Journal of Air Transport Management 132 (2026) 102950

Contents lists available at ScienceDirect

Journal of Air Transport Management

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

Shifting skies: A cross-country investigation of evolution of public perception toward urban air mobility through Twitter (X) discourse

Eeshan Bhaduri[1], Charisma F. Choudhury[*,2]

Institute for Transport Studies, University of Leeds, UK

ARTICLE INFO

Keywords:
Urban air mobility (UAM)
Public perception
Spatio-temporal heterogeneity
Social media data
Machine learning
Large language model
Dynamic topic modelling
Time series analysis
K-means clustering

ABSTRACT

Urban air mobility (UAM) is increasingly being recognised as a promising response to the challenges of rapid urban expansion and its negative externalities. While technological advancements in vertical take-off and landing (VTOL) aircraft have accelerated development in this space, the widespread adoption of UAM services hinges on societal acceptance driven by public perceptions. Understanding these perceptions, especially their variation across regions and over time, is critical for developing policies to maximise their adoption rate. This study leverages a large-scale and long-term Twitter dataset to discern the spatio-temporal evolution of public perceptions towards UAM. To this end, we employed a combination of machine learning (ML) and a large language model (LLM) for performing sentiment classification. Subsequently, sentiment polarities are integrated with time series analysis, indicating the prevalence of positive perception for most of the last decade, while detecting the effect of various real-world events. In terms of spatial K-means clustering results, it reveals four clusters of countries with distinct characteristics. For example, people in countries like the USA and Australia are observed to be highly opinionated towards UAM, while public discourse in Germany and India is more neutral. Finally, dynamic topic modelling coupled with an LLM-based representation uncovers underlying themes of public discourse. Topic model findings underline three major global themes: (1) industry innovation and testing, (2) unmanned aviation systems, and (3) mobility benefits. Furthermore, we identified in some cases that local themes driven by specific incidents have a more substantial effect in shaping the preferences than the generic global ones. The paper hence contributes to the literature by providing the first global-level dynamic spatio-temporal assessment of future UAM services. The insights are expected to offer valuable policy guidance for policymakers, regulators, and industry stakeholders aiming to improve the public acceptance of UAM technologies and consequently the uptake.

1. Introduction and motivation

Over half of the global population already resides in urban areas, and the United Nations forecasts a substantial rise of this percentage by 2050, ranging from 50% in low-income countries to 88% in high-income ones (United Nations, 2018). Unsurprisingly, such rapid urban expansion has been putting an increasing strain on transport infrastructure, exacerbating congestion and, in turn, causing economic inefficiencies. In this connection, urban air mobility (UAM) has emerged as a promising solution whose impact transcends the aviation sector, revolutionising the mobility systems and redefining urban planning paradigms (Pons-Prats et al., 2022). Although UAM involves both passenger and cargo transportation, the prevailing consensus from the UAM service providers indicates that the most viable (and near future) application is as an on-demand, shared or exclusive use, inter or intra-city air-passenger service (commonly known as *flying taxi* or *air taxi* and henceforth the terminology *UAM* denotes the passenger services only) (Long et al., 2023). On the technology side, the vertical take-off and landing (VTOL) aircraft, superior to their conventional counterparts (e.g., helicopter) in electric propulsion and autonomy, have garnered considerable public attention as the major facilitator of passenger UAM operation (Cohen et al., 2021). Consequently, the public attention would be the most critical precursor determining the *social acceptance* of UAM services and encompasses both enthusiasm and apprehension, each of which entails a multitude of perceptions such as service performance (Rothfeld et al., 2021; Samadzad et al., 2024) and disruption due to UAM operation (Garrow et al., 2025; Yedavali and Mooberry, 2019). In fact, Pons-Prats et al. (2022) pointed

* Corresponding author.
E-mail addresses: E.Bhaduri@leeds.ac.uk, eeshanbhaduri@gmail.com (E. Bhaduri), C.F.Choudhury@leeds.ac.uk (C.F. Choudhury).

[1] Research Fellow.
[2] Professor.

https://doi.org/10.1016/j.jairtraman.2025.102950
Received 17 May 2025; Received in revised form 16 October 2025; Accepted 28 November 2025
Available online 2 December 2025

out that more than the technological enablers, the societal acceptance dimension needs to be addressed to facilitate initial UAM adoption.

To systematically assess the public perception, many previous studies looked at it as a proxy of consumer demand. This led to perception being linked to a pre-defined set of mode-specific attributes (e.g., travel time, cost) and individual socio-demographic characteristics coupled with attitudes and lifestyle. Notwithstanding the relevance of these studies, the perception drivers for new modes could still be very different from existing ones. More importantly, perceptions are inherently dynamic, evolving over time and regions (Manca et al., 2023). In particular, it is speculated that the heterogeneity in the public's perception can lead to different willingness to pay, adoption rate, and, consequently, different behaviour towards the new unseen modes around the globe. Traditionally, attitudinal statements in revealed and stated preference surveys have been used to capture perceptions towards innovations and new technologies (Al-Haddad et al., 2020). These, however, can be limiting for their cross-sectional nature, as well as sample size and survey bias.

In contrast, the growing number of social media users has made data extracted from social media platforms an increasingly reliable source to capture public opinion in a passive manner. This massive data source could delineate different behavioural aspects of users, which can be tracked over time to understand the attitudes and perceptions of users towards certain events, topics, or concepts (Kwon et al., 2024). Furthermore, the global usage of social media platforms has made it possible to access data from users with different backgrounds in different countries, which is of particular interest when the heterogeneity among users might play a role in their attitudes and preferences (Rashidi et al., 2017). This motivates the current research to develop one of the first dynamic spatio-temporal assessments of public perception towards future UAM services.

In the present study, we collected a long-term geo-tagged Twitter dataset pertaining to UAM-related discussions. Such a tweet corpus, despite potential non-representativeness issues, enables us to capture both spatial and temporal heterogeneity in the evolution of public sentiments. Non-geotagged tweet datasets, albeit potentially having less bias, were excluded due to the missing spatial dimension of the sentiments. Hence, the analyses and findings are expected to advance the state-of-the-art knowledge in two distinct dimensions. *Firstly*, assessing the spatio-temporal heterogeneity in UAM-related sentiment polarities leveraging large language models (BERT) and machine learning (ML) classifiers, which can provide an understanding of potential global markets. *Secondly*, identifying the latent public discourse themes through a large language model (LLM)-based topic representation (BERTopic with OpenAI) and subsequently visualising the dynamicity. This can be useful to detect several policy levers (motivators and deterrents) and their relative effects within and across regions, leading to context-specific decision-making. To the best of our knowledge, this is the first of its kind in the transportation context, particularly for analysing perception towards futuristic alternatives.

The rest of this paper is structured as follows: Section 2 presents a summary of relevant studies in the literature, followed by Section 3 explaining the data collection and filtering, along with providing a preliminary description. Section 4 delineates the procedures employed for data analysis, encompassing data cleansing, annotation, and finally employing sentiment analysis techniques. Section 5 presents sentiment classification results and further examines the temporal and geographical heterogeneity. Subsequently, the aggregated results of the dynamic topic model across the regions have been presented. Finally, Section 6 offers comprehensive policy insights derived from the topic model findings, and future research directions are discussed in Section 7.

2. Literature review and research contribution

UAM technologies are associated with various uncertainties and challenges that extend beyond technical performance. In particular, public acceptance has been identified as a critical barrier and may constitute the principal impediment to the widespread deployment of UAM systems in urban environments (Straubinger et al., 2020). Therefore, in this section, we critically examine and assess the previous studies exploring potential UAM demand. Notably, most of these studies have been done in the context of the USA and Europe, while the datasets are largely cross-sectional, especially for survey-based ones (Long et al., 2023; Garrow et al., 2021). Drawing on differences in data sources and methodological approaches, existing demand assessment techniques can be broadly classified into two strands: *firstly,* public acceptance studies relying on survey-based assessment of individual perception, and *secondly,* global market studies based on macro-level economic and travel datasets. The *public acceptance* studies attempt to understand UAM adoption proclivity from the perspective of revealed and/or stated preferences, existing travel behaviour, perceptions, attitudes, and socio-demographic characteristics. Understandably, these studies are more granular in nature, though often are unable to shed light on evolving perception and differences across regions. In contrast, the macro-level market studies are inherently adaptable across various temporal and spatial contexts, but lack a nuanced understanding of public expectations of such emerging travel modes, posing the risk of inaccurate demand estimation.

2.1. Public acceptance studies

Given that Urban Air Mobility is still in its infancy, public acceptance studies predominantly relied on either stated-preference (SP) or opinion survey methods as critical tools for capturing prospective user behaviour and attitudes towards UAM services. SP survey approach elicits respondents' choices among hypothetical travel alternatives and their perceptions regarding key attributes, which are otherwise not well captured in high-level market studies. A critical examination of the previous studies reveals that user preferences are *assumed to be shaped* primarily by two overarching dimensions: *trip-specific* factors, such as expected travel time and service accessibility, and *alternative-related* perception, including notions of safety, noise exposure, excitement, and potential visual intrusion. The dual insights provided by SP surveys — capturing both choice behaviour and attitudinal dimensions — have therefore been valuable in characterising potential early adopters of UAM services.

Several studies conducted by academics, consulting firms, and vehicle manufacturers have sought to assess the importance of the above-mentioned factors on public adoption readiness (See Table 1). For example, Al-Haddad et al. (2020) in their multi-country study (Europe, USA, South America, and the Middle East; N=221), observed safety as the most critical attribute governing UAM adoption, followed by trip cost, time, service reliability, and operation characteristics. Also, the data and ethical concerns (fear of cybersecurity, data theft by third parties, and job loss) observed to have a negative influence on UAM adoption. At the same time, the authors highlighted the impact of cultural paradigms through the proxy of survey language; however, their inferences were limited to Germany only because of a sample representation issue. Pons-Prats et al. (2022) performed a qualitative review to point out several perception attributes (e.g., safety, noise, privacy, visual disruption) and trip-related factors (e.g., travel time, cost, last-mile connectivity) as critical public acceptance determinants in addition to technologies and implementation issues (e.g., airspace integration, low-altitude flights). Furthermore, they argued in favour of exploring the evolution of perception in the context of new and improved transportation services (e.g., autonomous vehicles, high-speed rails, mobility as a service). Another study done by Haan et al. (2021) attempted to measure the air taxi demand across 40 cities in the USA (N=1405) and observed that nearly one-third of air taxi demand is concentrated in three cities. Moreover, their findings also highlighted that uptake propensity depends on spatial factors such as existing ground infrastructure and congestion levels. Garrow et al. (2025) conducted

Table 1
Summary of selected previous studies on public acceptance of UAM services.

Study	Data source	Sample size	Spatial coverage	Time Period	Methodology	Key findings
Garrow et al. (2025)	Internet-based opinion survey	2499 individuals	5 USA cities	April–June 2018	Factor analysis and cluster analysis	• Enthusiasm and concern emerge as two key perception levers influencing eVTOL adoption. • Cluster analysis revealed that adoption likelihood differs based on individual travel behaviour, which is shaped by the supply of transport infrastructure.
Al-Haddad et al. (2020)	SP survey (convenience sample)	221 individuals	Europe, USA, South America, and Middle East	July–September 2018	Factor analysis and choice model	• Safety was perceived as the most important factor, followed by trip cost, trip duration, service reliability and operation characteristics. • Culture impacts perception, as German respondents expressed a lower interest in early adoption and a higher scepticism.
Samadzad et al. (2024)	SP survey	3 separate case studies with sample sizes 214, 323, and 239	Multiple cities in Iran	June–July 2022	Factor analysis and choice model	• UAM demand depends on trip purpose and geographical contexts, with the most viable operation segment being inter-city business trips. • Individual attitudes to existing transport modes affect their perception towards UAM services.
Yedavali and Mooberry (2019)	Online opinion survey	1540 individuals	Los Angeles, Mexico City, New Zealand and Switzerland	NA	Descriptive analysis	• Public perception relates to five factors: safety, noise, equity, visual pollution and privacy. • In addition to individual perception (particularly safety) and lifestyle attitudes, high-level operational factors influence public perception.
Wisk (2021)	Online interview and focus group survey	NA	30 DMAs in USA	NA	Descriptive analysis	• Public perception varies depending on regional factors such as city size and commute duration. • The excitement for UAM services in terms of variety-seeking and fun could play a key role in facilitating early adoption.

a follow-up study across five cities in the USA (N=2499), where the authors performed a latent segmentation of the potential UAM market based on individual socio-demographics, existing travel habits, and UAM-related enthusiasm and concerns. Notably, concerns surrounding eVTOL technology are primarily rooted in public scepticism regarding the safety of UAM and the potential risks associated with battery-operated aircraft. Nevertheless, they indicated that their results might not be generalisable beyond the regional boundary, subject to over-sampling (high income respondents) bias, and do not capture change in perception over time, which are critical for entirely hypothetical alternatives. Straubinger et al. (2021) argued that UAM service characteristics have a significantly higher decisive impact on corresponding mode choice and welfare effect as compared to the existing urban structure. Utilising a survey administered by Uber to residents of two USA cities (N=2419), Song et al. (2024) developed a latent class model to examine travel behaviour. Their findings revealed significant differences in adoption based on two individual variety-seeking natures, i.e., novelty-seeking and alternation, with the higher degree of the former leading to a greater likelihood of switching to UAM modes from existing ones. Samadzad et al. (2024) employed an online SP survey (N=776) to explore UAM demand and value of time for three different trip purposes (business, airport access, and tourism) and geographical contexts in Iran. Their results not only indicated the role of individual perceptions (e.g., interest in novel technologies) and UAM-specific concerns (e.g., safety and privacy) but also underlined how travel-related attitudes (e.g., preference for one's own car) influence one's opinion. Finally, Sadrani et al. (2025) employed an expert survey (N=18) in Germany and the USA to conclude that barriers to UAM adoption significantly differ across regions. In Germany, there are concerns regarding price affordability, investment uncertainty, and user acceptance, while in the USA, primary challenges relate to airspace utilisation, technical operations, and safety issues.

On the industry/ regulator/ policymaker side, the European Union Aviation Safety Agency (EASA) conducted an online poll with participants (N= 3690) from six European cities to find safety, noise, and security as top-most concerns. At the same time, Europeans are observed to show a greater interest in drone delivery services (about 64%) compared to passenger transport services (49%) (EASA, 2021). Airbus did a market survey (N=1540) to evaluate public opinions of UAM services among inhabitants of the USA, Mexico City, New Zealand, and Switzerland. Apart from individual socio-demographics and travel habits, the public perception was found to be linked with not only hypothetical ride-experiences (e.g., safety, privacy) but also with broad operation (e.g., noise and visual pollution due to high frequency and low altitude flights) (Yedavali and Mooberry, 2019). The other consumer interest study by Wisk (2021) suggested nearly 99% of respondents find air-taxi extremely or somewhat appealing, with 95% of them at least somewhat likely to use the same. While these very high numbers are possibly due to the sample being entirely high-income individuals, the study provides a preliminary assessment of interest amongst the target user group. Moreover, it also pointed out the role of emotional needs (e.g., uniqueness, luxury) in addition to the trip-specific and alternative-related perceptions. Upon reviewing the factors and methodology employed in the studies mentioned above, the following research gaps could be identified:

- The studies largely suffer from *a priori* assumptions regarding the set of influencing factors governed by research requirements, instead of being open-ended and flexible. This can restrict identifying region-specific factors, and more importantly, fail to assess their evolution over time.
- Most surveys to date have been carried out in developed nations and regions (particularly in the USA), which has resulted in a notable lack of studies focused on developing countries.
- Intuitively, the sample sizes for most of the studies are limited, making their inferences less applicable to broader socio-economic strata. While the findings are relevant for the initial assessment of the consumer base, wider adoption requires a method to discern public opinion on a larger scale.

In this regard, recent perception studies, although not specifically conducted within the UAM domain, have leveraged diverse types of social media datasets, such as tweets (Ding et al., 2021; Wang et al., 2022), Reddit comments (Bakalos et al., 2020), Facebook comments (Chen et al., 2021), and YouTube comments (Li et al., 2018).

Table 2
Summary of selected previous studies on global UAM market operation.

Study	Data source	Spatial coverage	Time Period	Methods	Variables	Major findings
Becker et al. (2018)	Undirected air passenger demand data (ADI)	Global (4435 cities)	Projection year: 2042 and base year: 2012	Gravity model	Independent variables: airfare, GDP, PPP, distance and population; Binary variables: language, tourism attraction, intercontinental connection, capital region	• The study identified 26 potential markets for the inter-urban mobility system, divided into two major sub-groups based on business activities and congestion level.
Robinson et al. (2018)	Uber Elevate and FAA obstacle database	Miami, USA	–	Geo-density analysis	Independent variables: urban sprawl and density, water bodies, transport network, population income, industry, congestion, existing airports, climate	• The study identified 9517 potential airpark locations and discussed the implications of higher geodensity on infrastructure requirements.
Wai et al. (2021)	Grab and Singapore Tourism Board	Singapore	–	Fare-time analysis	Independent variables: travel time, cost, dustance	• The study suggested that the longer the trip, the more skewed the fare-time ratio is.
NEXA Advisors (2019)	ArcGIS	Global (74 cities)	2023–2045	Interactive dashboard	Independent variables: city demographics, infrastructure costs, vehicle and supply chain, demand assumptions, regulatory and community constraints	• The study developed outputs related to business potential, direct and indirect economic impact, and tax revenues related to UAM operation.
Hamilton (2018)	USA census, BTS, FAA environment design tool	USA	–	Interactive dashboard	Independent variables: population, density, existing transport network and congestion, climate, urban infrastructure	• The study identified safety concerns and unmanned operations as major deterrents for UAM adoption. • UAM adoption faces infrastructure barriers and market competition, which vary across geographic regions.

Nevertheless, none of these studies simultaneously examines the temporal and geographical heterogeneity. Furthermore, the limitations of their dataset size (limited to a few countries and/or a few years) significantly restrict the depth of analysis. Not to mention, this could substantially hinder the comprehension of the differential impacts of public discourse themes across diverse spatio-temporal epochs, thereby hindering the contextualisation of policy interventions.

2.2. Global market operation studies

Global market operation studies look into identifying regions and assessing infrastructure capacity for initial operation, as well as focusing on predicting UAM market penetration, which is often dependent on the former component. Table 2 summarises the existing literature on UAM market operation across prospective cities and regions. For instance, Becker et al. (2018) applied a gravity model to forecast interurban air passenger demand for 2042, incorporating socioeconomic indicators to identify potential UAM (including UAM) markets. The study by Robinson et al. (2018) delineated prospective cities utilising qualitative criteria, including urban sprawl, population density, availability of water bodies, existing airports, affluence, presence of high-tech industries, ground transportation congestion, transportation patterns, and climatic conditions. In another study done in the Asian context (Singapore), Wai et al. (2021) assessed preliminary feasibility for UAM demand and subsequently evaluated the advantages for tourists based on the fare-distance-time trade-off. Apart from the academic studies, several commercial entities have also been involved in similar predictive analysis. For example, NEXA Advisors (2019) developed a comprehensive demand model for multiple UAM use cases based on a diverse set of inputs such as population size and density, gross domestic product (GDP) per capita, socio-demographics, levels of commercial and business aviation activity, and the presence of Fortune 1000 companies to inform demand projections. KPMG (Mayor and Anderson, 2019) projected UAM demand by incorporating variables such as city-level GDP and GDP growth, population size and growth rates, population density, forecasted income distribution, patterns of wealth concentration, and the availability of existing ground transport services. In a technical brief to NASA, Booz Allen Hamilton selected the focus urban areas based on attributes like surface traffic congestion, weather conditions, existing infrastructure and transport, legal and regulatory issues, and airline demand (Hamilton, 2018). As mentioned earlier, a few studies have also developed travel demand frameworks to predict the market penetration of UAM. For example, Haan et al. (2021) developed a UAM choice model to forecast the volume of prospective commuters, leveraging a composite dataset that integrated cell phone mobility data with census information as inputs for calibration. Another study by Balac et al. (2019) developed an agent-based framework utilising household travel survey data and census data, while the framework demonstrated demand sensitivity with technological and operational factors.

Upon reviewing the factors and methodology employed in the studies mentioned above, the following research gaps could be identified:

- There is a lack of consensus regarding the selection of key variables influencing UAM demand, which leads to the inclusion of a diverse set of factors, and in most cases, an incomplete set.
- Most of the market operation studies, if not all, potentially overlook the importance of public acceptance (e.g., excitement, safety, noise, privacy), perceived benefits (e.g., time saving, accessibility) and regulatory constraints.
- Calibrated choice models based on existing mode dataset (e.g., taxi) make strict assumptions, which might lead to erroneous projection of UAM demand. Also, these models are sometimes tightly hinged on a specific geographic region and therefore unable to appreciate spatial heterogeneity.

In summary, both the public acceptance and market operation studies suggest that the adoption of UAM services is influenced by several context-specific factors, while the set of factors also varies across studies from different regions. Although both findings can benefit the UAM community separately, they either miss out on the characterisation of socio-cultural underpinnings or the evolution over time and

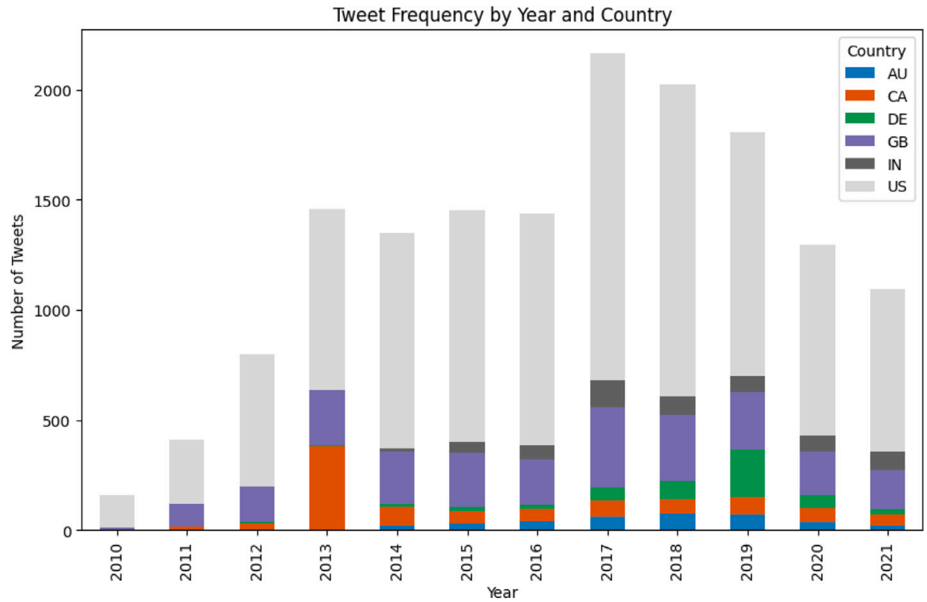

Fig. 1. Number of geo-tagged tweets across study countries during 2010–2021.

region. Acknowledging that both aspects are of utmost importance, the present study employs a (large-scale) data-driven exploratory approach to decipher sentiment polarities and the causal latent themes of public discourse in a joint manner.

3. Data

The Twitter data used in the research has been extracted based on specific keywords, location (if available), and timestamp using the Academic Research Twitter API (Application Programming Interface). Given that Twitter functions as a forum for the general public, we used terms used by the general public (e.g., flying car, flying taxi, air taxi, and air mobility) in addition to technical terms (e.g., UAM, VTOL, eV-TOL). In total, twenty-three keywords were used: flying taxi, flying car, flying vehicle, air taxi, advanced air mobility (AAM), aerial rideshare, urban air mobility (UAM), Urban Aeronautics, eVTOL, VTOL, flying ambulance, Urban Air Port, autonomous aircraft, autonomous plane, unmanned aircraft, unmanned aerial vehicle, Joby Aviation, Lilium Jet, Volocopter, VoloCity, CityHawk, CityAirBus, and Uber Air. Initially, we collected 16,701 tweets from twenty (20) countries; however, we removed countries with fewer than 300 qualified tweets. Finally, we filtered 15,442 English-language geo-tagged tweets from six (6) nations of interest (United States (US), United Kingdom (GB), Canada (CA), Germany (DE), India (IN), and Australia (AU)). Notably, the number of tweets about UAM peaked between 2017-18 and decreased afterwards, which has been discussed in Section 5 at a more granular level, along with possible stimulators (both positive and negative) events. Fig. 1 illustrates the distribution of retrieved tweets by country over the years. The chart indicates that tweets from accounts in the United States (10,539) constitute the largest share, followed by those from the United Kingdom (2513) and Canada (936). It is worth mentioning that the decline in the number of tweets could be a compounded effect of dwindling interest (or fading away of initial excitement) in UAM and an overall decrease in the number of active Twitter users.

4. Methodology

The present study explores two key dimensions of the public perception related to UAM- *polarity* and *underlying themes*. Therefore, it was necessary to concurrently employ both sentiment analysis and topic modelling methodologies. Evidently, natural language processing (NLP) techniques, including various machine learning (ML) models and large language models (LLMs), are most appropriate for both objectives. This section elucidates the step-by-step methodological approach (See Fig. 2) adopted by the current study.

4.1. Data processing and annotation

As a basic prerequisite, we executed a sequence of pre-processing procedures on the unstructured Twitter content, comprising data cleaning, tokenisation, stop-word elimination, and stemming and lemmatisation (Anderson et al., 2024). Notably, the first step, i.e., data cleaning, is critical to ensure success in assessing polarity (classification) as well as retaining meaningfulness (topic) of tweets. It involved four primary tasks: eliminating special characters and punctuation, converting text to lowercase, removing excess spaces, expanding abbreviations, and finally deleting hyperlinks and mentions.

The next immediate step was to perform data annotation. This phase saw 1000 randomly selected tweets annotated individually by three experts, reaching a moderate agreement level (Fleiss Kappa = 0.56) (Landis and Koch, 1977). It is worth noting that the aforementioned level of agreement is deemed acceptable, as a certain degree of annotation disagreement is expected when addressing hypothetical alternatives (Zapf et al., 2016). More importantly, this level of variation is not expected to significantly impact the performance of the language models developed in subsequent stages. Nevertheless, the initial annotated dataset was found to be imbalanced, i.e., a higher number of neutral (350) and off-topic (285) tweets relative to the other two categories (160 negative and 205 positive tweets). Importantly, such imbalances might lead to biased classification models, predominantly favouring the majority category while underperforming on the minority category (He and Ma, 2013). In the present study, we employed the *back-translation* method that generates supplementary instances of the minority class by translating the original text into another language and subsequently translating it back into the original language. This procedure may lead to nuanced differences in the text, facilitating an expansion of the minority class while maintaining the sentiment of the text (See Taheri et al. (2024) for details). Additionally, it contributed to the enhancement of the training data, which is essential given the limited size of the annotated dataset. The final dataset consisted of 1365 tweets, an increase from the 1000 tweets prior to the augmentation process. In the following subsections, we have presented the model

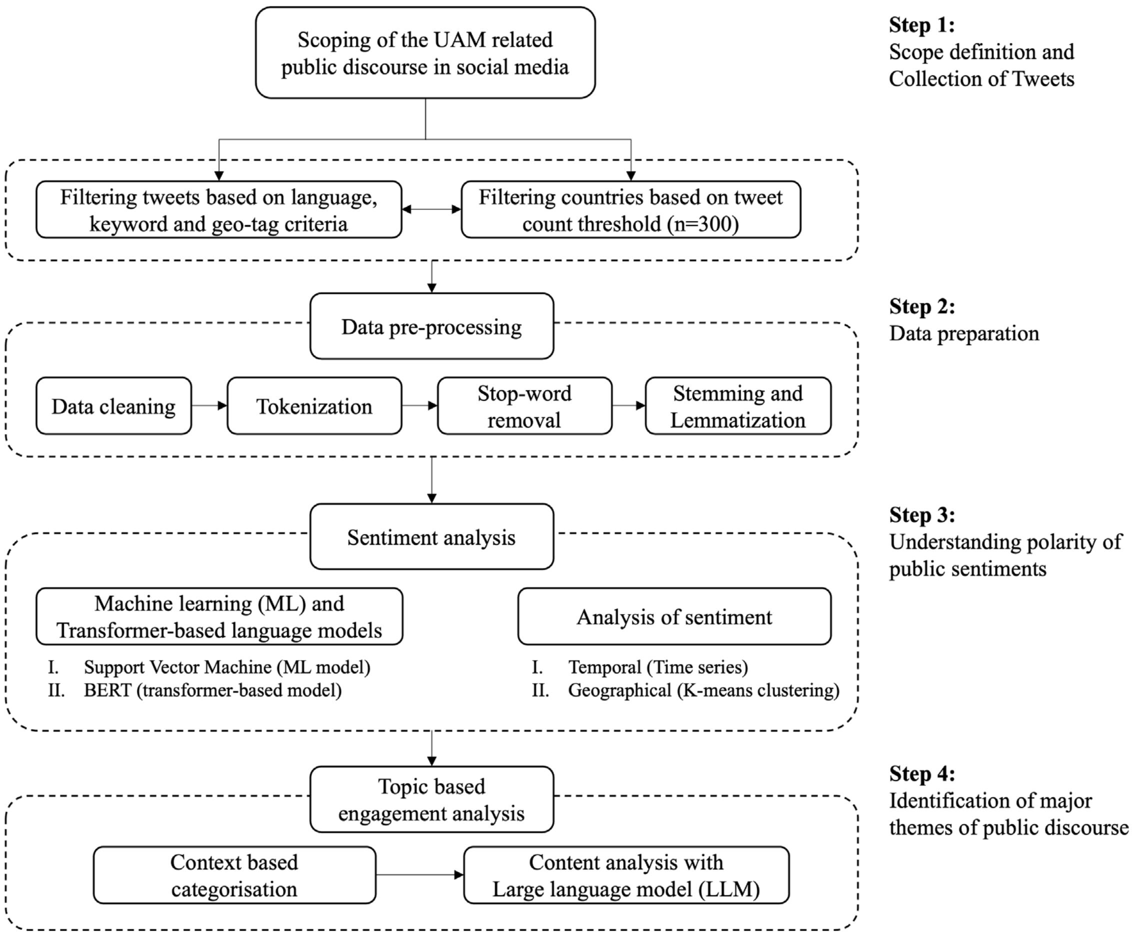

Fig. 2. Methodology flowchart.

outcomes only for the augmented (or balanced) dataset for brevity, as it outperformed the classification model with imbalanced data on every testing instance.

4.2. Sentiment analysis

Sentiment analysis involves systematically extracting and quantifying emotions, opinions, or attitudes from text or speech. By categorising expressions as positive, negative, or neutral, it offers an approximation of the sentiment conveyed (Qi and Shabrina, 2023). This section provides a brief overview and key methodological differences of the machine learning (ML) and large language model (LLM) techniques utilised in the present study.

4.2.1. Machine learning model

Machine learning models can be effectively trained on annotated datasets to automatically determine the sentiment expressed in tweets, typically classifying them into positive, negative, neutral, and off-topic categories. A range of well-established algorithms are commonly employed for this task, including Support Vector Machines (SVM), Random Forest (RF), and Naive Bayes (NB). These models learn to recognise patterns in textual features that correlate with different sentiment labels, making them suitable for large-scale analysis of social media data like Twitter. Among the three aforementioned ML techniques, we observed SVM to be the most superior one, and therefore, in the rest of the paper, we have solely focused on it for the sake of brevity. It is worth noting that SVM is a supervised learning algorithm widely used for classification tasks. It operates by finding an optimal hyperplane — either linear or non-linear — that best separates data points belonging to different classes, whether binary or multi-class (Hearst et al., 1998). The objective of SVM is to maximise the margin, defined as the distance between the hyperplane and the nearest data points from each class, referred to as support vectors. This optimal separation is known as the maximum margin hyperplane, which effectively partitions the dataset into distinct classes. In our study, we experimented with various regularisation parameters to strike an optimal balance between margin size, which reflects generalisation capability, and classification accuracy. A consistent train (80%) and test (20%) split of the dataset was maintained throughout.

4.2.2. Large language models

Bidirectional Encoder Representations from Transformers (BERT) is a state-of-the-art language model designed to pre-train deep bidirectional representations using both Masked Language Modelling (MLM) and Next Sentence Prediction (NSP) tasks (Devlin et al., 2018). It is the bidirectional nature that allows the model to jointly consider both left and right context in all layers, enhancing its ability to understand the meaning and nuances of language. BERT is based solely on the encoder component of the transformer architecture, which enables it to effectively capture long-range dependencies in text (Vaswani et al., 2017). Fine-tuning BERT involves appending task-specific layers, such as those trained on the annotated dataset, and further training the model on labelled data tailored to a specific application (UAM-specific tweets in our case). This process typically requires only a relatively small dataset, making it well-suited for scenarios involving limited manually annotated data. Moreover, once fine-tuned, the model can also be used to generate predictions on new texts. In this study, BERT was chosen as a representative large architecture, largely due to its accessible implementation libraries and strong community support at the time of model development.

4.3. Topic modelling

BERTopic initially generates document embeddings utilising a pre-trained language model (including LLMs), thereby maintaining semantic links. Subsequently, it diminishes the dimensionality of embeddings (e.g., UMAP) and categorises analogous articles into discrete subjects (e.g., HDBSCAN). BERTopic utilises a class-based TF-IDF methodology to enhance topic representation and ensure coherence. Moreover, BERTopic enables dynamic modelling by ascertaining topic representation at each time point without requiring several executions of the complete model (Grootendorst, 2022). In the current study, we initially applied a comprehensive BERTopic model to the entire corpus without temporal disaggregation, utilising the global representation of the primary subjects that may be discerned at different time intervals. Following that, the c-TF-IDF representation is calculated for each topic and time interval, resulting in unique subject representations at each time step and, in turn, creating a dynamic topic model. It is worth noting that we leveraged LLM (OpenAI) capabilities for both generating the embedding model (text-embedding-3-small) and creating the representation model (GPT-4-Turbo).

5. Results and analysis

This section begins by presenting the results of the sentiment classification and topic modelling exercises. It largely clarifies the efficacy of diverse machine learning and language models while emphasising the impact of augmenting an expert-annotated dataset. Subsequently, it presents key insights pertaining to the heterogeneity in sentiment polarities across countries and elucidates the clustering process to categorise them accordingly. Finally, cluster-specific dynamic BERTopic model outputs are presented, aiding in understanding the temporal evolution of topics over the study period.

5.1. Sentiment classification

As mentioned previously, we employed two distinct approaches (ML and LLM) to perform the sentiment analysis: SVM (Support Vector Machine) and BERT (Bidirectional Encoder Representations from Transformers). The following metrics are being evaluated: (1) *Precision* for a specific class is defined as the ratio of true positives relative to the sum of true and false positives; (2) *Recall* represents the ratio of true positives to the sum of true positives and false negatives; (3) *F1 score* is calculated as the harmonic mean of precision and recall; and (4) *Accuracy* indicates the proportion of correctly classified instances out of the total number of instances in the dataset.

Table 3 illustrates the performance outcomes of SVM and BERT, wherein we trained and/or fine-tuned with 80% and tested with 20% of the respective annotation datasets. Among these models, the BERT model (accuracy 0.751) was observed to perform significantly better than its ML counterpart (accuracy 0.725). Notably, a mostly similar performance pattern was observed across the other two evaluation metrics (precision, recall, and F1-score). At the same time, we found BERT to produce a more balanced production accuracy relative to SVM across four sentiment classes (See Fig. 3). Furthermore, we also performed a non-parametric test, i.e., Wilcoxon signed-rank test, on five-fold cross-validation results, to analyse whether the differences between BERT and SVM evaluation metrics are statistically significant or not. The test results suggested BERT is a significantly better model (at 0.10 significance level) than SVM, regardless of which metric (F1 scores for each class, accuracy, and weighted average) is used. The superior performance of BERT can be attributed to its deep contextual understanding of language, which allows it to capture subtle semantic nuances and relationships between words, even in short and informal texts like tweets. Unlike traditional ML models that rely on sparse, high-dimensional representations, BERT leverages pre-trained large embeddings that are highly effective in managing the variability, noise, and non-standardised language often present in social media content. Therefore, for all the subsequent analyses, we employed the BERT model on our whole tweet corpus to find the sentiment polarity across various time points and geographical regions.

Table 3
SVM and BERT model results.

Index	Precision	Recall	F1-score	Support
SVM classification model				
Negative	0.794	0.725	0.758	69
Neutral	0.667	0.828	0.738	58
Positive	0.708	0.829	0.764	82
Off-topic	0.762	0.500	0.604	64
Accuracy			0.725	
Macro avg	0.733	0.720	0.716	273
Weighted avg	0.734	0.725	0.719	273
BERT classification model				
Negative	0.864	0.760	0.809	75
Neutral	0.781	0.792	0.786	72
Positive	0.654	0.823	0.729	62
Off-topic	0.714	0.625	0.667	64
Accuracy			**0.751**	
Macro avg	0.753	0.750	**0.747**	273
Weighted avg	0.759	0.751	**0.751**	273

Note — The best values of the performance metrics have been bold-faced.

5.2. Spatio-temporal evolution of UAM sentiments

In the post-classification exercise, we attempted to discern the temporal change and spatial heterogeneity in the UAM-related sentiment polarity. To this end, we utilised time series analysis to achieve the former task while employing K-means clustering for the latter. Among various time series analysis methods, we relied on *frequency analysis* as it effectively analyses the popularity, sentiment trends, and event detection related to a specific topic over a longitudinal timeframe. Apart from obtaining absolute polarity values, we also defined the *sentiment index* as the difference between positive and negative tweets, facilitating a more intuitive understanding of polarity fluctuations.

The time series plot depicts positive, negative and neutral sentiments across six study countries (See Fig. 4(a)). Overall, the evolution trend of UAM-related tweets is distinctly different from the global trend of Twitter userbase, which grew exponentially from 54 million in 2010 to 305 million in 2015, and then continued to grow (albeit more slowly) to approximately 396 million by 2021(Turner, 2024). In contrast, the absolute count of UAM-related tweets indicates a steady increase in tweet frequency throughout the mid and late-2010s (peaking in 2017), followed by a consistent decline thereafter, largely attributable to the fading away of initial excitement rather than an inherent change in the Twitter user base. Another noteworthy observation is the high number of neutral tweets (including marketing promotions) during the 2017-20 period, which often overshadowed the actual human-generated ones. Also, the absolute numbers alone cannot reflect the relative change in polarity, which is more clearly illustrated by the sentiment index (see Fig. 4(b)). In addition to that, we have provided a list of major UAM-related developments to contextualise the major spikes in the time-series trend. (See Table 4). Notably, these findings align with the insights generated by the dynamic topic model presented in Section 6.

Broadly speaking, positive sentiments regarding UAM have prevailed over negative ones throughout most of the last decade, although the polarity has varied considerably across regions and time. For instance, during the early phase (2011–2012), technological demonstrations in both the United States and Germany, such as Joby Aviation's

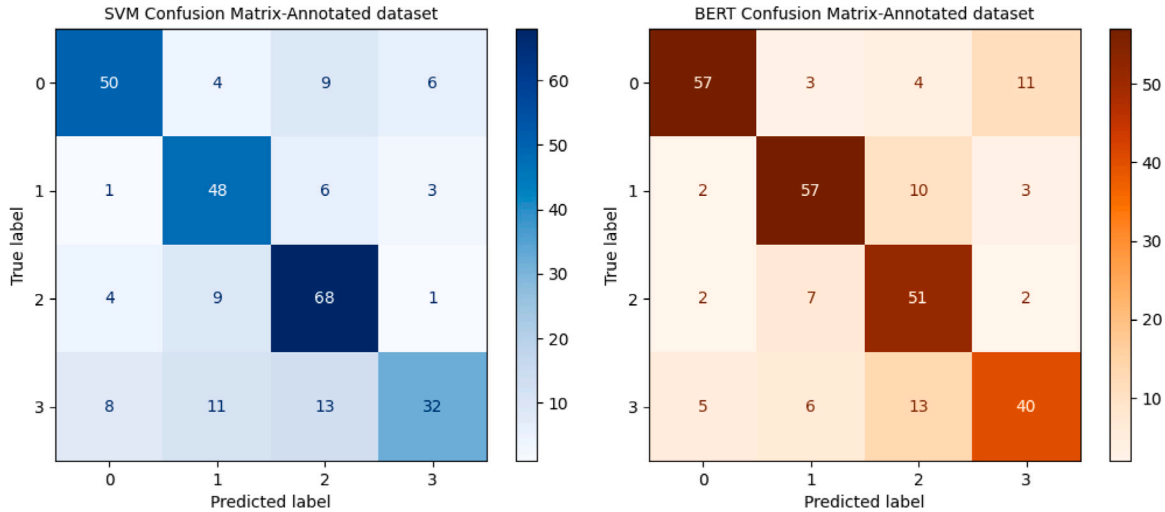

Fig. 3. Confusion matrices for (a) SVM and (b) BERT.

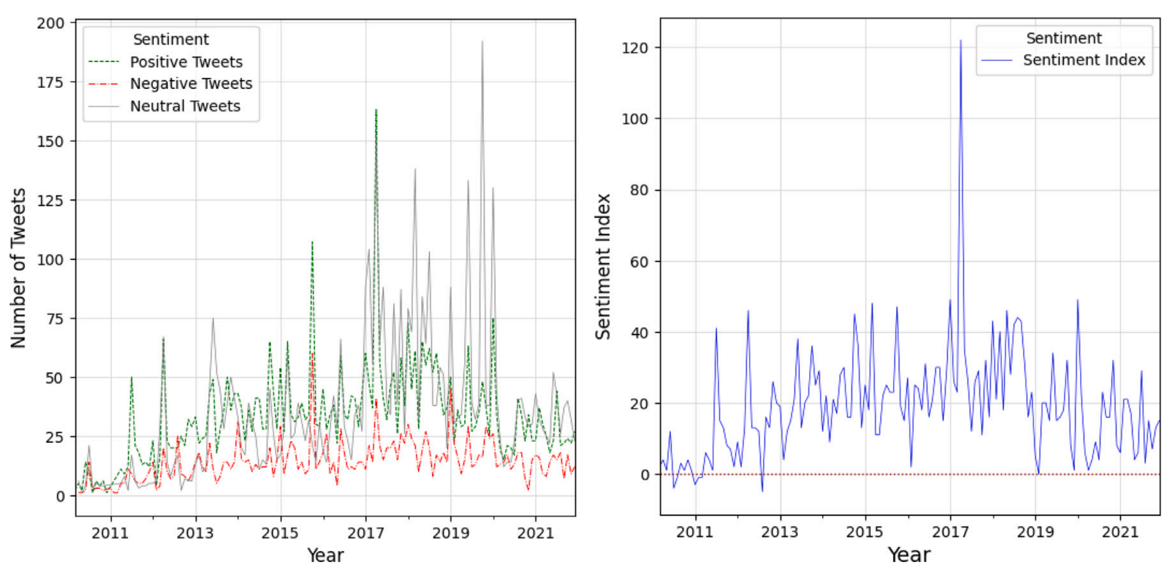

Fig. 4. UAM-related (a) sentiment polarity and (b) sentiment index over time.

early eVTOL prototypes and Volocopter's first manned electric multicopter flight, would be expected to generate positive sentiment, and our study corroborates this predominance of optimism. The subsequent period (2013–2014) showed a localised dip in positivity, particularly in Canada, which aligns with expectations given safety concerns surrounding the Maverick prototype crash in British Columbia. A resurgence of positive sentiment during 2015–2016 in the United States and Germany is consistent with anticipated reactions to heightened institutional engagement, such as NASA's initiation of structured UAM research and regulatory interest from European agencies in certifying eVTOL operations. Similarly, Uber's introduction of its on-demand air mobility concept (UberAir) would be expected to stimulate public enthusiasm, which our findings reflect. Conversely, during the same 2015–2016 period, a rise in negative sentiment, particularly evident in the United Kingdom and Germany, matches anticipated concerns related to drone-related safety and privacy incidents, as well as potential externalities such as noise, aerial congestion, and regulatory uncertainty. Finally, the major surge in overall UAM-related discussion during 2017–2019 coincided with intensified activity in the United States aerospace and mobility sectors. This spike is consistent with expected reactions to high-profile events, such as the FAA's drone regulation, Uber's Elevate Summit (2018), and major eVTOL prototype unveilings by companies like Boeing and Joby Aviation. Overall, these trends in sentiment align with observations from cross-sectional consumer research studies conducted by leading manufacturers, underscoring the heterogeneous and context-sensitive nature of public perception towards emerging UAM technologies (Wisk, 2021; Yedavali and Mooberry, 2019).

Turning to the spatial (or geographical) heterogeneity in the sentiment, we performed clustering with all three polarities (positive, negative and neutral) as it helps to bring the *opinionatedness* (or non-neutrality) dimension. Since off-topic tweets do not add to the understanding of sentiment polarity, we exclude them and focus solely on polarised ones. To this end, we calculated the country-specific normalised[3] polarity scores (varying between 0 and 1) for all three tweet sentiments and plotted the same along the X, Y, and Z axes, respectively. For example, the USA, with the highest share of negative tweets (14.63%) and the second highest share of positive tweets (31.79%) among the six countries, recorded both positive (0.943) and negative (1) normalised scores close (or equal) to 1. Subsequently, K-means clustering was employed to group countries with similar sentiment profiles resulting into identification of the following four clusters (See

[3] We used Min-Max scaling as part of the normalisation method.

Table 4
List of key time periods with corresponding UAM-related developments.

Period	Sentiment polarity	US	AU	GB	CA	IN	DE	Possible causes/events
2011–12	Positive	x						(1) Joby Aviation demonstrated early eVTOL prototypes; (2) Terrafugia Transition received FAA approval for 'roadable aircraft'; (3) Google co-founder Larry Page initiated funding of early UAM ventures (later evolved into Kitty Hawk and Opener projects).
2011–12	Positive						x	(1) Volocopter conducted the world's first manned flight of an all-electric multicopter.
2013–14	Negative				x			(1) The Maverick prototype crashed near an elementary school in British Columbia, Canada, on May 10, 2013, resulting in minor injuries to the two people aboard.
2015–16	Positive	x						(1) NASA formally initiates UAM-related research under its Aeronautics Research Mission Directorate; (2) Uber mentioned the vision of UberAir or on-demand air mobility.
2015–16	Positive						x	(1) First fully manned flight of the Volocopter VC200 approved under German aviation authority; (2) EASA and DLR start exploring integrating UAM airspace and certifying e-VTOL safety.
2015–16	Negative			x				(1) UK police recorded a significant increase in drone-related incidents (near misses and privacy concerns) in this period.
2015–16	Negative						x	(1) German authorities noticed about a dozen dangerous drone sightings near major airports.
2017–19	Positive	x						(1) FAA regulation (Part 107) opened up commercial use of small drones; (2) Uber Elevate Summit amplified public discussion around feasibility, challenges, and hype for UAM; (3) Heli-Expo 2018 and other trade shows triggered discussion among the public.

Fig. 5): (1) *Cluster A*: USA and Australia; (2) *Cluster B*: UK and Canada; (3) *Cluster C*: India; and (4) *Cluster D*: Germany. Further exploration reveals that cluster A countries (USA and Australia) exhibit a relatively high level of opinionatedness (normalised scores for positive and negative sentiments > 0.8) while cluster B countries (UK and Canada) largely fall in the middle (normalised score 0.4–0.6), except normalised score for negative sentiments in the UK > 0.8). In contrast, cluster C (India) and cluster D (Germany) demonstrate low opinionatedness (normalised score for negative and positive sentiments <0.3), which is equivalent to high neutrality. This effect is most pronounced for Germany, which has the highest number of neutral tweets amongst the study countries. It is essential to highlight that our results align closely with the observations from various multi-country studies conducted in industrial contexts (KPMG, 2022). However, a significant limitation of these cross-sectional studies has been their failure to identify potential sources of heterogeneity and evaluate how these factors evolved since the introduction of UAM, leading us to utilise dynamic topic modelling.

5.3. Topic model

The dynamic topic model results (See Table 5)[4] suggest the most prevalent themes (based on order and presence) across clusters are *industry innovation and testing, unmanned aviation system*, and *mobility benefits*. It suggests that public perception largely centres around technological innovation and real-life implementation and/or benefits of such emerging transport alternatives. Not only did we find a few topics, such as *safety and regulation* and *Maverick accident*, gaining traction in specific clusters, but also the absence of specific topics across clusters indicates the role of socio-cultural paradigm in shaping human perception. Also, we separated out the *fiction and fantasy* and *frustration from late implementation* topic to assess the *excitement* aspect in a better way and gauge the potential acceptance of UAM. It is worth noting that while developing the topic model, we excluded all the off-topic tweets (including marketing promotion contents from manufacturers and service providers) to achieve a clearer understanding of public sentiments. Finally, the topic model results add the temporal dynamicity to identify the trigger events required for developing policy instruments, as explained in the following sections.

[4] Refer to Table 6 in Appendix to find the list of keywords and their corresponding weights.

6. Public discourse themes and policy insights

Time series analysis facilitated the discovery of key time periods linked to sentiment shifts, whereas spatial analysis provided insights into its variation across countries. Building on those insights, we utilised the BERTopic model to discern predominant themes and their associated terms in tweets across four pre-defined clusters during the study period. It must be noted that the topic model discussion focuses exclusively on relevant and meaningful tweets leveraging BERTopic's superior semantic understanding. As stated earlier, the eight most pertinent themes (See Fig. 6) have been identified, with three present across all the clusters; however, the inferences for clusters 3 and 4 should be approached with additional caution due to their lower tweet volume. Of particular relevance to translating topic model results into actionable policies is the higher prominence of three themes (Industry innovation and testing, Unmanned aviation system, and Mobility benefits) and their contextualisation to key events. It may be noted that the policy insights are aimed at balanced uptake of UAM as opposed to focusing on sustainability alone.

6.1. Cluster 1: USA and Australia

The topic model results suggest that the *excitement* regarding UAM is the most frequent topic in cluster 1 countries (See Fig. 7), as it started growing since the early part of the last decade and peaked during the 2015-17 period. It indicates how the vision of 'flying cars' garners public attention in the mobility landscapes with higher car-dependency, low-density and decentralised urban planning (longer commutes) and relatively more frequent exposure to air travel. Furthermore, high individualism and low uncertainty avoidance play a key role in positioning these countries higher in terms of readiness for emerging technologies. Additionally, it is interesting to observe local alignments between *excitement* and *unmanned aviation system* (around 2015) and *industry innovation and testing* (around 2017), suggesting potential temporal associations between public sentiment and key developmental milestones. Finally, the initial excitement appears to have faded post-2016, possibly due to regulatory delays, perceived safety and noise concerns, equity issues, and limited visible progress towards implementation.

The *industry innovation and testing*, which primarily gained traction during the 2016–18 period, is observed to be the second most frequent topic. This is consistent with the USA's position as the global leader in technologies and innovations, while the timeframe largely coincides with the public demonstration of prototypes by leading manufacturers

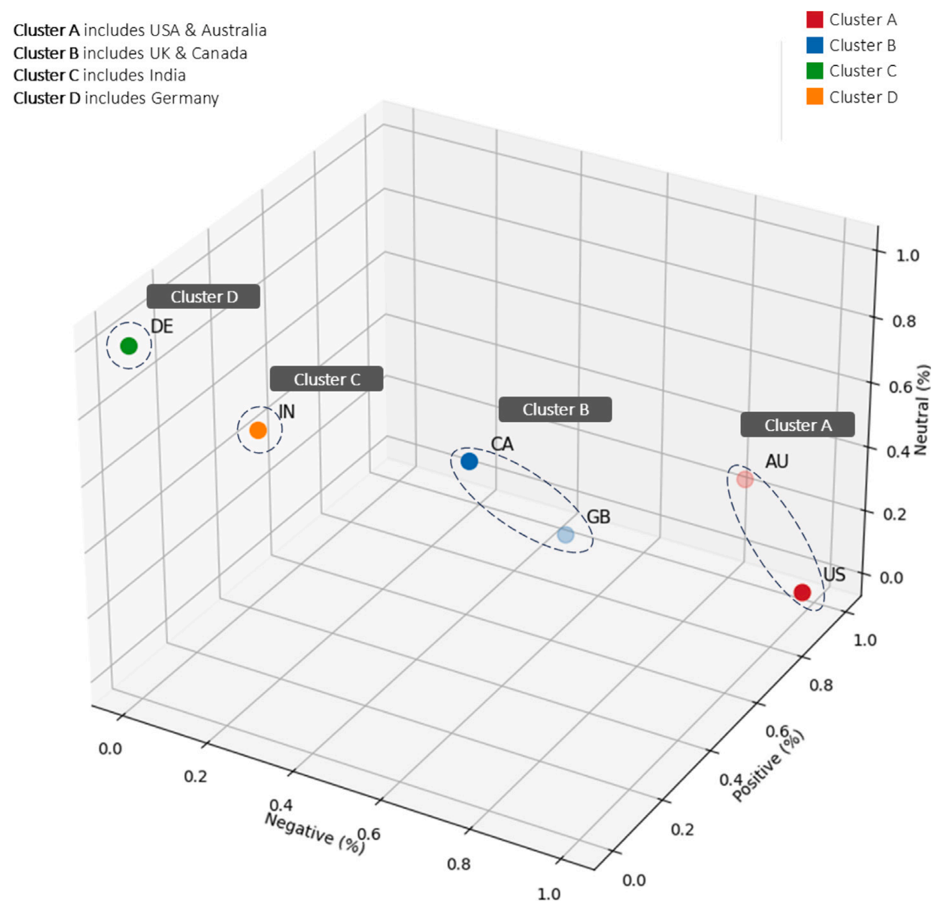

Fig. 5. Sentiment clusters of study countries- 4 clusters.

Table 5
List of topics (ordered) across four spatial clusters.

Topic rank	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1	Excitement	Industry innovation and testing	Industry innovation and testing	Industry innovation and testing
2	Industry innovation and testing	Excitement	Unmanned aviation system	Mobility benefits
3	Unmanned aviation system	Maverick accident	Mobility benefits	Unmanned aviation system
4	Mobility benefits	Technical operation	Excitement	
5	Technical operation	Unmanned aviation system		
6	Frustration from late implementation	Fiction and fantasy		
7	Fiction and fantasy	Mobility benefits		
8		Safety and regulation		

Fig. 6. Word cloud of observed topics across all clusters.

(e.g., Joby Aviation, Wisk). Intuitively, another topic, i.e., *technical operation*, shows partial convergence in certain intervals, highlighting the inter-relationship between these topics and suggesting that the public perception is increasingly based on technical nuances of UAM system operability. Of particular relevance is the steep curve (both rise and decline) points out that (1) public sentiments are highly reactive to on-ground implementation, (2) lack of solid policy and legislation erodes public confidence in near-term implementation, and (3) delayed commercial layouts fizzle initial interest. From a regulatory and industry perspective, clear certification pathways, well-communicated development timelines, and increased public consultation can add momentum to UAM adoption.

The *unmanned aviation system* emerges as the third most frequently discussed topic. Unmanned UAM development gained significant traction in the cluster 1 countries (particularly the USA), surging around 2014 and peaking during the 2015–18 period, driven by private sector innovation (e.g., Uber Elevate), federal research support (NASA, FAA), and growing investor and media interest. Upon closer examination of the keywords, we could find that the people perceive the topic through

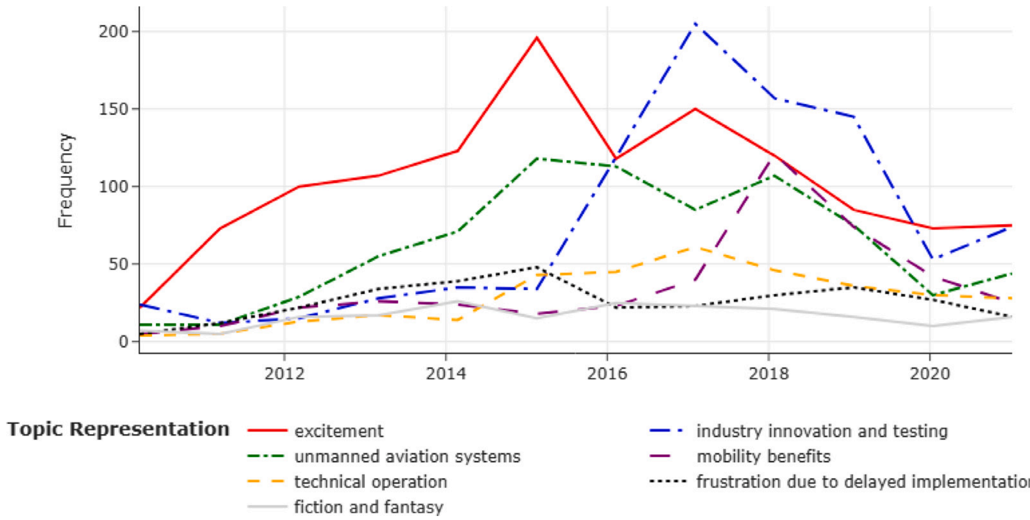

Fig. 7. Key topics of public discourse in cluster 1 (2010–21)

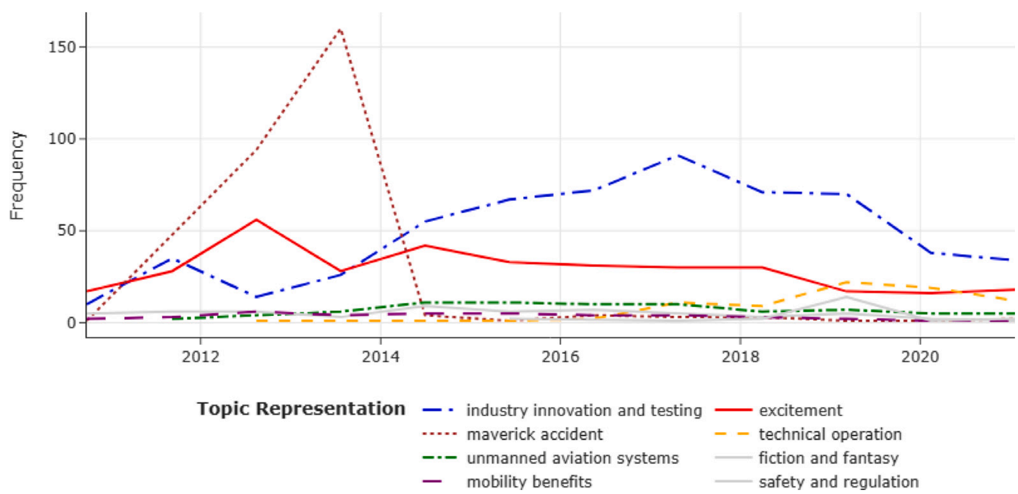

Fig. 8. Key topics of public discourse in cluster 2 (2010–21)

regulatory (e.g., FAA certification, airspace regulation) and operability (e.g., safety, reliability) lenses. It also underlines the institutional positioning of UAM technology as an extension of unmanned drone technologies. Nevertheless, public interest declined after 2018, which may be tied to regulatory bottlenecks and the industry focusing more on commercial usage of unmanned aerial vehicles (drones). From the service providers' perspective, for the UAM system to reach maximum popularity, the systems need to be positioned as a functionally (e.g., safe, saving time) and symbolically (e.g., environment-friendly, luxury, fun) attractive travel option, especially to the affluent early adopters. On the other hand, regulators can aid the endeavour of the service providers by focusing on the seamless integration of UAM into the existing transportation network.

Now shifting focus to the remaining three topics, we observe Twitter discussions centred around *mobility benefits*, *technical operation*, and *frustration from late implementation*. Expectedly, the trajectory of the former two displays local similarity with *industry innovation and testing*, underscoring increasing realism (excitement and scepticism) in public discourses as (USA) cities move closer to actual adoption. The discourse for the *mobility benefits* primarily pivots on the efficiency of UAM systems (vehicles and vertiports) in saving travel time lost in traffic congestion, while some focus on the connectivity of vertiports with major hubs. On the other hand, *technical operation* largely includes narratives about operational characteristics (including safety) and comparing (and relating) them with conventional fixed-wing or military aircraft. This suggests the presence of public misconceptions or limited understanding regarding the distinct nature and capabilities of UAM technologies. The last topic, i.e., *frustration from late implementation*, which is specific to the current cluster (particularly the USA), reflects negative sentiments associated with the delayed rollout of UAM technology, potentially discouraging enthusiastic adopters. This growing disillusionment may hinder market readiness and contribute to slower adoption when UAM systems eventually become available. In addition to the recommended policies for the first three topics, Original equipment manufacturers (OEMs) and industry stakeholders can prioritise alleviating operational concerns (safety, noise, frequency, etc.) through real-life demonstrations and phased implementation strategies to facilitate smoother public acceptance and trust.

6.2. Cluster 2: UK and Canada

A key point to consider in the topic model results for cluster 2 countries is the higher prominence of *industry innovation and testing* relative to *excitement* (See Fig. 8). Notably, the growth for *industry innovation and testing* is more gradual than its curve for cluster 1; however, in both instances, the peak occurred during the 2016–18 period. These findings, coupled with keyword analysis, highlight that the impact

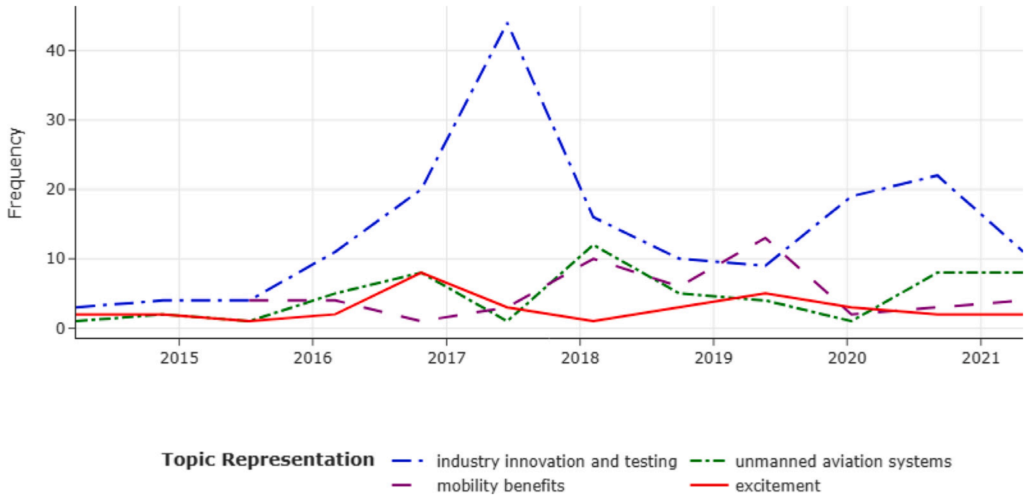

Fig. 9. Key topics of public discourse in cluster 3 (2010–21)

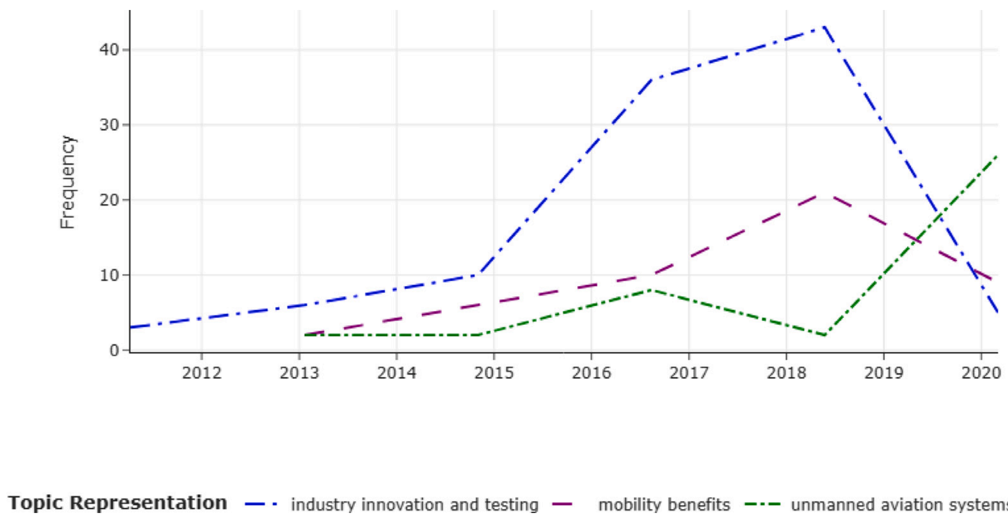

Fig. 10. Key topics of public discourse in cluster 4 (2010–21)

of on-ground implementation can transcend geographical boundaries. Nevertheless, such ripple effects are likely weak stimulators for public discourse. In contrast, a nearly flat-lined curve for the *excitement* indicates a lack of adoption enthusiasm, which can be tied to multiple social (e.g., risk-aversion), cultural (e.g., pragmatism, less risk-taking), planning (e.g., lower urban sprawl, better public transport network), and implementation (e.g., fewer prototype demonstrations) criteria. It essentially underlines that increased traction for industry innovation need not necessarily evoke adoption interest; instead, policymakers can complement ongoing innovation with public-facing initiatives, such as demonstration projects, community engagement campaigns, and educational outreach, that contextualise UAM within local mobility needs. At the same time, *maverick accident* being the third most frequently discussed (but with the highest peak) topic suggests public reaction (excitement and concern) has the highest intensity to such real-world demonstrations; however, failures can wane such interest too quickly. Hence, other than emphasising tangible benefits, addressing public concerns can be particularly effective for cluster 2 countries. Additionally, aligning with existing urban planning frameworks can help bridge the gap between technological advancement and increasing adoption in UAM.

The other three among the top-six most frequently discussed are *technical operation*, *unmanned aviation system*, and *fiction and fantasy*. Except for the last one, the other two were also observed in the top six for cluster 1 and may be linked with a rising sense of grounded realism in public discourse as we approach real-world operation. In this connection, their lower amplitude compared to cluster 1 could be attributed to the fact that the UK and Canada have not yet reached the level of real-world implementation seen in the USA. The other striking finding is that, unlike cluster 1, the public discourse about *technical operation* was found to be closely related to Internet of Things (IoT), artificial intelligence (AI), and robotics. This suggests a more cautious, systems-oriented view of UAM, potentially shaped by the emphasis on research, regulation, and interconnectivity over rapid deployment. Lastly, we observed targeted discussions centred around *safety and regulation*, a distinctive characteristic of this cluster. Keyword analysis suggests a broader narrative shaped by institutional oversight and prior exposure to regulations (particularly drone). From a policy standpoint, these findings highlight the importance of transparent safety frameworks and early engagement with national aviation authorities to build confidence and support for UAM integration.

6.3. Cluster 3: India

In line with the previous two clusters, *industry innovation and testing* was the primary theme of public discourse (See Fig. 9). Further examination of the temporal trend and keywords reinforces the notion regarding the ripple effect of on-ground demonstrations. At the

Table 6
Description of tweet contents (topics and top five keyword scores)

Cluster	Word 1	Word 2	Word 3	Word 4	Word 5
Topic: *Excitement*					
Cluster 1	want (0.082)	future (0.064)	get (0.052)	hoverboard (0.042)	need (0.037)
Cluster 2	want (0.096)	need (0.065)	lol (0.042)	wow (0.040)	get (0.039)
Cluster 3	finally (0.140)	world (0.112)	two (0.103)	gift (0.098)	pre order (0.098)
Cluster 4	*Not present*				
Topic: *Industry innovation and testing*					
Cluster 1	uber (0.077)	volocopter (0.043)	aviation (0.034)	terrafugia (0.033)	airbus (0.030)
Cluster 2	uber (0.047)	future (0.034)	first (0.034)	flight (0.034)	electric (0.028)
Cluster 3	first (0.068)	airbus (0.057)	tech (0.055)	future (0.053)	volocopter (0.047)
Cluster 4	volocopter (0.115)	jet (0.084)	uber (0.080)	flight (0.075)	innovation (0.074)
Topic: *Unmanned aviation system*					
Cluster 1	unmanned (0.163)	drone (0.111)	UAS (0.056)	UAV (0.040)	FAA (0.034)
Cluster 2	unmanned (0.217)	aerial vehicle (0.110)	drone (0.105)	UAV (0.095)	UAS (0.056)
Cluster 3	unmanned (0.171)	drone (0.103)	vehicle (0.099)	landing (0.091)	vertical (0.089)
Cluster 4	drone (0.324)	UAV (0.250)	pilot (0.061)	medical (0.048)	fun (0.048)
Topic: *Mobility benefits*					
Cluster 1	airport (0.080)	urban mobility (0.110)	traffic (0.074)	people cargo (0.051)	save time (0.051)
Cluster 2	traffic (0.379)	jam (0.178)	stuck (0.155)	need (0.102)	road (0.089)
Cluster 3	uber (0.218)	travel (0.101)	service (0.082)	airport (0.078)	cab (0.059)
Cluster 4	urban mobility (0.195)	city (0.083)	city airbus (0.081)	UAM (0.081)	air taxi (0.069)
Topic: *Technical operation*					
Cluster 1	plane (0.081)	airplane (0.039)	landing (0.035)	take (0.030)	osprey (0.027)
Cluster 2	mobility (0.129)	transport (0.096)	robotics (0.066)	AI (0.060)	IOT (0.054)
Cluster 3	*Not present*				
Cluster 4	*Not present*				
Topic: *Fiction and fantasy*					
Cluster 1	jetsons (0.235)	george (0.076)	harry (0.073)	potter (0.060)	like jetsons (0.041)
Cluster 2	london (0.213)	ron (0.123)	station (0.089)	tour (0.089)	group (0.059)
Cluster 3	*Not present*				
Cluster 4	*Not present*				
Topic: *Maverick accident*					
Cluster 1	*Not present*				
Cluster 2	driving (0.106)	taxiing (0.093)	takeoff (0.091)	landing (0.089)	passenger (0.081)
Cluster 3	*Not present*				
Cluster 4	*Not present*				
Topic: *Safety and regulation*					
Cluster 1	*Not present*				
Cluster 2	drone (0.241)	UK (0.182)	model (0.175)	aviation authority (0.144)	safety (0.143)
Cluster 3	*Not present*				
Cluster 4	*Not present*				
Topic: *Frustration from late implementation*					
Cluster 1	still (0.164)	waiting (0.163)	whereas (0.081)	year (0.074)	shit (0.042)
Cluster 2	*Not present*				
Cluster 3	*Not present*				
Cluster 4	*Not present*				

same time, India's huge potential market, increasing road congestion problem, and administrative support for entrepreneurial activities are expected to fuel such narratives. In contrast, *excitement* is the least prominent topic, diverging from patterns observed in other clusters. This likely reflects vastly different socio-cultural characteristics, transport infrastructure provisions, and urban planning of these regions. As a result, the idea of flying cars, often perceived as a high-income, aspirational technology, may resonate less with the general public, making widespread *excitement* less relevant in discourse. The other two topics, i.e, *unmanned aviation system* and *mobility benefits*, are characteristically (based on keywords) similar to the previous two clusters and mostly coincide with the global key events. In case of the former topic, we could detect a little surge in the post-pandemic period, which may be linked to policy discussions on drone policy frameworks, leading to the release of Drone Rules in 2021, liberalising the sector and sparking wider interest in unmanned aviation applications (including heavier payloads). To conclude, in India, focusing on inclusive innovation through affordable pilot programs and leveraging regulatory momentum to align UAM development with broader mobility needs can be the most effective way to encourage public acceptance and uptake.

6.4. Cluster 4: Germany

Analysis of trends in Germany depicts that *industry innovation and testing* significantly dominated the public discussions, with its growth curve similar to other clusters (See Fig. 10). Further exploration of the keywords reveals that indigenous aerospace industries such as Volocopter and Lilium remained prominent. Hence, Germany's strong innovation ecosystem and presence of indigenous players can be strategically leveraged by expanding research and development incentives, accelerating certification pathways, and promoting public demonstrations to enhance visibility, trust, and global competitiveness in UAM technologies. The other two topics are *mobility benefits* and *unmanned aviation system*, where the former gained traction during the 2017–19 period and the latter was observed to surge post-2018. Although the *mobility benefits* topic was found to have a similar broad narrative as in other countries, *unmanned aviation system* includes people discussing more about the utilitarian aspects of drones (e.g., medical, logistics) in addition to passenger transport. In order to maximise adoption, UAM service providers can strategically leverage this interest into the passenger UAM domain by demonstrating technological overlap (e.g., shared navigation, safety systems) and gradually transitioning to passenger transport.

7. Concluding remarks

Urban air mobility is poised to revolutionise the transportation sector; nevertheless, its widespread adoption is likely to be influenced by public perceptions, which are inherently dynamic and susceptible to a multitude of factors, including societal values, lifestyle preferences, and travel habits. Simultaneously, each of these factors is deeply embedded within the socio-cultural and geo-political frameworks, implying that they are anticipated to exhibit variations across time and geographical regions. Although there has been a growing body of cross-sectional studies analysing market segmentation based on *a priori* indicators of public acceptance, the data-driven detection of sentiment polarity and related spatio-temporal heterogeneity aspects remain poorly understood.

Based on the Twitter data collected from six countries across twelve years (2010–21), the current study provides strong evidence that considerable spatial and temporal heterogeneity exists in UAM-related public perception, which calls for context-specific understanding of *ease of acceptance*. In terms of methodologies, we not only employed state-of-the-art large and large language models for sentiment classification and topic modelling, respectively, but also enhanced the applicability of NLP results by integrating them with time series analysis and spatial K-means clustering. The *time series analysis* shows that UAM garnered a steady increase of public traction throughout the mid and late-2010s, while also suggesting growing neutrality during 2017–20. We also derived *sentiment index*, which aids in contextualising the spikes in polarity with real-world events (e.g., public demonstration, crashes) and thereby detecting their relative intensity (e.g., prototype testing creating higher impact) to stimulate public opinion. At the same time, our results empirically demonstrate that these effects differ across countries and could be attributed to differences in societal and cultural dimensions in addition to the disparity in existing transport infrastructure. The *clustering analysis* led to the identification of four country-level clusters, with cluster A countries exhibiting the highest opinionatedness in their perception, cluster B countries essentially having a balance between neutrality and (positively or negatively) polarised perceptions, and the rest (cluster C and D) demonstrating low opinionatedness. To achieve a more nuanced understanding of the underlying themes (or levers) of public discourse, we performed *dynamic topic modelling* across the previously identified cluster over the study period. As such, the global discussions are observed to pivot around three major themes: (1) industry innovation and testing, (2) unmanned aviation systems, and (3) mobility benefits. Besides, we could also detect local themes (e.g., maverick accident in Canada), which in cases have a greater stimulating influence than global ones. Importantly, the relative order of the themes and the keywords that comprise them reveals critical context-specific differences and provides a systematic direction for adopting localised policy interventions. For instance, excitement for UAM dominates public perception in the USA, possibly stemming from advances in prototype testing and public demonstration; however, it becomes secondary (cluster B and C) or even non-existent (cluster D) for other clusters. This implies that business operations and marketing campaigns related to UAM travel cannot take a *one-size-fits-all* approach; instead, they must be tailored to the specific socio-cultural contexts. We also note that while the actual benefits of UAM travel are growingly shaping public opinions, overall attraction wanes as on-field implementation gets delayed due to regulatory bottlenecks. Therefore, the initial market penetration of UAM services remains uncertain and warrants further examination. In this context, future extensions of this research could integrate both cross-sectional and data-driven approaches to gain more comprehensive insights into the long-term evolution of perception across diverse socio-demographic cohorts. Furthermore, future research could explore incorporating tweets in multiple languages to better account for native language usage and examine potential cultural and linguistic variations that may not be adequately captured in an English-only corpus.

To conclude, our findings indicate that societal acceptance of urban air mobility services differs across regions with varied socio-cultural contexts, and even within the same region, it changes over time. Therefore, fostering the successful uptake of UAM services necessitates moving beyond the cross-sectional approach, instead embracing large-scale data-driven approaches, which could facilitate tailor-made policy interventions. Such localised policy-making indicators could be equally useful for regulators and operators to ensure both sustainable and profitable integration of UAM into the existing transportation ecosystem.

CRediT authorship contribution statement

Eeshan Bhaduri: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Charisma F. Choudhury:** Writing – review & editing, Supervision, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgements

This study was supported by UKRI Future Leader Fellowship [MR/T020423/1 and MR/Y034384/1] and Towards Zero Carbon Aviation Project[EP/V000772/1]. The authors are grateful to Lucia Cebreiros Penzone and Olivia Barnes (University of Leeds), who assisted in the data annotation process, and Dr Arash Kalatian (Moneris) for his role in initial data collection.

Appendix. Cluster-specific keyword scores

See Table 6.

Data availability

Data will be made available on request.

References

Al-Haddad, C., Chaniotakis, E., Straubinger, A., Plötner, K., Antoniou, C., 2020. Factors affecting the adoption and use of urban air mobility. Transp. Res. Part A: Policy Pr. 132, 696–712. http://dx.doi.org/10.1016/J.TRA.2019.12.020.

Anderson, T., Sarkar, S., Kelley, R., 2024. Analyzing public sentiment on sustainability: A comprehensive review and application of sentiment analysis techniques. Nat. Lang. Process. J. 8, 100097. http://dx.doi.org/10.1016/J.NLP.2024.100097.

Bakalos, N., Papadakis, N., Litke, A., 2020. Public perception of autonomous mobility using ML-based sentiment analysis over social media data. Logistics 4 (2), 12.

Balac, M., Rothfeld, R.L., Hörl, S., 2019. The prospects of on-demand urban air mobility in Zurich, Switzerland. In: 2019 IEEE Intelligent Transportation Systems Conference. ITSC, IEEE, pp. 906–913.

Becker, K., Terekhov, I., Gollnick, V., 2018. A global gravity model for air passenger demand between city pairs and future interurban air mobility markets identification. In: 2018 Aviation Technology, Integration, and Operations Conference. p. 2885.

Chen, X., Zeng, H., Xu, H., Di, X., 2021. Sentiment analysis of autonomous vehicles after extreme events using social media data. In: 2021 IEEE International Intelligent Transportation Systems Conference. ITSC, IEEE, pp. 1211–1216.

Cohen, A.P., Shaheen, S.A., Farrar, E.M., 2021. Urban air mobility: History, ecosystem, market potential, and challenges. IEEE Trans. Intell. Transp. Syst. 22 (9), 6074–6087.

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Ding, Y., Korolov, R., Wallace, W.A., Wang, X.C., 2021. How are sentiments on autonomous vehicles influenced? An analysis using Twitter feeds. Transp. Res. Part C: Emerg. Technol. 131, 103356.

EASA, N., 2021. Study on the societal acceptance of urban air mobility in Europe. Eur. Union Aviat. Saf. Agency McKinsey & Co..

Garrow, L.A., German, B.J., Leonard, C.E., 2021. Urban air mobility: A comprehensive review and comparative analysis with autonomous and electric ground transportation for informing future research. Transp. Res. Part C: Emerg. Technol. 132, 103377. http://dx.doi.org/10.1016/J.TRC.2021.103377.

Garrow, L.A., Mokhtarian, P.L., German, B.J., Glodek, J.J.S., Leonard, C.E., 2025. Market segmentation of an electric vertical takeoff and landing (eVTOL) air taxi commuting service in five large U.S. cities. Transp. Res. Part A: Policy Pr. 191, 104267. http://dx.doi.org/10.1016/J.TRA.2024.104267, URL https://www.sciencedirect.com/science/article/pii/S096585642400315X?via%3Dihub.

Grootendorst, M., 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. http://dx.doi.org/10.48550/arXiv.2203.05794, arXiv preprint arXiv:2203.05794.

Haan, J., Garrow, L.A., Marzuoli, A., Roy, S., Bierlaire, M., 2021. Are commuter air taxis coming to your city? A ranking of 40 cities in the United States. Transp. Res. Part C: Emerg. Technol. 132, 103392. http://dx.doi.org/10.1016/j.trc.2021.103392.

Hamilton, B.A., 2018. Urban air mobility (UAM) market study. Tech. Out Brief Natl. Aeronaut. Space Adm. (NASA) 4, 2020, URL https://ntrs.nasa.gov/citations/20190000517.

He, H., Ma, Y., 2013. Imbalanced learning: foundations, algorithms, and applications.

Hearst, M.A., Dumais, S.T., Osuna, E., Platt, J., Scholkopf, B., 1998. Support vector machines. IEEE Intell. Syst. Appl. 13 (4), 18–28.

KPMG, 2022. Air Taxi Readiness Index 2022. Technical Report, URL https://kpmg.com/ie/en/home/insights/2023/03/air-taxi-readiness-index-2022.

Kwon, O.H., Vu, K., Bhargava, N., Radaideh, M.I., Cooper, J., Joynt, V., Radaideh, M.I., 2024. Sentiment analysis of the United States public support of nuclear power on social media using large language models. Renew. Sustain. Energy Rev. 200, 114570. http://dx.doi.org/10.1016/J.RSER.2024.114570.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159. http://dx.doi.org/10.2307/2529310.

Li, T., Lin, L., Choi, M., Fu, K., Gong, S., Wang, J., 2018. Youtube av 50k: an annotated corpus for comments in autonomous vehicles. In: 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAI-NLP). IEEE, pp. 1–5.
Long, Q., Ma, J., Jiang, F., Webster, C.J., 2023. Demand analysis in urban air mobility: A literature review. J. Air Transp. Manag. 112, 102436. http://dx.doi.org/10.1016/J.JAIRTRAMAN.2023.102436, URL https://www.sciencedirect.com/science/article/pii/S0969699723000790?via%3Dihub.
Manca, F., Pawlak, J., Sivakumar, A., 2023. Impact of perceptions and attitudes on air travel choices in the post-COVID-19 era: A cross-national analysis of stated preference data. Travel. Behav. Soc. 30, 220–239. http://dx.doi.org/10.1016/J.TBS.2022.10.006, URL https://www.sciencedirect.com/science/article/pii/S2214367X22001120.
Mayor, T., Anderson, J., 2019. Getting mobility off the ground. KPMG Insight.
NEXA Advisors, 2019. UAM study prospectus. https://Www.Nexaadvisors.Com/Uam-Global-Markets-Study. (Accessed September 3, 2020).
Pons-Prats, J., Živojinović, T., Kuljanin, J., 2022. On the understanding of the current status of urban air mobility development and its future prospects: Commuting in a flying vehicle as a new paradigm. Transp. Res. Part E: Logist. Transp. Rev. 166, 102868. http://dx.doi.org/10.1016/J.TRE.2022.102868, URL https://www.sciencedirect.com/science/article/pii/S1366554522002496.
Qi, Y., Shabrina, Z., 2023. Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach. Soc. Netw. Anal. Min. 13, 1–14. http://dx.doi.org/10.1007/S13278-023-01030-X/TABLES/3.
Rashidi, T.H., Abbasi, A., Maghrebi, M., Hasan, S., Waller, T.S., 2017. Exploring the capacity of social media data for modelling travel behaviour: Opportunities and challenges. Transp. Res. Part C: Emerg. Technol. 75, 197–211.
Robinson, J.N., Sokollek, M.-D.R., Justin, C.Y., Mavris, D.N., 2018. Development of a methodology for parametric analysis of STOL airpark geo-density. In: 2018 Aviation Technology, Integration, and Operations Conference. p. 3054.
Rothfeld, R., Fu, M., Balać, M., Antoniou, C., 2021. Potential urban air mobility travel time savings: An exploratory analysis of munich, paris, and san francisco. Sustainability 13 (4), 2217.
Sadrani, M., Adamidis, F., Garrow, L.A., Antoniou, C., 2025. Challenges in urban air mobility implementation: A comparative analysis of barriers in Germany and the United States. J. Air Transp. Manag. 126, 102780. http://dx.doi.org/10.1016/J.JAIRTRAMAN.2025.102780, URL https://www.sciencedirect.com/science/article/pii/S0969699725000432?via%3Dihub.
Samadzad, M., Ansari, F., Moez, M.A.A., 2024. Who will board urban air taxis? An analysis of advanced air mobility demand and value of travel time for business, airport access, and regional tourism trips in Iran. J. Air Transp. Manag. 119, 102636. http://dx.doi.org/10.1016/J.JAIRTRAMAN.2024.102636, URL https://www.sciencedirect.com/science/article/pii/S0969699724001017?via%3Dihub.
Song, F., Hess, S., Dekker, T., 2024. Uncovering the link between intra-individual heterogeneity and variety seeking: the case of new shared mobility. Transportation 51, 371–406. http://dx.doi.org/10.1007/S11116-022-10334-4/TABLES/8.
Straubinger, A., Rothfeld, R., Shamiyeh, M., Büchter, K.D., Kaiser, J., Plötner, K.O., 2020. An overview of current research and developments in urban air mobility – Setting the scene for UAM introduction. J. Air Transp. Manag. 87, 101852. http://dx.doi.org/10.1016/J.JAIRTRAMAN.2020.101852, URL https://www.sciencedirect.com/science/article/pii/S0969699719304302.
Straubinger, A., Verhoef, E.T., De Groot, H.L., 2021. Will urban air mobility fly? The efficiency and distributional impacts of UAM in different urban spatial structures. Transp. Res. Part C: Emerg. Technol. 127, 103124. http://dx.doi.org/10.1016/j.trc.2021.103124.
Taheri, A., Zamanifar, A., Farhadi, A., 2024. Enhancing aspect-based sentiment analysis using data augmentation based on back-translation. Int. J. Data Sci. Anal. 1–26, URL https://link.springer.com/article/10.1007/s41060-024-00622-w.
Turner, A., 2024. How many users does Twitter have? (2025). URL https://www.bankmycell.com/blog/how-many-users-does-twitter-have#section-1.
United Nations, 2018. World urbanization prospects: The 2018 revision. Popul. Div. U. N. 32, URL https://population.un.org/wup/.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. Adv. Neural Inf. Process. Syst. 30.
Wai, C.W., Tan, K., Low, K.H., 2021. Preliminary study of transport pattern and demand in Singapore for future urban air mobility. In: AIAA Scitech 2021 Forum. p. 1633.
Wang, S., Li, M., Yu, B., Bao, S., Chen, Y., 2022. Investigating the impacting factors on the public's attitudes towards autonomous vehicles using sentiment analysis from social media data. Sustainability 14 (19), 12186.
Wisk, 2021. Autonomous UAM: Taking Mobility to New Heights. Technical Report, URL https://wisk.aero/wp-content/uploads/2021/04/Wisk_Thought-Paper_ConsumerResearch-AutonomousUAM.pdf.
Yedavali, P., Mooberry, J., 2019. An Assessment of Public Perception of Urban Air Mobility (UAM). Technical Report, Airbus, URL https://www.airbus.com/sites/g/files/jlcbta136/files/2022-07/Airbus-UTM-public-perception-study%20-urban-air-mobility.pdf.
Zapf, A., Castell, S., Morawietz, L., Karch, A., 2016. Measuring inter-rater reliability for nominal data - which coefficients and confidence intervals are appropriate? BMC Med. Res. Methodol. 16, 1–10. http://dx.doi.org/10.1186/S12874-016-0200-9/TABLES/2.