Travel Behaviour and Society 43 (2026) 101197

Contents lists available at ScienceDirect

Travel Behaviour and Society

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

Extracting socio-psychological perceptions for analysis of travel behaviours

Yanyan Xu *, Panchamy Krishnakumari, Neil Yorke-Smith, Serge Hoogendoorn

Delft University of Technology, Delft, 2628CD, The Netherlands

ARTICLE INFO

Keywords:
Psychological determinants
Travel decision
Social media data
Dynamic topic modelling
Sustainable transport

ABSTRACT

This article proposes an evidence-based policy recommendation framework integrating social media data and natural language processing methods, to support inclusive and efficient transport policy-making. Given that current research underscores the crucial role of both external and psychological variables in individual travel decisions, psychological features – such as beliefs, attitudes or values – are frequently used as latent variables for travel behaviour interpretation and travel choice modelling. However, user-centric policy recommendations based on dynamic psychological variables are still limited. Most studies rely on survey data, which neglects the urgent dynamic trend of user perception change and its underlying relationship with travel behaviour. Hence there is a lack of illustration on how these psychological variables can be further used at specific temporal and spatial levels for travel behaviour interpretation. This would be valuable to identify priorities for more targeted (sustainability and other) policies and interventions. In this article, we utilize sentiment analysis and dynamic topic modelling to represent the spatial–temporal variance of psychological features. Integrating with corresponding travel behaviour, we illustrate how these dynamic psychological features can distinguish travel dissonance, identify key motivations, and reflect urgent social demands at precise spatial–temporal levels. We demonstrate these advances in a case study in New York City from 2019 to 2022 using Twitter (X) data. A comparison with existing travel-related policies in the case study validates the feasibility of our framework to support evidence-based policy recommendations. We conclude by discussing the potential of this framework to support sustainable transport promotion.

1. Introduction

Although numerous studies demonstrate the vital impact of socio-demographic, economic, policy and environmental factors on individual travel choices (An et al., 2023; Piligrimienė et al., 2020; Biehl et al., 2018; Zhao, 2018), the significant influence of psychological variables – such as attitude, satisfaction, beliefs, values, risk perception or environmental awareness – cannot be ignored (Rahman and Sciara, 2022; Dütschke et al., 2022). In many cases, the influence of external determinants on individual travel decisions varies significantly depending on personal psychological reactions (Susilo and Cats, 2014; Abenoza et al., 2017; Kroesen and Chorus, 2020; Arroyo et al., 2020). For instance, the success of implementing sustainable travel-related policies largely relies on personal attitudes or acceptance, particularly regarding one's motivation to shift to sustainable transport.

Existing studies use psychological variables as *latent variables* (Kroesen et al., 2017; Arroyo et al., 2020), mainly focusing on the quantitative aspects, such as 'positive' or 'negative' as attitudes, 'like' or 'dislike' as preferences, or satisfaction in numerical levels for travel choice modelling (Pineda-Jaramillo and Pineda-Jaramillo, 2022; Huan et al., 2022). However, few studies explore the intrinsic motivations behind these psychological features to enhance travel behaviour interpretation, social needs recognition (Anagnostopoulou et al., 2020), and evidence-based policy recommendation (Nikolić et al., 2021).

Simultaneously, these personal psychological variables are shaped by external circumstances, reflecting the underlying contextual determinants in travel decisions. For example, personal preference for public transport mainly determined by commuting efficiency, economic affordability, and user satisfaction (Liu et al., 2023). The willingness to adopt electric vehicles relies on the personal assessment of price, functionality, and eco-consciousness (Chinen et al., 2023). Therefore, a comprehensive understanding of the interaction between psychological variables and external contextual factors will reveal the key motivations and social demands behind travel decisions.

In particular, special periods such as COVID-19 have amplified the psychological impact on travel behaviour, bringing significant changes across multiple travel modes (Shamshiripour et al., 2020). Some of these changes even persisted after the pandemic, indicating the lasting

* Corresponding author.
E-mail addresses: Y.Xu-10@tudelft.nl (Y. Xu), p.krishnakumari@tudelft.nl (P. Krishnakumari), n.yorke-smith@tudelft.nl (N. Yorke-Smith), s.hoogendoorn@tudelft.nl (S. Hoogendoorn).

https://doi.org/10.1016/j.tbs.2025.101197
Received 29 November 2024; Received in revised form 16 October 2025; Accepted 19 November 2025
Available online 8 December 2025

psychological impact on travel habit reformulation. For instance, the sustained decline in subway usage in New York City was initially driven by heightened risk perception, which subsequently led to a weakening of trust in public transportation (Jain et al., 2022). By analysing the underlying drivers of pandemic-induced travel behaviour changes, this study aims to shed light on the interaction between dynamic behaviour and psychological variables in affecting dynamic travel decisions and demands, thus providing insights for policymakers and transport planners.

We recognize that key motivations are inherently dynamic, shaped by spatial and temporal variations in travel behaviour and psychological factors. It highlights the challenge of identifying timely and precise determinants to inform policy-makers of dynamic social needs and targeted interventions. Most existing studies rely on survey data to capture psychological variables and evaluate their impact on travel choices, given that surveys can provide information on user travel preferences and demands across diverse demographic groups, support controlled experimental designs, and contain relatively good data quality. There are limitations, however, in capturing the up-to-date psychological changes along with underlying motivations, which hinder the ability to identify timely and accurate social demands for informed policy recommendations. Even if some surveys are updated monthly or annually ((MPN), etc.), their data collection processes are often time-consuming to reflect real-time psychological shifting, particularly during urgent circumstances. In addition, surveys are typically conducted at relatively broad spatial scale, which may overlook the spatial variability and fine-grained psychological dynamics associated with localized social concerns and motivations. Meanwhile, survey questions are often formulated based on subjective perceptions of their designers, which may constrain the capacity to reflect all determinants in travel decisions. Furthermore, the survey datancollection process often costly and resource-intensive, which limits it efficiency for timely data acquisition.

Meanwhile, travel behaviour is not always consistent with psychological indicators in the real world, which refer to *travel dissonance* (De Vos, 2018). For instance, some low-income individuals might be compelled to use public transport due to economic constrains, whereas others might be obliged to drive because of the inadequate public transport infrastructure in their residential areas. According to Cognitive Dissonance Theory (CDT), individuals are likely to experience discomfort when encountering travel-related dissonance. As a result, there is an increased likelihood of individuals altering their travel behaviour or attitudes to fill the gap (Festinger, 1957). Recent studies utilize travel dissonance to interpret travel behaviour changes (De Vos, 2018), while fewer studies delve into the implicit reasons underlying 'behaviour-psychological' gaps which indeed revealing user concerns and urgent social demands. Understanding these drivers will enable policy-makers to implement more targeted interventions, enhance user travel experiences, and promote the adoption of sustainable transport.

In light of this situation, social media data and natural language processing (NLP) technologies offer opportunities to capture dynamic psychological variables along with their underlying determinants. As online platforms (Twitter/X, Facebook, Instagram, etc.) become common places to express public feelings, social media has emerged as a valuable resource for accessing real-time public opinions on a wide range of phenomena (Grant-Muller et al., 2015). Meanwhile, the embedded geo-information within social media data enables the analysis of spatial patterns in user-generated content. Integrated with NLP methods, social media data has been widely used to reflect public emotions and opinions via sentiment analysis and topic modelling (Qi et al., 2020; Schweitzer, 2014). Despite inherent biases – such as the tendency to express negative emotions or the unequal representation of populations – social media data still provides valuable insights of real-time psychological information. During COVID-19, it offered timely and cost-effective updates on public perceptions of travel policies and events. Leveraging these advantages, this study utilizes social media data integrated with NLP and recent Large Language Model (LLM) methods to capture dynamic travel-related psychological variables with underlying travel motivations.

This article aims to propose an evidence-based policy recommendation framework via social media data integrated with NLP and LLM methods. Demonstrating the critical role of psychological variables in identifying key motivations and recognizing social demands, we provide efficient and targeted policy-making support from a user perspective. Through sentiment analysis and dynamic topic modelling, we demonstrate dynamic user attitudes and key opinions in response to external contextual factors across multiple travel modes. By analysing the relationship between psychological variables and corresponding travel behaviour, we identify key determinants that imply dominant public concerns and social demands across spatial–temporal variations. We also point out *travel dissonance* at neighbourhood scale with related psychological concerns to indicate urgent social needs and improvement directions for policy-makers. The case study spans the pandemic period, investigating dynamic psychological variables to reveal key travel motivations associated social demands for cycling, subway and ride services in New York City from 2019 to 2022. Our contributions lie in the following aspects:

- We propose a general evidence-based policy recommendation framework at various spatial–temporal levels integrating social media data and NLP methods from a user-centric perspective.
- We demonstrate the extensive use of psychological variables in identifying travel motivations, travel dissonance, and social demands at various spatial–temporal levels.
- Our case study implements the framework spanning the pandemic period to illustrate the significant role of psychological variables in interpreting dynamic travel behaviour and identifying social demands among multiple travel modes in New York City (NYC). We also validate the results with prevalent policy interventions, indicating the potential use in other scenarios.

The remainder of the article is structured as follows: Section 2 reviews existing studies on evidence-based policy recommendation, the impact of psychological variables and external factors on travel decisions, and the use of social media data and natural language processing methods in transportation domain. Section 3 introduces the framework and methods, including data collection, sentiment analysis, dynamic topic modelling, and correlation analysis. Section 4 describes our case study in setting up psychological variables extraction, travel dissonance identification, and key motivations recognition in New York City from 2019 to 2022. Section 5 reports the outcomes with an explanation of the underlying reasons. Section 6 provides potential policy suggestions based on our findings, along with validation on existing policies and transport studies. Section 7 summarizes our research approach and discusses the potential use of the proposed framework in informing sustainable transport policy recommendations.

2. Background

This section briefly discusses evidence-based policy research in transportation, the external and psychological influences on travel decisions, and the current use of social media data in transport applications. We conclude the section by identifying the research gaps in exploring evidence-based policy recommendations from the user's perspective, particularly in terms of understanding user dynamic motivations and social demands at various spatial and temporal levels.

2.1. Evidence-based policy recommendation

The evidence-based policy recommendation in transportation started in the 1960s when researchers used empirical data to assess the effectiveness of various transportation systems (Meyer et al., 1965). The work highlights the feasibility of leveraging relevant data as evidence to improve and evaluate transportation systems. As data collection and analysis methods gradually improved, evidence-based approaches in transportation have become more formalized and efficient. Bones et al. (2013) examines the implementation of 'evidence-based design' (EBD) in transportation decision-making to mitigate innovation risk and support reliable policies development. Mulley and Reedy (2015) discusses the effective transmission from transport research to evidence-based public transport policies using New South Wales(NSW), Australia as a case study. Smith-Colin et al. (2014) demonstrates how various types of evidence from transportation asset management can support performance-based decision-making.

The emerging big data and artificial intelligence methods enable more comprehensive and robust evidence-based policy recommendation. Rodríguez-Mazahua et al. (2016) underscores the critical role of data-driven analysis in enhancing Intelligent Transportation Systems(ITS) performance and enabling more precise, evidenced-based policy interventions. Hackl et al. (2019) proposes a data-driven approach using survey data and statistical models to elucidate walking and cycling modal shares. The findings demonstrate the influence of key parameters in shaping travel behaviour and indicate the target spatial areas for specific policy interventions. Liu and Dijk (2022) illustrates how data strengthen evidence-based transport policy-making via two case studies in Maastricht and Groningen, which highlight the important role of 'data' in supporting short-term traffic regulations and policy adjustments. Meanwhile, evidence-based approaches are also essential for achieving sustainable transport goals – such as reducing carbon emissions, promoting public transport, and encouraging electric vehicles – by informing efficient and targeted interventions (Gota et al., 2019). Although these evidence-based approaches have shown their effectiveness in the transportation domain, few take advantage of big data from a user's perspective to provide in-time evidence-based recommendations.

2.2. External factors influence on travel decisions

Existing research demonstrates the impact of external contextual factors on individual travel decisions via static models and machine learning methods. Socio-demographic factors such as economic, education, gender, age, and culture have been shown to play vital roles in shaping travel mode choices (Shaer et al., 2021; Kim and Mokhtarian, 2018; He et al., 2018; Pisoni et al., 2022). Transport infrastructure, including its quality and accessibility, constitutes another critical variable in the travel decision-making process (Zhou et al., 2024; Yin et al., 2024). Yin et al. (2024), Bi et al. (2023) indicate that environmental factors, including weather, climate, and air quality, largely affect travel choices for seasonally sensitive modes such as cycling. Policy interventions also affect travel decisions. For instance, travel restrictions during the pandemic substantially reduced public transport usage (Chen et al., 2022), whereas incentives for electric vehicles have facilitated the adoption of sustainable vehicles (Wang et al., 2018). While these studies thoroughly analyse the direct effects of external contextual variables on travel decisions, it is equally important to investigate users' psychological reactions to these contextual factors and their further influence on travel decisions.

2.3. Psychological variables impact on travel decisions

There are multiple psychological factors, such as attitudes, values, beliefs, perceived behavioural control, and social norms, that have been clearly demonstrated their impact on travel decisions by theories such as reasoned action (TRA) (Fishbein, 1979) and theory of planned behaviour (TPB) (Ajzen, 1991). 'Attitudes' refers to the individual's positive or negative evaluation towards an object, events or behaviour. As one of the key psychological variables, attitude has been widely utilized to quantify the individual's sentiment towards specific travel modes. Early studies (Parkany et al., 2004; Beirão and Cabral, 2007; Scheiner and Holz-Rau, 2007) depict the correlation and causation between attitudes and travel decisions which demonstrate the feasibility of implementing 'attitude' as a latent variable in travel choice modelling by statistical and machine learning models (Bamberg et al., 2003; Beirão and Cabral, 2007; Pronello and Camusso, 2011; Molin et al., 2016; Moon, 2021). 'Social Norm' refers to the informal shared rules or expectations within social groups to guide or constrain behaviours (Burke and Young, 2011). 'Perceived behavioural control' refers to the individual's assessment of their ability to perform a behaviour (Ajzen, 1991). 'Values' refers to the long-term guiding principles which shaping people's behaviour that guide both attitudes and behaviour, while 'beliefs' refers to the cognitive convictions on the consequences of performing a behaviour (Ajzen, 2012). Compared to attitudes, beliefs and values tend to contain more stable and enduring impact on travel choices because they tend to deeply rooted in individual's culture which making them more resistant to change compared to other psychological variables such as attitudes (Lind et al., 2015; Paulssen et al., 2014; Arroyo et al., 2020). These studies underscore the importance of different psychological factors in shaping travel decisions, but they typically rely on static, survey-based data, which limits the ability to capture real-time changes with underlying motivations that drive dynamic travel behaviour.

Existing studies also demonstrate how psychological variables interact with external factors, thus influencing travel choices (Vu et al., 2019; Hasan and Ukkusuri, 2014; Bi et al., 2023). For example, although government incentive policies advocate sustainable transport, individuals' acceptance is largely influenced by their personal beliefs, opinions, and attitudes. Studies (Ru et al., 2018; Taube et al., 2018; Hamidi and Zhao, 2020) illustrate the impact of environmental awareness in driving individual mode shifting from private vehicles to green travel options. Wang et al. (2021) investigates the psychological impact on public bicycle system adoption and indicate that risk concern, accessibility, usage complexity, and observability are the key determinants of users' decision to shift into a sustainable public cycling mode. Chen (2016) demonstrates that positive attitude, perceived pleasure, and subjective norms are the key factors for personal preference in using sustainable transport. These studies underscore the interaction between external factors and psychological variables, highlighting their integrated role in explaining travel behaviour and discovering underlying motivations. Accordingly, this study emphasizes this interaction relationship to provide more robust and reliable evidence-based recommendations.

Some studies also point out the inconsistencies between psychological factors and actual travel choices, particularly the travel dissonance among individuals' attitudes and actual behaviour (Kroesen et al., 2017). Festinger (1957) asserts individuals are likely to experience discomfort when confronting travel dissonance, thereby enhancing the potential to adjust either attitude or behaviour to fill the gap. Consequently, some research implements this phenomenon to interpret or forecast travel behaviour change (De Vos and Singleton, 2020; An et al., 2022). Some studies compare travel dissonance among various travel modes to explore user satisfaction in a general approach. Humagain et al. (2021) states that active mode users have less satisfaction than private vehicle users when they experience travel choice dissonance. De Vos (2018) states that dissonant travellers have lower average satisfaction than consonant travellers in public transport. However, few studies investigate the underlying causes of these inconsistencies which shed light on urgent social demands and targeted intervention to improve travel experiences. Therefore, this study utilizes dynamic psychological variables to identify specific travel dissonances and analyse their determinants, thereby highlighting the particular social needs to enhance evidence-based recommendations.

2.4. Social media data and NLP in transportation

To better achieve this study approach in dynamic travel behaviour interpretation from the user perspective, and support evidence-based policy recommendation, we employ social media data and NLP techniques, which provide the feasibility to reflect real-time psychological variables and the underlying travel motivations.

Compared to conventional traditional surveys and questionnaires, social media platforms provide a time-efficient way to acquire real-time user-related information through contextual text, images, and geo-information (Liu et al., 2020; Chen et al., 2018; Dwityas and Briandana, 2017). Many studies demonstrated that social media tends to reflect near real-time user reactions, particularly in response to policies and events due to its inherent characteristic: rapid information sharing and ease of expression. These platforms are designed to prioritize breaking news, trending topics and real-time updates on policy announcements and events at near real-time speed compared to traditional media. Consequently, users are more easily exposed to new information when it occurs and are encouraged to interact immediately with emotional and attitudinal responses (Stieglitz and Dang-Xuan, 2013; Jost et al., 2018). Lorenz-Spreen et al. (2019) states that public opinions on social media can shift dramatically within hours following policy announcement or social events. Panagiotopoulos et al. (2016) demonstrates that social media enables rapid information sharing during emergencies, shaping public attitudes quickly based on real-time updates from authorities and media. Chen and Chen (2019) shows that daily attitudes on Twitter fluctuate sharply around major news events (e.g., natural disasters, political announcements). Yeung (2018) states that social media is ‘rapidly refreshed and constantly changing’, which enables the tracking of attitudes and behaviours in near real-time. Cinelli et al. (2020) demonstrates how COVID-19 news and government measures produced immediate sentiment shifts on Twitter within hours of announcement.

In transportation, social media has emerged as an important resource for user demand acquisition and system management evaluation. Integrated with context-mining and NLP methods (Ding et al., 2008; Blei et al., 2003; Grootendorst, 2022), recent researchers are able to acquire the latest psychological variables such as attitudes, concerns and demands (Park et al., 2022; Mo et al., 2023). For instance, Rahman et al. (2021), Borowski et al. (2020) demonstrate the ability of Twitter(X) in reflecting dynamic user attitudes towards active mobility via sentiment analysis. Rashidi et al. (2017), Vu et al. (2019), Shahriari et al. (2024) illustrate the integration of social media data and topic modelling in representing user concerns for travel behaviour interpretation. The recent emerging Large Language Models(LLMs) such as Llama (Touvron et al., 2023) and GPT-4 (Achiam et al., 2023) enhance the efficiency and scalability in representing elusive and dynamic travel-related psychological aspects by enabling fine-tuning on customized datasets (Yang et al., 2023; Barandoni et al., 2024). It provides a broader potential in formulating evidence-based recommendations, enabling policymakers and transport managers to adopt user-centric and targeted decision-making approaches.

According to these related works, we summarize the research gaps and the connections with our study as follows:

- Few studies leverage psychological determinants as feasible evidence for transportation policy recommendations. To address this gap, we propose a user-centred, evidence-based framework that integrates dynamic psychological insights to support targeted policy recommendations.
- Few studies investigate the underlying determinants of spatio–temporal travel-behaviour variations to clarify public travel demands and inform timely policy-making. In our study, we analyse how psychological variables interact with external contextual factors, revealing implicit reasons of dynamic travel-behaviour change and highlighting urgent social demands.
- Most studies rely on survey data that lacks real-time track of psychological trends, limiting their ability to explain rapid behavioural shifts during particular periods such as the pandemic. We take advantage of social media data with NLP technologies to track dynamic travel-related attitudes and opinions, interpret travel behaviour change throughout the COVID-19 period, thereby supporting targeted interventions.

3. Methodology

This study proposes an evidence-based policy recommendation framework integrating social media data with NLP methods to provide reliable and user-centric suggestions for policy-makers. This section introduces our methods in identifying travel motivations, user demands, and travel dissonances with underlying reasons at the variant spatial–temporal scales. Our methodological framework is shown in Fig. 1. We employ description analysis to represent dynamic travel behaviour via mobility data. Based on social media data (Twitter/X), we illustrate how to capture travel-related attitudes and user concerns via sentiment analysis and dynamic topic modelling. Integrating with travel patterns with psychological variables, we demonstrate travel dissonance recognition, key motivations and social demands identification, travel dissonance recognition with interpretation, thus providing reliable evidence for policy-making.

3.1. Descriptive analysis

We first implement *description analysis* to represent the dynamic travel behaviour based on historical ridership data. By aggregating historical daily ridership data, we display similarities and differences among at both spatial and temporal scales. We also compute the change ratio within specific time intervals at Eq. (1) to present dynamic travel patterns. To gain a deeper understanding of dynamic travel behaviour, we employ the origin–destination (OD) analysis at Eq. (2) to observe the trip patterns and purposes change.

$$C_{t,m} = \frac{P_{m,t} - P_{m,t-l}}{P_{m,t-l}}, \tag{1}$$

where t is the current time, l is the time interval, $P_{m,t}$ is the travel ridership at time t for travel mode m, $P_{m,t-l}$ is the travel ridership at time $t-l$ for travel mode m; $C_{t,m}$ refers to the change ratio for travel mode m from time $t-l$ to time t.

$$OD_{ij} = \frac{1}{T}\sum_{t=1}^{T} f(i,j,t), \tag{2}$$

where i is the origin, j is the destination, t is the travel time for each single trip, T represents the total time periods. OD_{ij} represents the daily average number of trips from location i to location j, $f(i,j,t)$ represents the trip flow from i to j during time interval t.

3.2. Sentiment analysis

We first illustrate the linguistic differences among each psychological variable to demonstrate their distinction in social media. For ‘attitude’, adjectives (e.g., great, terrible), and sentiment verbs (e.g., like, hate) reflect individuals’ opinions on certain travel modes that can be directly extracted from the posts. For social norms, the typical keywords emphasize others’ views (e.g., ‘people expect’, ‘my friends think’) rather than attitude keywords that focus on personal evaluation. In transportation-related posts, they display as ‘Most people take the car’, or ‘Everyone’s using bikes now’. Perceived behavioural control is normally identified by indirect cues that emphasize capability and feasibility, such as ‘It is impossible to bike in this city’, which indicates perceived low control, while ‘The metro is so easy to use’ suggests perceived high control. It is more related to the individual’s confidence in their own ability or external constraints to perform a behaviour,

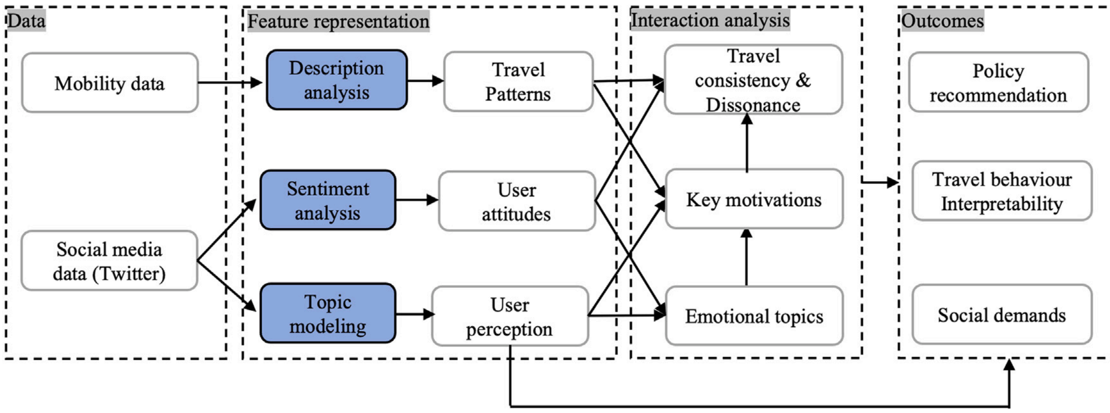

Fig. 1. Methodological framework.

rather than positive or negative evaluation of attitudes. Thus, the distinguishing key words for perceived behavioural control are different from attitude, which emphasizes 'can or cannot', 'possible or impossible'. For identifying values and beliefs, they are indirectly inferred from the posts, which require interpreting full sentences to detect whether they contain guiding principles, making it less commonly observed in short texts. For example, 'I value sustainability over convenience'.

As user attitudes tend to exhibit more real-time or short-term changes in response to the latest policies and events, they contain more content that reflects positive or negative sentiment towards specific issues or behaviours. We thus implement Natural Language Processing methods to distinguish these quantitative preferences. To acquire dynamic travel-related attitudes, we implement the BERT model to classify user sentiment based on tweets into two categories: positive and negative. To improve the classification performance on our dataset which contains the special pandemic period, we first fine-tune the BERT model with domain-specific data – COVID-19-related sentiments labelled tweets (Chen et al., 2020). After splitting the prepossessed tweets into training, validation and testing sets, adding an extra classifier, we train the model with a batch size of 32 and max length of 128 based on the 'bert-base-cased' model. The training and validation accuracy achieve 0.92 and 0.84, respectively, while the test accuracy reaches 0.86 with a precision of 0.85, recall of 0.85, and f1-score of 0.85. Finally, we test the fine-tuned BERT model on our own dataset with 600 rows of manually labelled tweets. It achieves an accuracy of 0.87, demonstrating the relatively good performance. Ultimately, we deploy the fine-tuned model to our travel-related tweets, with negative or positive labels for each tweet and visualize the results on a daily basis.

3.3. Dynamic topic modelling

To gain the underlying reasons behind dynamic travel behaviour, user attitudes, and travel dissonance, we implement dynamic topic modelling to present the main user psychological interactions with contextual factors. As an unsupervised learning method to extract topic representation from textual data, topic modelling such as Top2vec, Latent Dirichlet Allocation (LDA), and Latent Semantic Analysis (LSA) have been commonly used to represent human opinions (Blei et al., 2003; Angelov, 2020; Landauer et al., 1998). However, these methods require prior knowledge of topic numbers, imitated in interpretability, and do not work well on short texts (Jelodar et al., 2019), which is not suitable in our study based on short texts. While, *Dynamic Topic modelling*, as an unsupervised learning method based on the BERT model, possesses the capability to discern evolving topics and provide more coherent topic interpretation among sparse texts (Grootendorst, 2022) that is well-suited for our study. In particular, we embed each document (i.e., tweet) by the semantic similarity logic via the pre-trained language model Sentence-BERT (SBERT), convert each tweet into numerical representations via 'all-MiniLM-L6-v2' model (Reimers and Gurevych, 2019). Given that cluster models have difficulty in dealing with high-dimensional data, we apply Uniform Manifold Approximation and Projection (UMAP) (McInnes et al., 2018) to reduce high-dimension embedded data from the previous step while preserving both local and global features of the embedded documents. Then, we employ the density-based clustering method — HDBSCAN to aggregate the reduced embedding into distinct clusters with noise data labels as outliers (Campello et al., 2013). Since each cluster has different densities and shapes, we combine all clusters' documents into a single document to calculate word frequency to obtain cluster-level topic representation. To compute the importance scores of each word within a cluster, we use c-TF-IDF (Grootendorst, 2022) ((3)) to extract the most represented words. After capturing the overall topic within each cluster, we analyse the dynamic trends of each main topic temporally and spatially.

$$W_{t,c} = tf_{t,c} * log(1 + \frac{A}{tf_t}), \tag{3}$$

where: t, c is the term t in class c, in which c is the collection of documents into one single document for each cluster. $W_{t,c}$ is the weight of each term i in class c; $tf_{t,c}$ is the frequency of each term i in class c; A is the average amount of terms per class; f_i is the frequency of term i among all classes.

The latest (at the time of writing) large language models enable us to compute a more coherent topic representation rather than several single words for each cluster. Thus, we use 'ChatGPT-4o' to fine-tune our topic representation results from the generated topics from 'c-TF-IDF'. Our prompts are described in the following via the OpenAI API:

prompt = """" I have a topic that contains the following documents: [DOCUMENTS] The topic is described by the following keywords: [KEYWORDS] Based on the information above, extract a short topic label within ten words in the following format: topic: <topic label> """"

3.4. Travel dissonance identification

To identify the consistent travel behaviour and travel dissonance, we compute the normalized discrepancy between travel-related attitude and travel ridership at certain time periods. Eq. (4) describes the travel dissonance at both temporal and spatial levels, for which positive refers to a more favourable attitude than actual behaviour and negative indicates actual behaviour with unwilling sentiment. The closer the value is to 0, the more alignment between behaviour and attitude has been achieved.

$$D_t = \overline{S}_{t,k} - \overline{B}_{t,k} \tag{4}$$

where t is the time interval, k is the travel mode; $\overline{S}_{t,k}$ represents the normalized attitude at time t for travel mode k; $\overline{B}_{t,k}$ is normalized ridership volume at time t for travel mode k; D_t represents the discrepancy at time interval t.

3.5. Key determinants analysis

To identify which topics have a larger impact on actual travel behaviour, we implement correlation analysis between topic volumes and related travel ridership in Eq. (5). The value is between −1 to 1 for which −1 indicates a strong negative relation, 1 indicates a strong positive relation, and 0 refers to non-obvious correlations.

$$r = \frac{\sum(x_{t,o} - \overline{x}_o)(y_{t,n} - \overline{y}_n)}{\sqrt{\sum(x_{t,o} - \overline{x}_o)^2 \sum(y_{t,n} - \overline{y}_n)^2}} \tag{5}$$

where x is the volume of topic, y is related travel ridership, o is topic, t is the time point, n is the travel mode; $x_{t,o}$ is the topic c volume at time t, $y_{t,k}$ is the travel ridership at time t for travel mode n; $\overline{x}_o$ is the average volume of topic o and $\overline{y}_n$ is the average ridership volumes for travel mode n; r is the correlation coefficient.

The correlation results indicate three main groups: strongly positively correlated topics, strongly negatively correlated topics, and non-correlated topics. Although correlation does not equal causation, we cannot assert that strongly correlated topics determine travel choices. It still reflects the intense psychological motivations and public concerns associated with actual travel behaviour. Additionally, we examine attitudes towards key topics to identify strongly correlated topics with negative sentiment, highlighting certain aspects in urgent need of improvement.

4. Case study: Description

The case study aims to investigate the psychological impact and determinants in travel decisions among various travel modes in New York City from 2019 to 2022, spanning the COVID-19 period. As the main mobility options in New York City, these four travel modes experienced notable behaviour shifts during the pandemic. We first represent their travel pattern and trip purpose changes comparing before, within and after the pandemic. Then, we analyse travel-related attitudes and opinions to acquire public concerns from social media data. Analysing the relationship between psychological variables with related travel behaviour, we demonstrate key motivations and travel dissonance identification with explanation from user psychological perspective. The outcomes provide insights on priority neighbourhoods and periods that require targeted attentions and policy interventions. However, due to the lack of private vehicles data, we only collect and analysis psychological variables of private vehicles.

4.1. Data collection

Transportation-related Twitter data Travel-related social media data has been collected via Twitter (subsequently ‘X’) developer API from 2019 to 2022, which covers ‘pre-COVID-19’, ‘the outbreak of COVID-19’, and ‘post-COVID-19’ periods. To acquire specific mode-relevant ‘tweets’, we use keywords to filter the tweets by four categories: cycling, public transit (subway), ride services (taxi and ride-hailing), and private vehicles within the geo-boundaries of New York City. Notably, ‘cycling’ uses ‘bike’, ‘cycling’ and ‘bicycle’ search query terms; ‘ride services’ use ‘taxi’, ‘cab’, ‘Uber’ and ‘Lyft’ search query terms; ‘subway’ uses ‘subway’ and ‘metro’ search queries terms; ‘private vehicles’ uses ‘car’, ‘driving’, ‘drive’ and ‘private vehicles’ search query terms (summarized in Table 1). Given that travel-related tweets contain a certain amount of professional accounts (e.g., @NYCT Subway, @CitiBikeNYC) that primarily post operational updates or traffic information (e.g., car incidents, subway delays), these posts cannot be used to infer user psychological perceptions. Therefore, we separate the raw tweets into user tweets and agency tweets.

To reflect data transparency, we analyse the collected data characteristics, including geographic, temporal, and social features. Fig. 2 presents the geographic distribution and temporal scope of the sample data across the five boroughs of NYC with four different travel modes. Overall, the subway-related samples account for 34.2% of the total data, cycling-related samples account for 8.2%, ride services-related samples account for 12.4%, and the rest are the private vehicle-related samples. Subway-related samples are concentrated in Manhattan along transit-service areas with intensity decreases since 2020. Cycling-related samples also presents infrastructure-centred distribution in Manhattan and Brooklyn, with a noticeable increasing after 2020. Ride services and private vehicles related samples are relatively spatially sparse with high-demand areas in outer boroughs such as Queens and Brooklyn, reflecting the more reliance on cars in these areas.

Mobility data. We collect travel mobility data from NYC open data portals from 2019 to 2022 to represent dynamic travel behaviour among different travel modes. In particular, subway data is provided by NYC MTA data portal (MTA) which contains daily entrances and exits; Cycling data, available from Citi-bike official website (Citi-bike) that includes daily shared-cycling ridership with origins and destinations; Ride services data, accessible from Yellow taxi data portal (TLC) and For-Hire vehicle data portal (FHV), contains daily ride service ridership with origins and destinations. Due to the absence of a private vehicle dataset in New York City, we do not analyse its travel pattern, but still explore its related psychological variables using social media data.

4.2. Travel behaviour representation

We describe the dynamic temporal-spatial travel behaviour by aggregating daily ridership and calculating change ratios for subway, cycling, and ride services from January 2019 to December 2022. We aim to gain deeper insight into the potential relationships among different travel modes by comparing similarities and differences in ridership trends. Using geospatial origin and destination data, we analyse the OD patterns of cycling and ride services to understand variations in travel route preferences and trip purposes. By measuring annual OD similarity within Neighbourhood Tabulation Areas (NTAs), we demonstrate the evolution of high-frequency travel patterns and compare them with subway ridership to infer shifts in trip purposes.

4.3. Sentiment analysis and travel dissonance

To acquire dynamic travel-related attitudes towards different travel modes, we implement sentiment analysis based on prepossessed user tweets from January 2019 to December 2022. Aggregating on a daily basis, we analyse the dynamic trend of travel-related sentiment score between −1 to 1, the social engagement (tweet volumes), and related ridership to present potential relations, thereby identifying the psychological impact on travel decisions.

Meanwhile, since travel decisions are not always consistent with psychological indicators, we integrate sentiment analysis with travel behaviour representation to identify travel dissonance by computing the standardized differential between attitude and actual travel behaviour. Temporally, we use the moving average to plot the most disparities in time series for both short-term (daily) and long-term (monthly). Spatially, we calculate average ridership volumes on the neighbourhood level and set the threshold as 0.5 to extract the most disparate areas, emphasizing the priority areas for policymakers to pay attention to.

Table 1
Summary of travel-related tweets collection, in the case study.

Travel modes	keywords	Raw tweets	User tweets	Agency tweets
Cycling	bike, cycling, bicycle	69,906	35,788	34,106
Subway	subway, metro, public transit	160,735	149,096	116,21
Ride services	taxi, cab, uber, lyft	60,463	54,031	6428
Private vehicles	car, driving, drive, private vehicles	206,798	196,844	9927

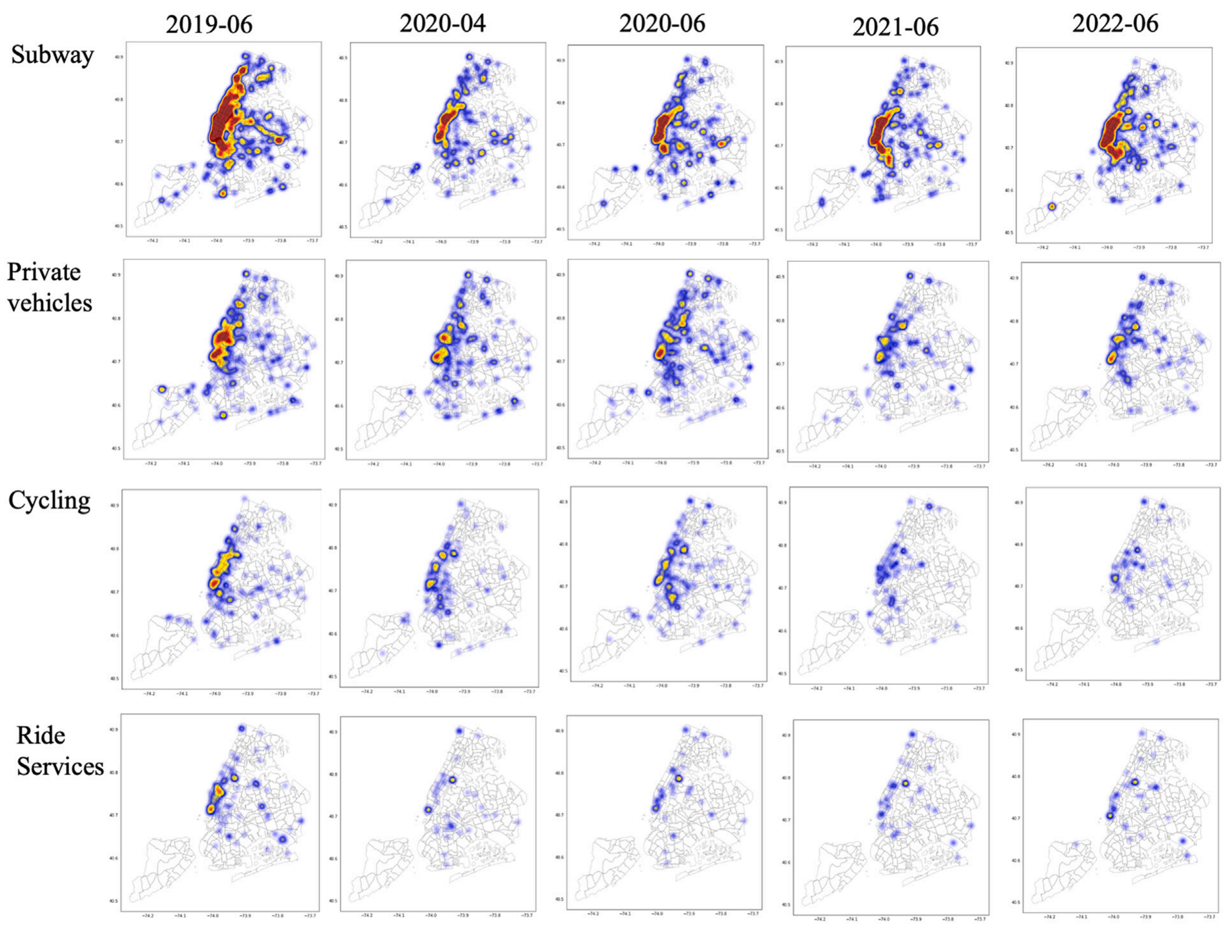

Fig. 2. Social media sample characteristics.

4.4. Key determinants analysis

To understand the underlying motivations behind dynamic travel behaviour, travel-related attitudes, and travel dissonance, we implement *dynamic topic modelling* via BERTopic and LLMs to acquire both primary topics and their spatial–temporal trends. First, we compute the overall topics and their frequency keywords from January 2019 to December 2022 to obtain a general view of the psychological reflections on the external contextual factors. Dynamic topic modelling elucidates temporal and spatial trends in key topics, providing valuable insights for policy-makers, highlighting areas and periods that require improvement. Furthermore, integrating topic modelling with travel dissonance implies the specific psychological concerns in precise areas and periods that enable policy-makers to implement more effective policies, strategies, and interventions. It also helps policy-makers understand the important psychological factors in promoting sustainable transportation.

We then conduct correlation analysis between key topics and corresponding travel patterns to identify the key determinants. This approach enables us to identify which topics are positively or negatively correlated with actual travel behaviour. We further examine attitude trends with highly correlated topics to distinguish public satisfaction.

5. Case study: Empirical results

This section demonstrates findings on the use of travel-related psychological variables in interpreting dynamic travel behaviour, highlighting public concerns and demands across different travel modes in the case study. Section 5.1 presents dynamic travel patterns throughout the pandemic in New York City, emphasizing the necessity of understanding long-term shifts in mobility behaviour from the user's perspective. Section 5.2 illustrates dynamic travel attitudes over the COVID-19 period and their utilization in travel dissonance identification. Section 5.3 reveals spatial–temporal variations of travel-related focus, demonstrates its use in key travel motivations recognition, travel behaviour and travel dissonance interpretation. Section 5.4 discusses and summarizes how these psychological factors interpret dynamic travel patterns change, impact travel behaviour and highlight urgent mobility requirements from a user-centred perspective.

5.1. Dynamic travel behaviour

Fig. 3 presents the temporal-spatial travel patterns for cycling, subway, and ride services (taxi and ride-hailing) from 2019 to 2022 based on NYC mobility open data. The upper figure describes temporal

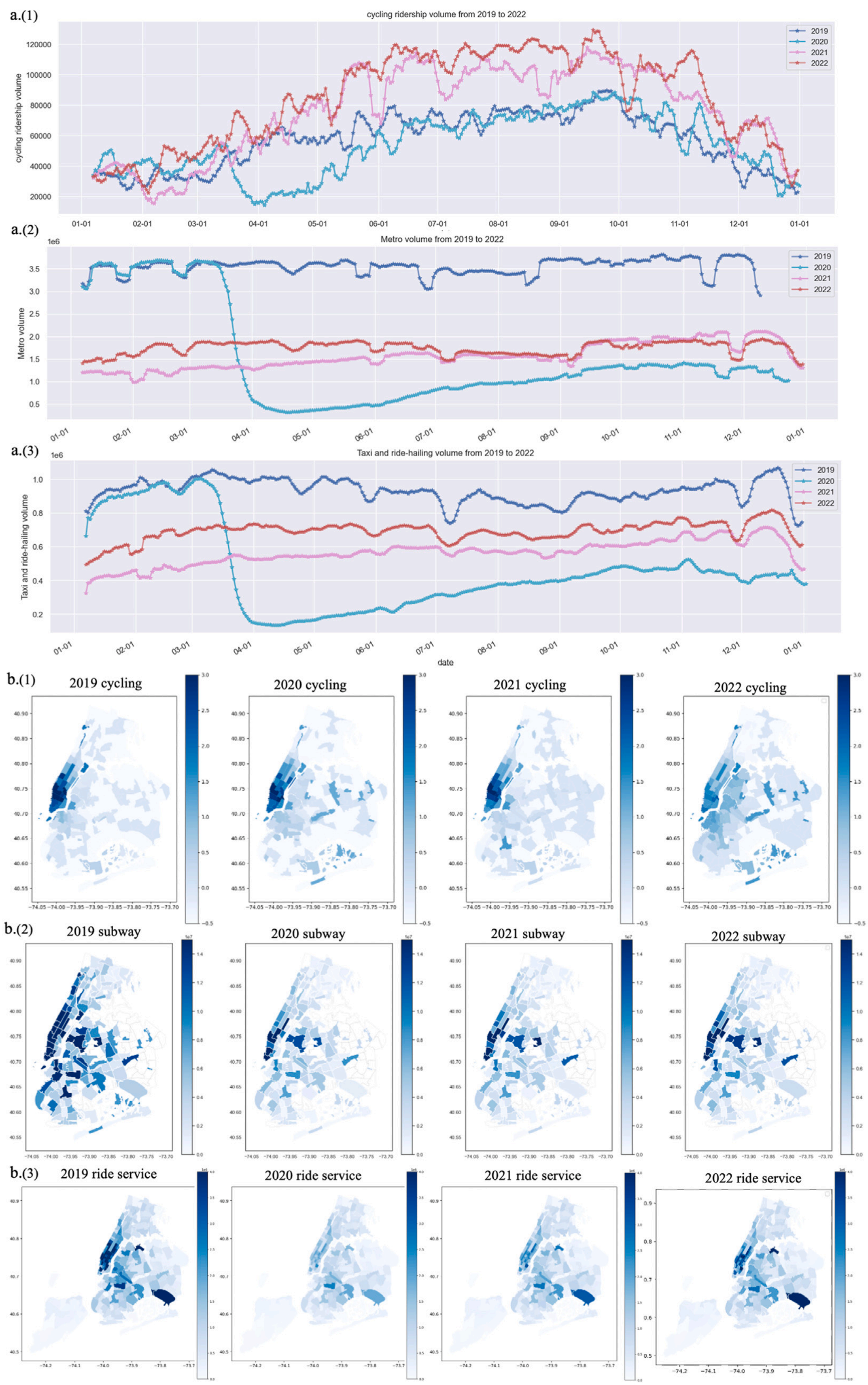

Fig. 3. Travel behaviour from 2019 to 2022: a(1) cycling temporal trend; a(2) subway temporal trend; a(3) ride services temporal trend. b(1) cycling spatial trend; b(2) subway spatial trend; b(3) ride services spatial trend.

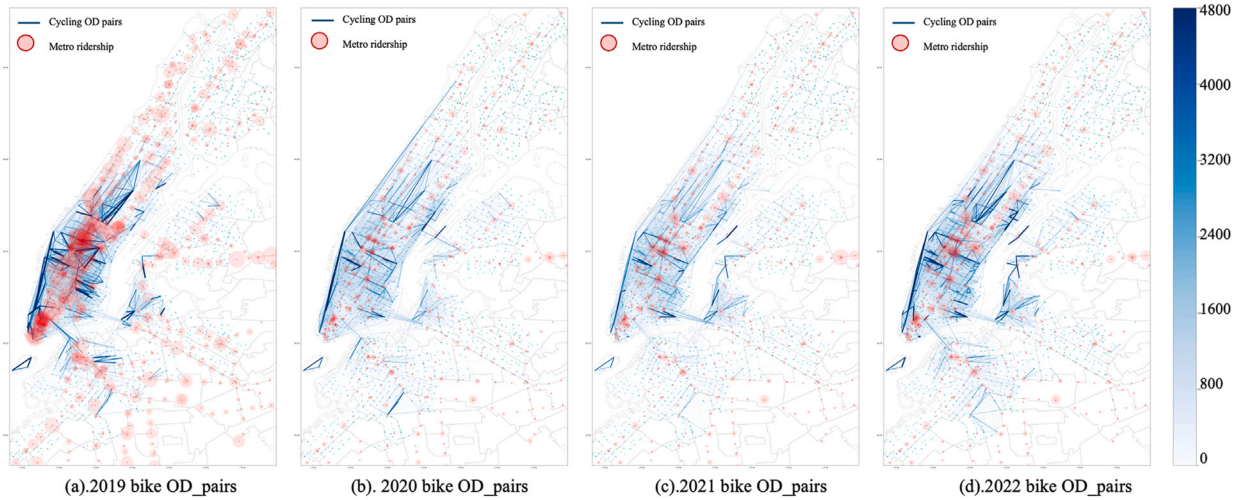

Fig. 4. Cycling OD analysis from 2019 to 2022.

patterns for which 2019 is in dark blue, 2020 is in light blue, 2021 is in pink, and 2022 is in red. Cycling ridership maintains a consistent seasonal trend over the four years, with peaks during the summer and gradually declines during the winter. It appears to be relatively unaffected by the pandemic. Subway and ride services display similar temporal trends, with higher ridership in 2019, followed by a dramatic decline in March 2020 due to the pandemic, and began to recover in 2021. Notably, neither subway nor ride services have reverted to pre-COVID-19 levels, implying that the pandemic may have brought lasting shifts in transportation behaviour. The lower part of Fig. 3 presents the spatial travel patterns. Cycling largely depends on the availability of shared-bike infrastructures, predominantly located in midtown and downtown Manhattan, with an expansion in 2022 to Brooklyn and Queens. Subway ridership is broadly distributed but experienced a dramatic decrease during the pandemic, with primary recovery observed in central Manhattan and Queens. Ride service is primarily concentrated in Manhattan and airport areas. In the post-pandemic period, airport-related services have shown the most substantial recovery, indicating a higher public demand for airport transportation.

We further analyse the origin–destination (OD) patterns for cycling and ride services, as well as their relations with subway ridership from 2019 to 2022. Fig. 4 depicts the cycling OD pairs in blue lines with varying weights indicating trip frequency, while red circles represent subway ridership at each station. Nearly 80% cycling trip destinations in 2019 were within the 500-metre service radius of subway stations with an average short trip distance of 625 m, implying the dominant 'last mile' cycling purpose. During and after the pandemic, leisure-purpose trips have gradually emerged. More destinations and origins appeared in 'High-line Parks', 'Central Parks' and 'Roosevelt Island', indicating the cycling motivation change. We also analyse the ride service OD pattern at Fig. 10 in Appendix, which shows highly airport-relevant travel behaviour. Even during and after the pandemic, airport-related trips were prioritized for recovery.

We observed several travel behaviour shifts throughout the pandemic: (1) Cycling volume exhibits identical seasonal patterns regardless of pandemic impact, but the purpose of trips has gradually shifted from commuting-related 'last-mile' to leisure. (2) Subway ridership experienced a substantial decline and did not recover pre-COVID level, indicating a long-term shift away from the use of public transportation. (3) Ride services consistently prioritized airport-related trips both before and after the pandemic, indicating the high demand driven by long-distance and time-sensitive travel purposes. Meanwhile, short-distance trips within the city centre experienced a relative decline, while remaining more resilient in areas with limited access to public transportation infrastructure. Studies state that these long-term shifts in mobility are influenced by various factors, including remote work, risk perception, and policy interventions (Jain et al., 2022). It is also worthwhile to understand the underlying psychological motivations behind these travel decisions, as they reflect the long-term mobility demands and inform future policy directions.

5.2. Attitudes analysis and travel dissonance

5.2.1. Travel-related attitudes

Fig. 5(a) describes four travel mode-related attitudes from 2019 to 2022. In general, cycling (green) and subway (blue) present relatively positive attitudes compared to private vehicles (red) and ride services (orange), with fluctuations over the years. At the beginning of the COVID-19, subway attitude experienced a significant decline while ride services attitude encountered an immediate increase. In contrast, cycling and private vehicles related attitudes remained relatively unchanged, indicating that risk perception played a crucial role in shaping short-term travel preferences. (b), (c) and (d) in Fig. 5 compare the trend of travel-related attitudes (red), psychological engagement referring the number of travel-mode related social media volumes (blue), and related travel ridership (black). It shows that besides the initial stage of the pandemic, travel-related attitudes do not always consistently with travel-mode related social media volumes, implying the actual travel behaviour is driven by complex factors beyond travel preferences. This discrepancy highlights the importance to investigate the underlying reasons to better understand public concerns and social demands. Therefore, we further explore the travel dissonance with potential causes at the various temporal and spatial levels.

5.2.2. Travel dissonance

Fig. 6 describes temporal and spatial cycling travel dissonance, highlighting the particular periods and areas that contain obvious psychological and actual behaviour gaps. The left figure illustrates temporal travel dissonance, characterized by higher ridership with lower attitudes, present in blue fillings between the normalized sentiment scores (red line) and the ridership volumes (blue line). The right figure demonstrates spatial travel dissonance in red areas for the four years from 2019 to 2022. Most cycling travel dissonance coincides with peak using periods in summer and is associated with intensive infrastructure areas in downtown and midtown Manhattan. We also present subway and ride services travel dissonance in Fig. 12 in Appendix. Next, we present the underlying causes of these dissonance areas and periods to support policy-makers for implementing targeted interventions to enhance traveller satisfaction.

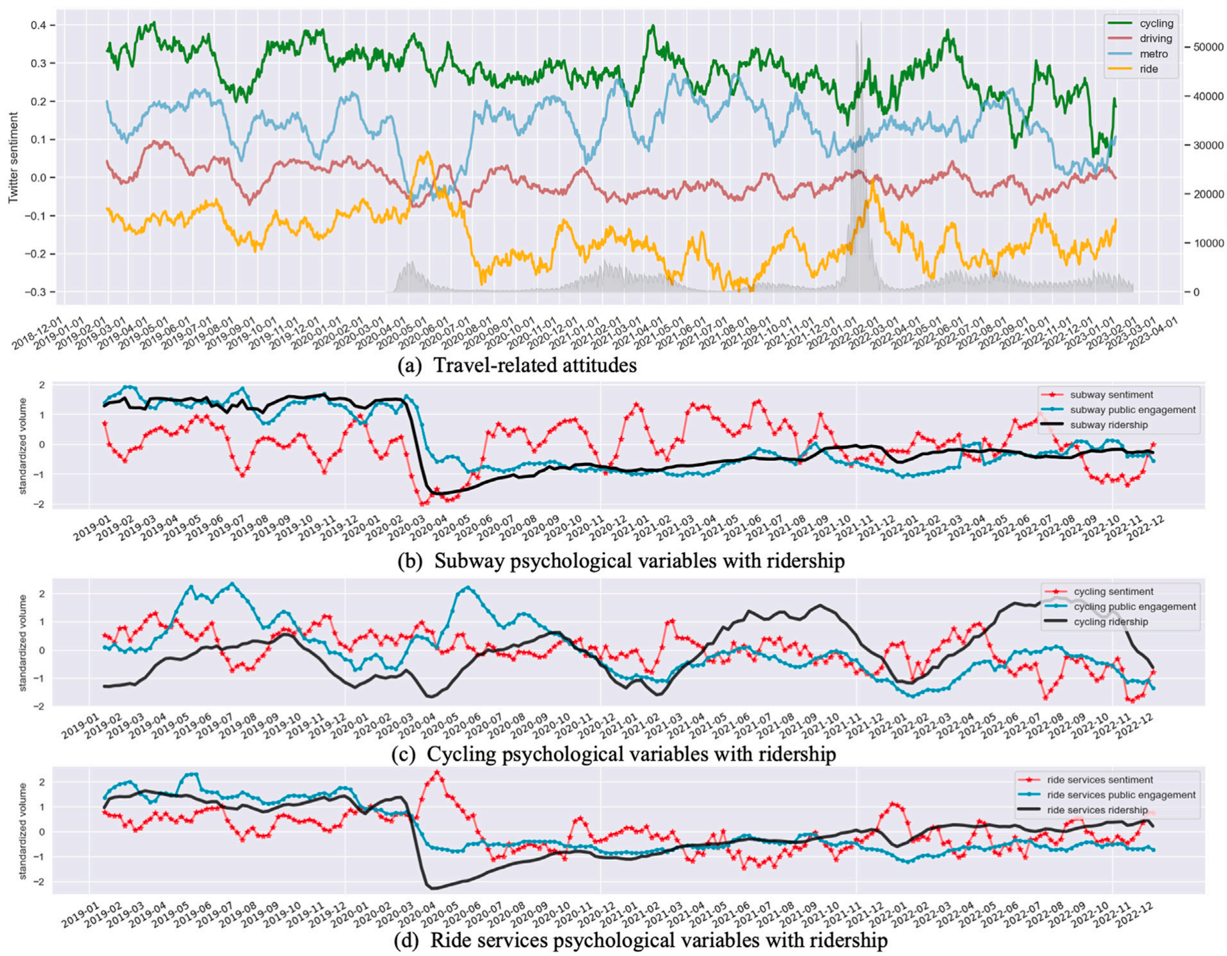

Fig. 5. (a) Four travel modes related attitudes: (b) Subway user engagement, sentiment compared with ridership; (c) Cycling user engagement, sentiment compared with ridership; (d) Ride services user engagement, sentiment compared with ridership.

5.3. Interpretable analysis

5.3.1. Travel-related top concerns

Fig. 7 presents the main topics among four travel modes from 2019 to 2022. We summarized the top 15 topics using 'c-TF-IDF' representation and fine-tuned large language model (LLM) representation at Tables 2 and 3 in Appendix. The result indicates the primary concerns for cycling concentrate on safety issues (cyclist safety, helmet wearing, lanes protection, stolen, and police), bike facilities (repair, path, routes conditions, parking) and cycling purpose (exercises, access to food); Subway main consideration focus on service problems (delay, platform conditions, indoor infrastructures and smells), price (fare, metrocard cost), transfer facilities (bus, shuttle, parking), and risk perception (police, crime, homeless and mask wearing); Topics related to ride services predominantly include services (ratings, drivers performance, communications, radio, smell), price (price comparison, cancellation fee, charged), ride purpose (airport, food delivery, home, and late-night rides), and the risk perception on pandemic (mask wearing, Covid-19); The main concerns regarding private vehicles encompass safety issues (accident,police), cost (insurance, price, license plates, electric cars, and rentals), driving experience (speed, navigation, weather conditions, smoking, noise, and night driving) and infrastructures (parking, road). Overall, 'safety' remains a major concern among the four travel modes with specific risk perception on the pandemic, particularly for subway and ride services. Sustainable transport (subway and cycling) users emphasize the service qualities and infrastructures as these factors substantially affect their travel experience. Meanwhile, 'cost' is a main concern for private vehicle and ride services users.

We also present the dynamic trends for these key topics within four years at Fig. 11 in Appendix. In general, the main topics for cycling include safety, lane protection, bike repairs, and trip purposes emerge each summer. For subway, service-related issues were predominant in 2019, but the pandemic shifted user focus to safety and mask-related concerns. In ride services, customer service and driver performance are major concerns, with increasing attention to food delivery during and after the pandemic. For private vehicles, driving experience, safety and costs remain the primary concerns.

5.3.2. Key motivations

Fig. 8 represents the correlation analysis between subway key topics with corresponding travel behaviour. The heat map on the left describes the correlation degree from −1 to 1, where red and blue color indicate the positive and negative relations with precise values. The strong positive correlated topics in red are 'topic 1 poor services, delay' (0.83), 'topic 2 subway transfer, bus and shuttle' (0.75), 'topic 4 commuters' (0.83), 'topic 7 subway fare' (0.77), 'topic 8 station, platform condition' (0.77), 'topic 9 subway infrastructures, seats, stairs' (0.80). The strongly negative correlated topic in blue is 'topic 6 wear mask, pandemic' (−0.49). The right figure depicts the dynamic trend of these prominent topics compared with subway ridership. This visualization demonstrates positively correlated topics, including subway services, indoor and platform infrastructures, cost, and subway transfer facilities (red,

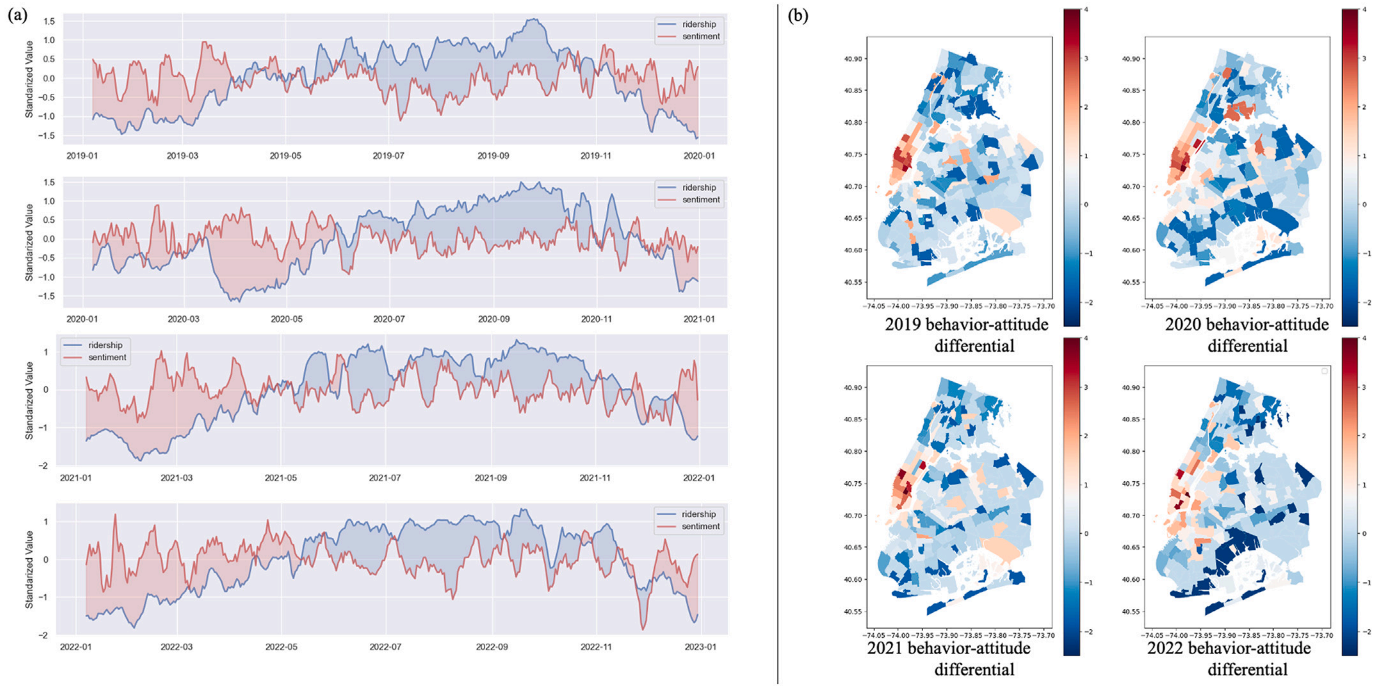

Fig. 6. Cycling attitude behaviour gaps.

pink, orange, purple), consistently associated with subway ridership (black). In contrast, risk perception (blue) during the pandemic, is negatively associated with subway usage.

Although we cannot definitively assert that these strongly correlated topics fully determine individual travel decisions, they do reflect key public concerns and urgent social needs. The sentiment analysis presented in the bottom right describes user attitude towards these correlated topics, highlighting some consistent positive correlated topics with relatively negative attitudes. It indicates the main concerns behind subway travel dissonance including 'topic 1 poor services, delay' in red and 'topic 7 fares, cost' in yellow that require the attention of policy-makers.

The ride services and cycling correlation results are shown in Figs. 13 and 14 in Appendix. For ride services, the positively correlated topics are 'topic 1 driver performance' (0.70), 'topic 2 ratings, customer services (0.67), 'topic 3 communication' (0.73), 'topic 7 airport services' (0.67), 'topic 8 radio' (0.69), 'topic 9 Uber pool' (0.65), 'topic 10 cancel, charged' (0.65), 'topic 15 cost, price' (0.65); The slightly negative topics are 'topic 5 communication' (−0.14) and 'topic 12 mask, pandemic' (−0.29). The sentiment analysis indicates that driver communications (pink), airport services (purple), and trip charges (orange) are consistently with negative attitude. For cycling, there is no significant correlation between cycling ridership and main topics over the four years, but there is a significant difference before and after the pandemic. The relatively consistently positive correlated topics before the pandemic were 'topic 1 cyclist safety', 'topic 3 lane protection', and 'topic 5 route condition'. After the pandemic, 'topic 2 bike repair', 'topic 8 exercise', and 'topic 11 bike shop, purchase' started to show stronger positive correlations with actual cycling behaviour. This trend supports the observed shift in cycling travel behaviour from a 'last-mile' purpose to a focus on individual leisure activities. The sentiment analysis of these key topics highlights that cyclist safety and lane protection are consistent concerns with negative attitudes that require attention from policymakers.

5.3.3. Interpretation on travel dissonance

Fig. 9 illustrates the implementation of dynamic topic modelling in subway travel dissonance interpretation from the psychological perspective at the neighbourhood level. Within the four years, the key concerns in the central areas of Manhattan related to services such as delays, lateness, and crowd issues. In the dissonant neighbourhoods of Queens and Brooklyn, the primary concern is related to price and cost. While in the Bronx, there are more concerns related to mask-wearing and shuttle transfer, especially after the pandemic in 2021 and 2022, indicating the remaining anxiety of COVID-19 and the demand for subway transfer facilities improvement. We also show cycling and ride services results in Fig. 15 in Appendix. For cycling dissonance areas, cyclist safety and route protection are the consistent concerns. While for ride services, there are more concerns related to drivers, cancellation, price, smell, and weather in airport areas; Drivers' performance-related compliance in the central Manhattan area. During and after the pandemic, food delivery and cost-related discussion have gradually emerged.

5.4. Summary on results

To get a deeper understanding of the underlying reasons behind the results, we summarize the role of psychological variables in travel behaviour explanation, travel dissonance identification and social demands recognition as follows:

- ***Cycling trip purpose from commuting to leisure***: The topic modelling result indicates a continuous increase related to topics on 'exercise', 'park' since the pandemic, suggesting an increased use of cycling for exercise purposes compared to the pre-pandemic period. The correlation results also demonstrated a stronger association between cycling ridership and the topic 'exercise' after the pandemic. Simultaneously, 'bike' topics in subway-related discussions have decreased since the pandemic began, suggesting a gradually weakened connection between subway use and bike sharing. These findings indicate with the rise of remote work

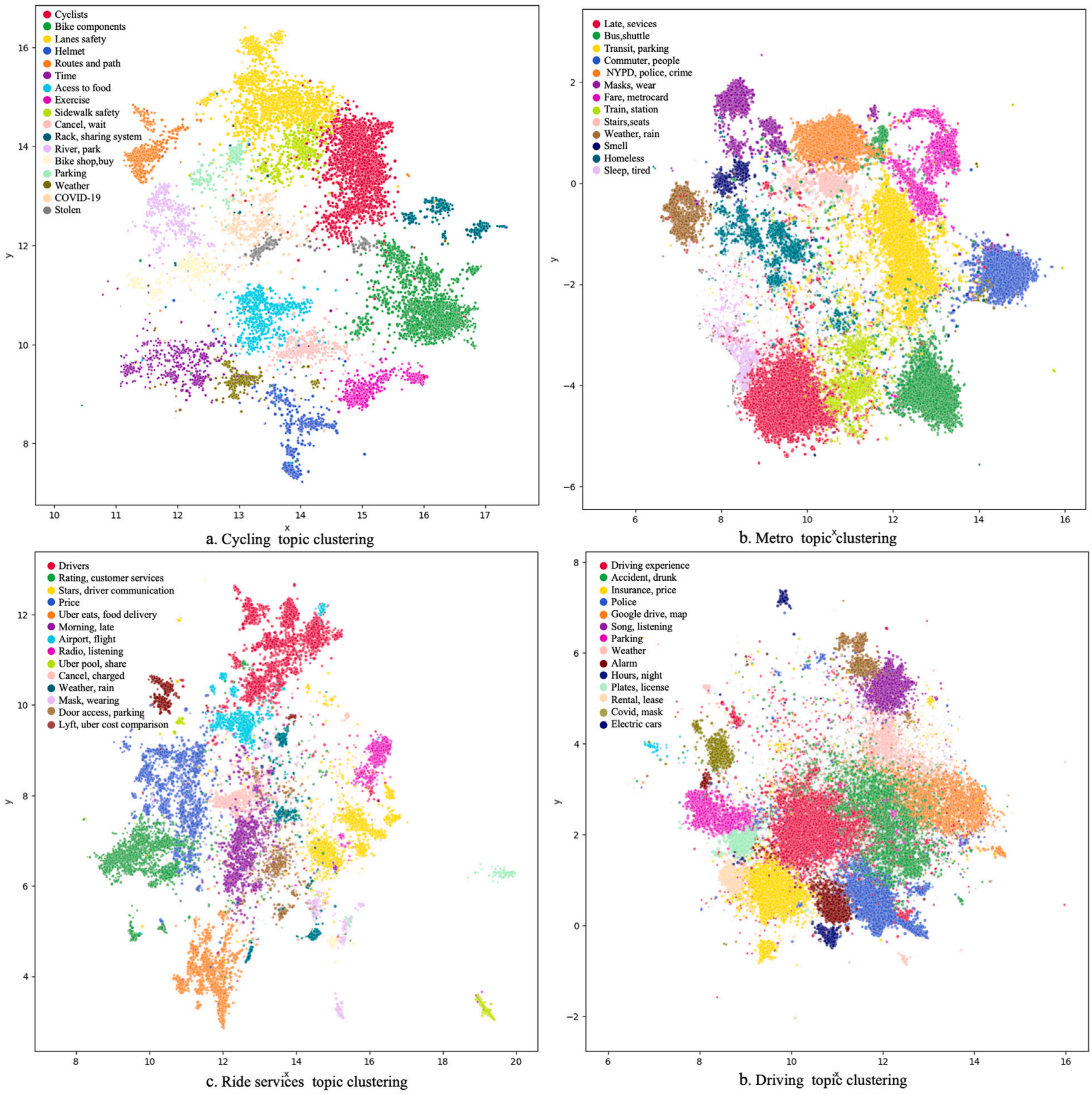

Fig. 7. Topic modelling for the four travel modes.

and health awareness, cycling emerged as a safe and flexible leisure activity. Supported by the improved bike infrastructure and culture shift to outdoor activities, this behaviour has evolved into a long-term mobility adoption.

- ***Subway noticeably decreasing since pandemic and not recovered to the pre-COVID period till now***: At the beginning of the pandemic, the sudden decrease in public transport use was mainly driven by lockdown policies and safety concerns. Topic modelling results reveal topics including 'mask', 'cleaning', and 'Covid' emerged in the same periods as the sentiment towards the subway decreased markedly. Evidence suggests a lasting shift in public transport habits as safety concerns on 'mask', 'homeless', 'crime', and 'policy' persist in the post-pandemic period.
- ***Airport-relevant trips consistently dominate ride service usage both before and after the pandemic***: Due to the time sensitivity for airport-related trips, concerns about delay on the subway system likely contribute to the preference for ride service. Meanwhile, the emerging COVID-related concerns in public transport further reinforce users' prioritization of ride services for these long-distance trips.
- ***Travel dissonance interpretability***: The integration of travel-related behaviour and attitude highlights the significant discrepancy in both spatial and temporal levels. Topic modelling results further point out public concerns in the most discrepancy areas that people reluctant to use, yet feel compelled to do so due to necessity. Particularly, safety concerns (lane protection, parking security, stolen) and route complaints mainly account for cycling dissonance areas; poor services (delay, crowd) and cost-related problems are the primary factors for subway dissatisfaction; driver performance and cost-related concerns are dominant for ride services dissonance. These various domain reasons for travel dissonance highlight the urgent improvement directions and areas.

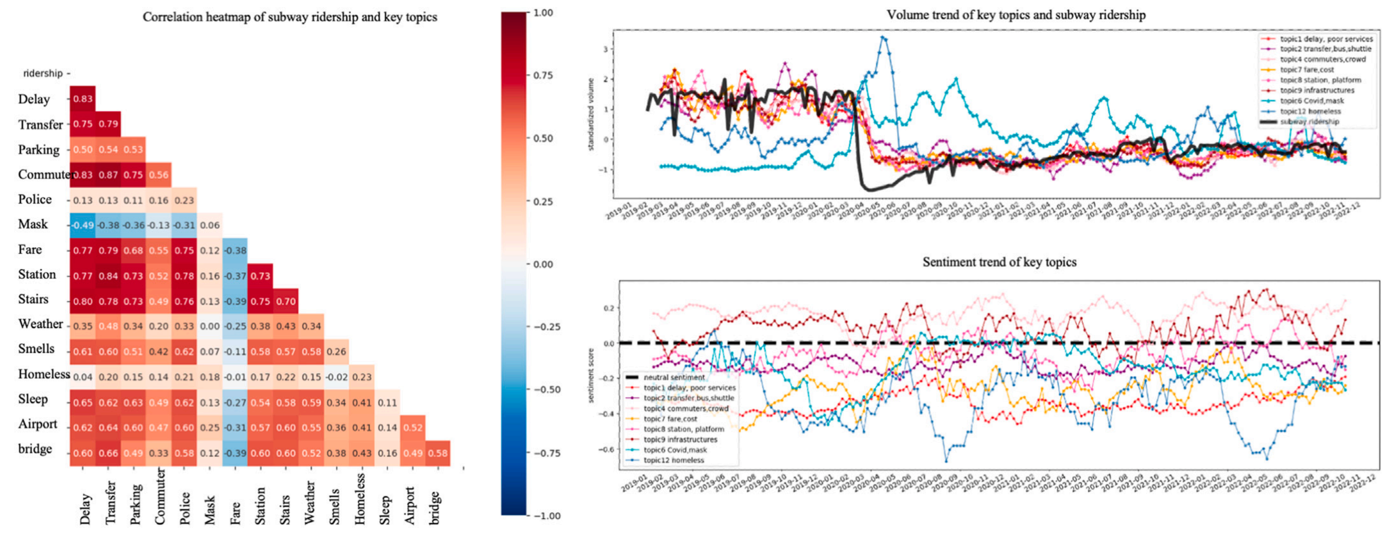

Fig. 8. Subway correlation analysis.

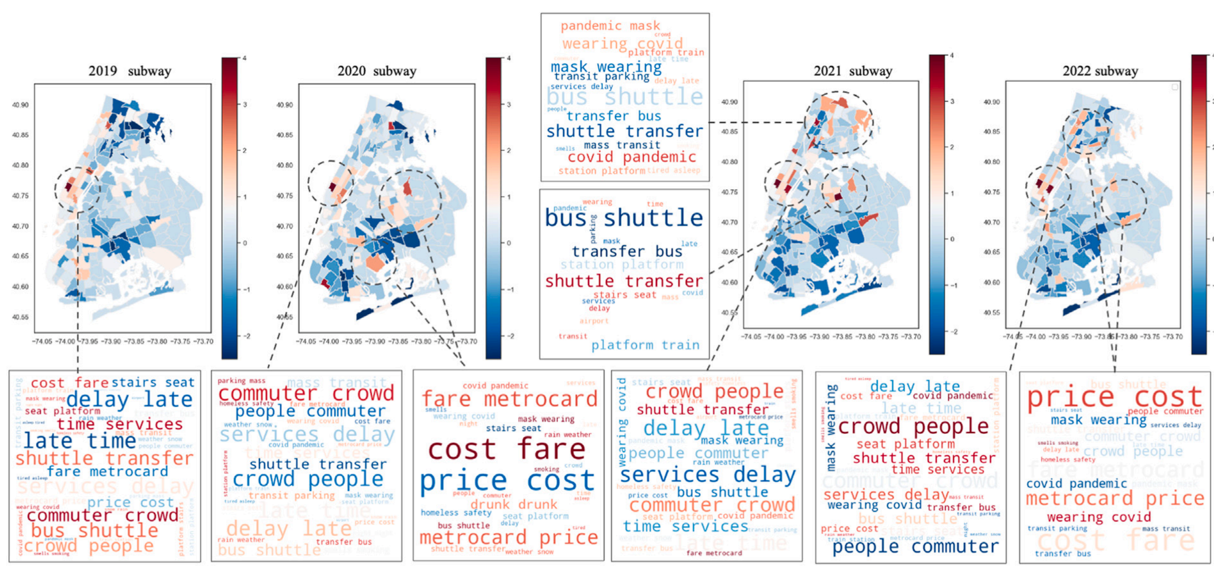

Fig. 9. Subway travel dissonance explanation at neighbourhood level.

6. Policy recommendation

To validate the effectiveness of the proposed framework and examine the representativeness of our findings from social media samples, we compare the outcomes with the New York City Department of Transportation's Citywide Mobility Survey and the existing studies with transport policy proposals that rely on more comprehensive representative samples, such as survey data. Several findings indeed provide quantitative evidences from the user perspective that consistent with these policies, indicating the applicability to support future transportation interventions. Despite the alignment with overall policy suggestions from our psychological results, the travel dissonance interpretation points out more targeted suggestions for policy-makers in precise neighbourhoods.

For cycling, multiple studies recommend enhancing bike infrastructures and adding digital support tools in response to the rapid growth of cycling ridership in New York City since the pandemic. These interventions include separating bike lanes from roads, expanding bike routes and providing real-time apps for weather and air quality to improve safety and user satisfaction (Möllers et al., 2022; Gong et al., 2024). Meanwhile, from the New York City Department of Transportation's Citywide Mobility Survey (New York City Department of Transportation, 2022), most riders aim at exercise or recreation (76%), while only a minority cycling for commuting to work or school (16%) after the COVID period. The main concerns about cycling concentrated on safety (42%), trip facility (17%) and insufficient bike storage (22%). These survey results are highly related with our findings from social media data. Our results also highlight that the safety concerns, including cyclist protection, lane safety, and parking security are the key priorities for cycling users. Additionally, some studies recommend bike infrastructures expansion in leisure areas rather than near subway stations, reflecting a post-pandemic trip purpose shift on leisure-oriented travel (Xin et al., 2022). Our results also demonstrate a strong correlation between cycling behaviour with psychological motivations related to 'exercise' and 'parks and trails' after the pandemic, indicating similar recommendations for policy-makers to enhance leisure-oriented trips and infrastructures.

Subway service improvements, including schedule optimization and real-time notification are the priorities from existing recommendations (Caspari et al., 2021; Wei et al., 2022). Our analysis also shows that service-related concerns have remained prevalent both before and after the pandemic, particularly from the complaints about delays, cancellations, and crowdedness. Our findings additionally indicate that poor service is one of the major complains in the high-travel dissonance neighbourhoods, implying many users reluctantly rely on public transport but with poor and discomfort travel experiences. It underscores the need for subway system improvement to enhance user satisfaction and travel experience. Safety is another key determinant for subway passengers, including crime and pandemic-related risk perceptions. According to the NYC DOT Citywide Mobility Survey (New York City Department of Transportation, 2022), 21% participants feel unsafe on public transit, which 20% related to COVID concerns that affect their use on public transit. Multiple studies advocate safety interventions by promoting mask-wearing, ensuring adequate social distancing, and enhancing security measures to mitigate public risk concerns, thus encouraging subway re-utilization after the pandemic (Olfindo, 2021). In addition to identifying 'mask', 'Covid' related risk concerns, our analysis highlights 'homeless' and 'police' as the significant negatively correlated factors with subway usage since the pandemic, indicating the need for targeted interventions to reduce passenger anxiety. Furthermore, existing research emphasizes the interconnectivity between the subway and other transport modes (e.g., buses, ride-hailing, shared-bikes, and private vehicles), suggesting improvements on transfer facilities such as shared-bike infrastructures, bus stations, or parking lots near subway stations to enhance subway usage efficiency (Ashraf et al., 2021; Wang and Noland, 2021). Similarly, our findings reveal key topics related to 'bus', 'bike', 'parking', and 'Uber', underscoring that multi-modal transport connectivity is a key factor for subway usage. Notably, in high travel-dissonance neighbourhoods in the Bronx, frequent discussions on 'shuttle bus' and 'transfer' imply these areas as priorities for establishing resilient public transportation transfer systems and facilities.

Multiple studies suggest improvements on airport-related ride services, considering the high demand and supply shortage. Recommendations include enhancing airport transport accessibility, incentivizing drivers to pick up, and improving passenger satisfaction through implementations such as providing frequent-airport server punch cards for taxi drivers, adding a time-specific ride-sharing program, and optimizing airport traffic flow management (Yazici et al., 2016; Zhai et al., 2019; Wang et al., 2024). Our case study points out that 'access', 'waiting', and 'drivers' are the main concerns for airport-related ride services, which aligns with the existing recommendations, suggesting that promoting ride-sharing and expanding public transport options (e.g., efficient subway systems) can help mitigate the demand shortages. Both existing studies (Zhai et al., 2019) and our case study point out the high frequency of inefficiency short-distance ride trips in high travel-dissonance areas (particularly central Manhattan and north Brooklyn). We also identify 'cost' as the primary concern in these neighbourhoods, implying the unwillingness for the high-cost ride services. Correspondingly, inadequate public transport infrastructures (Jin et al., 2019) have been stated as the major reason, which highlights the need to provide public transport facilities such as shared-bike stations, subway facilities to alleviate customers' discomfort due to economic problems. Time-sensitive topics ('time', 'morning', 'late') and cost-related topics are the major concerns in central Manhattan, implying that users opt for ride services to avoid delays and cancellation issues compared to the subway system, particularly in the peak hours, even suffering higher expenses. This underscores the need to improve public reliability and on-time performance.

Moreover, the NYC DOT Citywide Mobility Survey (New York City Department of Transportation, 2022) points out the main reasons for using private vehicles are convenience(40%), feeling unsafe on public transit (21%), and COVID-era safety concerns (20%), which indicates the long-lasting psychological impact on travel mode choices because of the pandemic. In addition, it states the key concerns of private vehicle users are the cost, parking facilities, and safety. These survey results are closely relevant to our topic modelling findings based on social media data, for which 'accident', 'price', and 'parking' are similarly the key topics from private vehicle-related posts. Both the mobility survey and our findings suggest the need of parking facility improvement and price reduction for private vehicles.

7. Conclusion and discussion

This article proposes an evidence-based policy recommendation framework from a user-centred perspective to support inclusive interventions according to public concerns and demands. Integrating social media data with NLP methods, this article demonstrates the ability of psychological variables to identify dynamic motivations, social demands, and travel dissonance with underlying reasons to enhance evidence-based transportation policy-making. The case study implements our proposed framework in New York City across the pandemic period. Outcomes illustrate the spatial–temporal variations in attitudes and concerns and provide thorough explanation for long-term travel pattern changing and trip purpose shifting among subway, cycling, and ride services. The integration analysis highlights key motivations that largely impact travel choices, while travel dissonance analysis points out emerging social demands and complaints. We summarize the novelty of this study as follows:

- ***Social media data with NLP methods provide an efficient and low-cost approach to acquire dynamic travel-related psychological variables, including user attitudes, concerns, and opinions.*** The sentiment analysis and topic modelling results present dynamic travel-related attitudes, public focus and concerns across multiple spatial and temporal levels. Compared to survey data, social media data enable nearly real-time public psychological trends tracking that provides valuable evidence from the user perspective to enhance transport policy-making. During special periods such as the pandemic, staying attuned to public sentiment and concerns is important as amplified public feelings largely impact travel behaviour. Meanwhile, integration with travel pattern data, these dynamic psychological variables enable travel dissonance identification, social demands recognition and behaviour shifting interpretation.
- ***Psychological variables provide valuable interpretation for dynamic travel behaviour including travel choices and trip purposes shift.*** Both the correlation analysis and comparison with other travel modes associated concerns help to identify key motivations for certain travel choices. In the case study, cycling trip purpose shifting from 'last-mile' to leisure since the pandemic is support by the gradually increased public discussion on 'exercise', 'parks' from cycling topics and the consistent risk perception from subway topics; The continuous public risk concern on 'mask' and 'safety' elaborate the remain lower usage of subway even after pandemic. The top concerns on the poor subway services, including delays, not only impact public sentiment on subway usage but also influence people's choice of ride service for airport-related trips.

- ***Public main focus and travel dissonance highlight social needs and concerns regardless whether they impact the actual travel behaviour, therefore implying target improvement aspects for policy-makers.*** The topic modelling results present the overall primary focus, while the main concerns in travel dissonance areas indicate the aspects and areas requiring intervention to enhance user travel experience and satisfaction. In the NYC case, poor services, safety issues, and metro costs are the main focus for the subway system; driver performance and high price have been mostly discussed for ride services; cyclist safety and route infrastructures attracted more attention for cycling. The travel dissonance results demonstrate actual travel choices with dissatisfaction and underscore the improvement aspects: for cycling, lane protection complaints occurred most in midtown Manhattan; For subway, delay and crowd have been the urgent problems in midtown area, while price and transport transfer received more complaints in Brooklyn and Queens neighbourhoods. For ride services, high prices and cancellations have been the main problems in airport areas, while customers expressed dissatisfaction with driver performance in other locations.

Even though this study aims to demonstrate the efficiency of social media data in representing timely user travel-related attitudes and concerns, it still contains data limitations. First, social media data cannot represent the entire population since only partial users express their opinions or emotions on these platforms. Although certain platforms include limited user profiles, their privacy policies and ethical restrictions limit the tracking of reliable demographic information at the user level, such as age or level of education (Hanlon and Jones, 2023; Böyük, 2024). Second, social media data contains spatial bias in larger cities such as New York City that have more users compared to smaller cities and rural areas. Incorporating other data resources, such as census demographic data and mobility survey data, will provide an enhanced approach for further work in these areas. Although partially validating the feasibility of our findings with mobility survey data and existing mobility studies that rely on more comprehensive representative samples, we state that social media data is not a replacement for other complete representative data resources, but as a complementary and real-time data source to provide timely insights and reactions to emerging events or policies.

Moreover, negative opinion expression is another bias as people are more likely to sharing complains and express risk perceptions. In this study, we take advantage of these negative psychological variables to reveal social concerns for policy-makers but might still neglect some key determinants. Finally, we cannot assert that highly correlated topics are critical determinations but point out the urgent social concerns. Future studies, such as causation analysis, will confirm the key determinants for each travel mode. Meanwhile, in this study, we discuss the psychological key motivations and concerns behind shared bikes, subways, and ride-sharing, future studies could apply similar methods to understand the psychological impact of other transport options such as electric vehicles and e-scooters, providing comprehensive suggestions for related policies and interventions.

CRediT authorship contribution statement

Yanyan Xu: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Panchamy Krishnakumari:** Writing – original draft, Supervision, Methodology. **Neil Yorke-Smith:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Serge Hoogendoorn:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition.

Declaration of competing interest

We declare this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We understand that the corresponding author is the sole contact for the editorial process. He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs.

Acknowledgements

This research was partially supported by TULIPS, a project funded by EU Horizon 2020 research and innovation programme under grant number 101036996.

Appendix

A.1. Additional figures and tables

See Figs. 10–15 and Tables 2 and 3.

Table 2
Summary of overall topic for cycling and subway.

Topics	Cycling		Subway	
	Representation	chatGPT-4o	Representation	chatGPT-4o
Topic 1	cyclist, people	cyclist safety and protection	minutes, service, delay	delay, late service issues
Topic 2	bike, tire	bike components and repair	bus, transfer	change bus and shuttle
Topic 3	lanes, parking	lane usage and parking issues	transit, parking	mass transit, parking issues
Topic 4	helmet, wear	helmet usage with cycling safety	commuter, people	commuters in subway
Topic 5	bridge, path	routes via bridges, pathways	nypd, police, crime	subway crimes
Topic 6	morning, time	different time to cycling	masks, wear	wearing masks in pandemic
Topic 7	tour, food	cycling to access food	fare, metrocard	metrocard fare and evasion
Topic 8	exercise, training	cycling for exercise, fitness	train, station	daily commuting
Topic 9	NYPD, police	NYPD parking enforcement, safety	stairs, seats	subway infrastructures
Topic 10	rack, dock	rack systems, sharing issues	weather, cold	weather conditions
Topic 11	shop, purchase	shops to purchases bikes	smells, smoking	odors issues, strong smells
Topic 12	parking	parking secure	homeless	homelessness and safety issues
Topic 13	weather, summer	seasonal weather conditions	sleep, tired	falling asleep, missing stops
Topic 14	pandemic, covid	protests during pandemic	ferry, airport	other transportation services
Topic 15	stolen lock	theft and stolen items	bridge tunnel, suicide	infrastructure, safety measures

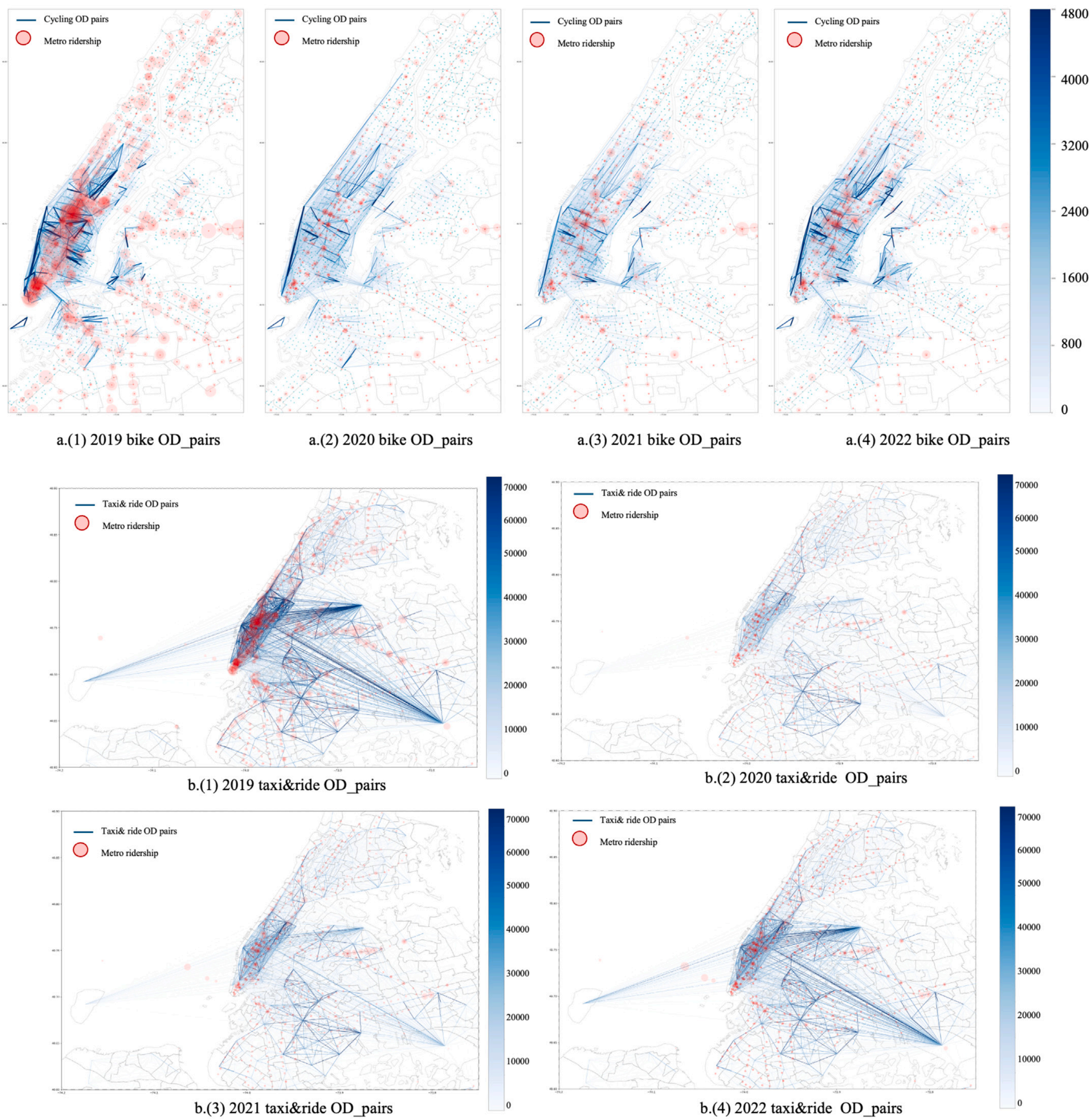

Fig. 10. Cycling and ride service OD patterns: upper image is the cycling OD pattern compared with subway ridership from 2019 to 2022; the above image is the ride services OD pattern compared with subway ridership.

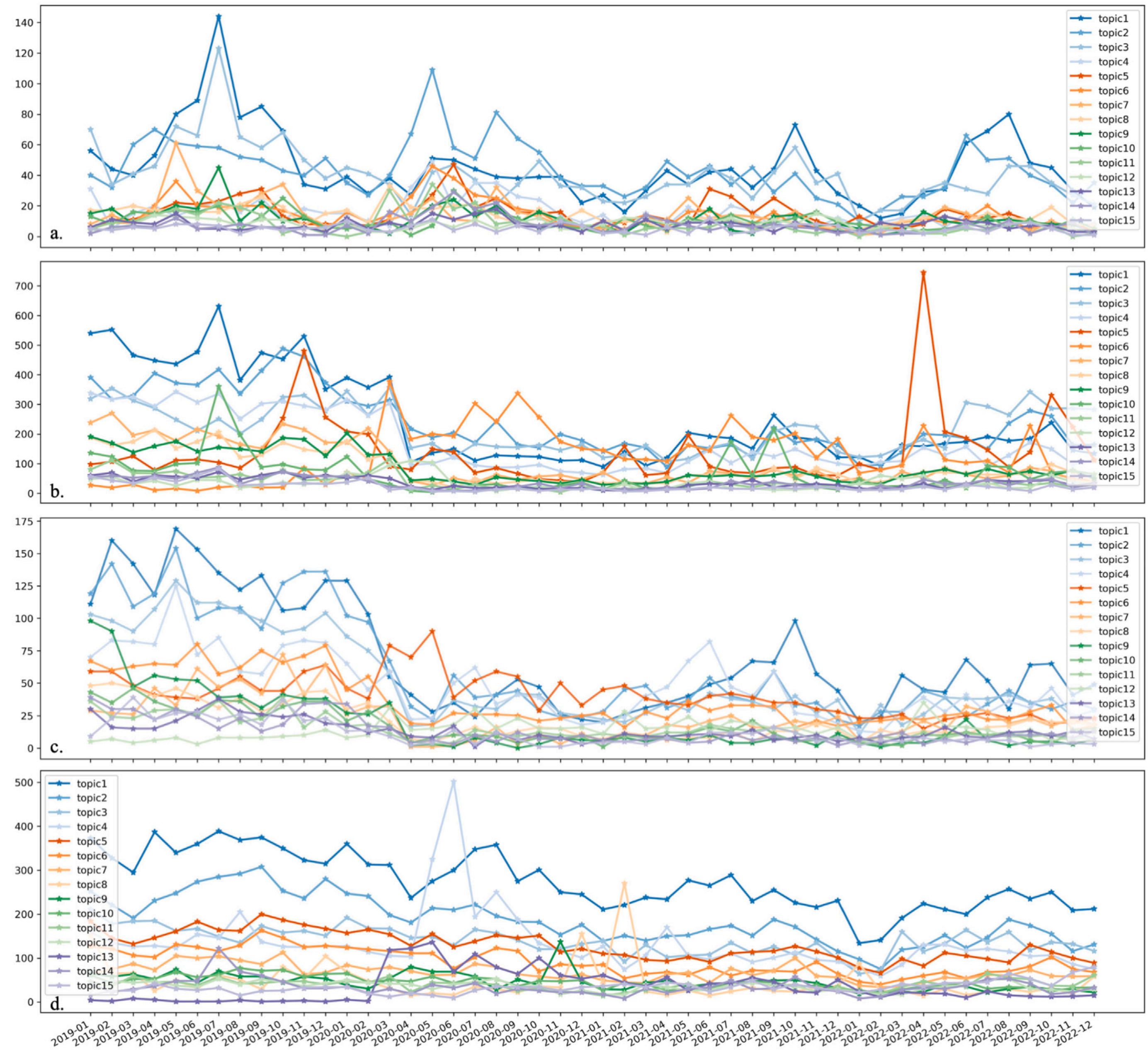

Fig. 11. Dynamic topic modelling for the four travel modes: a. cycling, b. metro, c. ride services, d. private vehicles.

Table 3

Summary of overall topic for ride services and private vehicles.

Topics	Ride-hailing		Private vehicles	
	Representation	chatGPT-4o	Representation	chatGPT-4o
Topic 1	drivers, taxi, cab	cab driver's performance	car, speed	cars and driving experiences
Topic 2	rating, customer	service ratings, customer feedback	accident, drunk	accidents and drunk driving
Topic 3	drivers, stars, talking	driver communication, languages	price, insurance	insurance costs and payments
Topic 4	price, cash	pricing and charges	police, NYPD	police deal incident
Topic 5	eats, delivery	Uber Eats food delivery orders	drive, maps	navigation by google maps
Topic 6	morning, late	morning commuting	song, radio	listening to songs while driving
Topic 7	airport, flight	services at airports	parking, parked	street parking issues
Topic 8	radio, listening	drivers playing radio music	snow, rain	driving in weather conditions
Topic 9	pools, uber	challenges of using Uber Pool	alarm, honk	noise, disturbance from honk
Topic 10	cancel, charged	racing at various times	hours, night	night driving, commute home
Topic 11	rain, slow	rides delayed in weather	plates, licence	car licence plates
Topic 12	mask, wearing	mask-wearing due to concerns	rental, lease	rental, leasing options
Topic 13	door, parking	problems with Uber access	covid, masks	mask-wearing driving
Topic 14	smell, smoke	car odors	bikes, cyclists	bikes and cars comparison
Topic 15	costs, cheaper	Uber, Lyft pricing comparisons	tesla, electric	electric cars

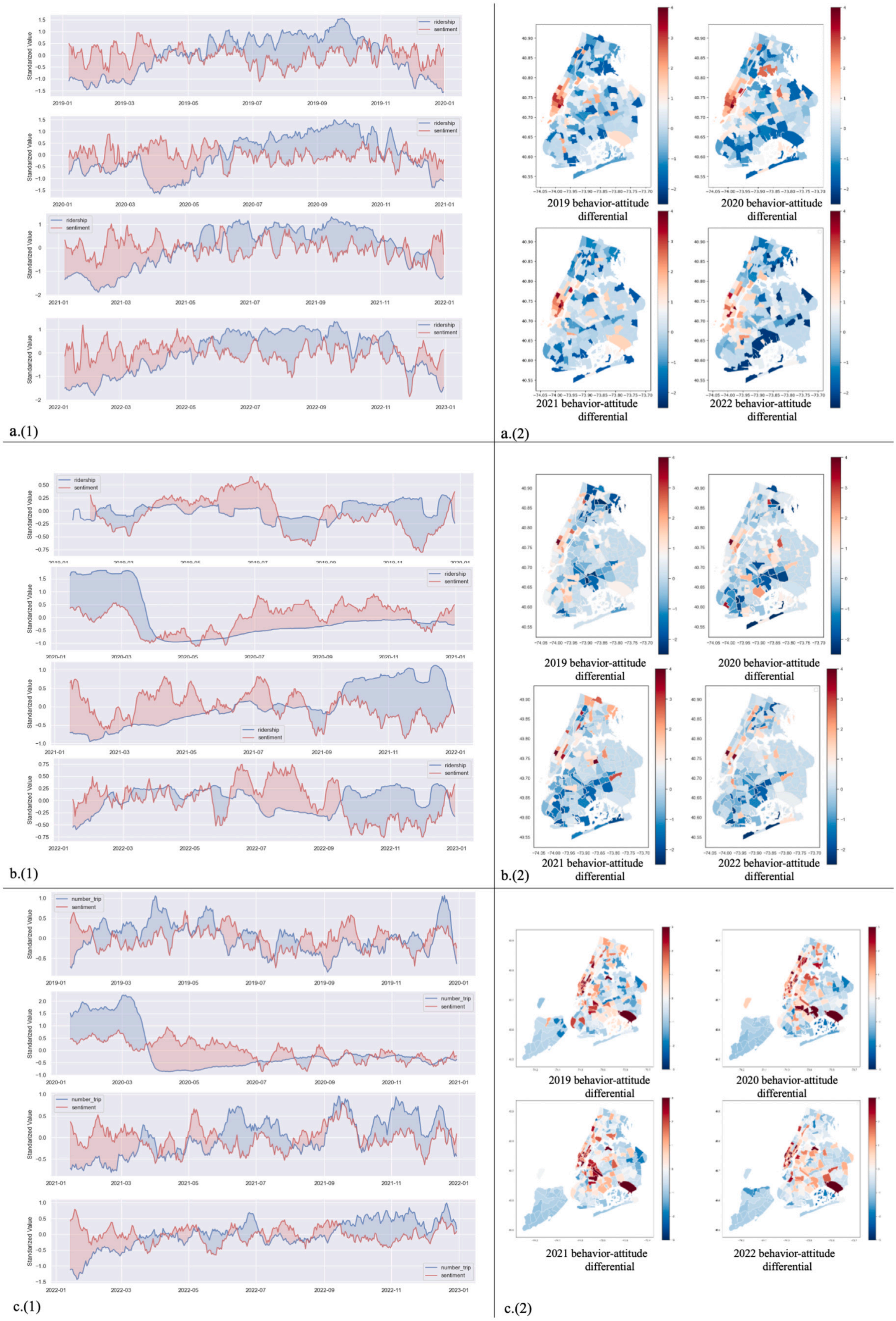

Fig. 12. Attitude behaviour gaps: (a) cycling; (b) subway travel; (c) ride services.

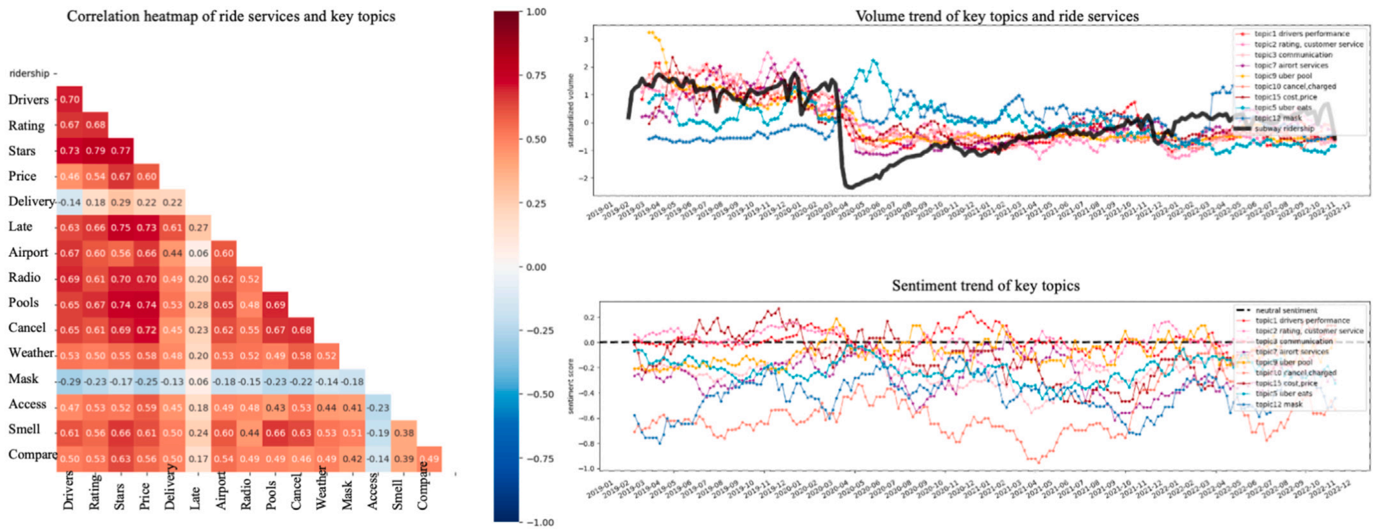

Fig. 13. Ride services correlation.

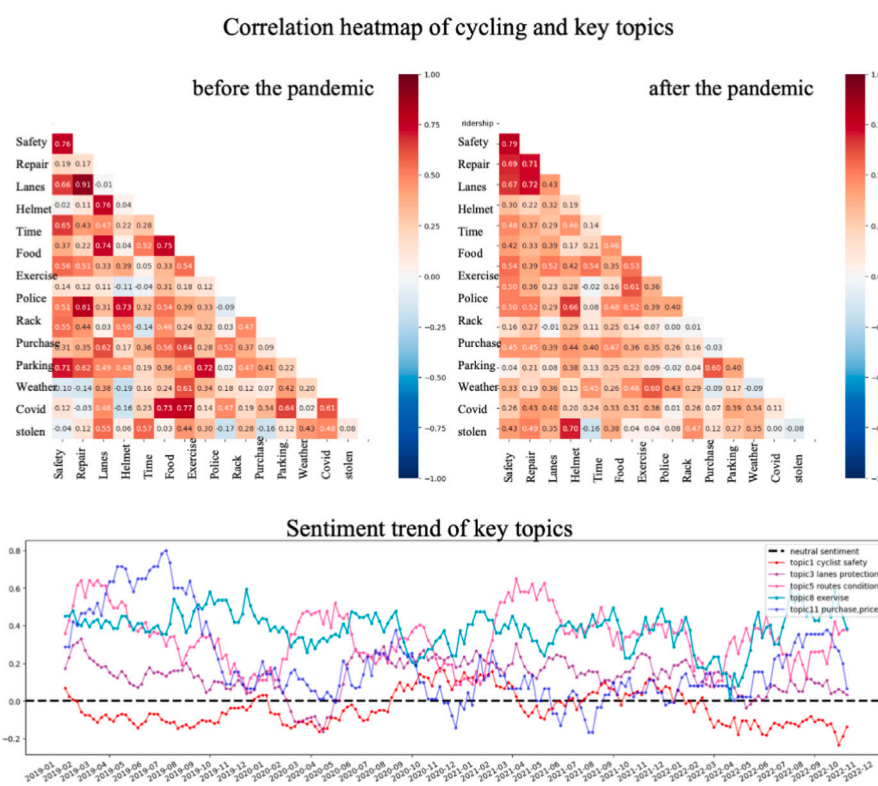

Fig. 14. Cycling correlation.

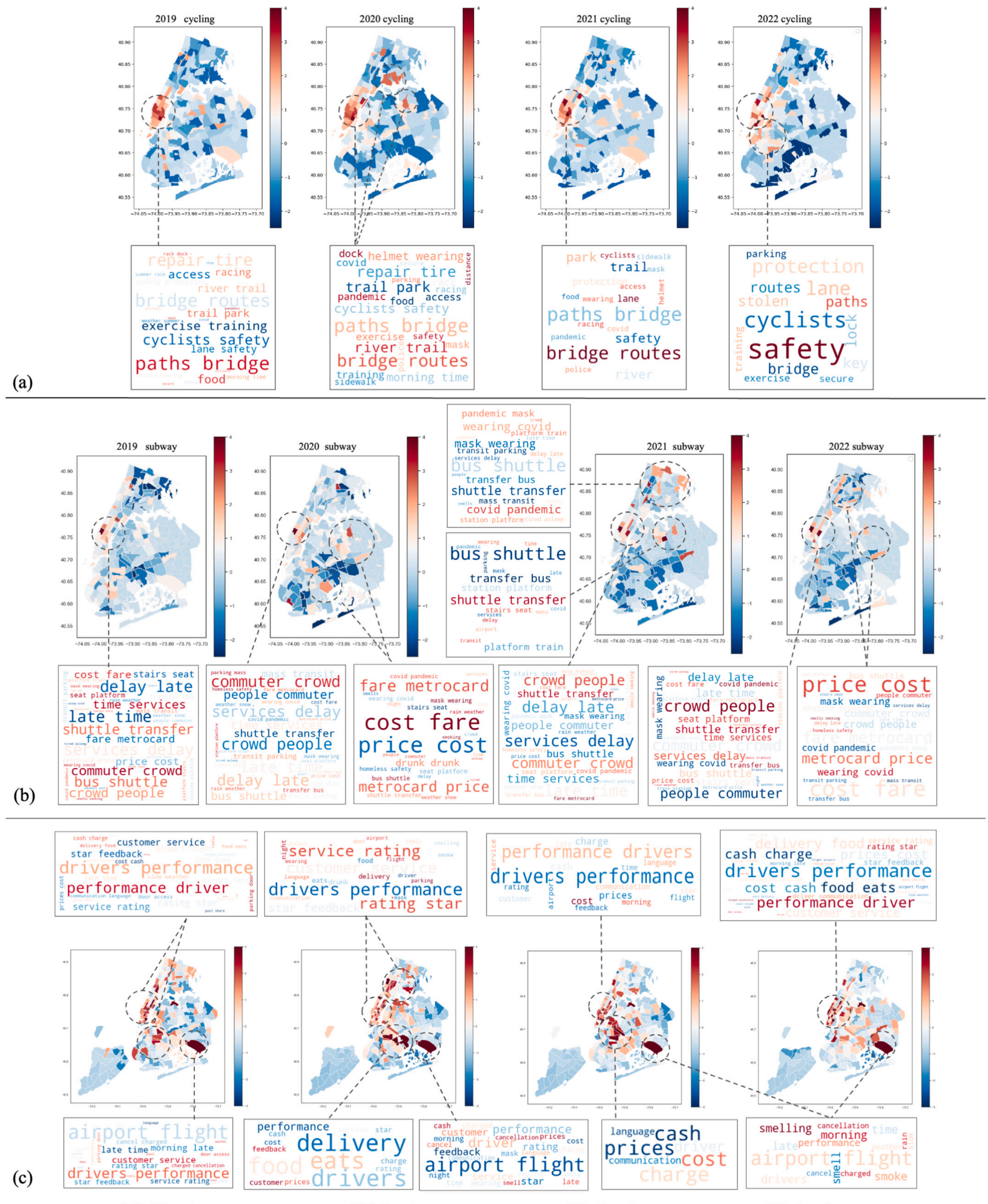

Fig. 15. Travel dissonance explanation in neighbourhoods level: (a) cycling travel dissonance and main topics; (b) subway travel dissonance and main topics; (c) ride services travel dissonance and main topics.

References

Abenoza, Roberto F., Cats, Oded, Susilo, Yusak O., 2017. Travel satisfaction with public transport: Determinants, user classes, regional disparities and their evolution. Transp. Res. Part A: Policy Pr. 95, 64–84.

Achiam, Josh, Adler, Steven, Agarwal, Sandhini, Ahmad, Lama, Akkaya, Ilge, Aleman, Florencia Leoni, Almeida, Diogo, Altenschmidt, Janko, Altman, Sam, Anadkat, Shyamal, et al., 2023. Gpt-4. technical report., arXiv preprint arXiv:2303.08774.

Ajzen, Icek, 1991. The theory of planned behavior. Organ. Behav. Hum. Decis. Process. 50 (2), 179–211.

Ajzen, Icek, 2012. Values, attitudes, and behavior. In: Methods, theories, and empirical applications in the social sciences. Springer, pp. 33–38.

An, Zihao, Heinen, Eva, Watling, David, 2022. Multimodal travel behaviour, attitudes, and cognitive dissonance. Transp. Res. Part F: Traffic Psychol. Behav. 91, 260–273.

An, Zihao, Heinen, Eva, Watling, David, 2023. The level and determinants of multimodal travel behavior: does trip purpose make a difference? Int. J. Sustain. Transp. 17 (2), 103–117.

Anagnostopoulou, Evangelia, Urbančič, Jasna, Bothos, Efthimios, Magoutas, Babis, Bradesko, Luka, Schrammel, Johann, Mentzas, Gregoris, 2020. From mobility patterns to behavioural change: leveraging travel behaviour and personality profiles to nudge for sustainable transportation. J. Intell. Inf. Syst. 54, 157–178.

Angelov, Dimo, 2020. Top2vec: Distributed representations of topics. arXiv preprint arXiv:2008.09470.

Arroyo, Rosa, Ruiz, Tomás, Mars, Lidón, Rasouli, Soora, Timmermans, Harry, 2020. Influence of values attitudes towards transport modes and companions on travel behavior. Transp. Res. Part F: Traffic Psychol. Behav. 71, 8–22.

Ashraf, Md Tanvir, Hossen, Md Amdad, Dey, Kakan, El-Dabaja, Sarah, Aljeri, Moathe, Naik, Bhaven, 2021. Impacts of bike sharing program on subway ridership in new york city. Transp. Res. Rec. 2675 (9), 924–934.

Bamberg, Sebastian, Ajzen, Icek, Schmidt, Peter, 2003. Choice of travel mode in the theory of planned behavior: The roles of past behavior, habit, and reasoned action. Basic Appl. Soc. Psychol. 25 (3), 175–187.

Barandoni, Simone, Chiarello, Filippo, Cascone, Lorenzo, Marrale, Emiliano, Puccio, Salvatore, 2024. Automating customer needs analysis: A comparative study of large language models in the travel industry. arXiv preprint arXiv:2404.17975.

Beirão, Gabriela, Cabral, J.A. Sarsfield, 2007. Understanding attitudes towards public transport and private car: A qualitative study. Transp. Policy 14 (6), 478–489.

Bi, Hui, Li, Aoyong, Zhu, He, Ye, Zhirui, 2023. Bicycle safety outside the crosswalks: Investigating cyclists' risky street-crossing behavior and its relationship with built environment. J. Transp. Geogr. 108, 103551.

Biehl, Alec, Ermagun, Alireza, Stathopoulos, Amanda, 2018. Modelling determinants of walking and cycling adoption: A stage-of-change perspective. Transp. Res. Part F: Traffic Psychol. Behav. 58, 452–470.

Blei, David M., Ng, Andrew Y., Jordan, Michael I., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3, 993–1022.

Bones, Emma J., Barrella, Elise M., Amekudzi, Adjo A., 2013. Implementation of evidence-based design approaches in transportation decision making. Transp. Res. Part A: Policy Pr. 49, 317–328.

Borowski, E., Chen, Y., Mahmassani, H., 2020. Social media effects on sustainable mobility opinion diffusion: Model framework and implications for behavior change. Travel. Behav. Soc. 19, 170–183.

Böyük, Mustafa, 2024. User data and digital privacy: Privacy policies of social media platforms. Turk. Online J. Des. Art Commun. 15 (1), 225–239.

Burke, Mary A., Young, H. Peyton, 2011. Social norms. In: Handbook of social economics, vol. 1, Elsevier, pp. 311–338.

Campello, Ricardo J.G.B., Moulavi, Davoud, Sander, Jörg, 2013. Density-based clustering based on hierarchical density estimates. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, pp. 160–172.

Caspari, Adam, Wood, Daniel, Campbell, Angel, Jefferson, Darian, Huynh, Tuan, Reddy, Alla, 2021. Using real-time data to detect delays and improve customer communications at new york city transit. Transp. Res. Rec. 2675 (7), 45–57.

Chen, Shang-Yu, 2016. Using the sustainable modified tam and tpb to analyze the effects of perceived green value on loyalty to a public bike system. Transp. Res. Part A: Policy Pr. 88, 58–72.

Chen, Mu-Yen, Chen, Ting-Hsuan, 2019. Modeling public mood and emotion: Blog and news sentiment and socio-economic phenomena. Future Gener. Comput. Syst. 96, 692–699.

Chen, Chao, Feng, Tao, Gu, Xiaoning, Yao, Baozhen, 2022. Investigating the effectiveness of covid-19 pandemic countermeasures on the use of public transport: A case study of the netherlands. Transp. Policy 117, 98–107.

Chen, Emily, Lerman, Kristina, Ferrara, Emilio, et al., 2020. Tracking social media discourse about the covid-19 pandemic: Development of a public coronavirus twitter data set. JMIR Public Health Surveill. 6 (2), e19273.

Chen, Ying, Mahmassani, Hani S., Frei, Andreas, 2018. Incorporating social media in travel and activity choice models: conceptual framework and exploratory analysis. Int. J. Urban Sci. 22 (2), 180–200.

Chinen, Kenichiro, Matsumoto, Mitsutaka, Liu, Chang, Tong, Pingsheng, Han, Yongliang Stanley, Endo, Hideki, 2023. Expanding electric-vehicle adoption beyond the national border: Insights for developing marketing policies for global electric-vehicle manufacturers. Int. J. Sustain. Transp. 17 (4), 393–405.

Cinelli, Matteo, Quattrociocchi, Walter, Galeazzi, Alessandro, Valensise, Carlo Michele, Brugnoli, Emanuele, Schmidt, Ana Lucia, Zola, Paola, Zollo, Fabiana, Scala, Antonio, 2020. The covid-19 social media infodemic. Sci. Rep. 10 (1), 16598.

Citi-bike, 2019–2022. Citi-bike: cycling data. URL https://citibikenyc.com/system-data.

De Vos, Jonas, 2018. Do people travel with their preferred travel mode? analysing the extent of travel mode dissonance and its effect on travel satisfaction. Transp. Res. Part A: Policy Pr. 117, 261–274.

De Vos, Jonas, Singleton, Patrick A., 2020. Travel and cognitive dissonance. Transp. Res. Part A: Policy Pr. 138, 525–536.

Ding, Xiaowen, Liu, Bing, Yu, Philip S., 2008. A holistic lexicon-based approach to opinion mining. In: Proceedings of the 2008 International Conference on Web Search and Data Mining. pp. 231–240.

Dütschke, Elisabeth, Engel, Lukas, Theis, Anna, Hanss, Daniel, 2022. Car driving air travel or more sustainable transport? socio-psychological factors in everyday mobility and long-distance leisure travel. Travel. Behav. Soc. 28, 115–127.

Dwityas, Nindyta Aisyah, Briandana, Rizki, 2017. Social media in travel decision making process. Int. J. Humanit. Soc. Sci. 7 (7), 193–201.

Festinger, L., 1957. A Theory of Cognitive Dissonance. volume 2 stanford University Press.

FHV, 2019–2022. Fhv: ride services data. URL https://www.nyc.gov/site/tlc/about/fhv-trip-record-data.page.

Fishbein, Martin, 1979. A Theory of Reasoned Action: Some Applications and Implications. University of Nebraska Press.

Gong, Wenjing, Rui, Jin, Li, Tianyu, 2024. Deciphering urban bike-sharing patterns: An in-depth analysis of natural environment and visual quality in new york's citi bike system. J. Transp. Geogr. 115, 103799.

Gota, Sudhir, Huizenga, Cornie, Peet, Karl, Medimorec, Nikola, Bakker, Stefan, 2019. Decarbonising transport to achieve paris agreement targets. Energy Effic. 12 (2), 363–386.

Grant-Muller, Susan M., Gal-Tzur, Ayelet, Minkov, Einat, Nocera, Silvio, Kuflik, Tsvi, Shoor, Itay, 2015. Enhancing transport data collection through social media sources: methods, challenges and opportunities for textual data. IET Intell. Transp. Syst. 9 (4), 407–417.

Grootendorst, Maarten, 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure. arXiv preprint arXiv:2203.05794.

Hackl, Roland, Raffler, Clemens, Friesenecker, Michael, Kramar, Hans, Kalasek, Robert, Soteropoulos, Aggelos, Wolf-Eberl, Susanne, Posch, Patrick, Tomschy, Rupert, 2019. Promoting active mobility: Evidence-based decision-making using statistical models. J. Transp. Geogr. 80, 102541.

Hamidi, Zahra, Zhao, Chunli, 2020. Shaping sustainable travel behaviour: Attitude, skills, and access all matter. Transp. Res. Part D: Transp. Environ. 88, 102566.

Hanlon, Annmarie, Jones, Karen, 2023. Ethical concerns about social media privacy policies: do users have the ability to comprehend their consent actions? J. Strat. Mark. 1–18.

Hasan, Samiul, Ukkusuri, Satish V., 2014. Urban activity pattern classification using topic models from online geo-location data. Transp. Res. Part C: Emerg. Technol. 44, 363–381.

He, Sylvia Y., Cheung, Yannie H.Y., Tao, Sui, 2018. Travel mobility and social participation among older people in a transit metropolis: A socio-spatial–temporal perspective. Transp. Res. Part A: Policy Pr. 118, 608–626.

Huan, Ning, Hess, Stephane, Yao, Enjian, 2022. Understanding the effects of travel demand management on metro commuters' behavioural loyalty: a hybrid choice modelling approach. Transportation 49 (2), 343–372.

Humagain, Prasanna, Vos, Jonas De, Singleton, Patrick A., 2021. Analyzing travel captivity by measuring the gap in travel satisfaction between chosen and alternative commute modes. Transp. Res. Part D: Transp. Environ. 97, 102965.

Jain, Taru, Currie, Graham, Aston, Laura, 2022. Covid and working from home: Long-term impacts and psycho-social determinants. Transp. Res. Part A: Policy Pr. 156, 52–68.

Jelodar, Hamed, Wang, Yongli, Yuan, Chi, Feng, Xia, Jiang, Xiahui, Li, Yanchao, Zhao, Liang, 2019. Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey. Multimedia Tools Appl. 78, 15169–15211.

Jin, Scarlett T., Kong, Hui, Sui, Daniel Z., 2019. Uber, public transit, and urban transportation equity: a case study in new york city. Prof. Geogr. 71 (2), 315–330.

Jost, John T., Barberá, Pablo, Bonneau, Richard, Langer, Melanie, Metzger, Megan, Nagler, Jonathan, Sterling, Joanna, Tucker, Joshua A., 2018. How social media facilitates political protest: Information, motivation, and social networks. Political Psychol. 39, 85–118.

Kim, Sung Hoo, Mokhtarian, Patricia L., 2018. Taste heterogeneity as an alternative form of endogeneity bias: Investigating the attitude-moderated effects of built environment and socio-demographics on vehicle ownership using latent class modeling. Transp. Res. Part A: Policy Pr. 116, 130–150.

Kroesen, Maarten, Chorus, Caspar, 2020. A new perspective on the role of attitudes in explaining travel behavior: a psychological network model. Transp. Res. Part A: Policy Pr. 133, 82–94.

Kroesen, Maarten, Handy, Susan, Chorus, Caspar, 2017. Do attitudes cause behavior or vice versa? an alternative conceptualization of the attitude-behavior relationship in travel behavior modeling. Transp. Res. Part A: Policy Pr. 101, 190–202.

Landauer, Thomas K., Foltz, Peter W., Laham, Darrell, 1998. An introduction to latent semantic analysis. Discourse Process. 25 (2–3), 259–284.

Lind, Hans Brende, Nordfjæ rn, Trond, Jørgensen, Stig Halvard, Rundmo, Torbjørn, 2015. The value-belief-norm theory personal norms and sustainable travel mode choice in urban areas. J. Environ. Psychol. 44, 119–125.

Liu, Chang, Bardaka, Eleni, Paschalidis, Evangelos, 2023. Sustainable transport choices in public transit access: Travel behavior differences between university students and other young adults. Int. J. Sustain. Transp. 17 (6), 679–695.

Liu, Xu, Dijk, Marc, 2022. How more data reinforces evidence-based transport policy in the short and long-term: Evaluating a policy pilot in two dutch cities. Transp. Policy 128, 166–178.

Liu, Xuerui, Mehraliyev, Fuad, Liu, Chun, Schuckert, Markus, 2020. The roles of social media in tourists' choices of travel components. Tour. Stud. 20 (1), 27–48.

Lorenz-Spreen, Philipp, Mønsted, Bjarke Mørch, Hövel, Philipp, Lehmann, Sune, 2019. Accelerating dynamics of collective attention. Nat. Commun. 10 (1), 1759.

McInnes, Leland, Healy, John, Melville, James, 2018. Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.

Meyer, John Robert, Kain, John F., Wohl, Martin, 1965. The Urban Transportation Problem. Harvard University Press.

Mo, Baichuan, Xu, Hanyong, Zhuang, Dingyi, Ma, Ruoyun, Guo, Xiaotong, Zhao, Jinhua, 2023. Large language models for travel behavior prediction. arXiv preprint arXiv:2312.00819.

Molin, Eric, Mokhtarian, Patricia, Kroesen, Maarten, 2016. Multimodal travel groups and attitudes: A latent class cluster analysis of dutch travelers. Transp. Res. Part A: Policy Pr. 83, 14–29.

Möllers, Alessa, Specht, Sebastian, Wessel, Jan, 2022. The impact of the covid-19 pandemic and government intervention on active mobility. Transp. Res. Part A: Policy Pr. 165, 356–375.

Moon, Sun-Jung, 2021. Investigating beliefs and attitudes and intentions regarding green restaurant patronage: An application of the extended theory of planned behavior with moderating effects of gender and age. Int. J. Hosp. Manag. 92, 102727.

MPN, 2019–2022. Mpn: Dutch mobility data. URL https://www.mpndata.nl/.

MTA, 2019–2022. Mta: subway data. URL https://new.mta.info/open-data.

Mulley, Corinne, Reedy, Liz, 2015. Research into policy: A case study of improving the research evidence base for transport policy makers in nsw, australia. Case Stud. Transp. Policy 3 (2), 215–221.

Nikolić, Tatjana Mamula, Pantić, Sanja Popović, Paunović, Ivan, Filipović, Sanja, 2021. Sustainable travel decision-making of europeans: Insights from a household survey. Sustainability 13 (4), 1960.

Olfindo, Rosechin, 2021. Transport accessibility, residential satisfaction, and moving intention in a context of limited travel mode choice. Transp. Res. Part A: Policy Pr. 145, 153–166.

Panagiotopoulos, Panos, Barnett, Julie, Bigdeli, Alinaghi Ziaee, Sams, Steven, 2016. Social media in emergency management: Twitter as a tool for communicating risks to the public. Technol. Forecast. Soc. Change 111, 86–96.

Park, June Young, Mistur, Evan, Kim, Donghwan, Mo, Yunjeong, Hoefer, Richard, 2022. Toward human-centric urban infrastructure: Text mining for social media data to identify the public perception of covid-19 policy in transportation hubs. Sustain. Cities Soc. 76, 103524.

Parkany, Emily, Gallagher, Ryan, Viveiros, Phillip, 2004. Are attitudes important in travel choice? Transp. Res. Rec. 1894 (1), 127–139.

Paulssen, Marcel, Temme, Dirk, Vij, Akshay, Walker, Joan L., 2014. Values attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. Transportation 41, 873–888.

Piligrimienė, Žaneta, Žukauskaitė, Andželika, Korzilius, Hubert, Banytė, Jūratė, Dovalienė, Aistė, 2020. Internal and external determinants of consumer engagement in sustainable consumption. Sustainability 12 (4), 1349.

Pineda-Jaramillo, Juan, Pineda-Jaramillo, Daniel, 2022. Analysing travel satisfaction of tourists towards a metro system from unstructured data. Res. Transp. Bus. Manag. 43, 100746.

Pisoni, Enrico, Christidis, Panayotis, Cawood, E. Navajas, 2022. Active mobility versus motorized transport? user choices and benefits for the society. Sci. Total Environ. 806, 150627.

Pronello, Cristina, Camusso, Cristian, 2011. Travellers' profiles definition using statistical multivariate analysis of attitudinal variables. J. Transp. Geogr. 19 (6), 1294–1308.

Qi, Bing, Costin, Aaron, Jia, Mengda, 2020. A framework with efficient extraction and analysis of twitter data for evaluating public opinions on transportation services. Travel. Behav. Soc. 21, 10–23.

Rahman, Mashrur, Sciara, Gian-Claudia, 2022. Travel attitudes, the built environment and travel behavior relationships: Causal insights from social psychology theories. Transp. Policy 123, 44–54.

Rahman, Rezaur, Shabab, Kazi Redwan, Roy, Kamol Chandra, Zaki, Mohamed H., Hasan, Samiul, 2021. Real-time twitter data mining approach to infer user perception toward active mobility. Transp. Res. Rec. 2675 (9), 947–960.

Rashidi, Taha H., Abbasi, Alireza, Maghrebi, Mojtaba, Hasan, Samiul, Waller, Travis 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.

Reimers, Nils, Gurevych, Iryna, 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. arXiv preprint arXiv:1908.10084.

Rodríguez-Mazahua, Lisbeth, Rodríguez-Enríquez, Cristian-Aarón, Sánchez-Cervantes, José Luis, Cervantes, Jair, García-Alcaraz, Jorge Luis, Alor-Hernández, Giner, 2016. A general perspective of big data: applications, tools, challenges and trends. J. Supercomput. 72, 3073–3113.

Ru, Xingjun, Wang, Shanyong, Chen, Qian, Yan, Shuai, 2018. Exploring the interaction effects of norms and attitudes on green travel intention: An empirical study in eastern china. J. Clean. Prod. 197, 1317–1327.

Scheiner, Joachim, Holz-Rau, Christian, 2007. Travel mode choice: affected by objective or subjective determinants? Transportation 34, 487–511.

Schweitzer, Lisa, 2014. Planning and social media: a case study of public transit and stigma on twitter. J. Am. Plan. Assoc. 80 (3), 218–238.

Shaer, Amin, Rezaei, Meysam, Rahimi, Behnam Moghani, Shaer, Fatemeh, 2021. Examining the associations between perceived built environment and active travel, before and after the covid-19 outbreak in shiraz city, iran. Cities 115, 103255.

Shahriari, Siroos, Abbasi, Alireza, Rashidi, Taha Hossein, 2024. Exploiting social media to better understand travel behaviour and mobility. In: Handbook of Travel Behaviour, vol. 12, (219).

Shamshiripour, Ali, Rahimi, Ehsan, Shabanpour, Ramin, 2020. And abolfazl kouros mohammadian. How is covid-19 reshaping activity-travel behavior? evidence from a comprehensive survey in chicago. Transp. Res. Interdiscip. Perspect. 7, 100216.

Smith-Colin, Janille, Fischer, Jamie Montague, Akofio-Sowah, Margaret-Avis, Amekudzi-Kennedy, Adjo, 2014. Evidence-based decision making for transportation asset management: Enhancing the practice with quality evidence and systematic documentation. Transp. Res. Rec. 2460 (1), 146–153.

Stieglitz, Stefan, Dang-Xuan, Linh, 2013. Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. J. Manage. Inf. Syst. 29 (4), 217–248.

Susilo, Yusak O., Cats, Oded, 2014. Exploring key determinants of travel satisfaction for multi-modal trips by different traveler groups. Transp. Res. Part A: Policy Pr. 67, 366–380.

Taube, Oliver, Kibbe, Alexandra, Vetter, Max, Adler, Maximilian, Kaiser, Florian G., 2018. Applying the campbell paradigm to sustainable travel behavior: Compensatory effects of environmental attitude and the transportation environment. Transp. Res. Part F: Traffic Psychol. Behav. 56, 392–407.

TLC, 2019–2022. Tlc: yellow taxi data. URL https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page.

Touvron, Hugo, Lavril, Thibaut, Izacard, Gautier, Martinet, Xavier, Lachaux, Marie-Anne, Lacroix, Timothée, Rozière, Baptiste, Goyal, Naman, Hambro, Eric, Azhar, Faisal, et al., 2023. Llama: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971.

New York City Department of Transportation, 2022. New York City Congestion Management System Report. Report, New York City Department of Transportation (NYC DOT, URL https://www.nyc.gov/html/dot/downloads/pdf/2022-cms-report.pdf. (Accessed 25 October 2023).

Vu, Huy Quan, Li, Gang, Law, Rob, 2019. Discovering implicit activity preferences in travel itineraries by topic modeling. Tour. Manag. 75, 435–446.

Wang, Yacan, Douglas, Matthew, Hazen, Benjamin, 2021. Diffusion of public bicycle systems: Investigating influences of users' perceived risk and switching intention. Transp. Res. Part A: Policy Pr. 143, 1–13.

Wang, Haoyun, Noland, Robert B., 2021. Bikeshare and subway ridership changes during the covid-19 pandemic in new york city. Transp. Policy 106, 262–270.

Wang, Shanyong, Wang, Jing, Li, Jun, Wang, Jinpeng, Liang, Liang, 2018. Policy implications for promoting the adoption of electric vehicles: do consumer's knowledge, perceived risk and financial incentive policy matter? Transp. Res. Part A: Policy Pr. 117, 58–69.

Wang, Zhiqi, Zhang, Yufeng, Jia, Bin, Gao, Ziyou, 2024. Comparative analysis of usage patterns and underlying determinants for ride-hailing and traditional taxi services: A chicago case study. Transp. Res. Part A: Policy Pr. 179, 103912.

Wei, Keji, Vaze, Vikrant, Jacquillat, Alexandre, 2022. Transit planning optimization under ride-hailing competition and traffic congestion. Transp. Sci. 56 (3), 725–749.

Xin, Rui, Ai, Tinghua, Ding, Linfang, Zhu, Ruoxin, Meng, Liqiu, 2022. Impact of the covid-19 pandemic on urban human mobility-a multiscale geospatial network analysis using new york bike-sharing data. Cities 126, 103677.

Yang, Hongyang, Liu, Xiao-Yang, Wang, Christina Dan, 2023. Fingpt: Open-source financial large language models. arXiv preprint arXiv:2306.06031.

Yazici, M. Anil, Kamga, Camille, Singhal, Abhishek, 2016. Modeling taxi drivers' decisions for improving airport ground access: John f. kennedy airport case. Transp. Res. Part A: Policy Pr. 91, 48–60.

Yeung, Douglas, 2018. Social media as a catalyst for policy action and social change for health and well-being. J. Med. Internet Res. 20 (3), e94.

Yin, Ganmin, Huang, Zhou, Fu, Chen, Ren, Shuliang, Bao, Yi, Ma, Xiaolei, 2024. Examining active travel behavior through explainable machine learning: Insights from beijing, China. Transp. Res. Part D: Transp. Environ. 127, 104038.

Zhai, Wei, Bai, Xueyin, Peng, Zhong-ren, Gu, Chaolin, 2019. A bottom-up transportation network efficiency measuring approach: A case study of taxi efficiency in new york city. J. Transp. Geogr. 80, 102502.

Zhao, Chunli, 2018. Thomas alexander sick nielsen, anton stahl olafsson, trine agervig carstensen, and xiaoying meng. Urban form, demographic and socio-economic correlates of walking, cycling, and e-biking: Evidence from eight neighborhoods in beijing. Transp. Policy 64, 102–112.

Zhou, Yuyang, Wang, Peiyu, Zheng, Shuyan, Zhao, Minhe, Lam, William H.K., Chen, Anthony, Sze, N.N., Chen, Yanyan, 2024. Modeling dynamic travel mode choices using cumulative prospect theory. Transp. Res. Part A: Policy Pr. 179, 103938.