortúzar, J.de D., 2016. Valuing crowding in public transport: implications for cost-benefit analysis. Transp. Res. Part A Policy Pract. 91, 358–378. https://doi.org/10.1016/j.tra.2016.06.025.

Belkahla Driss, O., Mellouli, S., Trabelsi, Z., 2019. From citizens to government policy-makers: social media data analysis. Gov. Inf. Q. 36, 560–570. https://doi.org/10.1016/j.giq.2019.05.002.

Benini, P.B.J., 2020. A pandemic playbook for transportation agencies. 10.17226/25993.

Blair, S.J., Bi, Y., Mulvenna, M.D., 2020. Aggregated topic models for increasing social media topic coherence. Appl. Intell. 50, 138–156. https://doi.org/10.1007/s10489-019-01438-z.

Blei, D.M., Ng, A.Y., Jordan, M.T., 2003. Latent dirichlet allocation. Adv. Neural Inf. Process. Syst. 3, 993–1022.

Bokányi, E., Kondor, D., Vattay, G., 2021. Scaling in words on twitter. R. Soc. Open Sci. 6, 190027 https://doi.org/10.1098/rsos.190027.

Burchell, J., Ison, S., Enoch, M., Budd, L., 2019. Implementation of the workplace parking levy as a transport policy instrument. J. Transp. Geogr. 80, 102543 https://doi.org/10.1016/j.jtrangeo.2019.102543.

Chen, Y., Wang, Y., Wang, H., Hu, Z., Hua, L., 2020a. Controlling urban traffic-one of the useful methods to ensure safety in Wuhan based on COVID-19 outbreak. Saf. Sci. 131, 104938 https://doi.org/10.1016/j.ssci.2020.104938.

Chen, Y., Silva, E.A., Reis, J.P., 2020b. Measuring policy debate in a regrowing city by sentiment analysis using online media data: a case study of Leipzig 2030. Reg. Sci. Policy Pract. 675–692 https://doi.org/10.1111/rsp3.12292.

Cuello-Garcia, C., Pérez-Gaxiola, G., Amelsvoort, L., 2020. Social media can have an impact on how we manage and investigate the COVID-19 pandemic. J. Clin. Epidemiol. 127, 198–201. https://doi.org/10.1016/j.jclinepi.2020.06.028.

Dandannavar, P.S., Mangalwede, S.R., Deshpande, S.B., 2019. A proposed framework for evaluating the performance of government initiatives through sentiment analysis, Advances in Intelligent Systems and Computing. Springer Singapore. 10.1007/978-981-13-0617-4_32.

Deng, Q., Gao, Y., Wang, C., Zhang, H., 2020. Detecting information requirements for crisis communication from social media data: an interactive topic modeling approach. Int. J. Disaster Risk Reduct. 50, 101692 https://doi.org/10.1016/j.ijdrr.2020.101692.

El-Araby, A., Sabry, I., El-Assal, A., 2022. A comparative study of using MCMD methods integrated with entropy weight method for evaluating facility location problem. Oper. Res. Eng. Sci. Theory Appl. 5, 121–138. https://doi.org/10.31181/oresta250322151a.

Fakhruddin, B., Blanchard, K., Ragupathy, D., 2020. Are we there yet? The transition from response to recovery for the COVID-19 pandemic. Prog. Disaster Sci. 7, 100102 https://doi.org/10.1016/j.pdisas.2020.100102.

Fang, H., Wang, L., Yang, Y., 2020. Human mobility restrictions and the spread of the Novel Coronavirus (2019-nCoV) in China. J. Public Econ. 191, 104272 https://doi.org/10.1016/j.jpubeco.2020.104272.

Gal-Tzur, A., Grant-Muller, S.M., Kuflik, T., Minkov, E., Nocera, S., Shoor, I., 2014. The potential of social media in delivering transport policy goals. Transp. Policy 32, 115–123. https://doi.org/10.1016/j.tranpol.2014.01.007.

Hasan, S., Ukkusuri, S.V., 2014. Urban activity pattern classification using topic models from online geo-location data. Transp. Res. Part C Emerg. Technol. 44, 363–381. https://doi.org/10.1016/j.trc.2014.04.003.

Heinrich, G., 2009. Parameter estimation for text analysis [WWW Document]. accessed 7.22.2022 Technical report, University of Leipzig, Germany, version 2, 5. http://www.arbylon.net/publications/text-est.pdf.

Islam, T., 2019. Yoga-Veganism: Correlation mining of twitter health data. ArXiv, abs/ 1906.07668. 10.48550/arXiv.1906.07668.

Khan, S.M., Ngo, L.B., Morris, E.A., Dey, K., Zhou, Y., 2017. Social media data in transportation. Data Analytics for Intelligent Transportation Systems. Elsevier Inc. 10.1016/B978-0-12-809715-1.00011-0.

Kim, K.E., 2022. Ten takeaways from the covid-19 pandemic for transportation planners. Transp. Res. Rec. J. Transp. Res. Board 2019, 036119812210905. 10.1177/03611981221090515.

Kumar, D., Ukkusuri, S.V., 2020. Enhancing demographic coverage of hurricane evacuation behavior modeling using social media. J. Comput. Sci. 45, 101184 https://doi.org/10.1016/j.jocs.2020.101184.

Leavey, J., 2013. Social media and public policy: what is the evidence? [WWW Document]. URL http://www.politicipublice.ro/uploads/Social-Media-and-Public-Policy.pdf (accessed 7.22.2022).

Li, Z., Wang, C., Emrich, C.T., Guo, D., 2018. A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods. Cartogr. Geogr. Inf. Sci. 45, 97–110. https://doi.org/10.1080/15230406.2016.1271356.

Li, D., Zhang, Y., Li, C., 2019. Mining public opinion on transportation systems based on social media data. Sustain. 11 (15), 4016. https://doi.org/10.3390/su11154016.

Lian, Y., Liu, Y., Dong, X., 2020. Strategies for controlling false online information during natural disasters: the case of Typhoon Mangkhut in China. Technol. Soc. 62, 101265 https://doi.org/10.1016/j.techsoc.2020.101265.

Liu, Z., Li, R., Wang, X., Shang, P., 2018. Effects of vehicle restriction policies: analysis using license plate recognition data in Langfang, China. Transp. Res. Part A Policy Pract. 118, 89–103. https://doi.org/10.1016/j.tra.2018.09.001.
Liu, Z., Li, R., Wang (Cara), X., Shang, P., 2020a. Noncompliance behavior against vehicle restriction policy: a case study of Langfang, China. Transp. Res. Part A Policy Pract. 132, 1020–1033. https://doi.org/10.1016/j.tra.2020.01.005.
Liu, S., Shuai, P., Zhang, X., Chen, S., Li, L., Liu, M., 2020b. Fine-tuned transformer model for sentiment analysis. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 12275 LNAI, 336–343. https://doi.org/10.1007/978-3-030-55393-7_30.
Lucini, F.R., Tonetto, L.M., Fogliatto, F.S., Anzanello, M.J., 2020. Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. J. Air Transp. Manag. 83, 101760 https://doi.org/10.1016/j.jairtraman.2019.101760.
Ma, T., Zhao, Y.W., Zhou, H., Tian, Y., Al-Dhelaan, A., Al-Rodhaan, M., 2019. Natural disaster topic extraction in Sina microblogging based on graph analysis. Expert Syst. Appl. 115, 346–355. https://doi.org/10.1016/j.eswa.2018.08.010.
Malandrino, A., Demichelis, E., 2020. Conflict in decision making and variation in public administration outcomes in Italy during the COVID-19 crisis. Eur. Policy Anal. 6, 138–146. https://doi.org/10.1002/epa2.1093.
Mogaji, E., Adekunle, I., Aririguzoh, S., Oginni, A., 2022. Dealing with impact of COVID-19 on transportation in a developing country: insights and policy recommendations. Transp. Policy 116, 304–314. https://doi.org/10.1016/j.tranpol.2021.12.002.
Monmousseau, P., Marzuoli, A., Feron, E., Delahaye, D., 2020. Impact of COVID-19 on passengers and airlines from passenger measurements: managing customer satisfaction while putting the US Air Transportation System to sleep. Transp. Res. Interdiscip. Perspect. 7, 100179 https://doi.org/10.1016/j.trip.2020.100179.
Nguyen, K.M., 2021. The role of social media in policy formulation improvement in California. Master's Projects. 981. 10.31979/etd.jt87-cb9v.
Ogie, R.I., Clarke, R.J., Forehead, H., Perez, P., 2019. Crowdsourced social media data for disaster management: lessons from the PetaJakarta.org project. Comput. Environ. Urban Syst. 73, 108–117. https://doi.org/10.1016/j.compenvurbsys.2018.09.002.
Pathik, N., Shukla, P., 2020. Simulated annealing based algorithm for tuning LDA hyper parameters. In: Pant, M., Kumar Sharma, T., Arya, R., Sahana, B.C., Zolfagharinia, H. (Eds.), Soft Computing: Theories and Applications. Springer Singapore, pp. 515–521. https://doi.org/10.1007/978-981-15-4032-5_47.
Pullano, G., Valdano, E., Scarpa, N., Rubrichi, S., Colizza, V., 2020. Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: a population-based study. Lancet Digit. Heal. 2, e638–e649. https://doi.org/10.1016/S2589-7500(20)30243-0.
Qiang, J., Qian, Z., Li, Y., Yuan, Y., Wu, X., 2020. Short text topic modeling techniques, applications, and performance: A survey. IEEE Trans. Knowl. Data Eng. 14, 1–1. 10.1109/tkde.2020.2992485.
Qiu, X.P., Sun, T.X., Xu, Y.G., Shao, Y.F., Dai, N., Huang, X.J., 2020. Pre-trained models for natural language processing: a survey. Sci. China Technol. Sci. 63, 1872–1897. https://doi.org/10.1007/s11431-020-1647-3.
Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., Liu, P.J., 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. ArXiv, abs/1910.10683. 10.48550/arXiv.1910.10683.
Rashidi, T.H., Abbasi, A., Maghrebi, M., Hasan, S., Waller, T.S., 2017. Exploring the capacity of social media data for modelling travel behaviour: opportunities and challenges. Transp. Res. Part C Emerg. Technol. 75, 197–211. https://doi.org/10.1016/j.trc.2016.12.008.
Röder, M., Both, A., Hinneburg, A., 2015. Exploring the space of topic coherence measures. WSDM 2015 - Proc. 8th ACM Int. Conf. Web Search Data Min. 399–408. 10.1145/2684822.2685324.
Rodrigues, D.O., Boukerche, A., Silva, T.H., Loureiro, A.A.F., Villas, L.A., 2018. Combining taxi and social media data to explore urban mobility issues. Comput. Commun. 132, 111–125. https://doi.org/10.1016/j.comcom.2018.10.004.
Rogers, J., 2019. The use of social media and its impact for research. BioResources 14, 5022–5024. https://doi.org/10.15376/biores.14.3.5022-5024.
Simon, H., 2014. Administrative behavior. Free Press, New York.
Singh, J.P., Dwivedi, Y.K., Rana, N.P., Kumar, A., Kapoor, K.K., 2019. Event classification and location prediction from tweets during disasters. Ann. Oper. Res. 283, 737–757. https://doi.org/10.1007/s10479-017-2522-3.
Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C.D., Ng, A., Potts, C., 2013. Recursive deep models for semantic compositionality over a sentiment treebank, in: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, Seattle, Washington, USA, pp. 1631–1642.
Sun, L., Yin, Y., 2017. Discovering themes and trends in transportation research using topic modeling. Transp. Res. Part C Emerg. Technol. 77, 49–66. https://doi.org/10.1016/j.trc.2017.01.013.
Teixeira, J.F., Lopes, M., 2020. The link between bike sharing and subway use during the COVID-19 pandemic: the case-study of New York's Citi Bike. Transp. Res. Interdiscip. Perspect. 6, 100166 https://doi.org/10.1016/j.trip.2020.100166.
Tian, H., Gao, C., Xiao, X., Liu, H., He, B., Wu, H., Wang, H., Wu, F., 2020. SKEP: Sentiment knowledge enhanced pre-training for sentiment analysis. abs/2005.05635, 4067–4076. 10.18653/v1/2020.acl-main.374.
Vega-Carrasco, M., O'sullivan, J., Prior, R., Manolopoulou, I., Musolesi, M., 2020. Modelling grocery retail topic distributions: Evaluation, Interpretability and Stability. ArXiv, abs/2005.10125. 10.48550/arXiv.2005.10125.
Wang, Y., Li, H., Lin, C., 2019. Modeling sentiment evolution for social incidents, in: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM '19. Association for Computing Machinery, New York, NY, USA, pp. 2413–2416. 10.1145/3357384.3358136.
Wang, J., Fan, Y., Palacios, J., Chai, Y., Guetta-Jeanrenaud, N., Obradovich, N., Zhou, C., Zheng, S., 2022. Global evidence of expressed sentiment alterations during the COVID-19 pandemic. Nat. Hum. Behav. 6, 349–358. https://doi.org/10.1038/s41562-022-01312-y.
Wuhan Bureau of Statistics, 2020. Wuhan Statistical Yearbook [WWW Document]. URL http://tjj.hubei.gov.cn/tjsj/sjkscx/tjnj/gsztj/whs/202102/P020210202610060079335.pdf (accessed 3.30.2023).
Yang, S., Ning, L., Jiang, T., He, Y., 2021. Dynamic impacts of COVID-19 pandemic on the regional express logistics: evidence from China. Transp. Policy 111, 111–124. https://doi.org/10.1016/j.tranpol.2021.07.012.
Yue, C., 2017. Entropy-based weights on decision makers in group decision-making setting with hybrid preference representations. Appl. Soft Comput. J. 60, 737–749. https://doi.org/10.1016/j.asoc.2017.07.033.
Zhan, X., Ukkusuri, S.V., Zhu, F., 2014. Inferring urban land use using large-scale social media check-in data. Networks Spat. Econ. 14, 647–667. https://doi.org/10.1007/s11067-014-9264-4.
Zhang, J., 2020. Transport policymaking that accounts for COVID-19 and future public health threats: a PASS approach. Transp. Policy 99, 405–418. https://doi.org/10.1016/j.tranpol.2020.09.009.
Zhang, J., Hayashi, Y., Frank, L.D., 2021. COVID-19 and transport: findings from a world-wide expert survey. Transp. Policy 103, 68–85. https://doi.org/10.1016/j.tranpol.2021.01.011.
Zhang, Z., He, Q., Gao, J., Ni, M., 2018. A deep learning approach for detecting traffic accidents from social media data. Transp. Res. Part C Emerg. Technol. 86, 580–596. https://doi.org/10.1016/j.trc.2017.11.027.
Zhao, W., Chen, J.J., Perkins, R., Liu, Z., Ge, W., Ding, Y., Zou, W., 2015. A heuristic approach to determine an appropriate number of topics in topic modeling. BMC Bioinformatics 16, S8. https://doi.org/10.1186/1471-2105-16-S13-S8.
Zhu, P., Tan, X., 2021. Is compulsory home quarantine less effective than centralized quarantine in controlling the COVID-19 outbreak? Evidence from Hong Kong. Sustain. Cities Soc. 74, 103222 https://doi.org/10.1016/j.scs.2021.103222.
Zipf, G.K., 1949. Human behavior and the principle of least effort. Addison-Wesley Press, Oxford, England.