



HAL
open science

Animating Hypothetical Trips to Communicate Space-Based Temporal Uncertainty on Digital Maps

Morgane Koval, Yvonne Jansen, Fanny Chevalier

► **To cite this version:**

Morgane Koval, Yvonne Jansen, Fanny Chevalier. Animating Hypothetical Trips to Communicate Space-Based Temporal Uncertainty on Digital Maps. IEEE Transactions on Visualization and Computer Graphics, In press, pp.1-11. 10.1109/TVCG.2024.3388517. hal-04552326

HAL Id: hal-04552326

<https://inria.hal.science/hal-04552326>

Submitted on 19 Apr 2024




HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Animating Hypothetical Trips to Communicate Space-Based Temporal Uncertainty on Digital Maps

Morgane Koval , Yvonne Jansen , Fanny Chevalier 

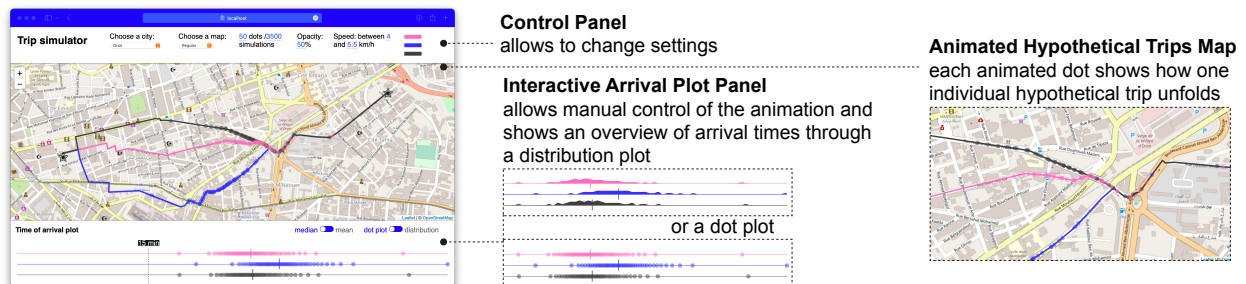


Fig. 1: Our tool is composed of three main components: (1) a control panel used to change visual settings or model parameters; (2) a map with embedded animated hypothetical trips, each displayed as a moving dot along a route; and (3) an interactive arrival plot panel which shows the arrival times of each hypothetical trips either through a dot plot or a distribution plot, along with the mean or the median, and which allows manual control of the animation in order to replay specific time ranges or pause on a particular timestamp.

Abstract—This paper explores a novel approach to communicating plausible space-based temporal variability of travel durations. Digital maps most often only convey single numerical values as the estimated duration for a path and this piece of information does not account for the multiple scenarios hidden behind this point estimate, nor for the temporal uncertainty along the route (e.g., the likelihood of being slowed down at an intersection). We explore conveying this uncertainty by animating hypothetical trips onto maps in the form of moving dots along one or more paths. We conducted a study with 16 participants and observed that they were able to correctly extract and infer simple information from our uncertainty visualizations but that identifying moving dots' changes in speed is a more complex task. We discuss design challenges and implications for future visualizations of space-based temporal uncertainty.

Index Terms—Geospatial Visualization, Uncertainty Visualization, Animation, Hypothetical Outcome Plot

1 INTRODUCTION

Trip durations are generally uncertain to varying degrees. Yet, we are often presented with a single numerical value of the overall estimated time a trip would take. This leaves out the details about the variability behind this estimate. Without such details, relevant questions would be difficult to answer: What are the estimates in a best or worst case scenario? What are the chances I arrive around the overall estimate? Where may I find myself slowed down? To support such tasks, temporal uncertainty must be conveyed relative to spatial information.

There exist many representations which visualize space and time dimensions [9]. One such approach consists of distorting space, for example, using Transmogrification [18], so as to create linear cartograms where the distance between two points represents a duration. Since they distort a path's geometry, such as turns or curvature, these techniques may occlude valuable information in the context of navigation tools [35]. An alternative is to embed temporal data onto the space, that is, to directly visualize temporal information on a path's geometry through ticks as time units [13, 48], or visual properties of the path's stroke [7]. While these are viable options for encoding single time

values, communicating also some form of time uncertainty for a path would clutter the map and render it illegible.

Communicating uncertainty visually is a challenge [36]. A review of prior work [47] suggests that approaches which use a frequency framing of uncertainty—such as using 1 in 10 instead of 10%—are an efficient way to communicate uncertainty. This is the case, for example, for quantile dot plots [26, 40] and hypothetical outcome plots [37]. These types of visualizations can be used to show *global* temporal uncertainty of an event (e.g., when a bus may arrive [26, 40]) by visualizing a representative sample of plausible outcomes of the event. However, they do not address *local* temporal uncertainty, i.e. what happens during a trip. Information such as where and how long one might plausibly get stuck can be important when deciding which route to take.

To address these challenges, we propose visual means to communicate temporal uncertainty in relation to space. This includes both global (i.e., how long the trip might take overall) and local temporal uncertainty (i.e., whether a specific location influences travel time). We describe an approach which visualizes plausible trip instances along a route, through (i) an animated hypothetical trips map showing how each plausible trip would unfold in space and through time, and (ii) an arrival plot showing for each route the distribution of the overall duration of the hypothetical trips (see Figure 1).

Combining concepts from work on hypothetical outcome plots [37, 39] and on traffic flow visualization [46], we render simulated instances of a trip as animated dots that stop, slow down or accelerate along a path on a map. This technique calls for a large number of simulations—at least 500—to be run to approximate a representative sample, but displaying all of them can hinder viewers' ability to extract useful information. Building on Kay et al.'s *quantile* dot plots [40], we similarly use quantiles of trips' overall duration to select a smaller amount of dots that are still representative of the variability between trips.

- Morgane Koval is with Inria Centre at the University of Bordeaux, France. E-mail: morgane.koval@inria.fr.
- Yvonne Jansen is with Univ. Bordeaux, CNRS, Inria, LaBRI, France. E-mail: yvonne.jansen@cnrs.fr.
- Fanny Chevalier is with University of Toronto, Canada. E-mail: fanny@cs.toronto.edu.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxx

We chose animation because it helps with the problem of occlusion on maps, which are already information-rich. Still, animated maps can be cognitively demanding [33] and make it difficult to extract information such as the average time a route might take. We designed interactive arrival plots to provide a quick visual overview of the underlying distribution of arrival times, and to allow control over the animated trips, e.g. hovering over the 10 minutes timestamp on the plot would show where hypothetical trips would then be on the map.

We conducted a study with 16 participants to evaluate (1) whether viewers could understand and extract data from such animations and visualizations, (2) to what extent adding space-based temporal uncertainty about a path could be valuable, and (3) whether such a feature would hinder the usual map experience. We observed that our solution did not restrict map exploration as participants were still able to perform typical tasks such as retrieving a path’s average duration, or construct and evaluate an itinerary. Participants were also able to extract new kinds of information and compare paths according to outliers (e.g., which path would be the slowest in the worst-case scenario) using the arrival plots. Additionally, our results show that they were able to reason about both space and time making use of the interaction between the hypothetical trips map and the arrival plots.

Our contributions are (1) a general approach to space-based temporal uncertainty communication; (2) quantitative and qualitative empirical data evaluating different aspects of the approach through a specific use case; and (3) recommendations for future work on animated maps aiming to showcase space-based temporal uncertainty.

2 RELATED WORK

Our work explores how to convey space-based temporal uncertainty through the use of animation of hypothetical trips. We discuss previous work in relation to space and time visualization, animation, and uncertainty visualization, which informed our design process.

2.1 Space and Time Visualizations

Visualizing time and space dimensions in conjunction is a long-standing challenge [3]. The difficulty stems from presenting three information-rich dimensions, also referred to as the *space-time cube* [2, 5, 9, 28, 43], onto a 2D canvas to support spatio-temporal reasoning. To support diverse analytic tasks, researchers have proposed techniques which slice the space-time cube to show a point in time (e.g. [8, 30]), drill it to extract the timeline of a point in space (e.g. [59]), flatten it to show timelines onto space (e.g. [4, 6, 8, 51]), decouple space and time as separate synchronized views (e.g. [15]), and more. Bach et al.’s survey [9] provides a general space-time cube framework and operations for spatio-temporal visualizations and analysis.

Specifically, we are interested in visualizing entities that move over space and through time, i.e., how hypothetical trips would unfold in the context of a geographical map. Doing so for one instance is feasible, e.g. by encoding trip progression as ticks along a route [13] or by varying visual properties of the path locally, such as thickness or saturation [48]. When many such trajectories exist, 2D and 2.5D static visualizations are prone to visual clutter [2, 5], or offer low spatial resolution when they rely on slicing and spreading time as small-multiples [17]. Animation of particles along their respective motion path is a popular alternative approach to depict such data with representative examples being traffic flow visualizations, e.g. [20, 46, 53], and animated migration maps [57]. Animated maps have been shown to outperform small-multiples for visually identifying space-time clusters [30], and identifying geographical local events and changes between subsequent timestamps [17], and are thus a promising approach in our context.

Small-multiples, animated maps, embedded visualization [4], and 2.5D visualizations [5] all put an emphasis on the spatial dimensions as the primary substrate onto which temporal data is displayed. Another approach is to relax this spatial substrate and instead leverage the spatial dimensions of the canvas to encode time. For instance, linear cartograms encode time as the distance between two points [14, 38], hence distorting the original space substrate (i.e. the map). Such technique can be beneficial to retrieve information about a route duration, but comes at the cost of accurately representing shape and geometry

of a path [35]. If distorting space must be avoided, then the line can be curved in a sinusoidal manner so that duration matches the path length [19]. However, these curves cannot always espouse the shape of a route, making spatial reasoning difficult.

Space and time visualization is a rich and active field, with many approaches to bear inspiration from. In this work, we are specifically interested in communicating space-based temporal uncertainty, an aspect that has not been empirically investigated to the best of our knowledge.

2.2 Animation

Bartram et al. defined motion as a distinctive aspect of display, emphasizing its unique potential for perceptual efficiency and interpretive complexity [10, 12]. As such, animation has been leveraged in visualization as a powerful technique to encode data [21], with a variety of purposes including revealing relationships in data (e.g. [52]), encoding emotion (e.g. [11, 58]), encoding object attributes (e.g. [25]), and conveying uncertainty or randomness (e.g. [32, 37]).

Animations are particularly well-suited for visualizing geo-temporal data. By mapping the data time dimension to the time dimension of the visual representation, animated maps “replay” the temporal evolution of space-based information, without relying on visual variables which would create visual clutter. For instance, animated flow maps [20, 46, 53] is a popular approach to depicting moving entities, like migrating birds or aerial traffic. Similarly, animated choropleth maps [34] are effective at emphasizing changes between timestamps, with examples including depicting the relative variability of temperatures, populations, or rainfalls over time. However, they could also make it difficult to see change in patterns, as each frame is displayed only very briefly [27, 29].

Specifically, our work is interested in visualizing the outcomes of simulations, as a means to communicate space-based temporal uncertainty of hypothetical trips. One of our inspirations is the work of Hullman et al. which uses animation to display how different hypothetical alternatives result in different summary visualization outcomes (e.g. boxplots) [37]. For maps, animation has been suggested as a way to display spatial variability between different simulations to facilitate analysis, including for land elevation [23] or land usage [16]. On the example of land usages, participants in one study found the animation useful [24], whereas in other studies, participants thought it was great as a first approach to uncertainty but preferred static visualizations [1, 16].

Our work leverages animation’s joint qualities as a tool to visualize geo-temporal visualization, and communicating uncertainty, and studies their potential for trip planning.

2.3 Uncertainty Visualization

Uncertainty can refer to different phenomena (see work by Spiegelhalter [54] for a deeper discussion), but this article is concerned with the visualization of aleatoric uncertainty (i.e., calculable risk given a mathematical model to make predictions of future outcomes). In more details, the common problem we aim to address relates to space-based temporal uncertainty, that is, a geo-temporal data set with inherent time variability across different points.

Temporal uncertainty can be due to a lack of information about when an event starts, ends, or what is its overall duration. If the goal of the visualization is to help read durations, then Geschwandtner et al. suggest that using a lighter color or error bars is efficient, but that gradient colors may be more appropriate to help judge probabilities [31].

Kay et al. proposed quantile dot plots to communicate when a bus might arrive (i.e., uncertainty about the start of an event) and report better transit-related decision-making than when using density plots [26, 40]. This visualization was also studied in the context of predictive uncertainty around task duration estimates, and, combined with other debiasing methods, it induced less optimistic estimates [42]. Quantile dot plots are a way to show large data sets by selecting a representative subset, which applies well to predictive visualization problems as predictions generally require large numbers of simulation to be run. On top of cluttering the space, large data sets—and especially large spatial data sets from predictions—can result in visualizations that prompt erroneous judgments from viewers who mistake uncertainty for another attribute. Liu et al. addressed this specific problem in the

context of ensemble visualizations of hurricane path predictions by selecting representative samples of a data set preserving spatial attributes [44]. Similarly, hypothetical outcome plots (HOPs) animate sequentially between a set of hypothetical outcomes. Hullman et al., and Kale et al. compared HOPs to conventional static visualizations (violin plots, error bars, or line ensembles) and both reported more accurate judgments about underlying distributions [37, 39].

In our case, we do not solely consider *global* temporal uncertainty about the overall duration, but also *local* temporal uncertainty due to spatial components of the temporal data, such as the likelihood of stopping at an intersection.

The field of GIS (geographic information system) has also studied how to visualize uncertainty in geo-temporal data sets, for example, in the context of meteorology where both experts and the general public want to see data visualized. Multiple techniques to include temporal uncertainty in geospatial visualizations have been explored, such as using smooth versus abrupt edges in 3D, blurring areas, or applying gradients [45]. Depending on the field, experts might be more comfortable with a different visualization [41]. In a study on land allocation issues, participants (who were experts) thought that communicating uncertainty was favorable as it helped clarify the information without making it harder to understand [1]. Designing uncertainty visualizations for non-experts can be different as they are less used to seeing the data, and it often has to be simplified. Yet, concealing information could result in errors of judgments. Displaying uncertain pollution data on a map, instead of a single numerical value, was shown to prompt more conservative decisions (i.e., less risky) [49].

When designing for uncertainty visualizations, it is important to consider where the uncertainty is coming from, what must be communicated and to whom. In our case, the uncertainty is coming from statistical variability (predictions), and we aim to convey how traveling along a route may unfold to a general audience.

3 CONCEPT AND IMPLEMENTATION

We propose a new technique to visualize how uncertainty concerning traveling along a geospatial path may unfold, by animating the movement of simulated instances—hypothetical trips—shown as dots, along the path. We thus visualize uncertainty *implicitly* [22] by showing a subset of plausible hypothetical trips—a similar approach to quantile dot plots [26, 40]—using animation—inspired by hypothetical outcome plots (HOPs) [37, 39]—to avoid the need for an additional encoding channel on an already information dense map. We describe here the visualization pipeline from the data source to the animation and interaction. Note that the main contribution of this article is the concept, and that any concrete implementation of the concept requires to make some choices along the pipeline as detailed in this section.

3.1 Data Source: Means of transportation

Our technique requires time-stamped geospatial paths, where paths are routes taken by drivers, bike riders, or pedestrians. For each of these modes of transportation, a different set of route characteristics and factors are relevant to consider with respect to how traveling along this path may unfold. These characteristics and factors become parameters in our data model. Drivers, for example, can be stuck for a long time in a traffic jam and they cannot run through red lights. Pedestrians rarely take roads where there is no sidewalk, but, depending on local culture, they may choose to cross at a red light even when formally forbidden. People on bikes generally prefer avoiding steep hills.

While our approach is generalizable to various transportation means, our implementation focuses on pedestrians as a use case. This allows us to consider only factors relevant to this type of transportation. However, the tool design and associated visualization can be equally applied to simulations built from models tailored for driving, biking, maybe also train rides, or combinations thereof.

3.2 Generating a Sample of Hypothetical Trips

Travel uncertainty can be driven by numerous factors such as stoplights, traffic, intrinsic speed, and more. In order to generate a set of plausible travel scenarios for a given path, we built a model to simulate plausible

walking progressions through the path. For instance, some people may be lucky at an intersection and arrive when the light is green, whereas others have to stop for various amounts of time. Our model is not an accurate depiction of what would really happen in real life, but an approximation to produce realistic-enough looking simulations. It can easily be exchanged for a more sophisticated model.

We use OSRM¹ to fetch itineraries: we input multiple stops (such as a starting point and a destination) and retrieve between one and three different routes with detailed characteristics for each route segment, which we combine with data from the Overpass API on crossing and traffic lights. For each segment, we define random variables for events such as the likelihood to slow down (e.g., a crossing with no traffic light) or to stop (e.g., a big intersection where one needs to cross). The routes from OSRM include speeds for each route segment to factor in changes of speed specific to the segment (e.g., elevation). To account for the variability in walking speed between people, the model additionally includes a random variable for (walking) speed, with a minimum and a maximum speed (e.g., from 2.5 to 5 kilometers per hour), which we use to draw an average speed for the overall trip, and adjust it for each road segment with OSRM data. We then run a Monte Carlo simulation to generate a set of hypothetical trips from this model: for each draw, we generate a random number for each random variable in our model to determine how that hypothetical trip unfolds, e.g., whether one is stopped or slowed down at each turn or crossing (and for how long or how much speed is lost). A hypothetical trip is thus composed of multiple segments, and for each segment, a start, an end, and a duration are attached. Additionally, we compute the sum of all segments to get the overall duration for each hypothetical trip.

A more detailed description of our simulation model and sampling process is available in the [supplementary material](#) (on OSF).

3.3 Subsetting the Sample of Hypothetical Trips

We need to draw enough hypothetical trips from our model to sufficiently approximate the many ways in which a trip may unfold², but showing all these hypothetical trips on a map would be visually overwhelming. Therefore, we need to select a subset of these hypothetical trips. We use the overall duration of each hypothetical trip as a proxy to select a somewhat representative subset. Similar to work on quantile dot plots [26, 40], we select our subset based on quantiles of overall durations. First, we sort all hypothetical trips by their overall duration, then we select the extrema (shortest and longest), and finally, we select another $n - 2$ hypothetical trips so that they are evenly spaced in the sorted list (n being the size of the subset we want to have). Consequently, ranges with more hypothetical trips of similar duration are included more frequently than sparsely present ranges.

3.4 Visualization and Animation of Hypothetical Trips

We visualize each hypothetical trip in the generated subset as a dot (filled circle) on the map and animate it moving along a polyline determined by the route's geometry and the values for the random variables for each route segment. All dots depart at the same time and spread out as differences between hypothetical trips accumulate, e.g., if in a given hypothetical trip, there is a stop at an intersection for 25 seconds, then the associated dot will stop at that location (see [Figure 2](#)). A viewer should thus infer the space-based temporal uncertainty by observing dots behave differently locally. However, depending on the factor with which the animation is sped up, a stop of a single dot may not be perceptible.³

Once all dots have reached the destination, the animation restarts from the beginning. Selecting a larger subset to visualize, e.g., more than 10 dots, should be combined with a reduced opacity/alpha setting to make overlaps between dots more visible.

¹Open Source Routing Machine

²We set 100 and 10,000 draws as lower and upper limits.

³For example, with a factor of 1,000 (about 16 trip minutes per second), a 25 sec stop would only last 2.5 msec in the animation.



Fig. 2: Hypothetical trips—moving from top to bottom—along a path getting stopped at an intersection. The three arrows point towards three distinctive groups (the two extrema and the middle pack). Left: The three groups are evenly spaced on the path. Middle: The first dot has to stop at the intersection and the middle pack thus catches up. Right: The middle pack is now stopped, the first dot moves on, and the last dot catches up.

3.5 Arrival Plots

Retrieving information from a dynamic visualization—the animated map—can be challenging to viewers as it requires one’s constant attention. The animation only transiently shows a subset of the information at a given time and can make it difficult to form a global view of the underlying distribution of trip lengths. To alleviate these difficulties, we complement the animated map with arrival plots—a timeline which provides an overview for each route of when each hypothetical trip arrives at the destination, similar to the concept of slit-tear visualizations by Tang et al. [56]—either as a dot plot or a distribution plot (see Figure 1). Additionally, the arrival plots also show median and mean values of plausible arrival times for each route as a vertical bar on the plots. In our implementation, we display the timeline for each route in our tool, in a bottom panel which the viewer can easily refer to at any time. These plots are minimal (do not have explicit axes) so as to minimize clutter, only the current playback elapsed time is displayed.

3.6 Interactive Features

We allow for four different types of interaction within our tool:

Model parameters. Different model parameters can be exposed so that viewers can adjust them: the walking speed (a range), and the number of simulations generated for each route (between 100 and 10,000).

Path creation. When our tool is launched, no route is displayed. One first needs to (1) select a starting point by clicking anywhere on the map, (2) select a destination by clicking a second time, and, optionally, (3) add stops in-between by clicking again. When adding a new stop, new routes are generated so that one can compare the detour to the regular suggested route (the more straightforward itinerary). To add a stop to a path that already exists on the map, one needs to drag that path and release when the cursor is over the location of the desired new stop. Holding **Shift** while clicking on any element (a path, a detour marker, a starting or destination point) removes it.

Animation control. When hovering over the arrival plots, the animation playback syncs with the x-axis of the plots (time) to show where each simulation would be at that point in time. As such, one can replay specific parts of the animation on demand. By pressing the **Spacebar**, the animation can be (un)paused. For each route on the map, an associated color box is available on the right of the upper parameters bar and hovering over it shows the length of the route directly on the map.

Personalization of the visualization. Our tool is supplemented with a bottom bar (see Figure 1) on which one can switch between the dot plot view and the distribution plot view, and also choose between seeing the median or the mean duration in the arrival plots. In the control panel at the top, we also allow for different map tiles to be selected, namely the regular view provided by [OpenStreetMap \(OSM\)](#), a satellite view, a toner view, and a topography view. Viewers can also change the number of dots per route (i.e., displayed hypothetical trips — between 4 and either 100 or 300 divided by the number of routes on the map). The opacity of the dots can also be changed (from 0% to 100%).

Traditional pan-and-zoom interaction is also enabled.

A demo version of our tool is available on a [GitHub page](#)⁴.

⁴This web page is also provided in the [supplementary material](#). Note that

4 STUDY DESIGN

We set out to understand whether people can effectively understand and decode the information within our tool (**readability**), and whether there is added value to presenting uncertainty in path predictions (**usefulness**). We therefore outline the following research questions:

RQ1 To which extent can people read, understand, and make sense of the visualizations used in the tool?

- A. Can people read, understand, and make sense of the information displayed through our timeline visualizations (dot plot and distribution)?
- B. Can people extract information from the animated hypothetical trips displayed onto the map?

RQ2 To which extent do people acknowledge travel uncertainty when route planning with our approach?

- A. Do people communicate about uncertainty when seeing our visualizations?
- B. Do people reason with the uncertainty that is conveyed by our visualizations?
- C. Do people take into account and/or leverage uncertainty when planning for a path?

RQ3 Do our visualizations hinder the map experience?

4.1 Readability

By readability, we mean that we aim to investigate whether people are able to extract information from our interface, and more specifically from the arrival plots and the animation. We thus designed our tasks so as to focus on two aspects:

1. Are people able to interpret and extract information conveyed by the uncertainty visualizations? (RQ1)
2. To which extent do they acknowledge and communicate about this uncertainty? How extensively? What do they focus on? (RQ2)

4.1.1 Visualizing uncertainty

To answer the first aspect of readability (whether people are able to extract information), we designed a set of questions focusing on *temporal* information and *spaced-based temporal* information, which requires decoding the visualization, understanding the meaning of the animation, and in some cases, comparing between multiple hypothetical trips. Specifically, to correctly answer questions participants had to:

1. read and compare extreme values (outliers) on the arrival plots (temporal uncertainty);
2. read and compare median values on the arrival plots (temporal uncertainty);
3. read and interpret the underlying distribution shown by the arrival plots (temporal uncertainty);
4. use both the animated hypothetical trips map and the arrival plots to extract information (space-based temporal uncertainty)
5. interpret the animated hypothetical trips map and identify visual patterns in the movement at specific locations (space-based temporal uncertainty).

The complete phrasing of the questions that we devised for each item is available in [Table 1](#). All of these questions have ground truth answers (i.e., they are not based on participants’ interpretation or preferences) as the goal of these tasks is to evaluate people’s ability to extract correct information. To prompt answers, some of these questions required

both version (the [GitHub page](#) and [html page](#) from supplemental) require a working Internet connection as the tool will make requests to OSRM API. No data from users is logged or retrieved. The demo uses the described model to simulate walking trips for pedestrians but is not the version that was used for our study, as it needs a local server running on a machine. As a result, this version uses OSRM public default routing which only returns driving itineraries (and might result in illogical detours for pedestrians).

participants to select a path (e.g., which path is the fastest on average) among propositions (i.e., multiple-choice question). To make sure they actually extracted the correct information rather than choosing a path randomly, we then asked them to select a time duration among 3 propositions to answer “How long would it take [for the fastest path on average]?” For the task requiring participants to locate an area along a path with the help of the animated map, they only had to click somewhere on the map to indicate the chosen location. For each question, participants were always asked to verbally state how they arrived at the answer, which would allow us to understand the strategies they used to address each question. **Participants’ output from these tasks are used to answer RQ1-A and RQ1-B.**

4.1.2 Understanding and communicating about uncertainty

The second aspect of readability we aim to evaluate is how much of the uncertainty information participants are taking into account and communicating about. We asked them questions that required to:

1. compare different routes globally via uncertainty visualizations (animated map and/or arrival plots) in terms of unreliability;
2. read from the arrival plots and communicate about the extracted information.

A list of the questions associated with each item is available in Table 1. These tasks are meant to evaluate whether people acknowledge uncertainty while using or seeing the output of our tool. As a result, the associated questions we asked them do not have a wrong or right answer. Participants were asked to answer and explain their reasoning (out loud) for us to understand and analyze their strategies and examine whether they considered space-based temporal uncertainty. **Their answers are used to answer RQ2-A and RQ2-B.**

RQ1	[BC]	In the <u>best</u> -case scenario, which path would take the least amount of time (to travel)? New Bedford/Valparaiso →
	[WC]	In the <u>worst</u> -case scenario, which path would take the most amount of time (to travel)? Boston/Ottawa →
	[AL]	On <u>average</u> , which path takes the least amount of time? Wellington/Pau →
	[AM]	On <u>average</u> , which path takes the <u>most</u> amount of time? Mont-de-Marsan/Saint-Louis →
	[Goal]	Assuming you want your trip to take between X-Y minutes: which path would you take to maximize your chances of meeting this <u>goal</u> ? Algiers/Oklahoma City →
	[Bridge]	In the worst-case scenario, how long would it take you to reach the <u>bridge</u> that crosses the river/water? Christchurch/Viña del Mar → □
	[BNecks]	Identify (click on) a part of the path where you are very likely to experience <u>bottlenecks</u> (such as being stopped or slowed down). Auckland/Tulsa □
RQ2	[MU]	In your opinion, which of these paths looks the <u>most unreliable</u> ? Berlin/Dakar → □
	[CT]	Suppose you are taking this path to meet your friend: how would you communicate the <u>time</u> it would take for you to arrive? Kitchener/Touba →

Table 1: List of the questions that were asked to participants. Each question was asked twice (two different curated scenarios accessible by clicking on the link associated to the city’s name or in the supplemental material: `Question_Name_of_the_city.html`), in an order dependent on the group each participant was assigned to. → represents the questions that require looking at the interactive plot panel, and □ at the map and dots (minimal strategy).

4.1.3 Curated scenarios

For each task and associated question (see Table 1), we curated two distinct scenarios as repetitions with diverse examples. Pairs of scenarios were selected so that they covered diverse characteristics (e.g., different type of city architecture, different temporal patterns). A description of our selection is detailed in the supplementary material.

4.1.4 Interface setup

We used a custom limited version of our tool for the readability tasks. We fixed the numbers of simulations to 500 since each scenario was pre-loaded and thus stored locally. We did not want participants to be able to change that value (it would re-generate simulations) because we wanted the exact same simulations to be shown to everyone. We also removed control of parameters and fixed values for the path colors (we picked colors that are color-blind friendly) and average value (we picked the median, and we de-activated the toggle button allowing to switch modes), and we set the speed to be between 3 and 4 km/h.

For the other parameters, we set defaults values which we deemed reasonable to answer questions without having to modify the settings, but let participants modify these values at will. We set the number of dots to 50 (out of 500 simulations), the opacity of the dots to 50%, the arrival plots panel to show the dot plot, and the map tile to be the regular type (provided by *OpenStreetMap*) by default (see Figure 1).

4.1.5 Experimental groups

To avoid order effects, we randomized the orders of the tasks and scenarios so as to create four distinct groups. As a result, two scenarios (e.g., Christchurch for the [Bridge] question and Ottawa for the [WC] question) were never repeated in the same order among the four groups and the two same tasks (e.g., [BNecks] questions) never followed each other. Only the [CT] questions were manually put in each group so that every participant would start the readability part with a first iteration of the [CT] question, and end with the other one. The order of the tasks for the four groups is provided in the supplementary material.

4.2 Usefulness

We were also interested in evaluating how our tool would be used and whether and how the map experience would be impacted by the additional visualizations our tool offer (RQ3). To this end, we introduced the following scenario:

Imagine that you have 3 hours to spend in Toronto and you would like to spend them walking around the city. Use *Google Maps* and/or our tool to plan out your walk.

The city we chose was the one where the study took place so that participants would be familiar with the location. We allowed access to *Google Maps* so that participants could also use a more common tool, should they prefer that. This also acted as a baseline for them to compare routes and what they are seeing (e.g., they could compare the estimated duration provided by each tool for the same path).

For that task, participants would be given around 6 minutes. Towards the end, the experimenter would introduce a plot twist:

A friend of yours called to ask you to meet at [location chosen by the experimenter based on participants’ choice of itinerary]. How would you make that change of plan fit in your current devised itinerary?

We use participants’ output for this task to answer our research questions RQ2-C and RQ3.

4.3 Protocol

Participants were first handed an information sheet informing them that the experiment required audio and screen recording, and asking for their consent. They were then prompted to fill in a demographics form to provide their age range, self-reported gender, and occupation. The experimenter then explained that this was a study on animated maps and instructed participants to externalize their thinking process out loud. They were also told that the interaction with the tool would be happening on the computer screen in front of them, and the instructions and form would be displayed on the tablet device placed below the computer screen. Participants were informed that they could not go back to a previous answer to modify it, and that they should not change the window’s size and location. Additionally, they were notified that they could take a break at any point and that they could ask any question during the study if something was unclear.

All participants started with a short [tutorial](#)⁵ to learn how to use the tool. Upon completion of the tutorial, participants were asked whether they had any questions before starting the experiment.

Participants were presented with questions and scenarios of the readability tasks in one of four pre-defined orders (see section 4.1.5), followed by the usefulness task (see section 4.2). At the end of the study, they were asked to answer three sets of follow-up questions in a semi-structured interview: (1) What do you think of the tool? Do you have any comments or suggestions? Can you think of any strengths or weaknesses? (2) How would you describe this tool to a friend? and (3) What do you think one dot represents? How do you think the dots are picked among the simulations?

The experiment lasted about 44 minutes overall, and participants were compensated with a \$15 CAD Amazon gift card after completion.

5 RESULTS

We recruited 16 participants (4 for each group) through convenience sampling, using word of mouth at our institution and snowballing. Due to a technical issue, the audio and video recordings of one participant was compromised and therefore excluded from our qualitative analysis.

We retrieved the demographics forms, the answers forms, the log files (containing the click positions from [BNecks]) and the screen+audio recordings. We used the answers forms and the log files to retrieve participants' responses to the different questions (click and multiple choice questions) for the quantitative analysis. For the qualitative analysis, we produced transcripts of the audio files and used them with the screen recordings to determine participants' strategies and thinking processes, and to retrieve general comments.

5.1 Demographics

Out of our 16 participants, ten were between 18-24 years old, five were between 25-34 years old, and one was between 45-54 years old. Eleven participants (69%) identified as men, and five as women. All participants worked in the Computer Graphics or Human-Computer Interaction fields: one was a research scientist, one was a post-doctoral researcher, and the others (14/16) were (PhD or graduate) students. None had prior knowledge of the project and what it aimed to achieve.

5.2 Understanding, Reading and Making Sense of Visualizations (RQ1)

This research question was addressed through the **readability** tasks and questions requiring to read from the uncertainty visualizations (see section 4.1 and Table 1). We report participants' responses and strategies in the following subsections and discuss the implication on their ability to understand and extract pieces of information. An illustrated overview of participants' answers (correct or incorrect) is available in the supplementary material.

5.2.1 Extracting information from arrival plots (RQ1-A)

We asked participants questions that required to read and compare values provided by the arrival plots (see [BC], [WC], [AL], [AM], [Goal] in Table 1).

For the [WC] and [BC] questions, all participants except P12 selected the correct path and time duration. P12 realized that they had looked at the central tendency instead of the leftmost dot on the arrival plot. They could not correct their previous answer, but they verbally acknowledged their mistake. Overall, we found that participants were able to correctly retrieve extreme values and compare them across multiple routes.

For the [AL] and [AM] questions, all participants selected the correct path but P4 then chose the wrong time duration (28 minutes instead of 27). From the transcript we retrieved that P4 stated they first wanted to answer "around 35 minutes [...] because [they] just want to be safe." It was the second question they had to answer, that is, the first question after answering [CT] which involves personal preferences. P4's mistake

⁵The tutorial is available at this [GitHub](#) page as well as in the supplementary material ([tutorial.html](#)). Please note that an internet connection is needed for both versions as the actual route calculation is done with [OSRM API](#). The tool does not track nor log any data or interaction with it.

might thus be due to a misunderstanding of the question as they tried to choose a more conservative time duration. Overall, results again suggest that participants were able to read our arrival plots to retrieve the central tendency showcased by a vertical bar on the plots, and to compare this value across multiple hypothetical trips.

We examined the different strategies participants used to address these questions (i.e., what information was leveraged to answer the questions). With no surprise, they either looked at the average value that was given, or at the extremum value (either right or left depending on the question), which is consistent with the few errors participants made on these questions. P3 and P13 were the only participants to increase the number of dots to make sure the plots were representative enough to answer the questions. As we did not explain beforehand how we selected the simulations that we displayed on the map and plots, this strategy suggests that these participants wanted to make sure they had all the information available to answer these questions correctly.

Out of the two instances of the [Goal] questions—which required identifying the path that increases the likelihood to arrive within x to y minutes—only that in Algiers produced incorrect answers. P4, P10 and P13 chose the pink path over the correct yellow one. P4 dismissed the correct path and the incorrect blue paths because their worst-case hypothetical trip had an arrival time greater than the $[x-y]$ range, which means that they considered arriving later than the goal to be worse than arriving before. Similarly, P10 chose the faster path, because "it was fine time-wise and then it was mostly along the same street" which indicates that they looked at both the map and the arrival plot and then only focused on how to get to the destination before the range. P13 chose the fastest path "because even in the worst case of pink, it's 55 minutes and the average of arrival time is also the smallest".

The preferred strategy for these questions was looking at the central tendency to see whether it was within the range or else, then which path had a central tendency closest to the range. However, seven participants (P1, P3, P5–P8 and P11) mentioned looking at least once at the whole arrival plot (and thus the underlying distribution) to get a better sense of how likely being in that duration range would be: "I think from looking at the distribution plot, I would choose the yellow one, because I think the majority of the points falls within 50 to 55." (P6)

5.2.2 Extracting information from the animated map (RQ1-B)

To address this research question, we examine the response to questions that required looking at and reasoning with the animation (i.e. the [Bridge] and [BNecks] questions).

The [Bridge] questions required participants to visually locate a specific animated dot and use the arrival plots to retrieve a time duration. With the exception of P2 and P13, all participants answered the instance in Viña del Mar correctly. P2 and P13 both looked at the arrival plots and retrieved the *overall* duration in the worst-case scenario (the rightmost dot) instead of looking at the map. P2 also answered incorrectly for the second instance of this question (in Christchurch): P2 looked at the dots in the map and retrieved the time when the *first* group of dots reached the bridge, instead of the last one. All other participants selected the correct time of arrival.

Hovering over the arrival plot in order for the animation playback to be at the exact moment the last dot reached the bridge and then reading the time value was the preferred strategy by far. Instead of hovering, P7 paused the animation, and P15 just read the value directly without interacting nor pausing. To our surprise, P13 noticed the bridge in Christchurch was at a distance of about one third of the whole path so they retrieved the worst-case total duration and divided it by 3. Opting for such mental calculation means that one assumes the speed was constant throughout the whole trip. It is unclear if this was P13's assumption. It is also unclear whether they opted for this approach over the other way as employed by other participants, or simply did not figure out that such alternative would help arrive to the answer.

For the [BNecks] questions, participants had to select an area on the map where they would be likely to experience bottlenecks. To detect such pattern, one had to watch the animation and identify when dots would start to slow down or stop (see an example in Figure 2). The correct areas (i.e., the one where a change of speed could be

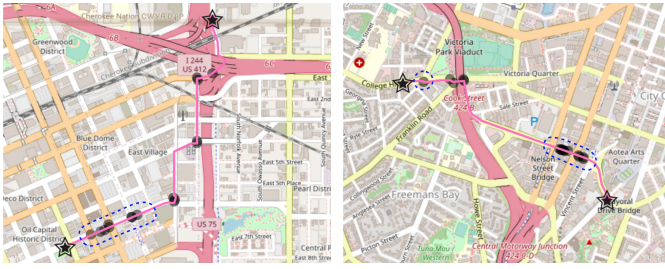


Fig. 3: Locations that participants selected as creating bottlenecks on the map (black circles) in Tulsa (left) and Auckland (right). The dotted blue ellipses represent the areas we considered as correct. Answers outside of these were labeled as incorrect.

visually detected) along with the locations selected by participants are shown in Figure 3. In Tulsa, while 8 participants did select correct areas, the others chose intersections where no sign of slowing down nor stopping was visible. Similarly, in Auckland, four participants chose an intersection different from than the ones we consider correct.

P2, P5 and P10 selected an incorrect location twice while P3, P4, P6, P8, P12, P14 and P16 selected a correct location twice. This question was described as difficult by multiple participants, even by those who were able to select the correct area. Those who chose an incorrect area reasoned by looking at the path because they could not visually identify where dots were slowing down or stopping. They thus relied on heuristics concerning features such as highways or big turns and deduced that they would probably cause bottlenecks: *“I cannot tell from the visuals. I will say just speaking from experience I will say around, under the, under the highway that will be a slow point.”* (P9) This suggests that, even if the animation was not informative enough for all participants, many were still able to read the map and reason with it (no occlusion from the animation). However, it also indicates that changes in speed are hard to infer from our animation alone.

The responses we retrieved from participants and the strategies they followed revealed that they were able to effectively read and extract relevant information from the arrival plots. The (few) errors that were made using these plots can be attributed to oversights when reading the instructions. One aspect of our visualizations that was more complicated was reasoning with the underlying distribution. Regarding the animation and thus the spaced-based temporal information type of questions, two observations emerged: (1) participants were able to read from the arrival plots while looking at the dots in the map thanks to the interactive features binding the arrival plots and animation replay, but (2) drawing conclusions based on the animation alone was judged as complex by participants as visually detecting changes in speed was the most difficult task.

5.3 Acknowledging Uncertainty (RQ2)

We address RQ2 by examining participants' answers from the remaining of the readability tasks (questions [MU] and [CT]) and checking whether participants used uncertainty lexicon (e.g., using words such as “approximately” or “around”). From here onwards, P16 is not included in the report as their audio/video recording was compromised.

5.3.1 Communicating about uncertainty (RQ2-A)

Participants were asked twice (once at the beginning of the readability tasks, once at the end) to verbally state how they would communicate their time of arrival to a friend when looking at the output of our tool for a fixed route (questions [CT])⁶. Three out of 15 participants (P1, P11, P13) chose to only communicate a single numerical value, such as *“it's going to be 36 minutes”* (P11). Among the remaining participants, three (P4, P5, P6) gave a single numerical value the first time (at the beginning of the study) and then communicated more uncertainty the second time (after completing the readability tasks). Participants who employed uncertainty lexicon within their answers used words such as

⁶Collected responses are detailed in the [supplementary material](#).

“approximately” (4 participants), “about” (2), “around” (5), “somewhere near” (1), or added “-ish” (2) to their estimate. Eight participants even included ranges such as *“you can expect me to come in 36 minutes, but if there's traffic then it won't take longer than 50 minutes”* (P7).

What stems from these responses is the participants' focus on the central tendency, even when combined with words that express uncertainty. We looked for the different reasoning mentioned by participants who did not choose to use solely the central tendency to build their answers, and two strategies emerged. Four participants (P7, P8, P13, P15) talked about or considered longer times of arrival because they acknowledged their propensity to be late, or they overshot because *“the penalty for being late is higher [than for being early]”* (P13). P14, on the other hand, looked at the first dot on the left (the best-case scenario) because they are *“the fastest walker [they] know.”* This illustrates how different concerns or beliefs held by participants shape the way they integrate the available information in their reasoning. If most of them only relied on the central tendency, there are some cases where it made sense to participants to also incorporate additional data.

5.3.2 Reasoning with the uncertainty (RQ2-B)

Without providing participants with a definition for the word ‘unreliable’, we asked them to *“Select the path that looked the most unreliable in [their] opinion”* ([MU] questions). Since the visualizations they had before them included uncertainty, we were interested in seeing whether they would incorporate it in their reasoning (that is, whether they would value and consider that piece of information). In the Berlin scenario, all participants chose the yellow path as the most unreliable. In Dakar, P10 was the only one to choose pink over yellow as most unreliable. This does not match the participant's answer for they viewed reliable as a path that does not display many crossings (meaning that, in their book, the yellow path should be the most unreliable). One common mistake that was corrected by most participants before selecting their answers was thinking they needed to select the path they would normally choose instead of the one they viewed as most unreliable. As such, P10's audio-video recording hints at that participant mistakenly selecting the path they viewed as most reliable.

For the [MU] questions, we identified strategies during which participants would look at the plot:

1. **Unreliable means slower:** Participants selected the path that had the highest central tendency (that is, the slowest on average). *“It would have to be the yellow path because of the significantly higher average time”* (P2) – Used by P*1, 2, 4, 8, 9, 10-15.
2. **Unreliable means highly variable:** Participants looked at the underlying distribution and selected the path displaying the most variability. *“I would think the yellow path because the dots seem to be more spread out”* (P6) – Used by P*1, 3-9, 12, 15.

– or look at the map and animated dots:

1. **Unreliable means intricate:** Participants selected the path which had the most turns or which looked the more complicated in terms of navigation. *“Because it's like... makes a lot of turn”* (P3) – Used by P*1, 2, 7, 9-11, 13, 14.
2. **Unreliable means visually scattered:** Participants selected the path with the animated dots that looked the most inconsistent in terms of overall speed. *“It seems to really spread out towards the end, where pink stays like very, very close together with the dots that are varying in speed.”* (P15) – Used by P15.

Only P3, P4, P5, P6 and P10 used a single strategy for these tasks, whereas the others used mixed strategies (such as looking at both the central tendency and the path's geometry). Two participants (P2 and P7) used the visual features of the paths (e.g., turns) to try to rationalize the underlying distribution given by our arrival plots: *“I don't trust the larger roads because there can be traffic so [...] that's intuitively the reason why this might have a larger range.”* (P7) This is interesting because it can indicate that providing this kind of information to viewers does not necessarily prevent them from making their own assumptions and trying to figure out on their own if the uncertainty visualization they see aligns with their internal representation of the path.

P4 and P10 based their answer solely on the central tendency shown on screen. P2, P11 and P14 looked at both the central tendency and the path (how it looked: for example, whether it had a lot of turns). The remaining participants used information that they extracted from the uncertainty visualization (either the animated map or the underlying distribution showcased by the arrival plots), indicating that they indeed considered the variability of the simulations.

5.3.3 Leveraging uncertainty when planning (RQ2-C)

We use the usefulness task and analyze participants' screen and audio recordings, looking for instances where they would mention time values other than the central tendency from the arrival plots, and more generally, vocabulary hinting at temporal uncertainty (i.e. words such as "approximately" or "about" applied to trip durations).

About half of the participants (P1, P4, P6, P8, P9, P13–P15) only looked at the central tendency but P6, P8, P9 and P13 used words such as "around" and "approximately" to nuance their time estimates. This can be explained by people's tendency to round values [55] rather than a clear indication of them expressing temporal uncertainty. P15 stated, at the beginning of the usefulness task, that they wanted to use our tool to specifically look at the underlying distribution ("I feel like getting a sense of the spread will be very, very important there, right? Especially if I didn't want to be late for that particular thing.") but ended up not doing so. Among the remaining participants, none of them looked at the best-case scenario but considered the worst case. They included that value in their decision process either (1) as a limit (e.g., if it exceeds 3 hours then it is risky and too long), or (2) to get an idea of the worst duration and then gauge whether they can still fit in one more stop.

We gathered evidence from participants which suggests that having access to uncertainty has some value: depending on the situation, one may want to consider the worst-case scenario for example. However, temporal uncertainty was not consistently acknowledged by all participants throughout the study. There is an emphasis to be made on the fact that uncertain temporal information is not for everyone—some people care less about temporal uncertainty but will look at a path's geometry to make a decision.

5.4 Impact on the Map Experience (RQ3)

We categorized comments collected from the end-of-study semi-structured interview regarding the weaknesses and strengths that were mentioned. To comment on the map experience, we also analyzed the transcript and video recordings from the usefulness task, evaluating whether participants were able to construct their own itinerary with the tool, which entails exploring different routes and detours, and reading and reasoning with the available information. The relevant excerpts from participants this part are provided in the supplementary material.

5.4.1 Weaknesses

We specifically looked for clues indicating participants struggled to perform an action (e.g., they mention not being able to find the information they are looking for, do not understand how to perform an action).

During the usefulness task, all participants experimented with the tool for at least a bit. The action most participants struggled with was adding detours: they attempted to click sequentially on multiple locations. As doing so adds new paths in our tool, it caused frustration⁷.

While this limitation was problematic for the usefulness task, it does not relate to the proposed animated map nor the arrival plots (but is rather an easy-to-solve UI problem). Participants also mentioned the choice of colors for each path as not well-suited: the yellow dots and line would often blend with the built-in colors of the map.

Another weakness is the time it takes to fully reveal the arrival plots. Since our animation runs at a fixed speed, longer hypothetical trips take longer to arrive at their destination. P12 requested to skip animation playback, instead of having to hover over the arrival plots.

⁷The proper action would be to drag the existing path and release where to add a stop.

5.4.2 Strengths

We looked for instances where participants would bring up something they enjoyed about our tool⁸. It was up to the participants if they wanted to compare our tool to *Google Maps* or not. Some did, some did not. When asked how they would describe our tool to a friend, P2 said "*Google Maps on steroids?*" and P3 "*would describe it as a better Google Maps.*" In more details, the features introduced by our tool that were mentioned by participants can be categorized as follow:

1. **A sense of reliability:** P11 stated that "*It's like Google Maps, but you get a better estimate*" whereas P7 viewed our tool as more reliable. A possible explanation could be that our default speed was closer to the average speed of those participants, and thus our tool indicating an average time closer to what they think it usually takes. What can also provide a sense of reliability is the fact that we are showing multiple simulations, summarized through our arrival plots, instead of a singular value that may be inaccurate. "*It's kind of common annoyance [...] with Google Maps where it says, oh, it's going to take 10 minutes, but actually, it's almost never correct how long it takes, at least for me, when you walk somewhere.*" (P3) – Mentioned by P*3, 5-7, 11.
2. **Ranges of values:** Some participants were vocal about their preferences towards having a range of possibilities as an available piece of information. The range we provide is actually composed of multiple values/simulated durations being presented on a plot, which means one can look at any specific value they deem relevant. It gives viewers more choice on what to take into account (for example, P4 and P7 view the worst-case scenario value as useful). – Mentioned by P*4, 5-7, 9, 11, 12, 15.
3. **The animation:** Animating the simulations with dots on the map shows what can happen at specific locations along a path, as P7 reports: "*I can, from animation, I can see which part is going slowly.*" One aspect the animation brings, which we did not anticipate, was that it also shows the direction one has to go, which P10 found helpful. – Mentioned by P*7, 9, 10, 12, 13, 15.
4. **Customizability:** There are two aspects behind customizability that we address with our tool. We allow for (1) a custom speed to be selected and (2) path refinement (i.e., users can explore multiple options and change the suggested routes to match their needs based on the available information). Regarding the speed, having that control over this parameter can help get estimates that are more personalized. "*Ofentimes [...] something like Google Maps... Usually, the timing that it gives isn't necessarily [...] customized to me.*" (P15) – Mentioned by P*5, 7, 9, 10, 12, 15.

We did not ask participants to speculate on whether they would use our tool in their everyday lives but P2 still mentioned "*[they] would use this tool for tourism reasons because [they] could plan [their] getaways better*" and P11 "*liked it more than Google Maps for planning trips.*"

The general impression was positive overall with no indication that our technique made basic tasks (e.g., retrieving a path's average duration) more complicated, or that it impaired the map experience. At worst, a map reader might consider uncertainty visualizations useless, but it seems to not prevent people from extracting relevant pieces of information. At best, our method provides a more custom experience and more information for participants to base their decisions on.

5.5 General comments

We report here comments made by participants that do not answer our research questions but are relevant to the **usefulness** of our technique.

Some participants liked how accurate our tool seemed to be (providing ranges, simulations), but P6 inquired about the need for accuracy in such a context: "*Maybe the estimations, like this accuracy might not be something that [...] I care about the most when I try to navigate places.*" Additionally, while having speed ranges was seen as a plus, many mentioned they did not know their own speed: "*It would be even more useful if I knew how quickly I actually walk but I don't.*" (P15)

⁸We do not report here comments that are directly related to the UI.

We never specified to participants that our use case was focused on pedestrians, and instead let them figure it out on their own. As a result, four participants (P5, P8, P9 and P11) thought it was meant for driving/cars. This became clearer when we mentioned, for the usefulness task, to pretend they were going for a walk.

During the usefulness task, P1 was selecting an itinerary similar to the one they used to walk everyday and realized that they were taking the least optimal route: *"It's actually [...] helping me because I usually walk this route. But I do the worst one."* This is further evidence that our tool provides understandable, and potentially useful information which may not have been available using traditional tools.

6 DISCUSSION

Our work tackles the difficult challenge of visualizing outcomes of trip prediction along one or multiple routes in such a manner that viewers can understand space-based temporal uncertainty when following such routes. The problem entails communicating space-time information about trips along a route (a non-trivial problem in and on itself), as well as communicating uncertainty associated with accounting for multiple hypothetical trips (which is another complex problem of its own). We identify the potential of animating particles in maps to support these dual goals, for animation's properties as a space-efficient representation (borrowed from traffic flow animated maps) which also acts as a frequency framing of temporal uncertainty (extending on HOPs [37, 39]). We propose an approach that supports reading and understanding both global temporal uncertainty, as well as local space-based temporal uncertainty, using these animated maps in combination with arrival plots, which present an overview of the distributions.

Our study results suggest that people were able to read, evaluate, and correctly interpret information presented through our approach. We found that participants were able to reason about hypothetical trips being display, with regard to overall temporal estimates (average, best, and worst case), as well as local space-based temporal estimates (e.g. likelihood to reach a particular location within a given time). Participants had more difficulties identifying granular changes in speed (i.e. where and to which extent one would be slowed down).

Our work provides evidence that animating a representative sample of plausible outcomes resulting from a predictive/simulation model are effective tools for implicit communication of uncertainty applied to spatio-temporal reasoning. A similar strategy could be applied to other contexts, especially in instances where the representation of a single instance (e.g. in our case one trip) is already visually-rich. Future research should look into combining (i) animated visualizations (akin to our animated hypothetical trips map, or the hypothetical outcome plots [37]) with (ii) representations of distributions (akin to our arrival plots, or the quantile dot plots [40]) to support tasks involving both global and local uncertainty reasoning.

Arguably, the application domain we chose to study these ideas (i.e. simulating trips) may not make the most compelling case as to the usefulness of our general approach for real-world tasks such as trip planning. We found that in spite of being successful at answering study questions (i.e. readability), most participants may not really care that much about uncertainty information in the first place in these scenarios (i.e. usefulness). This does not invalidate our results. Rather, we see an opportunity for more research as to what other domains our general approach would apply. Our visualization pipeline is directly applicable to hurricane forecasts, wind flows simulations, population flows, for example. It would be interesting to investigate what application domains may most benefit from support for space-based temporal uncertainty. Specifically, the context in which this approach is to be used will influence the non-trivial task of selecting appropriate colours: an issue we faced with map tiles showcasing a lot of different color contrasts.

Another interesting insight from our study is that participants relied on heuristics and beliefs that are specifically pertaining to the application domain when approaching some tasks. For instance, we found that participants were more inclined to consider worst-case estimates in their uncertainty reasoning than best-case estimates, which stems from lateness being less socially acceptable than being early. Similarly, the notion of "reliability" (which requires understanding uncertainty)

bears a particular meaning in our context, which would differ in other contexts. Future work could explore more in depth the impact of the application domain on the roles, meaning, and usage of uncertainty information when such information is exposed to viewers.

Our visualization pipeline is generalizable. Here, we demonstrated a full instantiation for the specific case of hypothetical trips, but all building blocks could utilize a different method. We provided participants with a complete solution comprising of multiple components, including an automatic animation, user-controlled playback animation through interaction with the slider, as well as static visualizations providing global distribution views. It would be interesting, in future work, to tease apart the role of each such components in supporting space-based temporal uncertainty reasoning. Additionally, we created our own model to make up for the lack of existing data and to generate hypothetical trips. Future implementations could incorporate empirical models of walking patterns [50] to showcase even more realistic movements of dots, or other types of models. Similarly, our sampling technique stems from prior literature [40], but other sampling methods may be more suited to other scenarios. Finally, the visualizations per se can be adapted to the underlying data to be visualized. In fact, there are many features that could be added and evaluated such as a textual explanation of local uncertainty (e.g., red lights, roadworks) or additional interactions with the map to retrieve each moving dot's current elapsed time. Still, *our recommendation is to employ animation of representative outcomes, whose playback is bound to interactive control of an overview distribution plot.*

Finally, our study has some limitations, which future work could help address. First, our participants were all scientists, and thus potentially better-versed in visualization and statistical literacy, which could have skewed our results. Future studies should explore whether exposure to, and experience with uncertainty representations impact people's ability to make sense of animated hypothetical trips. Second, we evaluated our approach on high-resolution monitors. It is unclear how our approach (or future ones building on our work) would cater to smaller displays such as mobile phones. Our participants already struggled in identifying speed variations. Reducing the display space would only make things worse, and might call for alternative representations which are less space demanding and which require finding a way to compact the displayed information. Future work could also investigate other variables to better encode and show a change of speed (e.g., color, size), or the benefits of allowing refined control over the playback speed to provide a more precise exploration of the data.

7 CONCLUSION

We proposed a novel approach to communicate space-based temporal uncertainty: animated hypothetical trips on a map and interactive arrival plots allowing to control the animation playback at will. We conducted a study with 16 participants to evaluate our technique's readability and usefulness and we observed that overall, participants were able to retrieve values such as the central tendency of the estimated duration for a path, extrema and outliers, and to compare these between different routes. Our animation only helped some participants detect potential bottlenecks, suggesting that there is a need for other parameters to be investigated in order to showcase that change in speed more clearly (e.g., size of dots, color, playing speed). Our technique is a promising step towards communication of space-based temporal uncertainty to lay people as participants were able to construct their own itinerary and extract the information they deemed useful for that task. Overall, our work provides evidence that our generalizable visualization pipeline is worth exploring in other contexts where reasoning about global and local uncertainty is important, and hope it will provide a useful framework for future exploration in this space.

ACKNOWLEDGMENT

We thank Mian Wei, Anna Offenwanger and the members of the DGP lab for their help and feedback, and LaBRI for their travel grant. This work is part of the ANR Ember project, supported by a grant from Agence Nationale de Recherche (ANR-19-CE33-0012), and was also supported in part by a grant from NSERC (RGPIN-2018-05072).

REFERENCES

- [1] J. C. J. H. Aerts, K. C. Clarke, and A. D. Keuper. Testing Popular Visualization Techniques for Representing Model Uncertainty. *Cartography and Geographic Information Science*, 30(3):249–261, 2003. doi: 10.1559/152304003100011180 2, 3
- [2] F. Amini, S. Rufiange, Z. Hossain, Q. Ventura, P. Irani, and M. J. McGuffin. The Impact of Interactivity on Comprehending 2D and 3D Visualizations of Movement Data. *IEEE Transactions on Visualization and Computer Graphics*, 21(1):122–135, 2015. doi: 10.1109/TVCG.2014.2329308 2
- [3] G. Andrienko, N. Andrienko, U. Demsar, D. Dransch, J. Dykes, S. I. Fabrikant, M. Jern, M.-J. Kraak, H. Schumann, and C. Tominski. Space, time and visual analytics. *International Journal of Geographical Information Science*, 24(10):1577–1600, 2010. 2
- [4] N. Andrienko and G. Andrienko. Interactive visual tools to explore spatio-temporal variation. In *Proceedings of the working conference on Advanced visual interfaces*, pp. 417–420, 2004. 2
- [5] N. Andrienko and G. Andrienko. Visual analytics of movement: An overview of methods, tools and procedures. *Information Visualization*, 12(1):3–24, 2013. doi: 10.1177/1473871612457601 2
- [6] B. Bach, N. Henry-Riche, T. Dwyer, T. Madhyastha, J.-D. Fekete, and T. Grabowski. Small multiples: Piling time to explore temporal patterns in dynamic networks. *Computer Graphics Forum*, 34(3):31–40, 2015. doi: 10.1111/cgf.12615 2
- [7] B. Bach, C. Perin, Q. Ren, and P. Dragicevic. Ways of Visualizing Data on Curves. In *TransImage 2018 - 5th Biennial Transdisciplinary Imaging Conference*, pp. 1–14. Edinburgh, United Kingdom, Apr. 2018. doi: 10.6084/m9.figshare.6104705 1
- [8] B. Bach, E. Pietriga, and J.-D. Fekete. Visualizing dynamic networks with matrix cubes. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 877–886, 2014. 2
- [9] B. Bach, C. Shi, N. Heulot, T. Madhyastha, T. Grabowski, and P. Dragicevic. Time Curves: Folding Time to Visualize Patterns of Temporal Evolution in Data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):559–568, 2016. doi: 10.1109/TVCG.2015.2467851 1, 2
- [10] L. Bartram. Perceptual and interpretative properties of motion for information visualization. In *Proceedings of the 1997 workshop on New Paradigms in Information Visualization and Manipulation*, pp. 3–7, 1997. 2
- [11] L. Bartram and A. Nakatani. Distinctive parameters of expressive motion. In *CAe*, pp. 129–136, 2009. 2
- [12] L. R. Bartram. *Enhancing information visualization with motion*. PhD thesis, Simon Fraser University, 2001. 2
- [13] J. Bertin. *Semiology of graphics*. University of Wisconsin press, 1983. 1, 2
- [14] S. Bies and M. van Kreveld. Time-Space Maps from Triangulations. In W. Didimo and M. Patrignani, eds., *Graph Drawing*, pp. 511–516. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. 2
- [15] V. Bilbily, E. Huynh, K. Singh, and F. Chevalier. Space, Time, and Choice: A Unified Approach to Flexible Personal Scheduling. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, UIST ’21, pp. 484–497. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3472749.3474764 2
- [16] S. Blenkinsop, P. Fisher, L. Bastin, and J. Wood. Evaluating the Perception of Uncertainty in Alternative Visualization Strategies. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 37(1):1–14, 2000. doi: 10.3138/3645-4V22-0M23-3T52 2
- [17] I. Boyandin, E. Bertini, and D. Lalanne. A qualitative study on the exploration of temporal changes in flow maps with animation and small-multiples. *Computer Graphics Forum*, 31(3pt2):1005–1014, 2012. doi: 10.1111/j.1467-8659.2012.03093.x 2
- [18] J. Brosz, M. A. Nacenta, R. Pusch, S. Carpendale, and C. Hurter. Transmogrification: Causal Manipulation of Visualizations. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST ’13, pp. 97–106. Association for Computing Machinery, New York, NY, USA, 2013. doi: 10.1145/2501988.2502046 1
- [19] K. Buchin, A. van Goethem, M. Hoffmann, M. van Kreveld, and B. Speckmann. Travel-Time Maps: Linear Cartograms with Fixed Vertex Locations. In M. Duckham, E. Pebesma, K. Stewart, and A. U. Frank, eds., *Geographic Information Science*, pp. 18–33. Springer International Publishing, Cham, 2014. 2
- [20] S. Buschmann, M. Trapp, and J. Döllner. Real-time animated visualization of massive air-traffic trajectories. In *2014 International Conference on Cyberworlds*, pp. 174–181. IEEE, 2014. 2
- [21] F. Chevalier, N. H. Riche, C. Plaisant, A. Chalbi, and C. Hurter. Animations 25 years later: New roles and opportunities. In *Proceedings of the international working conference on advanced visual interfaces*, pp. 280–287, 2016. 2
- [22] S. Deitrick and E. A. Wentz. Developing Implicit Uncertainty Visualization Methods Motivated by Theories in Decision Science. *Annals of the Association of American Geographers*, 105(3):531–551, 2015. doi: 10.1080/00045608.2015.1012635 3
- [23] C. R. Ehlschlaeger, A. M. Shorridge, and M. F. Goodchild. Visualizing spatial data uncertainty using animation. *Computers & Geosciences*, 23(4):387–395, 1997. doi: 10.1016/S0098-3004(97)00005-8 2
- [24] B. J. Evans. Dynamic display of spatial data-reliability: Does it benefit the map user? *Computers & Geosciences*, 23(4):409–422, 1997. doi: 10.1016/S0098-3004(97)00011-3 2
- [25] J.-D. Fekete, N. Elmqvist, and Y. Guiard. Motion-pointing: target selection using elliptical motions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 289–298, 2009. 2
- [26] M. Fernandes, L. Walls, S. Munson, J. Hullman, and M. Kay. Uncertainty Displays Using Quantile Dotplots or CDFs Improve Transit Decision-Making. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI ’18, pp. 1–12. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3173718 1, 2, 3
- [27] C. Fish, K. P. Goldsberry, and S. Battersby. Change Blindness in Animated Choropleth Maps: An Empirical Study. *Cartography and Geographic Information Science*, 38(4):350–362, 2011. doi: 10.1559/15230406384350 2
- [28] P. Gatalsky, N. Andrienko, and G. Andrienko. Interactive analysis of event data using space-time cube. In *Proceedings. Eighth International Conference on Information Visualisation, 2004. IV 2004.*, pp. 145–152, 2004. doi: 10.1109/IV.2004.1320137 2
- [29] K. Goldsberry and S. Battersby. Issues of Change Detection in Animated Choropleth Maps. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 44(3):201–215, 2009. doi: 10.3138/carto.44.3.201 2
- [30] A. L. Griffin, A. M. MacEachren, F. Hardisty, E. Steiner, and B. Li. A comparison of animated maps with static small-multiple maps for visually identifying space-time clusters. *Annals of the Association of American Geographers*, 96(4):740–753, 2006. 2
- [31] T. Gschwandtner, M. Bögl, P. Federico, and S. Miksch. Visual Encodings of Temporal Uncertainty: A Comparative User Study. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):539–548, 2016. doi: 10.1109/TVCG.2015.2467752 2
- [32] P. K. Han, W. M. Klein, B. Killam, T. Lehman, H. Massett, and A. N. Freedman. Representing randomness in the communication of individualized cancer risk estimates: effects on cancer risk perceptions, worry, and subjective uncertainty about risk. *Patient education and counseling*, 86(1):106–113, 2012. 2
- [33] M. Harrower. The Cognitive Limits of Animated Maps. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 42(4):349–357, 2007. doi: 10.3138/carto.42.4.349 2
- [34] M. Harrower. Unclassed animated choropleth maps. *The Cartographic Journal*, 44(4):313–320, 2007. 2
- [35] S. Hong, R. Kocielnik, M.-J. Yoo, S. Battersby, J. Kim, and C. Aragon. Designing interactive distance cartograms to support urban travelers. In *2017 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 81–90, 2017. doi: 10.1109/PACIFICVIS.2017.8031582 1, 2
- [36] J. Hullman. Why authors don’t visualize uncertainty. *IEEE transactions on visualization and computer graphics*, 26(1):130–139, 2019. 1
- [37] J. Hullman, P. Resnick, and E. Adar. Hypothetical Outcome Plots Outperform Error Bars and Violin Plots for Inferences about Reliability of Variable Ordering. *PLOS ONE*, 10(11):1–25, Nov. 2015. doi: 10.1371/journal.pone.0142444 1, 2, 3, 9
- [38] C. Kaiser, F. Walsh, C. J. Q. Farmer, and A. Pozdnoukhov. User-Centric Time-Distance Representation of Road Networks. In S. I. Fabrikant, T. Reichenbacher, M. van Kreveld, and C. Schlieder, eds., *Geographic Information Science*, pp. 85–99. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. 2
- [39] A. Kale, F. Nguyen, M. Kay, and J. Hullman. Hypothetical Outcome Plots Help Untrained Observers Judge Trends in Ambiguous Data. *IEEE transactions on visualization and computer graphics*, Aug. 2018. doi: 10.1109/tvcg.2018.2864909 1, 3, 9

- [40] M. Kay, T. Kola, J. R. Hullman, and S. A. Munson. When (Ish) is My Bus? User-Centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pp. 5092–5103. Association for Computing Machinery, New York, NY, USA, 2016. doi: [10.1145/2858036.2858558](https://doi.org/10.1145/2858036.2858558) 1, 2, 3, 9
- [41] C. Kinkeldey, A. M. MacEachren, and J. Schiewe. How to Assess Visual Communication of Uncertainty? A Systematic Review of Geospatial Uncertainty Visualisation User Studies. *The Cartographic Journal*, 51(4):372–386, 2014. doi: [10.1179/1743277414Y.0000000099](https://doi.org/10.1179/1743277414Y.0000000099) 3
- [42] M. Koval and Y. Jansen. Do You See What You Mean? Using Predictive Visualizations to Reduce Optimism in Duration Estimates. In *CHI 2022 - Conference on Human Factors in Computing Systems*. New Orleans, United States, Apr. 2022. doi: [10.1145/3491102.3502010](https://doi.org/10.1145/3491102.3502010) 2
- [43] M.-J. Kraak. The space-time cube revisited from a geovisualization perspective. In *Proc. 21st international cartographic conference*, pp. 1988–1996. Citeseer, 2003. 2
- [44] L. Liu, A. P. Boone, I. T. Ruginski, L. Padilla, M. Hegarty, S. H. Creem-Regehr, W. B. Thompson, C. Yuksel, and D. H. House. Uncertainty visualization by representative sampling from prediction ensembles. *IEEE Transactions on Visualization and Computer Graphics*, 23(9):2165–2178, 2017. doi: [10.1109/TVCG.2016.2607204](https://doi.org/10.1109/TVCG.2016.2607204) 3
- [45] A. M. MacEachren. Visualizing Uncertain Information. *Cartographic perspectives*, 13:10–19, 1992. 3
- [46] T. Nagel and B. Groß. Shanghai metro flow: Multiple perspectives into a subway system. *Leonardo*, 50(5):511–512, 2017. 1, 2
- [47] L. Padilla, M. Kay, and J. Hullman. *Uncertainty Visualization*, pp. 1–18. John Wiley & Sons, Ltd, 2021. doi: [10.1002/9781118445112.stat08296](https://doi.org/10.1002/9781118445112.stat08296) 1
- [48] C. Perin, T. Wun, R. Pusch, and S. Carpendale. Assessing the Graphical Perception of Time and Speed on 2D+Time Trajectories. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):698–708, 2018. doi: [10.1109/TVCG.2017.2743918](https://doi.org/10.1109/TVCG.2017.2743918) 1, 2
- [49] A. Preston and K.-L. Ma. Communicating uncertainty and risk in air quality maps. *IEEE Transactions on Visualization and Computer Graphics*, 29(9):3746–3757, 2023. doi: [10.1109/TVCG.2022.3171443](https://doi.org/10.1109/TVCG.2022.3171443) 3
- [50] I. Rhee, M. Shin, S. Hong, K. Lee, S. J. Kim, and S. Chong. On the levy-walk nature of human mobility. *IEEE/ACM Transactions on Networking*, 19(3):630–643, 2011. doi: [10.1109/TNET.2011.2120618](https://doi.org/10.1109/TNET.2011.2120618) 9
- [51] G. Robertson, R. Fernandez, D. Fisher, B. Lee, and J. Stasko. Effectiveness of animation in trend visualization. *IEEE transactions on visualization and computer graphics*, 14(6):1325–1332, 2008. 2
- [52] H. Rosling and Z. Zhang. Health advocacy with gapminder animated statistics. *Journal of epidemiology and global health*, 1(1):11, 2011. 2
- [53] R. Scheepens, C. Hurter, H. Van De Wetering, and J. J. Van Wijk. Visualization, selection, and analysis of traffic flows. *IEEE transactions on visualization and computer graphics*, 22(1):379–388, 2015. 2
- [54] D. Spiegelhalter. Risk and Uncertainty Communication. *Annual Review of Statistics and Its Application*, 4(1):31–60, Mar. 2017. doi: [10.1146/annurev-statistics-010814-020148](https://doi.org/10.1146/annurev-statistics-010814-020148) 2
- [55] J. Talbot, V. Setlur, and A. Anand. Four experiments on the perception of bar charts. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2152–2160, 2014. doi: [10.1109/TVCG.2014.2346320](https://doi.org/10.1109/TVCG.2014.2346320) 8
- [56] A. Tang, S. Greenberg, and S. Fels. Exploring Video Streams Using Slit-Tear Visualizations. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '08, pp. 191–198. Association for Computing Machinery, New York, NY, USA, 2008. doi: [10.1145/1385569.1385601](https://doi.org/10.1145/1385569.1385601) 4
- [57] W. Thompson and S. Lavin. Automatic generation of animated migration maps. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 33(2):17–28, 1996. 2
- [58] L. Xie, X. Shu, J. C. Su, Y. Wang, S. Chen, and H. Qu. Creating emordle: Animating word cloud for emotion expression. *IEEE Transactions on Visualization and Computer Graphics*, 2023. 2
- [59] N. Yusof, R. Zurita-Milla, M.-J. Kraak, and B. Retsios. Interactive discovery of sequential patterns in time series of wind data. *International Journal of Geographical Information Science*, 30(8):1486–1506, 2016. 2