



ISSN: 1567-7141 http://ejtir.tudelft.nl/

Choosing Suitable Indicators for the Assessment of Urban Air Mobility: A Case Study of Upper Bavaria, Germany

Christelle Al Haddad¹

Department of Civil, Geo and Environmental Engineering, Technical University of Munich, Germany

Mengying Fu² Bauhaus Luftfahrt e.V., Germany

Anna Straubinger³ Bauhaus Luftfahrt e.V., Germany

Kay Plötner⁴ Bauhaus Luftfahrt e.V., Germany

Constantinos Antoniu⁵

Department of Civil, Geo and Environmental Engineering, Technical University of Munich, Germany

I echnological advances are disrupting mobility patterns and transport technologies, both on the ground and in the air. The latter has been recently observed in the research community of urban air mobility (UAM). Research in this area has studied several areas of its implementation, such as vehicle concepts, infrastructure, transport modeling, or operational constraints. Few studies however have focused on evaluating this service as an alternative among existing transportation systems. This research presents an approach for the selection of indicators for a multi-criteria analysis for the assessment of UAM, in a case study of Upper Bavaria, Germany. A 5-stage approach is showcased including an expert assessment for the relevance and feasibility of indicators, based on two rating scales. A threshold for selection is presented, applied and validated for both scales. The results included a list of indicators for assessing the potentials of UAM integration to existing public transportation systems; the chosen indicators were then compared against existing ones for sustainable urban mobility. A high match between resulting indicators and previous ones further validate the results, and suggest that there is a need for an iterative approach in the assessment of disruptive transport technologies.

Keywords: assessment, key performance indicators, multi-criteria analysis, urban air mobility

¹ A: Arcisstraße 21, 80333 Munich, Germany E: christelle.haddad@tum.de

² A: Willy-Messerschmitt-Straße 1, 82024 Taufkirchen, Germany E: Mengying.Fu@bauhaus-luftfahrt.net

³ A: Willy-Messerschmitt-Straße 1, 82024 Taufkirchen, Germany E: Anna.Straubinger@bauhaus-luftfahrt.net

⁴ A: Willy-Messerschmitt-Straße 1, 82024 Taufkirchen, Germany E: Kay.Ploetner@bauhaus-luftfahrt.net

⁵ A: Arcisstraße 21, 80333 Munich, Germany E: c.antoniou@tum.de

1. Introduction

Disruptive transport technologies are shaping the future of mobility by imposing new system requirements and challenging the way conventional transportation systems operate. In the coming decade, fully autonomous vehicles are likely to be launched, promising safe and comfortable transportation (Bimbraw, 2015). Ground shared autonomous mobility has accordingly gained a wide research interest (Fagnant & Kockelman, 2014; Narayanan, et al., 2020). Technological advances have led to the exploration of the third dimension in transportation: the sky. Urban air mobility (UAM) or aerial vehicle concepts for passenger transportation promise, according to Airbus (2017) a safer, more reliable, and more environmental alternative to alleviate congestion on transport networks. This research area has been witnessing an increase in activities and initiatives, worldwide. In Europe, manufacturers have been developing several concepts for UAM. Among others, (Airbus, 2017) unveiled the self-piloted multi-passenger "CityAirbus" and the singlepassenger "Vahana" concepts. In Southern Germany, startups are also actively involved in prototyping vehicle concepts, like (Lilium, 2018) and (Volocopter, 2018). To study the challenges and opportunities of such systems, regional and national governments are increasingly investing in this area. In Europe, the Commission (2018) is supporting the UAM initiative, as part of the European Innovation Partnership in Smart Cities and Communities (EIP-SCC). In the US, different partners and stakeholders are joining their efforts with the US National Aeronautics and Space Administration (NASA), to develop a framework for the integration of UAM airspace research (Thipphavong, et al., 2018).

Besides technological developments in terms of battery storage, electrical power transmission and distributed propulsion systems (Shamiyeh, et al., 2017), the passenger market of this new service has also been a research focus (Porsche Consulting, 2018; Uber Elevate, 2016), as well as the user perceptions and mode choice for this service (Fu, et al., 2019), and user adoption and use of this system (Al Haddad, et al., 2020). System requirements and operation constraints have been carefully researched, including aspects related to regulations, infrastructure availability, air traffic control, environmental impacts, but also community acceptance (Vascik, 2017). Rothfeld et al. (2020) gave an overview on these aspects by focusing on main factors of acceptance, potential passengers' value of time, modeling approaches, and possible spatial and welfare effects of UAM implementation.

Vertiport prototypes have been studied by (Cohen, 1996) to optimize land use, site selection, and community acceptance. From a land-use perspective, Rothfeld et al. (2020) focused on evaluating the service implications on the inhabitants and the changes on the city. An initial analysis of the system performance was also of interest and was researched by (Rothfeld, et al. (2018) in an agentbased simulation. As prototypes are built, crashes during test flights (O'Connor, 2019) demonstrate the need for a rigorous assessment methodology, to take into account system failures when evaluating the service among other transport systems. To the best of the authors' knowledge however, no studies have tackled the evaluation of UAM as a transport alternative, in other words the evaluation of different transport scenarios with and without UAM. The aim of this research is to fill this gap, and provide a first insight into assessing this service, based on a developed criteria system, to be applied to the assessment tool selected. For the case study in this paper, a multicriteria decision analysis (MCDA) was selected to make a choice between several alternatives; another motivation was to combine conflicting objectives (environmental, socio-economic, transport) including qualitative measures that would be difficult to express in monetary values as required in a cost-benefit analysis. While previous studies presented indicator frameworks for MCDA, few (if any) have focused on criteria selection for disruptive transport technologies, and most focused on sustainability plans for already-existing modes of transport. The main objectives of this research are therefore both methodological and practical; the former helps setting a framework to select and weigh suitable indicators for assessment methods of systems with high levels of uncertainty (such as UAM), while the latter offers a practical contribution into the selected indicators for UAM assessment. In the remainder of this paper, the multi-criteria assessment method used is presented, including its key aspects, typical methods used, applications, and a formulated gap analysis. After that, a methodology is proposed, focusing on the selection of indicators and their assessment based on relevance and feasibility, i.e. the possibility to measure them (quantitatively or qualitatively) within the project scope. In this method, the rating system is also investigated, by looking at two rating scales. A threshold method is provided for the selection of the indicators. As the multi-criteria analysis can be hierarchical, the main indicators can be defined according to the desired objectives; i.e. environmental, socio-economic, transport related; on the other hand, the sub-indicators are the key performance indicators selected following the expert assessment. The methodology is then applied for a case study of UAM in Upper Bavaria, Germany. The application is presented, and the selected indicators and appropriate weights given. After that, a discussion is given and selected indicators are compared with indicators from existing frameworks. Finally, a conclusion is drawn with insights for researchers and policymakers.

2. Multi-Criteria Decision Analysis

2.1 Steps and methods

The multi-criteria decision analysis (MCDA) is a decision support tool that assesses problems to select the best alternative following a set of conflicting criteria, related to the defined objectives (Yeh, et al., 1999). Dodgson et al. (2009) summarize the steps of an MCDA as follows.

- Define the aim of the MCDA, along with the selection of decision-makers for the assessment.
- Identify the alternatives to be assessed. ٠
- Outline objectives and criteria, which could have a hierarchy including high-level • objectives and low-level ones; main criteria and sub-criteria.
- Score alternatives based on the criteria defined.
- Weigh the criteria to discern between their relative importance in the decision. •
- Combine the weights with the scores to reach an overall value for each alternative.
- Assess the results.
- Conduct a sensitivity analysis for the obtained results. •

Decision-making problems can be categorized under full-aggregation methods and outranking methods (Guarini, et al., 2018). Full aggregation methods (multiple-attribute utility theory [MAUT], analytical network process [ANP], analytical hierarchy process [AHP]) aggregate different indicators into one global score. MAUT is based on the utility theory (Keeney & Raiffa, 1976), while AHP by Saaty (2008) develops priorities derived through pairwise comparison of criteria, and relies on expert judgment for deriving priority scales, where inconsistencies are measured and improved. ANP, an extension of the AHP, also assumes interdependency among various decision levels and criteria. When criteria can be assumed to be preferentially independent from each other, linear additive models can be used (Dodgson, et al., 2009) by multiplying criteria scores by the weight, to obtain an overall performance value. Conversely, outranking approaches (Preference Ranking Organization METHod for Enrichment of Evaluations [PROMETHEE] and the ELECTRE method [From French: ELimination Et Choix Traduisant la REalité, meaning elimination and choice translating reality]) evaluate a partial or total ordering by pairwise preferences among alternatives (Guarini, et al., 2018), evaluated along several attributes (Boucher, 2016).

When looking specifically at weighing indicators, these can be grouped under subjective and objective weighting methods (Zardari, et al., 2015). Subjective weighting methods depend on subjective judgments for weight assignments and include methods like ranking, point allocation,

pairwise comparisons (including AHP), swing method, and the Delphi method (Linstone, et al., 1975). On the other hand, objective weighting methods establish weights based on the analysis of the initial data, and include methods like the entropy method, the critic method, and the standard deviation method. In the Critic method, conventional statistical measures (means and standard deviation) are used, allowing two fundamental concepts of the decision process to be represented: the contrast intensity of the alternatives' performances in each single criterion and the conflict of the evaluation criteria with each other (Diakoulaki, et al., 1995). Statistical methods have been used by Wang and Luo (2010) to integrate correlations with standard deviations to determine attribute weights in multiple attribute decision making. Attributes are removed where correlations are very high as they have little impact on the overall assessment. Sensitivity analysis is also performed to control for the stability of the best decision alternative.

2.2 Applications for transport

Several methods have been used for multi-criteria analysis in transport projects, among which additive methods like MAUT and AHP. According to Tsamboulas et al. (1999), additive methods are very straightforward and rational in treating transportation decision problems and are usually linear. MAUT can be used with cardinal and qualitative data and has two main assumptions: error terms are not considered and the utility function is linear; in this method, appropriate utility transformations are done to treat uncertainty (Keeney & Raiffa, 1976) and the alternative with the higher rank is usually chosen.

MCDAs have been used for a wide range of transport applications, including road vehicles (Tzeng, et al., 2005), bicycle planning (Guy & Urli, 2006), port systems (Guy & Urli, 2006), bus rapid transits (BRTs) (Stamos & Triantafyllos, 2012), intermodal projects (Stoilova, et al., 2017), clean energy vehicles (Li, et al., 2019), bridge planning, and autonomous vehicles (Mathieu & Nongaillard, 2018; Owczarzak & Żak, 2015; Pickering, et al., 2018). Tzeng et al. (2005) applied MCDA using AHP to determine the relative weights of evaluation criteria in selecting alternative-fuel buses for public transportation. In this study, experts belonging to different decision-making groups performed the multiple attribute evaluation of alternative vehicles. In China, AHP and the VIKTOR (from Serbian: VIseKriterijumska Optimizacija I Kompromisno Resenje, meaning multi criteria optimization and compromise solution) optimization techniques were used to evaluate clean energy vehicles, taking into account multiple criteria like energy performance, energy cost, vehicular emission, market acceptance and energy security (Li, et al., 2019). For that, thirty-five experts were interviewed, including government officials, academic researchers and industrial executives to select and rank the criteria. In an intermodal project evaluation, Stoilova et al. (2017) developed a hierarchical MCDA using a main set of criteria including environmental, technological, social, and economic. Owczarzak & Żak (2015) used MCDA for assessing their concept of on-demand public transportation, based on autonomous vehicles. The alternatives included solutions based on driverless vehicles, but also traditional forms of passenger transportation. Criteria importance and sensitivity of the decision-maker towards change of the criteria values were crucial in defining a model of preferences for the assessment. Selected criteria included travel time, travel cost, comfort of travel, timeliness, availability, reliability, environmental friendliness, and safety. Moreover, eight variants were defined for the assessment. The final ranking of the alternatives was then applied using the ELECTRE III/IV method. MCDA was also used by Claussmann et al. (2018) to optimize functions such as safety, legal rules, preferences and comfort of passengers or energy consumption, and by Pickering et al. (2018) to support the evaluation of ethical decisions in autonomous vehicles collisions.

Despite the wide use of MCDA in transportation projects, a review by Macharis and Bernardini (2015) stressed on the importance of a stakeholder consideration, where a multi-actor approach could be more suitable for the assessment of transport projects by integrating the criteria defined by multiple stakeholders into one comprehensive evaluation process. The definition of criteria or indicators selection is a crucial step that paves the way to the evaluation of alternatives and is often

based on stakeholder participation. Existing frameworks for indicators selection often focus on sustainable urban mobility. Castillo and Pitfield (2010) defined a framework to identify and selected a small subset of sustainable transport indicators, including stakeholder participation and a consideration of environmental, social, and economic issues. The criteria formulated in the framework were then weighted according to sustainable transport objectives and change according to context. Experts were asked to perform a pairwise comparison using AHP for different indicators. In drawing policies for sustainable urban mobility, participative procedure was also used by Marletto and Mameli (2012) in the indicator selection process, starting from a framework based on expert consultation, followed by the engagement of citizens and stakeholders for a deliberative MCA.

In an effort to focus on sustainable urban mobiliy, the European commission set out in its Urban Mobility Package a concept for Sustainable Urban Mobility Plans (European Commission, 2013), which emerged from a broad exchange between stakeholders and planning experts across the European Union. Following these plans, Sustainable Urban Mobility Indicators (SUMI) have been drafted as tools for cities and urban areas to identify the strengths and weaknesses of their mobility system and to focus on areas for improvements resulting from new mobility practices or policies (European Commission, 2017). A sustainable urban mobility plan (SUMP) should start from a goal and objective clearly drawn. To translate it into easier to implement guidelines, briefings have been provided (https:// www.eltis.org/mobility-plans/practitioner-briefings) for cycling, walking, parking, ITS, electrification, shared mobility, etc.

2.3 Gap analysis

As mentioned in Section 1, and to the best of the authors' knowledge, no studies have tackled the assessment of UAM as a transport alternative. The development of an indicator system for this purpose is therefore crucial. While the presented methods and applications in Section 2 provide better insights into decision-making including transport applications, few focus on the selection of indicators and existing frameworks often address plans of sustainable urban mobility. A gap therefore arises for the assessment of disruptive transport technologies, based on different goals including but not limited to sustainability plans. Such modes, including urban air mobility, do not yet exist, and an assessment of different UAM scenarios would therefore need to evaluate whether these modes would bring positive contributions to existing systems (comparison with/without), in order to select the best alternative/scenario. The selection of indicators would therefore need to follow similar iterative processes as done in previous studies (Castillo & Pitfield, 2010; European Commission, 2017; Marletto & Mameli, 2012), yet in a more realistic and feasible way, not necessarily including a pairwise comparison/ an AHP process, which can be highly demanding/unrealistic, given a high number of starting indicators. The methodology would therefore aim to provide this set of indicators for UAM, and could be generalized to select and weigh indicators for systems associated with a high level of uncertainty, which have not been taken into account in existing indicator frameworks; for instance, practitioner briefings in SUMP do not include UAM or similar disruptive transport technologies.

3. Methodology

In this section, a methodology is formulated to develop an indicator system for the assessment of UAM, which could be generalized to systems with high levels of uncertainty, for which existing indicator systems might not be available. The method focuses on the selection and weighting of suitable KPIs (Key performance indicators), and the thresholds set for this selection. The advantage for using MCDA is the ability to combine conflicting objectives using indicators that include qualitative measurements that would be difficult to express in monetary value, as required in the cost-benefit analysis. Another motivation for the proposed approach is to make the method realistic, in terms of the specific context where the assessment takes place; therefore focusing on

the feasibility of each indicator (realistic indicators as pointed out by out by Tafidis et al. (2017)). Indicators resulting from the proposed approach can then be compared with SUMI (European Commission, 2017).

As indicated by the literature, an MCDA includes three main steps, namely:

- Step 1: Criteria selection
- Step 2: Criteria weighting
- Step 3: Criteria aggregation

Steps 1 and 2

As assessing future transport systems is often subject to a high level of uncertainty, especially when evaluating alternatives for short and long-term applications, subjective methods become inevitable, as uncertainty associated with the future rises. Subjective evaluation based on expert panel assessment might be therefore highly beneficial, as suggested in the Delphi method (Linstone, et al., 1975). When a large number of indicators is selected, pairwise comparison become less feasible and/or reliable, due to the high number of pairs to be compared. Therefore, to decrease the subjectivity of weighting assignment, these methods can be combined with objective methods using simple statistical analysis including mean and standard deviation, for a consistency check between experts. The scale used for the rating can vary and could be of significance to the selection process. An iterative 5-stage method is proposed for the KPI selection and weighing, based on the reasoning above. The method is presented in Figure 1 below.



Figure 1. 5-stage assessment method for KPI selection

The stages are described as follows.

- Stage 1: Experts from the related field are gathered and asked to propose key performance indicators that could be useful for assessing the transport project according to the required objective.
- Stage 2: Experts are asked to rate each proposed indicator based on its relevance (Re) and feasibility (Fe). Relevance indicates the perceived importance of an indicator ("how assessment?"), whereas important is it for the feasibility refers to the availability/accessibility of an indicator ("how feasible is it to obtain that indicator in the project/exercise scope?"). Experts are also given the option to provide comments on the KPI measurement.
- Stage 3: A preliminary analysis first eliminates the KPIs that are not applicable to the MCDA or that have duplicates. Further, the relation between relevance and feasibility is investigated.
- Stage 4: The suitable KPIs are selected according to their means and standard deviations. Some are removed and others are kept for a further qualitative analysis.
- Stage 5: The selected indicators are controlled by removing the ones with high correlation, since they have little impact on the overall assessment as suggested by Wang and Luo (2010). Then, the resulting KPIs are compared with findings from the literature. In particular, in the context of sustainable urban mobility, indicators can be compared with SUMI (European Commission, 2017).

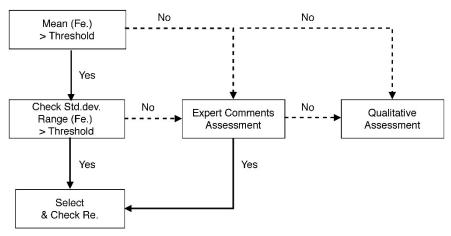


Figure 2. Thresholds for KPI selection

Based on the experts' rating, thresholds are set for selecting the KPIs according to both relevance and feasibility; these are detailed in Figure 2. The proposed threshold method stems from the diversity of the experts' opinions and backgrounds, and attempts to select indicators taking into account these differences in a consistent manner. Unlike the multi-actor multi-criteria analysis proposed by Macharis and Bernardini (2015), this method does not give different weights for different stakeholders; it aggregates the inputs of experts (who are also different stakeholders) by taking the means of different inputs. Therefore, each expert opinion has the same weight, and the experts/stakeholders are concerned with the same overall objective. The experts will therefore rate the indicators based on this understanding. Starting from the assessment of the feasibility of the indicator, a first comparison between the feasibility mean and the given threshold (depending on the used rating scale) is done. If the mean meets the required value, the standard deviation (stdv) is then checked, as well as the feasibility range ([mean - stdv ; mean + stdv]); if the latter meets the threshold, the KPI is selected, and its relevance is then checked (and used for the weighting). If threshold requirements are not met (feasibility mean is lower than the threshold), expert comments are reviewed; in case an input on measurement is provided, the indicator is selected. Otherwise, the KPI is kept for a further qualitative assessment (provided its relevance is above threshold).

The methodology is proposed for criteria selection, and weighting, using two rating scales: 1-3-9 and a 1-3-5-7-9. The former is based on the quality function deployment, where 1 indicates a low impact, 3 a medium, and 9 a high impact. Hunt (2015) suggest that this method is effective by clearly "popping out" some projects above the others; in this case some criteria. Therefore, the ones receiving a low rating of 1 could be easily removed. The latter scale is based on the AHP scale, where 1 represents the lowest impact and 9 the highest.

Step 3

After selecting and weighting the indicators, they are aggregated based on the MAUT. In transport projects, the multi-criteria decision analysis is often hierarchical in two or more layers. The upper layer often concerns main indicators like environmental, socio-economic, transport. The subindicators help to fulfill the goals of each main indicator, and thereby to reach the overall objective. In the proposed methodology, the main indicators are defined by the project objectives, with weights varying according to a sensitivity analysis. By fixing a higher weight for the environmental indicator, the project success would then highly depend on environmental outcomes. The subindicator weights however result from the relevance of the KPIs, according to the methodology proposed in Figures 1 and 2.

4. Setting the Scene

The methodology discussed in this paper was developed in the context of a project, funded by the Bavarian Ministry of Economic Affairs, Regional Development and Energy. The project aims to assess the long-term application potential of UAM as a complementing service offer to public transport on the use case of Upper Bavaria. In this context, UAM is defined as the transport opportunity enabling air passenger transportation, with a special focus on complementing existing public transport systems. In the course of the project, an agent-based modeling approach was used to quantify the impact of different UAM system configurations

For this purpose, the simulation was fed with possible UAM vertiport locations, service networks, operational concepts, vehicle concepts and developments of the transport service of other modes of transport. It was therefore necessary to tackle various topics in order to obtain numerical results enabling an assessment:

- Definition of city development over time (population growth, income development, expansion of urban area)
- Definition of transport system changes over time (infrastructure investments, changes in public transport schedules, changes in vehicle capacity, propulsion technology, operational concepts of ground transportation)
- Operational concepts of UAM (on-demand or scheduled, pricing, fleet size, level of autonomy)
- Vehicle concepts (energy usage, range, cruising speed, vertical distance, energy demand)
- UAM infrastructure (vertiport density, vertiport capacity, UAM flight routes, • communication infrastructure)

These factors can be partially evaluated independently. Some factors, like pricing or fleet size, are influenced by others like certification, market regulation, business models or regulation on efficient integration with public transport. A holistic evaluation of several different parts was thus necessary to achieve the overall goal of this research and deliver an assessment of the integration of UAM in public transport. The findings of this work can give guidance to potential stakeholders in the field of UAM. First profitable, social or environmentally-friendly system set-ups can be identified, thereby giving insight into open topics to address. Policy recommendations can promote efficient service designs that foster welfare and enable an efficient connection of UAM to existing transport systems.

Scenario Description

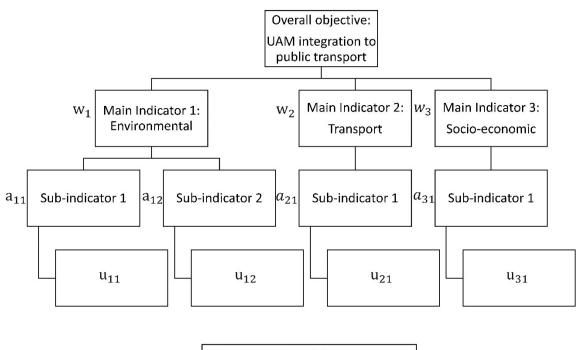
The impact of UAM overall performance was simulated using MATSim, the multi-agent transportation simulation (Horni, et al., 2016), with a UAM extension (Rothfeld, et al., 2018). Considering the effects of potential business models, several scenarios, including a baseline case without UAM and other cases with UAM, were formed based on the variation of a list of parameters. Sensitivity analysis was then conducted from the perspectives of UAM stations, UAM networks, UAM routing, and UAM vehicles, to better understand the impact of those pre-defined factors and how they may affect mobility behaviors. The variation of single parameters gave insight into relevant factors to set up efficient operational concepts. Combining different levels of the various parameters gave guidance on potentially profitable business models. Possible business models could include set-ups similar to current limousine services or a provision of UAM similar to current public transport services. Five scenarios were then defined, ranging from the lowest demand scenario, to the highest demand one, and are described in details in Fu et al. (2020). In this paper, we present the assessment framework performed for the three scenarios presented in Table 1, selected as the lowest, middle-ground, and highest demand scenarios. For each of these alternatives (Scenarios 1, 2, 3) projected for 2030, simulations were ran, in order to understand the potentials brought by this mode, and select the best scenario with respect to the desired goals, as described in Section 3; the assessment therefore had four total alternatives for the year 2030, for which simulations were runs: one alternative without UAM in Upper Bavaria, and three alternatives (Scenarios 1, 2, 3) for different UAM configurations.

UAM scenarios	Scenario 1	Scenario 2	Scenario 3
Network	Low density	Medium density	High density
Speed [km/h]	50	100	300
Base price [C/trip]	10	5	0
Km price [C/km]	5	2	1
Fleet size [veh/station]	10	50	100
Process times (preflight+ postflight) [min]	30	20	10
Seat capacity [PAX/veh]	2	2	4
Mode share [% of trips]	0.03%	0.62%	1.29%
Mode share [% of km]	0.05%	0.93%	1.60%

Table 1. UAM scenarios for the case study

5. Application and Results

The methodology proposed in this paper is applied to the case study presented above for the selection and weighting of suitable indicators. As explained, the aim of this assessment is to find the alternative most suitable in meeting the project objective: finding the potentials of UAM integration to public transportation systems in Upper Bavaria. This assessment includes, as described, an alternative without UAM and three others with UAM. The scenarios were assessed based on different levels; based on main indicators: environmental, socio-economic and transport based, and sub- indicators, resulting from the KPI selection process. The multi-criteria analysis applied to this case study is therefore hierarchical, as illustrated in Figure 3, where the main indicators used are displayed, and the sub-indicators are only indicatively represented.



Overall Utility: $\sum_{i,j} w_i a_{ij} u_{ij}$

Figure 3. Hierarchical Multi–Criteria Analysis

The main indicators were weighted based on the desired achieved objectives (environmental, socio-economic, or transport-related), following a sensitivity analysis to understand the impact of the associated weights on the alternative selection. The upper level weighting assignment (w1, w2, w3) shown in Figure 3) is not given in the following sections, as it goes beyond the scope of this paper. The lower level, or the sub-indicators weighting (a11, a12, a21, a31) stem from the methodology proposed in Figures 1 and 2; it is computed as the relative weight of the subindicators, and explained below. The case study outcomes allow therefore an appropriate selection of criteria and weighting. After simulating the proposed scenarios defined in Table 1, the KPI results were obtained, and then converted to utility points depending on the desired target for each (u11, u12, u21, u31). It was therefore crucial to set a desired target for each indicator, i.e. minimize or maximize. Finally, the overall utility was computed according to the formula given in Figure 3 (linear additive model) and the best-case alternative was chosen (highest score, as in full aggregation methods). In the below, the indicators selection and weighting for the case study are detailed according to the five-stage approach, and in two rounds, for each of the proposed rating scales for feasibility and relevance: 1-3-9 and 1-3-5-7-9.

Stage 1: Expert focus group

A focus group consisting of ten experts, including university and industry researchers, and representatives from the Munich Airport, were asked to provide KPIs they believed to be useful for the project assessment. These experts were stakeholders who were part of the funding project for this study. The initial expert focus group resulted in 48 KPIs, as presented in Table 5 in the Appendix.

Stage 2: KPI Rating

Experts were then asked to rate each indicator based on relevance and feasibility. This stage was done twice, for each of the rating scales. Experts' ratings were then aggregated by taking the means of relevance and feasibility values obtained.

Stage 3: Preliminary Analysis

First, a pre-selection of the indicators reduced the number of KPIs to 36, by removing some that were irrelevant or not compatible for the MCDA. This included for instance indicators that were specific to the UAM alternatives, for which no output could be obtained for the alternative without UAM. After that, the relationship between relevance and feasibility was investigated for both scales. The plots for these comparisons are shown in Figures 4(a) and 4(b). Though not directly feeding into the next steps, this analysis was useful to understand the assessment, and to compare both rating scales.

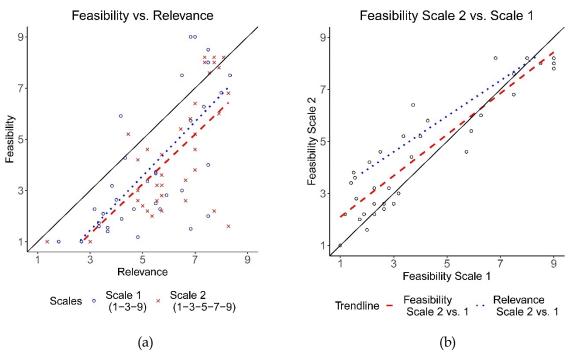


Figure 4. Feasibility and relevance analysis including a 45-degree solid line

A plot of feasibility and relevance is shown in Figure 4(a) shows the trend in this relation for both rating scales: 1 (1-3-9) and 2 (1-3-5-7-9), where feasibility is almost lower than relevance; the former scale is based on the method by Hunt (2015), while the latter aims to give a rather uniform ordinal scale with the same range (1–9). It can be noted that the linear regression for the first scale (solid line) is higher than that of the second scale (dashed line); this is mostly observed for the higher values and is fairly understandable as the first scale (1-3-9) only has three values, and therefore indicators that are estimated as important could be rated as 9, whereas in the second scale (1-3-5-7-9) they could also be weighted as 7.

Figure 4(b) shows that the second scale (1-3-5-7-9) is linearly related to the first scale (1-3-9), for both feasibility and relevance, which could indicate some consistencies in the ratings. Moreover, the trendline or regression line of relevance is above that of feasibility, which shows again that relevance is mostly higher than feasibility, according to the experts' rating.

Stage 4: KPI Selection

Following the approach given in Figure 2, the rating scale of 1-3-9 was given a threshold of 3 and the scale of 1-3-5-7-9 a threshold of 5, as these were the mid-points of medians of the scales. For both approaches, this resulted in 20 selected KPIs for the MCDA, 13 for a further qualitative analysis (Table 3), and 3 that were removed. Both rating scales resulted in the same KPI selection. However, as the second scale is richer in range and therefore provides more rating options, it was used to determine the criteria weight. For each sub-indicator, the relative weight was computed by simply calculating the ratio of its relevance over the sum of relevance of the corresponding main indicator.

Stage 5: Review

The selected KPIs were directly compared with related work in the literature. The summarized results of this 5-stage assessment are provided in Table 2. Resulting KPIs were grouped under the main indicators: environmental, transport, socio- economic (Table 2), and additionally land-use for the qualitative indicators (Table 3). Indicator descriptions for Tables 2 and 3 are given in details in Table 6 of the appendix.

Main indicator	Sub-indicator	Target	Relevance	Relative weight (%)
Environmental	Energy consumption	Minimize	7.7	22.9*
Environmental	Air emissions (CO_2 , $NO_{x,}$)	Minimize	7.6	22.4
Environmental	Noise emissions	Minimize	8.3	24.5
Environmental	Visual pollution	Minimize	5.7	17.0
Environmental	Average space required per transported passenger	Minimize	4.5	13.2
Transport	Total travel time saved	Maximize	8.1	9.8
Transport	Congestion (ground)	Minimize	7.4	8.9
Transport	Inconvenience (Access, egress, waiting	Minimize	7.9	9.6
-	time)			
Transport	On-time performance	Maximize	7.0	8.5
Transport	Induced demand	Minimize	6.8	8.3
Transport	Public transport modal share	Maximize	7.7	9.4
Transport	Total number of passenger trips	Maximize	7.9	9.6
Transport	Total passenger kilometres	Maximize	7.4	8.9
Transport	Investment costs	Minimize	7.0	8.5
Transport	Operating costs	Minimize	7.0	8.5
Transport	Safety (Number of Accident)	Maximize	8.3	10.0
Socio-economic	Privacy	Maximize	5.4	22.3
Socio-economic	Equity	Maximize	7.0	29.2
Socio-economic	Affordability (percentage of household	Maximize	5.2	21.6
	income devoted to transport)			
Socio-economic	Accessibility to employment	Maximize	6.5	26.9

Table 2. Final set of KPIs for the case study

* 22.9 = [7.7/(7.7 +7.6 +8.3 +5.7 +4.5)]*100

Table 3. KPIs for the qualitative assessment of the case study

Main indicator	Sub-indicator	Relevance
Environmental	Impact of construction	4.8
Transport	Security	7.7
Transport	Scalability	5.6
Transport	Flexibility	5.7
Socio-economic	Quality of life	5.2
Socio-economic	Social inclusion	5.7
Socio-economic	Life cycle (sustainability)	6.6
Socio-economic	Job impact	5.7
Land-use	Housing relocation	5.7
Land-use	Urban sprawl	5.6
Land-use	Income distribution over space	5.0
Land-use	Housing cost	4.8
Land-use	Population density	4.8

6. Discussion and Conclusions

6.1 Main findings

In this paper, the proposed methodology has been applied to the case study of UAM in Upper Bavaria, leading to the selection of 20 KPIs and 13 for a further qualitative assessment, out of an initial suggestion of 49 indicators. This is comparable to the study by Radović and Stević (2018), who go from around 62 KPIs to 20, in an assessment involving 19 decision-makers.

Both rating scales led to the selection of the same KPIs for this case study. Both were therefore useful in the 5-stage assessment, which could suggest that the weighting and threshold approaches were rather consistent. Still, the second scale (1-3-5-7-9) was used in computing the sub-indicators

weighting as it was assumed to be more precise due to the higher range it provides. The approach combining feasibility and relevance could prove to be of high usefulness for projects with a high level of uncertainty, mostly dealing with future system assessments, as it goes from the principle of selecting based on feasibility, since relevance is taken into account by given weights (unless relevance is "below than threshold"). The threshold assessment for the KPI selection could also be promising for analyzing expert assessment associated with one stakeholder only (in this case: the government), to ensure the selection of suitable indicators. Complementing the analysis with a qualitative analysis can ensure a comprehensive methodology, where indicators that cannot be included in the MCDA are taken into account through a balancing and discussion of the advantages and disadvantages they bring to the results. Furthermore, the assumption of independence between the criteria is tackled in the correlation analysis following stage 5, where indicators with high levels of correlation are removed. The proposed methodology therefore attempts to combine subjective methods (rating) with more objective statistical methods (standard deviation and mean).

In this case study, the simulation of scenarios given in Table 1 led to a matrix with values for each of the KPIs in Table 2. More details on the demand scenarios can be found in Fu et al. (2020). The KPI values and assessment results go beyond the scope of this study which aims to provide an assessment framework, rather than establish assessment results. However, for illustration purposes, we present final results of the scenario assessments, for which we take equal weights for the main indicators: environmental, transport, and socio-economic (Table 4). In this case, the best scenario (with highest overall utility) is scenario 3 (with the highest demand), followed by the alternative without UAM, then scenario 1, then 2. Although the numerical utility values are not key here, it is clear that for each main indicator, one scenario performs best. For instance, by looking only at the environmental aspect, scenario 1 outperforms the rest; scenario 3 outranks the other for each of the transport and socio-economic indicators. This only demonstrates that, beyond the choice of KPIs and their careful weighting, the overall assessment would still highly depend on the overall stakeholder objectives and views. A sensitivity analysis would therefore be crucial in determining the priorities of the assessment: environmental, transport, or socio-economic, or a careful balance of the three.

Main indicator	Main weight	No UAM	Scenario 1	Scenario 2	Scenario 3
Environment	1/3	96.57	99.99	41.41	0
Transport	1/3	54.00	52.49	31.42	66.65
Socio-economic	1/3	12.34	0.88	38.82	100
Total	1	54.30	51.12	37.22	55.55

Table 4. MCDA example results for the case study

The proposed method can therefore serve as a tool to draw indicator sets for disruptive transport technologies, as few studies have explored their assessment among other transport alternatives. This was the case for this study; using the proposed methodology, indicators were highlighted that would otherwise not have been captured (using existing indicator frameworks). The contributions of this work are therefore twofold; first practical by providing a framework for the selection of suitable indicators for the assessment of UAM, and second methodological by developing a methodology that could be further extended to disruptive systems with high uncertainty.

6.2 Comparison with related work

Despite the lack (to date) of studies in UAM assessment, similar research could give meaningful insights on the validity of this work. Owczarzak & Żak (2015) for instance presented the following measures as evaluation criteria for assessing public transportation solutions based on autonomous vehicles : travel time, travel costs, travel comfort, reliability, timeliness, availability, environmental friendliness, safety. These were partially observed in the proposed indicator system.

In order to assess the validity of the selected indicators, we compared them with the ones from the literature, notably SUMI (European Commission, 2017), which are indicators in the context of sustainable urban mobility plans. The indicators presented in Tables 2 and 3 highly overlap with SUMI indicators on sustainability like energy consumption, air emissions, noise emissions, average space per passenger, total travel time, congestion, inconvenience, modal share, safety, equity, affordability, accessibility, security, and quality of life. While the resulting indicators do not necessarily have the same exact formulation, they map out the main ideas of the indicators; for example the closest indicator on average space per passenger (Table 2) would be the "space usage indicator" in SUMI. Some SUMI indicators were not reflected in the selected indicators like multimodal integration indicator, active mobility indicator, and urban functional diversity indicator. On the other hand, the results captured additional indicators (not in SUMI), such as visual pollution (more UAM specific), but also transport indicators like on-time performance, induced demand, total passenger kilometres, total number of passenger trips, costs (operation, investments). Other qualitative indicators were also highlighted by the study including construction impact, scalability, flexibility (of vertiport and infrastructure), job impact (due to automation), and land-use indicators.

6.3 Limitations

Still, the approach does not come without limitations. Subjective indicator weighting leads inevitably to human biases, mostly in providing subjective judgments and preferences, even when consistency is checked; this can also be a result of the very selective group of experts in terms of number and expertise. Due to the high number of indicators, this uncertainty increases. However, some methods become inconceivable with the increasing number of indicators, such as the AHP, which would require a very high number of pairwise comparisons. An alternative could be to use the best worse case method (Rezaei, 2015), which relies on best and worst criteria identified by decision – makers and pairwise comparison only between the best and worst – case indicators and the rest, to compute the criteria weights and obtain the consistency ratios.

6.4 Future research

Findings from this study pave the way to possible extensions of urban mobility frameworks to include recent mobility trends. Policymakers should look into the inclusion of disruptive transport technologies such as UAM, in existing plans and indicator systems, to further facilitate their assessment among conventional transport modes. Future research can validate both practical and methodological contributions of this paper by validating the resulting UAM indicators and applying the proposed method to further disruptive transport technologies.

Acknowledgements

This research has been partly funded by the Bavarian Ministry of Economic Affairs, Regional Development and Energy (Grant no.: LABAY101A).

References

Airbus, Urban 2017. Rethinking Air Mobility. [Online] Available at: https://www.airbus.com/newsroom/stories/rethinking-urban-air-mobility.html

Al Haddad, C. et al., 2020. Factors affecting the adoption and use of urban air mobility. *Transportation* Research Part A: Policy and Practice, Volume 132, p. 696-712.

Bimbraw, K., 2015. Autonomous Cars: Past, Present and Future - A Review of the Developments in the Last *Century, the Present Scenario and the Expected Future of Autonomous Vehicle Technology.* s.l., s.n., p. 191–198.

Boucher, P., 2016. 'You Wouldn't have Your Granny Using Them': Drawing Boundaries Between Acceptable and Unacceptable Applications of Civil Drones. Science and Engineering Ethics, 01 10, Volume 22, pp. 1391_-1418.

Castillo, H. & Pitfield, D. E., 2010. ELASTIC-A methodological framework for identifying and selecting sustainable transport indicators. Transportation Research Part D: Transport and Environment, Volume 15, p. 179-188.

Claussmann, L. et al., 2018. Multi-Criteria Decision Making for Autonomous Vehicles Using Fuzzy Dempster-Shafer Reasoning. s.l., s.n., p. 2195–2202.

Cohen, M. M., 1996. The Vertiport as an Urban Design Problem, s.l.: s.n.

Commission, E., 2018. Commission Welcomes European Cities Joining the Urban Air Mobility Initiative. [Online]

Available at: https://ec.europa.eu/transport/media/news/news/2018-05-30-commission-welcomeseuropean-cities-joining-urban-air-mobility-initiative_en

Diakoulaki, D., Mavrotas, G. & Papayannakis, L., 1995. Determining Objective Weights in Multiple Criteria Problems: The Critic Method. Computers & Operations Research, Volume 22, p. 763-770.

Dodgson, J. S., Spackman, M., Pearman, A. & Phillips, L. D., 2009. Multi-criteria analysis: a manual.

European 2013. Sustainable Urban Mobility Plans. [Online] Commission, Available https://ec.europa.eu/transport/themes/urban/urban-mobility/urban-mobilityat: actions/sustainable-urban_en

European Commission, 2017. Sustainable Urban Mobility Indicators. [Online] Available at: <u>https://ec.europa.eu/transport/themes/urban/urban_mobility</u>/sumi_en

Fagnant, D. J. & Kockelman, K. M., 2014. The Travel and Environmental Implications of Shared Autonomous Vehicles, Using Agent-Based Model Scenarios. Transportation Research Part C: Emerging Technologies, Volume 40, p. 1-13.

Fu, M., Rothfeld, R. & Antoniou, C., 2019. Exploring Preferences for Transportation Modes in an Urban Air Mobility Environment: a Munich Case Study. Transportation Research Record: Journal of the Transportation Research Board, Volume 2673, pp. 427-442.

Fu, M., Straubinger, A. & Schaumeier, J., 2020. Scenario-based Demand Assessment of Urban Air Mobility in the Greater Munich Area. In: AIAA AVIATION 2020 FORUM. s.l.:s.n.

Guarini, M. R., Battisti, F. & Chiovitti, A., 2018. A methodology for the selection of multi-criteria decision analysis methods in real estate and land management processes. Sustainability, Volume 10, p. 507.

Guy, E. & Urli, B., 2006. Port Selection and Multicriteria analysis: An Application to the Montreal-New York Alternative. Maritime Economics & Logistics, Volume 8, p. 169–186.

Horni, A., Nagel, K. & Axhausen, K. W., 2016. The multi-agent transport simulation MATSim. s.l.: Ubiquity Press London.

Hunt, D. O., 2015. Popping a Project Using a 1-3-9 Ranking System in Selecting a Project. [Online] Available at: https://davidhuntpe.files.wordpress.com/2015/10/use-of-a-1-3-9-scoring-system-forevaluating-project-alternatives.pdf

Keeney, R. L. & Raiffa, H., 1976. Decision Analysis with Multiple Conflicting Objectives. Wiley& Sons, New York.

Li, C. et al., 2019. Multi-Criteria Analysis of Policies for Implementing Clean Energy Vehicles in China. *Energy policy,* Volume 129, p. 826–840.

Everyone. Aircraft [Online] Lilium, 2018. The for Available at: https://lilium.com/mission/

Linstone, H. A., Turoff, M. & others, 1975. The Delphi Method. s.l.:Addison-Wesley Reading, MA.

Macharis, C. & Bernardini, A., 2015. Reviewing the Use of Multi-Criteria Decision Analysis for the Evaluation of Transport Projects: Time for a Multi-Actor Approach. Transport Policy, Volume 37, p. 177-186.

Marletto, G. & Mameli, F., 2012. A participative procedure to select indicators of policies for sustainable urban mobility. Outcomes of a national test. European transport research review, Volume 4, p. 79-89.

Mathieu, P. & Nongaillard, A., 2018. Effective Evaluation of Autonomous Taxi Fleets. s.l., s.n.

Narayanan, S., Chaniotakis, E. & Antoniou, C., 2020. Shared autonomous vehicle services: A comprehensive review. Transportation Research Part C: Emerging Technologies, Volume 111, p. 255–293.

O'Connor, 2019. eVTOL K., Boeing Prototype Crashes. [Online] Available https://www.avweb.com/recent-updates/evtols-urban-mobility/boeing-evtolat: prototype-crashes/

Owczarzak, Ł. & Żak, J., 2015. Design of Passenger Public Transportation Solutions Based on Autonomous Vehicles and Their Multiple Criteria Comparison with Traditional Forms of Passenger Transportation. Transportation Research Procedia, Volume 10, p. 472–482.

Pickering, J. E. et al., 2018. Model-to-Decision Approach for Autonomous Vehicle Convoy Collision Ethics. s.l., s.n., p. 301-308.

Porsche Consulting, 2018. The Future of Vertical Mobility: Sizing the Market for Passenger, Inspection, and Good Services until 2035. s.l.:s.n.

Radović, D. & Stević, Ž., 2018. Evaluation and selection of KPI in transport using SWARA method. The International Journal of Transport & Logistics, Volume 18, p. 60-68.

Rezaei, J., 2015. Best-Worst Multi-Criteria Decision-Making Method. Omega, Volume 53, p. 49-57.

Rothfeld, R., Balac, M., Ploetner, K. O. & Antoniou, C., 2018. Initial Analysis of Urban Air Mobility's Transport Performance in Sioux Falls. Atlanta, s.n.

Rothfeld, R. et al., 2020. Urban air mobility. In: Demand for Emerging Transportation Systems. s.l.:Elsevier, p. 267-284.

Saaty, T. L., 2008. Decision Making with the Analytic Hierarchy Process. International Journal of Services Sciences, Volume 1, p. 83-98.

Shamiyeh, M., Bijewitz, J. & Hornung, M., 2017. A Review of Recent Personal Air Vehicle Concepts. Bucharest, Romania.

Stamos, I. & Triantafyllos, D., 2012. Bus Rapid Tranist as a Contribution to Sustainable Transportation. The Case of Athens, Greece. Journal of Environmental Protection and Ecology, Volume 13, p. 611-619.

Stoilova, S., Kunchev, L. & others, 2017. Study of Criteria for Evaluation of Transportation with Intermodal *Transport.* s.l., s.n., p. 24–26.

Tafidis, P., Sdoukopoulos, A. & Pitsiava-Latinopoulou, M., 2017. Sustainable urban mobility indicators: policy versus practice in the case of Greek cities. Transportation research procedia, Volume 24, p. 304–312.

Thipphavong, D. P. et al., 2018. Urban Air Mobility Airspace Integration Concepts and Considerations. s.l., s.n., p. 3676.

Tsamboulas, D., Yiotis, G. S. & Panou, K. D., 1999. Use of Multicriteria Methods for Assessment of Transport Projects. Journal of Transportation Engineering, Volume 125, p. 407–414.

Tzeng, G.-H., Lin, C.-W. & Opricovic, S., 2005. Multi-criteria analysis of alternative-fuel buses for public transportation. Energy policy, Volume 33, p. 1373–1383.

Uber Elevate, 2016. Fast-Forwarding to a Future of On-Demand Urban Air Transportation.

Vascik, P. D., 2017. Systems-Level Analysis of On Demand Mobility for Aviation, s.l.: s.n.

Volocopter, 2018. Reinventing Urban Air *Mobility*. [Online] Available at: https://www.volocopter.com/en/

Wang, Y.-M. & Luo, Y., 2010. Integration of Correlations with Standard Deviations for Determining Attribute Weights in Multiple Attribute Decision Making. Mathematical and Computer Modelling, Volume 51, p. 1-12.

Yeh, C.-H., Willis, R. J., Deng, H. & Pan, H., 1999. Task Oriented Weighting in Multi-Criteria Analysis. European journal of operational research, Volume 119, p. 130-146.

Zardari, N. H., Ahmed, K., Shirazi, S. M. & Yusop, Z. B., 2015. Weighting Methods and Their Effects on Multi-Criteria Decision Making Model Outcomes in Water Resources Management. s.l.:Springer.

Appendix A

Main indicator	Sub-indicator
Environmental	Energy consumption
Environmental	Air emissions (CO2, Nox,)
Environmental	Noise emissions
Environmental	Visual pollution
Environmental	Average space required per transported passenger
Environmental	Battery life-cycle
Environmental	Impact of construction
Environmental	Animal impact
Environmental	High frequency /5G influence
Environmental	Immission
Transport	Total travel time saved
Transport	Congestion (ground)
Transport	Inconvenience (Access, egress, waiting time)
Transport	On-time performance
Transport	Induced demand
Transport	Public transport modal share
Transport	Total number of passenger trips
Transport	Total passenger kilometers
Transport	Investment costs
Transport	Operating costs
Transport	Average travel speed per mode
Transport	Availability
Transport	Flexibility
Transport	Scalability
Transport	Vehicle-km travelled
Transport	Lead time to the market
Transport	Utilization rate (time used vs. time on the ground)
Transport	Efficiency (empty flights)/fleet management
Socio-economic	Privacy
Socio-economic	Equity
Socio-economic	Affordability (percentage of household income devoted to transport)
Socio-economic	Accessibility to certain areas
Socio-economic	Accessibility to employment
Socio-economic	Job impact
Socio-economic	Sustainability
Socio-economic	Willingness to pay
Socio-economic	Travel time budget
Socio-economic	Quality of life/welfare
Socio-economic	Social inclusion
Socio-economic	Seamless mobility
Socio-economic	Impact of doubters
Socio-economic	Restrictions (luggage and group size)
Land-use	Relocation of companies and households
Land-use	Urban sprawl
Land-use	Population density
Land-use	Income distribution over space
Land-use	Housing cost

Table 5. Initial set of KPIs for the case study

Appendix B

Table 6.	Description of selected KPIs (for both the MCDA and qualitative assessment)

KPI	Description
Energy consumption	Total energy consumption (kWh)
Air emissions	Total air emissions (CO ₂ , NO _x)
Noise emissions	Index based on dB emitted at a zone divided by the population density
	in this zone
Visual pollution	Index based on kilometres travelled above a zone divided by the population density in this zone
Average space required per	For UAM: Vertiport area (Sqm)/Passenger
transported passenger	for or him. Veruport area (squit), i assenger
Total travel time saved	Total travel time saved compared with status-quo (minutes)
Congestion	Congestion on the ground (hours or vehicle-kilometres travelled)
Inconvenience	Total time spent in access, egress, and waiting time (minutes)
On-time performance	Reliability or on-time performance of the alternative
Induced demand	Induced demand for transport (increase in number of trips)
	resulting from the chosen alternative
Public transport modal share	Percentage share of public transport
Total number of passenger trips	Total number of trips travelled
Total passenger kilometres	Total kilometres travelled
Investment costs	For UAM: investment costs of the chosen alternative
Operating costs	For UAM: operating costs of the chosen alternative
Safety	Number of accidents
Privacy	For UAM: sum of affected dwellings due to take-offs
	and landings in a buffer area around each vertiport
Equity	Equity in the incomes of UAM users
	(percentage difference in incomes of UAM users vs. non-UAM users)
Affordability	Percentage of household income devoted to transport
Accessibility to employment	Average accessibility to employment per zone
Impact of construction	Environmental impact resulting from construction or infrastructure needed
Security	Perceived security of passengers
Scalability	Extent to which alternative is scalable (in the case of UAM, a change ir
	vertiport configuration without big efforts)
Flexibility	Flexibility in the network design
Quality of life	Quality of life or welfare of the community
Social inclusion	Social inclusion of groups with reduced mobility
Life cycle	Life cycle or sustainability of an alternative (how much does it last, or is
Life cycle	it reusable?)
Job impact	Impact of alternative on jobs (due to automation in the case of UAM)
Housing relocation	Housing relocation as an impact of land-use change from UAM
Urban sprawl	Urban sprawl as an impact of land-use change from UAM
	Change of income distribution over space as an impact of land-use
Income distribution over space	change from UAM
Housing cost	Change in housing cost as an impact of land-use change from UAM
Population density	Change in population density as an impact of land-use change from UAM