



**HAL**  
open science

## A Comparison of Geographical Propagation Visualizations

Vanessa Peña-Araya, Anastasia Bezerianos, Emmanuel Pietriga

► **To cite this version:**

Vanessa Peña-Araya, Anastasia Bezerianos, Emmanuel Pietriga. A Comparison of Geographical Propagation Visualizations. CHI '20 - 38th SIGCHI conference on Human Factors in computing systems, Apr 2020, Honolulu, United States. pp.223:1–223:14, 10.1145/3313831.3376350 . hal-02510038

**HAL Id: hal-02510038**

**<https://inria.hal.science/hal-02510038v1>**

Submitted on 17 Mar 2020

**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.

# A Comparison of Geographical Propagation Visualizations

Vanessa Peña-Araya

vanessa.pena-araya@inria.fr

Anastasia Bezerianos

Université Paris-Saclay, CNRS, Inria, LRI  
91405, Orsay, France

anastasia.bezerianos@lri.fr

Emmanuel Pietriga

emmanuel.pietriga@inria.fr

## ABSTRACT

Geographical propagation phenomena occur in multiple domains, such as in epidemiology and social media. Propagation dynamics are often complex, and visualizations play a key role in helping subject-matter experts understand and analyze them. However, there is little empirical data about the effectiveness of the various strategies used to visualize geographical propagation. To fill this gap, we conduct an experiment to evaluate the effectiveness of three strategies: an animated map, small-multiple maps, and a single map with glyphs. We compare them under five tasks that vary in one of the following dimensions: propagation scope, direction, speed, peaks, and spatial jumps. Our results show that small-multiple maps perform best overall, but that the effectiveness of each visualization varies depending on the task considered.

## Author Keywords

Animation; small-multiples; geo-temporal data; propagation.

## CCS Concepts

•Human-centered computing → Geographic visualization; Empirical studies in visualization;

## INTRODUCTION

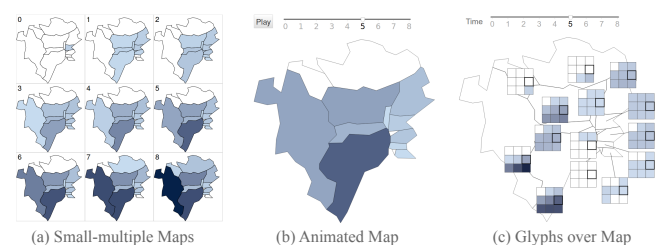
Propagation phenomena occur in a diverse domains, from epidemiology to cybersecurity. Examples include disease spread, keyword tags on social media, packet replication in communication networks, software viruses, *etc.* In many scenarios *geography* is key in understanding and analyzing the phenomenon. Consider the Ebola epidemic which occurred in West Africa from 2013 to 2016, causing more than 11,300 deaths and major socio-economical disruption to the affected countries [53]. The analysis of its propagation revealed that the virus tended to disperse more frequently among geographically-close regions, and mainly within countries [26].

Gaining such insights from analyzing propagation patterns is extremely valuable. In the example above, it can help both make predictions about dispersion during the crisis and inform decisions taken by health authorities to contain future outbreaks. But epidemiology is not the only domain where geovisualization plays a key role. For instance, the geolocation

of users can be an important factor in social media analysis (*e.g.*, tag or meme transmission [35]), that can complement the information given by the network's topology.

Propagation can be categorized as a movement pattern [24], and as such it has some resemblances with other spatio-temporal patterns. For example, spatial autocorrelation has been found between influenza and commuting paths [15]. However, propagation is not induced by the motion of entities in space, but by the *replication* of entities (*e.g.*, virus infection, meme or hashtag reuse). This replication and the context in which it develops makes propagation different from other movement patterns in two respects. First, although diseases propagated continuously in early human epidemics, nowadays' connectivity allows geographical jumps [47]. This differs from movement trajectories of individual entities (*e.g.*, migrating birds) or vector fields (*e.g.*, water flow) that are continuous. Second, the replication of entities allows propagation to have multiple synchronous strong peaks in different and distant geographical places [63]. This is not possible when analyzing the trajectories of vehicles or living beings as their discrete nature constrains the number and strength of synchronous peaks.

Which geovisualization techniques best support experts in the understanding of propagation patterns remains an open question. There is little empirical data about the effectiveness of different techniques in geo-temporal tasks that involve movement more generally, and none that considers tasks specific to propagation analysis. In this paper, we identify three main visualization strategies: small-multiple maps, animated maps, and maps with glyphs (illustrated in Figure 1). We empirically compare them using five tasks derived from both a general taxonomy of movement patterns [24] and from the literature on propagation analysis. Our results show that small-multiple maps perform best overall, but that both animation and maps with glyphs outperform them for some tasks, and that self-reported confidence is higher with animated maps.



**Figure 1.** The three propagation visualization strategies evaluated, illustrated with a simplified dataset featuring only a small region and 9 time-steps: (a) small-multiple maps; (b) a single animated map; (c) a single map with glyphs overlaid over each region - each cell in a glyph encodes the value of one time-step for the corresponding region.

Vanessa Peña-Araya, Anastasia Bezerianos & Emmanuel Pietriga. A Comparison of Geographical Propagation Visualizations. In *CHI '20: Proceedings of the 38th Annual ACM Conference on Human Factors in Computing Systems*, 14 pages, ACM, April 2020.

©ACM, 2020. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version will be published in *CHI '20, April 25–30, 2020, Honolulu, HI, USA*.

<https://doi.org/10.1145/3313831.3376350>

## RELATED WORK

We first review related work on the formal characterization of spatio-temporal dynamics. We then give an overview of visualization techniques for spatio-temporal dynamics, and of the perception studies that evaluate them.

### Characterizing Spatio-Temporal Dynamics

Movement patterns are usually defined by the sequential trajectories that discrete entities follow when they change position over space and time [32]. The analysis of these patterns has been used for understanding how people commute in cities [30], analyzing air traffic [1] and describing birds migration [64], among many other applications [4]. In order to define some common grounds for their analysis and visualization, researchers have worked on the formalization of movement patterns (*e.g.*, [2, 24, 54]). In particular, Dodge *et al.* [24] introduce a taxonomy with a general and flexible categorization of movement patterns by level of abstraction and description of their parameters. For example, in their taxonomy, *periodicity* is defined as a primitive pattern, that could be considered as one of the components of the higher level pattern of *migration*.

However, spatio-temporal dynamics are not always about the movement of individual entities. For instance, when analyzing the propagation of memes on Twitter, Kamath *et al.* [35] consider the replication of digital content, the memes, through connections among users, not the movements of users themselves. This particular type of propagation can be modeled by a network of entities, whose links among entities enable exchange patterns over space and time. These models have been used for instance to analyze disease spread [7, 41, 50] and diffusion on social media [27].

In this paper we focus on spatio-temporal patterns of *propagation* using disease spread models. We build upon work that formalizes movement patterns in general, and adapt it to our purposes by describing two new analytic tasks derived from real-world cases studies described in propagation literature.

### Visualizing Spatio-temporal Dynamics

Two main strategies exist for visualizing the spatio-temporal dynamics of moving entities: movement can be represented *explicitly* (*e.g.*, flow maps [22, 1, 58] and origin-destination maps [67, 68]), or *implicitly* (*e.g.*, individual points changing position over time [60, 28] or areas changing color based on changes in the number of entities [42, 46]). Our focus, however, is not on techniques that visualize the movements of entities, but rather on techniques that represent the spatio-temporal dynamics themselves. We organize them according to Peña-Araya *et al.*'s categorization [55], in which geo-temporal visualizations either *juxtapose time* or *juxtapose location*.

**Juxtaposing Location.** Small-multiples and animation are the most commonly used strategies to represent change over time [59]. Small-multiples juxtapose the frames corresponding to each time-step on screen, filling all available screen space. They provide a complete, static overview of change over the analyzed time interval. As the number of time-steps increases, individual frames have to be made smaller leading

to a loss of detail. Examples of use include the visualization of mobility dynamics [65] and linguistic changes over a country [28]. Animations take a different approach: the time-steps are organized into an animated sequence. Frames are displayed one after the other, which means that each frame can fill the entire space available on screen. This sequential presentation gives a better sense of continuity, but can neither give a static overview of change over the time interval, nor enable direct visual comparison between non-contiguous frames. Animation is a very popular technique in geovisualization for presentation purposes (*e.g.*, [34, 44, 45]) and it is also used in more data analysis-oriented contexts (*e.g.*, [14, 19]). Based on neither small-multiples nor animation, Motion-Rugs [13] can nevertheless be considered as a particular case of juxtaposing location, as it reduces the spatial dimension into sequential one-dimensional slices that compose a simplified representation of a group of entities' movements.

**Juxtaposing Time.** Space-time cubes [40] fall in this category. For example, Kjellin *et al.* [38] depict infection cases as discrete cubes, and Andrienko *et al.* [3] traffic trajectory attributes on a city. Although more abstract, glyphs overlaid on top of maps can also be used to visualize movement data. For instance, Chen *et al.* [16] use such glyphs to display movement trajectories extracted from social media data.

### Visual Perception of Spatio-Temporal Dynamics

Being the most commonly-used general strategies to represent change over time, animation and small-multiples have been empirically compared in multiple contexts. Griffin *et al.* [31] evaluate their performance for visual identification of moving clusters and observe that participants performed better with animation. The study by Boyandin *et al.* [11] corroborate their findings for analyzing flow maps, showing that animation should be preferred for sudden-change detection and that small-multiples allow for comparisons across many or arbitrarily-distant time-steps, as they reduce the load on short memory. Comparative evaluations of the two strategies beyond spatio-temporal visualization yield similar results about sudden changes in network visualization [5, 6] or trend analysis with scatterplots [57, 12].

The results reported by Koussoulakou and Kraak [39] seem to contradict the above findings. They found animation more effective for long term analysis in a variety of map reading and interpretation tasks. Some details about experimental conditions in this study are not clear though (number of time-steps, level of interactivity). Variations in map and task complexity as well as display resolution could explain these discrepancies.

Regardless of the diversity of contexts in which animation and small multiples have been compared, there is no study in the literature about the effectiveness of both techniques at supporting the analysis of geotemporal movement: available studies do not consider geography and use fewer time steps ([5, 6, 31, 12, 57]), or focus on tasks and data not related to spatio-temporal dynamics ([11, 39]). Furthermore, studies that compare other geo-temporal visualizations (*e.g.*, [37, 43, 55]) focus on representing trends over space and time, and are therefore less related to our work. Detecting change over

time is a common task on these studies, but their results cannot be directly applied to propagation as analysis goals and movement patterns differ. For example, being aware of adjacency between locations is essential for spotting jumps when studying the propagation of a disease, but not when analyzing migration flows. And propagation may continue to increase and peak across all locations, which is not possible when a predefined number of individual entities move. Our work aims to fill this gap by evaluating three geo-temporal visualizations for the specific characteristics of propagation dynamics.

## STUDY RATIONALE

We first describe the three strategies considered in the study to represent propagation, and the specific design choices we make for each corresponding visualization technique. We then motivate our tasks and formulate hypotheses associated with them. Finally, we explain how we have generated the synthetic datasets used in the study (maps and propagation patterns).




### Selection and Design of Visualization Techniques

Although several of the spatio-temporal visualization techniques discussed earlier can handle propagation phenomena, making the study tractable implies limiting the set of techniques to compare. We selected these based on the following criteria: the techniques should be a) representative of the different strategies overall and b) in widespread use, readily able to encode geographical propagation, or c) variations on techniques from the literature requiring minimal extensions.

We considered a broad range of techniques, several of which we discarded based on inherent limitations. For instance, space-time cubes work well for a limited number of trajectories as the cubes remain sparse, but will typically yield cluttered views in the case of replication-based patterns on propagation maps, requiring users to break down the cube into slices. We also discarded techniques for which we were unable to identify a design that was not contrived. For example, techniques such as OD-Maps [67] support particular types of spatial-temporal data but require extension to encode propagation. Finally, we did not consider visualizations containing visual elements that required *a priori* analysis of the studied phenomenon and focused on those that show raw data instead. For instance, we did not consider elaborate animations such as that by Dudas *et al.* [26] or propagation graphs (*e.g.* [66, 21]) as they explicitly show computationally extracted propagation trajectories, making assumptions about what are important/interesting paths. Instead, we chose techniques that use raw data, as they are applicable to more analysis contexts since they leave the interpretation of the pattern to the analyst.

We eventually selected: a small-multiple map and an animated map as representatives of the **Juxtapose Location** strategy, and glyphs overlaid on a map as a representative of the **Juxtapose Time** strategy. Multiple variations are still possible for each of the three techniques, and further design choices have to be made. As noted by Tversky *et al.* [62] in their survey of studies comparing static *vs.* animated representations, one pitfall to avoid is the introduction of differences between conditions in terms of information conveyed or additional procedures. Minimizing such confounding factors is key to the

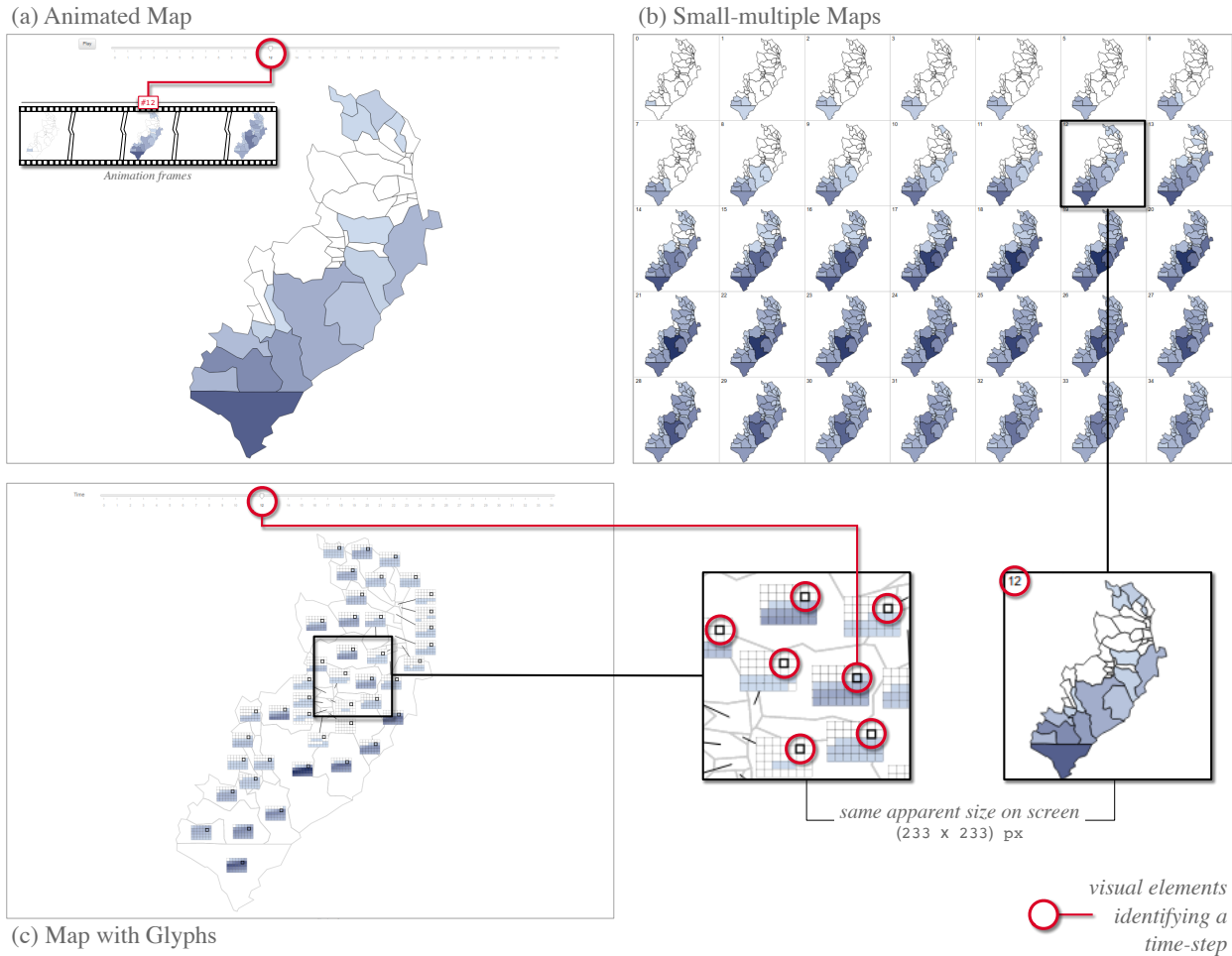
interpretation of the study's results, so we paid attention to enforce the following across techniques: 1) provide the same information using equivalent visual encoding channels; and 2) provide equivalent levels of interactivity.

All three techniques are illustrated in Figure 2 using the same propagation dataset with 35 time-steps. All encode the number of occurrences of the phenomenon in each geographical region using a sequential color scale, generated with an interpolator in the CIE L\*a\*b color space. The color corresponding to the minimum number of occurrences is a light shade of blue (#c6dbef  in RGB hexadecimal notation); the one corresponding to the maximum a dark shade of blue (#08306b ). Geographical regions that do not feature any occurrence of the phenomenon are colored white (#ffffff ). The light shade of blue corresponding to the minimum number of occurrences is set to be perceptually easy to distinguish from white (absence of occurrence) to avoid unnecessary ambiguities in the reading of values, which could introduce noise in our results.

**Animated Map:** this technique represents each time-step as a choropleth map displayed fullscreen, one after the other. Each time-step corresponds to one frame of the animation. Frame  $n + 1$  replaces frame  $n$  instantaneously, without interpolation between them, as in a classic cartoon. Playback speed is set to 10 frames per second. The literature emphasizes the importance of interactivity to aid comprehension [6, 31, 62]. In addition to a button that toggles playback, this technique features an interactive slider that indicates which time-step is currently displayed (Figure 2-a). Users can grab the knob and move it to any time-step, to quickly browse through the frames of the animation. Users can also click on any time-step, the knob instantaneously jumps to this time-step and the corresponding map frame is displayed. If the animation is running, interacting with the slider stops playback. To provide equivalent levels of interactivity across techniques, two additional features help users relate geographical regions over time-steps: hovering over a region with the cursor highlights it (increased stroke-width); clicking inside it keeps it highlighted throughout all time-steps even if the cursor exits the region.

**Small-multiple Maps:** this technique also represents each time-step as a choropleth map, but it displays all maps simultaneously by juxtaposing them in a grid pattern. A number in the top-left corner of each map indicates the time-step (Figure 2-b). Being all displayed simultaneously, the maps are necessarily much smaller than in the other two conditions. To provide equivalent interactivity across techniques, the following interactions help relate geographical regions over time-steps: hovering over a region on one of the small-multiple maps highlights it (increased stroke-width) on all small-multiple maps simultaneously (brushing [56]); clicking inside it keeps it highlighted across all maps even if the cursor exits the region.

**Map with Glyphs:** this third technique superimposes glyphs on top of a base map. Each glyph is associated with a geographical region, and encodes time-steps as small cells organized in a grid pattern (Figure 2-c). The same color mapping is used to encode the number of occurrences. In terms of layout, the technique is similar to a small-multiple-with-gaps configuration [49] that optimizes distances and directions rather than



**Figure 2.** The evaluated techniques, illustrated on the same map, showing the same data, and emphasizing the same time step (#12): (a) Animated Map, (b) Small-multiple Maps, (c) Map with Glyphs. Annotations in red (○) highlight elements that help users identify time-steps, in each visualization.

compactness and alignment, also relaxing grid constraints. Because it shows all time-steps simultaneously, this technique does not need to animate the representation. Because it juxtaposes time instead of locations, this strategy does not need to display many small maps, but can instead show a single map, fullscreen. While it does have its own weaknesses, this strategy has interesting properties, as discussed in our hypotheses, which led us to include it in the study even though it is not in widespread use compared to the other two. To provide equivalent levels of interactivity across techniques, the following interactive features help users relate time-steps over geographical regions: users can select a particular time-step using a time slider similar to that on the Animated Map, which highlights the corresponding time cell in all glyphs on the map; and conversely, clicking a particular time cell in any of the glyphs highlights it in all glyphs. To be consistent with the other techniques, hovering over a region highlights it; clicking inside it keeps it highlighted even if the cursor exits the region.

Some of our design choices warrant further discussion. First, we could have used cartograms instead of choropleth maps. While some cartogram techniques can be useful when encoding more than one variable per geographical feature [52, 55], they also have weaknesses that make them ill-suited to our con-

text (at least as a default choice): the abstract representation approximates topology, can be disorienting, and makes animated maps very unstable when mapping data to area (which would have been the main reason to investigate cartograms in the first place), no matter the type of cartogram [51].

Another choice we made concerns the design of glyphs overlaid on a map. While other visual mappings could have been used to encode occurrences over time (*e.g.*, a line chart, a bar chart), we chose a glyph design based on color, to minimize differences in visual encoding between techniques. Indeed, we are interested in comparing three strategies, and introducing unnecessary differences in the way data are visually encoded would have represented a possible confounding factor.

### Task Motivation

The goal of our study is to compare the effectiveness of these visualizations for the understanding and analysis of propagation phenomena. To identify and test possible propagation patterns, we first looked at the taxonomy of object movement patterns by Dodge *et al.* [24]. They categorize propagation as a generic compound pattern that occurs *when one object starts to show a certain movement parameter value and little by little other objects start adopting the same pattern.*

From the possible parameters that define a propagation movement, we focused on the ones referred to as *primary derivatives* in the taxonomy [24]: *propagation distance*, *direction*, *speed* and *duration*. We selected these parameters instead of simple primitive characteristics (like exact position or time), as they are more complex and represent real-world analysis tasks, such as finding out whether a disease is spreading quickly across a continent or is contained in specific regions (e.g., [17, 26]). We constructed the first three experimental tasks by varying one of these movement patterns.

In a fictitious scenario of a disease propagating across multiple regions in a country, participants were asked to:

**Scope:** determine if a disease affects the whole country or is constrained to a region near the origin (binary answer). This task considers the propagation *distance* parameter. We asked participants for a high level characterization of scope (local vs. global) instead of the exact extent of the propagation, as this would result in long trials for animation and small multiples since participants would have to consider each region of the map separately across time frames.

**Direction:** determine if the propagation follows a particular direction or if it spreads in multiple directions (binary answer). This task considers the propagation *direction*. Again we asked participants for a high level characterization of direction (if one exists or not) instead of identifying the specific direction, as the generated propagation does not necessarily follow compass directions and expressing direction in a form that is more free would require imprecise input (e.g., sketching).

**Arrival:** determine how much time it takes for the phenomenon to propagate to a particular region (numeric answer). We used this task as a proxy for the *speed* parameter. On one hand, calculating the actual speed is tedious as an experimental task, as participants have to count the amount of infected regions at each time-step. On the other hand, a higher-level characterization based on the notion of fast/slow speed is subjective. As judging speed requires understanding when the phenomenon reaches specific locations, we consider this core subtask instead, which is more amenable to operationalization.

We do not include the last of the four original parameters in the taxonomy [24]: the propagation *duration*. Indeed, it is of much lesser relevance for the study of visualization technique efficiency, as the start and end of a propagation phenomenon can be identified automatically by pre-processing the data.

However, as the movements considered in the taxonomy are mainly focused on a set of objects changing position, we believe the parameter list is incomplete for our purposes. Based on the literature on disease propagation and social media analysis, we consider two additional movement parameters.

The first parameter is the presence of *spatial jumps/hops*. It indicates whether the propagation happened only between adjacent locations or not. It allows us to differentiate between, for example, disease-dispersion in geographically contiguous locations (e.g., [17, 47, 9]), and dispersion that includes sporadic long-range jumps, as those generated by air transport (e.g., [26]) or information transmission in social media that has

short and long geographical range (e.g., [35, 28]). The second parameter refers to the identification of the propagation's *peak*. For example, it is common in epidemiology to look for the peak number of infection cases (e.g., [26]). As mentioned in the introduction and related work, these two tasks represent particular characteristics of propagation that are not shared with other movement patterns.

We introduce two tasks based on these movement parameters: **Hops:** determine if propagation takes place between adjacent locations, or if geographical jumps occur (binary answer). We asked participants to report the existence (or lack) of hops, instead of requesting them to identify the specific time-steps or number of hops that occur. We avoided the first because this would be reduced to a combination of Arrival tasks and the second because the number of hops would differ between trials making their difficulty hard to control.

**Peak:** determine when the propagation reaches its maximum value in a single region (numerical answer). Instead of asking for the peak across all regions, we asked participants to find the region with the peak amount of cases and report the corresponding time-step. This simplification is made, again, for the sake of task operationalization. Indeed, visually aggregating across all regions to determine peak values is a difficult task without precise numerical values. For instance, when comparing two time-steps, one where all regions seem to have a high value, and another where all regions seem to have medium values except one with a very high value, it is unclear in which time-step the peak appears.

## Hypotheses

We formulate our hypotheses based on the following properties of each visualization technique and type of task:

**Animation** displays time-steps as a sequence of frames, allowing the human visual system to easily detect sudden changes between consecutive frames [11, 31]. However, it is impossible to get an overview of the phenomenon at a glance. Users need to remember or revisit past frames in order to make judgments. Finally, animations require time to play.

**Small Multiples**, on the other hand, allow users to see the complete sequence at a glance, by juxtaposing all frames. They provide an overview of the entire phenomenon in one glance. Their main disadvantage is that, in order to fit on screen, frames have to be small. They lose detail, and are more difficult to read. Small-multiples reduce the load on short-term memory and allow for comparison across many, not necessarily contiguous, time frames [11].

**Glyph Map** also provides an overview at a glance, this time organized by geographical region. This aids comparing time-steps in the same region. Comparison across regions is possible, but requires visually relating time-steps. Interactive highlighting (e.g., brushing) is thus needed. Glyphs have to be small enough to fit on the map without overlapping and without overly obscuring the underlying base map. Time cells are necessarily small, making color comparisons challenging [61].

Tasks **Direction** and **Hops** require participants to compare the map across consecutive time steps: to determine if the

direction of propagation is consistent (for **Direction**), and if a non-contiguous region gets infected (**Hops**). Prior studies suggest that **Animation** performs well for analyses involving sudden changes [5, 6, 11, 31]. We thus expect similar results for tasks that require comparing contiguous frames such as **Hops**. **Small Multiples** should be less efficient, but the benefits of temporal juxtaposition in the **Animation** condition might actually get canceled by the cost of playing the animation or interacting with the timeline. These two tasks should prove particularly difficult for **Glyph Map** as the information will have to be gathered from across multiple glyphs.

**Task Scope** requires participants to find out if the disease appeared in most regions over time at least once. This task should be easier to perform with **Glyph Map** as participants just need to check that most glyphs feature at least one colored cell. The task should be more challenging with **Small Multiples**, as participants will have to check that most regions are colored in at least one time-step. We expect **Animation** to perform worst as participants may have to remember what regions have been colored over time, depending on the propagation pattern.

**Task Arrival**, here **Glyph Map** should perform best, as it requires participants to only focus on a specific glyph and identify the time cell in which the virus appears. We expect **Small Multiples** to perform worse, as participants will have to scan for that specific region across maps. Finally, **Animation** should perform worst as participants will have to wait for the animation to reach the frame when the region gets colored.

**Task Peaks** requires participants to identify which region becomes darkest (on the sequential color scale) over time. **Small Multiples** should perform best, as they enable a quick scan of all maps at a glance. However, the small size of individual maps might counter this advantage. The performance of **Animation** and **Glyph Map** will likely depend on the propagation pattern. If candidate time-steps for peaks are temporally close (*i.e.*, participants only have to determine which of two consecutive frames is darker), **Animation** should fare better. In such cases color comparisons between the small time cells in **Glyph Map** should prove difficult. If candidates for peaks are temporally distant (*e.g.*, one region has a high value early in time, and another near the end), **Glyph Map** may fare better. Although time cells are small, they are all visible at the same time. Using **Animation**, participants will have to remember previous frames, or go back-and-forth between them.

The above observations give rise to the following hypotheses, which were formulated before data collection:

**H1:** For task **Scope**: **Glyph Map** should perform best, followed by **Small Multiples**, and **Animation**.

**H2:** For task **Direction**: **Animation** should perform best, followed by **Small Multiples**, and **Glyph Map**.

**H3:** For task **Arrival**: **Glyph Map** should perform best, followed by **Small Multiples**, and **Animation**.

**H4:** For task **Hops**: **Animation** should perform best, followed by **Small Multiples**, and **Glyph Map**.

**H5:** For task **Peaks**: **Small Multiples** should perform best. It is unclear if **Animation** or **Glyph Map** will perform best.

## Dataset Generation

We generate artificial datasets to control for the different propagation parameters and ensure homogeneous task difficulty across conditions. Disease models have been used for decades in health research (*e.g.*, [7, 36, 50]). They make it possible to generate realistic simulations of propagation phenomena. We use the widespread SIR compartment model [50], which is easy to setup, fast to compute, freely available, and whose underlying network can take geography into account.

The model works as follows. Each individual in a population can have three possible states: *S*, susceptible to a disease; *I*, infected, can spread the disease to others; or *R*, recovered or died, and therefore no longer part of the simulation. The population itself is modeled as a network of individuals. A subset of them is set as seeds of the infection. Once the simulation starts, infected persons (*I*) can pass the disease with a defined probability to susceptible individuals (*S*) connected to them. Infected individuals (*I*) can also recover (*R*) with a defined probability. The infection and recovery probabilities are set between 0.1 and 0.5, depending on how fast or slow we want the disease to propagate. The simulation is executed as a stochastic Gillespie process [29], considering continuous time. It stops when it reaches an equilibrium. We use the epidemic library [23] as the implementation of this mode. We use **NetworkX** [33] to build the population network.

To avoid uncontrolled bias due to geographic familiarity, we use three distinct maps<sup>1</sup> from regions that participants are unlikely to identify, as Beecham *et al.* [8] did. Training trials use a map of the Ica region in Peru, which contains 43 administrative divisions. Main trials use a map of the Biobio region in Chile (51 divisions after island removal), or of the Zacatecas state in Mexico (56 divisions). The maps are chosen to have a range of diverse configurations in terms of division count and density. The simulations generate a continuous timeline, from which we extract 35 time-steps by sampling the simulation results using variable-length intervals. This is the same number of time-steps as in Boyandin *et al.* [11], representing a relatively large time interval for this type of study.

Finally, we generate the simulated population and its connectivity network for each map as follows. We first generate a number of individuals based on a density measure. To keep the computation tractable, the density of each map (that can number in the millions) is subdivided by a fixed factor to ensure that the maximum population remains in the range of a few thousands. With this, instead of generating millions of individuals per map we generate a few thousands only. We then create as many nodes as the chosen number of individuals and position them homogeneously in each region. To connect two nodes (*i.e.*, propagate), we use a geographical threshold graph model [48] featured in **NetworkX**. The probability that two nodes get connected is given by their distance: the closer they are (spatially), the higher the probability.

*Task specific aspects:* We intervene in the simulation in the following ways to produce desired patterns. First, for **Scope** and **Direction** tasks, in order to prevent the disease from

<sup>1</sup>All study material and results, including maps, logs and analysis scripts, are available at <http://ilda.saclay.inria.fr/propagvis>.

spreading to some particular locations, we add a new state for populations called vaccinated (*V*) that, as the name suggests, cannot be infected. Additionally, for *Direction* tasks, we also make sure that some regions eventually get rid of the virus, so as to avoid participants being able to characterize direction by looking at a single time-step. Second, for tasks that involve identifying a particular time-step (*Arrival* and *Peaks*), we create variations for the different repetitions where the key event (peak or arrival time) happens either at the beginning or at the end of the simulation (see rationale in Section Experimental Design). To do so, we follow two strategies: increase/decrease the probability of infection, or increase/decrease the width of the simulation-sampling time-interval. Third, in order to create jumps for task *Hops*, we manually insert connections between non-contiguous regions by adding edges between randomly-selected individuals. We connect at least three non-adjacent regions that result in 3 hops for each repetition in the main trials and 1 hop for training.

## STUDY DESIGN

Our study evaluates the effectiveness of 3 visualization techniques (*Animation*, *Small Multiples*, *Glyph Map*) for 5 tasks (*Scope*, *Direction*, *Arrival*, *Hops* and *Peaks*).

### Experimental Design

We use a within-subjects design where all participants are exposed to each visualization. For each technique, a participant has to perform 10 training trials (2 repetitions per task), and 5 measured tasks  $\times$  6 repetitions = 30 main trials. In total, the experiment consists of 18 participants  $\times$  3 visualizations  $\times$  5 tasks  $\times$  6 repetitions = 1620 main trials.

For tasks that call for a binary answer (*Direction*, *Scope* and *Hops*), half of the repetitions (3) are allocated to each answer. For those that call for a numeric answer (*Arrival* and *Peaks*), we ensure that the correct answer is in the first half of the timeline for half of the repetitions (3), and in the second half for the other three. This ensures that participants cannot predict where the answer is (and thus ignore parts of the time-steps, skipping ahead). It also allows us to balance between answers that can be given fast with a sequential search vs. and those that take longer (*e.g.*, first vs. second half of the timeline for *Animation* or *Small Multiples*).

Despite the care taken to generate datasets of equivalent complexity, uncontrolled differences can introduce noise in our observations as we are dealing with relatively elaborate tasks. We thus reuse the same datasets across repetitions for the different visualization techniques. To avoid participants realizing this and identifying a previously-seen dataset, we create 3 variations that yield different answers to the question asked. These variations mirror the original maps (vertical and horizontal flips), also time-shifting the propagation by one or two steps.

We counterbalance technique presentation order using a Latin square, as well as datasets. Task type presentation order is randomly assigned for each participant, as are repetitions.

### Apparatus and Participants

We conducted the experiment on a Macbook Pro with a 27" Apple Thunderbolt Display (2560x1440 pixels). The

study's UI is Web-based, and visualizations are generated with D3 [10]. We ensure that each visualization gets equivalent screen real-estate. The three techniques cannot have the exact same size and aspect ratio, so we keep width constant, yielding: *Animation* and *Glyph Map* at  $1663 \times 1406$  pixels, and *Small Multiples* at  $1663 \times 1256$  pixels. Form input & widgets are located on the left of these visualization panels.

We decided on 18 participants (enabling counterbalancing of technique presentation order) before starting the experiment. We recruited participants until we arrived at this pre-defined number. Exclusion criteria included participants failing to complete all conditions or failing any of the 3 training sessions.

From the final 18 participants (10 male and 8 female), one reported color blindness in differentiating blue and red, but such a distinction was not needed in our tasks. All had normal or corrected-to-normal vision. Age ranged from 23 to 47 ( $M = 28.6$ ,  $SD = 6.7$ ) and most of them were students (11/18) from either a PhD or a Masters' program. Their backgrounds were mainly HCI, Computer Science and Visualization. They were all volunteers, and did not receive any monetary compensation.

### Procedure

The experiment lasts approximately one hour. Participants first sign a consent form, are told that they can withdraw at any time, and fill in a demographics questionnaire. They are then given an overview of the experiment, with a description of the visualizations and tasks. For each visualization, they go through a detailed explanation and training phase, followed by the main trials, and finally a post-hoc questionnaire:

**1) Training.** The experimenter first explains the technique in depth, allowing participants to interact with it and familiarize themselves with the interface. Next, participants perform the training trials in the presence of the experimenter and are encouraged to ask clarification questions. They can repeat a trial or the entire training session, until they feel comfortable with the visualization and tasks. If they make no errors and have no further question, they can proceed with the main trials.

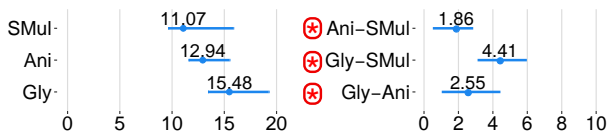
**2) Main trials.** For each of the 30 main trials, the question is displayed, along with the interactive visualization, and either the possible answers (binary choice tasks *Direction*, *Scope* and *Hops*) or an input field (quantitative answers for tasks *Arrival* and *Peaks*). After submitting their answer, participants are asked to report their confidence on a Likert scale. While tasks are presented in random order, all repetitions of a given task are grouped together. An intermediate page appears before moving to the next task, enabling participants to rest.

Once all main trials are completed for a given visualization, participants fill in a post-hoc questionnaire about the perceived difficulty of answering each type of task with this technique. They are also asked to verbally indicate if they used a particular strategy or any other comments. Finally, after all trials are completed, participants fill-in a form to rank the three visualizations based on preference and justify their ranking.

### Measures

We define three primary measures (2 objective, 1 subjective) that apply to all five tasks:





**Figure 3.** Left: Mean Completion Time in seconds for each visualization, for all tasks. Right: Pairwise comparisons for each visualization. Error bars represent 95% Bootstrap confidence intervals. Evidence of differences are marked with a ⊕ (the further away from 0 and the tighter the CI, the stronger the evidence).

- Completion Time: measured from the moment participants see questions until they answer. Averaged over repetitions (6).
- Error Rate: computed as the number of incorrect answers per task over the total number of repetitions (binary answers); or by normalizing according to the minimum and maximum values in the dataset (numerical answers).
- Self-reported Confidence: on a 3-point Likert scale (*highly confident, confident, not confident/random selection*).

Additionally, we record the following secondary measures:

- Self-reported Easiness to Complete Task: on a 5-point Likert scale from very easy (5) to very difficult (1). This measure is collected for each individual task, per visualization.
- Overall Ranking: numerical ordering of the three techniques (1-best), collected once at the end of the experiment. Accompanied by a text field for additional explanations.

## RESULTS

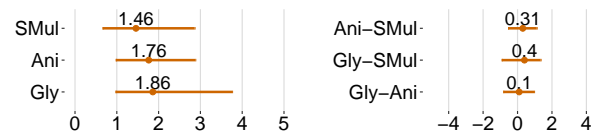
We report and interpret all our results using interval estimation instead of p-values [20, 25]. Collected data, analysis scripts and dataset generation code are available as supplementary material to this submission, as mentioned earlier.<sup>1</sup> All analyses were planned before collecting the data, and the study was preregistered [18] with the Open Science Framework.<sup>2</sup>

### Overall Results across Tasks

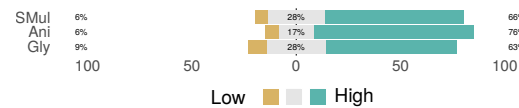
**Completion Time:** Figure 3 shows completion time for all tasks collectively, mean times per technique on the left, mean differences on the right. Mean times are shorter for Small Multiples (11.07s), followed by Animation (12.94s) and Glyph Map (15.48s). There is strong evidence that Small Multiples is faster, by 1.86s on average for Animation and by 4.41s on average for Glyph Map. There is also evidence that Animation is faster than Glyph Map by 2.55s.

**Error Rate:** Figure 4 shows error rates for all tasks collectively. Mean error rates per technique are on the left, mean differences on the right, mean error rates are lower for Small Multiples (1.46%), followed by Animation (1.76%) and Glyph Map (1.86%). There is no evidence that mean error rates are different across techniques.

**Confidence:** Figure 5 shows the self-reported confidence for each visualization, for all tasks. Confidence is high for all three visualizations in more than half the trials, although more so for Animation (76% of trials) than for Small Multiples (66%) and Glyph Map (63%).



**Figure 4.** Left: Mean Error Rate in % for each visualization, for all tasks. Right: Pairwise comparisons for each visualization. Error bars represent 95% Bootstrap confidence intervals.



**Figure 5.** Self-reported confidence across all tasks per visualization.

### Results per tasks

We now break down the results of Completion Time, Error Rate and Confidence per task. The gray texts next to task names refer to the associated hypothesis and summarize the expected outcome. Figure 6 shows the actual results, with CIs and the direction of our hypotheses (gray background).

#### Scope - (H1: Gly>SMul>Anim)

There is no strong evidence of differences among the three techniques, neither for Completion Time nor for Error Rate. It is possible that Small Multiples is less error prone than the other techniques, but evidence is not conclusive. Confidence was high for more than 80% of the trials for the three techniques, with Animation having the highest mean value. Therefore, **H1** is not supported.

We tentatively attribute this lack of evidence of a difference to participants doing the task approximately. Instead of inspecting all regions, participants went with an estimate of whether a big chunk of the map was filled or not, and thus did not need to visually scan all regions or wait for the animation to finish.

#### Direction - (H2: Anim>SMul >Gly)

There is evidence that Completion Time was different across all visualizations, with Animation being indeed the fastest (9.2s), followed by Small Multiples (12.18s) and Glyph Map (15.92s). We observe that on the contrary Animation is more error prone (7.41%) than Glyph Map (2.78%), although there is no evidence that these techniques were different from Small Multiples (4.63%). Confidence was high for more than 50% of trials for all techniques, Animation being the one with the highest confidence (71% of trials), followed by Glyph Map (65%) and Small Multiples (52%).

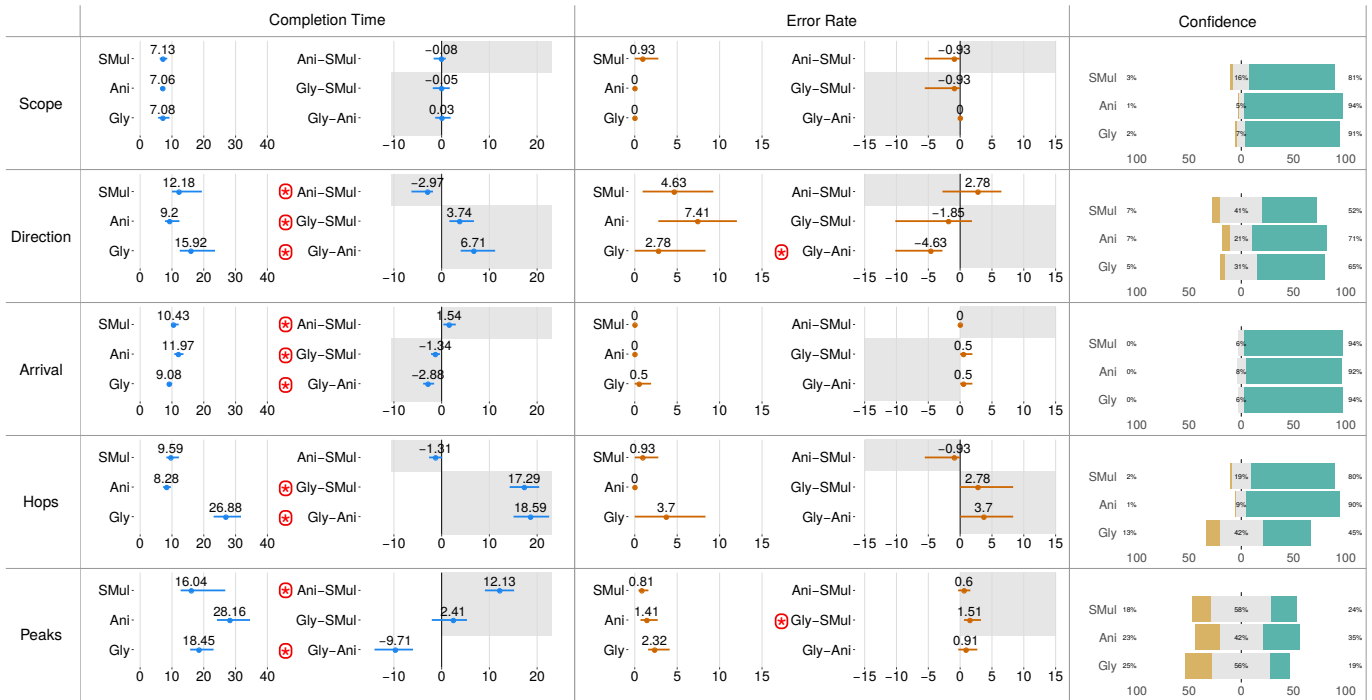
Our results only partially support **H2**. Animation is indeed the fastest, but participants seem to have traded speed for accuracy. As they found the task easy with Animation (indicated by the high confidence) they may have paid less attention with that technique, lowering their overall accuracy.

#### Arrival - (H3: Gly>SMul>Anim)

There is indeed evidence that Glyph Map is faster than Small Multiples (by 1.34s on average) and Animation (by 2.88s). And that Small Multiples is faster than Animation by 1.54s. There is no evidence of a difference in Error Rate.

Our results support **H3**, with the performance difference between techniques observed in Completion Time.

<sup>2</sup><https://osf.io/tdby5>



**Figure 6. Results for Completion Time (sec), Error Rate (in %) and self-reported confidence for each task. In each row (task) for the first two measures, mean values per visualization are seen on the left and means of pairwise differences on the right. Error bars represent 95% Bootstrap confidence intervals. Gray rectangles indicate the direction of our hypotheses. Evidence of differences are marked with a ⊕ (the further away from 0 and the tighter the CI, the stronger the evidence). The third column shows the percentage of trials that participants reported being highly confident (■) about, confident (◐) and not confident/random selection (◑).**

#### Hops - (H4: Anim > SMul > Gly)

There is strong evidence that Glyph Map is slower than both Small Multiples (by 17.29s on average) and Animation (by 18.59s). Our results suggest that Animation is faster and less error-prone than Small Multiples, but without strong evidence. There is no strong evidence of a difference in Error Rate between techniques. Results only suggest that Small Multiples is the least error-prone followed by Animation. Regarding confidence, participants reported high confidence in more than 80% of the trials with Animation and Small Multiples, and only in 45% of trials for Glyph Map.

Our results partially support **H4**. As expected Glyph Map performed worst. Nevertheless, we did not observe differences between Small Multiples and Animation. It seems that in our setup the cost of playing the Animation is equivalent to the cost of side-by-side comparison in Small Multiples.

#### Peaks - (H5: SMul > [Anim | Gly])

There is strong evidence that Small Multiples is faster than Animation (by 12.12s on avg.), and that Glyph Map is faster than Animation (by 9.71s). There is also evidence that Small Multiples is less error-prone than Glyph Map. Confidence was more divided in this task. Participants reported having average confidence for most trials (58% of trials for Small Multiples, 56% for Glyph Map, 42% for Animation) with the remaining trials divided between high and low confidence.

Our results confirm that (**H5**) Small Multiples is the best technique for this task, in terms of Completion Time when compared to Animation, and in terms of Error Rate when

compared to Glyph Map. Glyph Map also outperformed Animation in terms of Completion Time. This is likely because Glyph Map provides an overview of all time cells, easing their visual comparison despite their small size, a task that seems to be easier than remembering candidate colors or interactively moving back and forth between frames in animations. This was overall a harder task, as evidenced by several trials featuring low confidence levels across techniques.

## SUMMARY AND DISCUSSION

Our study compared different geo-temporal visualizations for the purpose of understanding and analyzing propagation phenomena. It involved a set of five tasks aimed at characterizing different propagation parameters. Small-multiple maps (Small Multiples) performed best overall, but animated maps (Animation) and maps augmented with glyphs (Glyph Map) outperformed them when determining propagation Direction and Arrival time, respectively. Looking more closely at individual measures, Small Multiples is faster than Animation and Glyph Map overall. But there is no difference among techniques in terms of error rate, at least when all tasks are considered collectively. As hypothesized, technique performance depends on the task considered. Interestingly, objective metrics are not aligned with subjective ones. Animation obtained the highest self-reported confidence score for almost all tasks, regardless of performance. This preference also shows in the overall ranking, as animation was preferred by 2/3 of all participants. Next we discuss our results disaggregated by the main task characteristics, drawing insights from our observations.

### *Comparison of consecutive time-steps.*

The characterization of Hops and propagation Direction involves such local comparisons. We hypothesized that Animation would perform better than Small Multiples for Hops (H4). But our results only suggest a trend toward such an effect. We tentatively attribute this to the fact that our study includes a larger number of time-steps than previous studies [5, 6, 11, 31]. Animations thus take longer to play.

As predicted, Animation was the fastest technique for Direction tasks (H2). However, it also proved more error-prone than Glyph Map. There seems to be a speed/accuracy trade-off at play, possibly related to the complexity of Glyph Map, which likely requires more effort to interpret. Our results confirm that Glyph Map, the only technique that juxtaposes time [55] in this study, performs worst for tasks involving the comparison of consecutive time-steps across regions (H2, H4). Indeed, the information is dispersed over the entire representation, making the tasks cognitively demanding.

*Summary of findings:* Animated maps perform better than small multiples when characterizing propagation direction. But there is no clear evidence that they perform better for tasks that require detecting sudden changes (*e.g.*, spatial jumps).

### *Search in large time intervals.*

Tasks Scope and Arrival require participants to perform a quick search over the whole sequence of time-steps. We hypothesized that Small Multiples and Glyph Map, which provide an overview of time-steps, would perform better than Animation (H1, H3). Glyph Map, followed by Small Multiples, was indeed the most efficient for Arrival tasks, which only involve a localized region. But this was not the case for Scope tasks, which involve multiple regions, with our results only suggesting a trend for Small Multiples to be less error-prone. When multiple regions need to be taken into account, the benefits of Glyph Map are attenuated and Small Multiples performs better.

*Summary of findings:* The overview offered by a map with glyphs, and to a lesser extent by small multiple maps, best supports tasks that involve long time intervals focused on a single region (*e.g.*, determining the arrival of a phenomenon). When multiple regions have to be considered (*e.g.*, determining the scope of a phenomenon), this advantage disappears.

### *Global search followed by local comparison.*

Identifying Peaks typically involves performing a search over the whole time interval, followed by comparisons of a few consecutive time-steps. Confirming hypothesis (H5), Small Multiples were faster than Animation and less error-prone than Glyph Map. We tentatively attribute the poor performance of Glyph Map to the small size of cells, which can make visual comparison of subtle differences in color shade difficult. Glyphs using other encoding channels likely suffer from the same resolution problem, though this remains to be tested. Animation was the slowest technique across the board, lacking an overview of time-steps and forcing users to go back and forth inside the time interval to compare candidates.

*Summary of findings:* small-multiple maps best support tasks that require both global and local temporal comparisons.

## **Limitations and Future Work**

To be tractable, experiments that involve multiple tasks and techniques necessarily require making choices of factors to study. As this is, to our knowledge, the first study on propagation visualization, we focus on the main strategies for visually organizing the data, while covering a representative set of tasks. Notable factors we did not include are: the number of time steps and geographical regions, and the level of interactivity.

We did not vary these factors, as they all relate in some way to the available screen real-estate. Increasing the number of time-steps will have varying impact on the visualizations. It will decrease the size of individual small-multiple maps, increase visual occlusion on maps with glyphs, but have no direct impact on the appearance of animated maps. However, it will make navigation more difficult in the latter, putting a heavier load on memory. Increasing the number of geographical regions will also impact visualizations differently. For instance, it may make individual small-multiple maps difficult to read more quickly than the other two because of the limited screen real-estate dedicated to each map. Adding interactive features like pan-and-zoom, could compensate for some of the above issues, but to a different extent depending on the visualization considered, number of time-steps and map complexity. All three factors are very likely to interact – as *experimental* factors – with one another. While we did not vary them, we carefully controlled them to have both equivalent and reasonable conditions. Our study involved datasets consisting of 35 time-steps, a larger number than previous studies [5, 6, 31]. We used maps with geographical region counts in the same range (43 to 56). With these choices all three visualizations could be accommodated without the need for pan-and-zoom. These factors should be studied further, and we hope our findings can inform the design of future studies.

To control for equivalent complexity and patterns we used a synthetic dataset generated by simulation models from epidemiology, thus capturing the nature and properties of propagation. Nevertheless, using real-world data will yield complementary empirical insights and remains future work.

Finally, from the five tested tasks, two have numeric answers (identifying a time in Arrival and Peak) and three are binary (identify the existence of a movement Direction or of Hops, and if Scope was global or local). Binary responses are not as expressive and may not reveal subtle differences between techniques when it comes to the interaction between the temporal and spatial dimensions. For example, characterizing Scope in more detail (*e.g.*, identify all affected regions or the compass direction of propagation) may yield accuracy differences between techniques depending on the number of regions involved, that we would not have been able to detect. Nevertheless, with the exception of Scope, our tested measures were indeed enough to tease out differences between visualizations. Results from this first investigation could inspire future work that considers the design of more nuanced measures in order to investigate more subtle differences across visualizations.

## **ACKNOWLEDGMENTS**

We thank Gytis Dudas for kindly replying to our questions about the analysis of disease propagation.

## REFERENCES

- [1] G. Andrienko, N. Andrienko, G. Fuchs, and J. M. C. Garcia. 2018. Clustering Trajectories by Relevant Parts for Air Traffic Analysis. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (Jan 2018), 34–44. DOI : <http://dx.doi.org/10.1109/TVCG.2017.2744322>
- [2] Gennady Andrienko, Natalia Andrienko, and Marco Heurich. 2011. An event-based conceptual model for context-aware movement analysis. *International Journal of Geographical Information Science* 25, 9 (2011), 1347–1370. DOI : <http://dx.doi.org/10.1080/13658816.2011.556120>
- [3] Gennady Andrienko, Natalia Andrienko, Heidrun Schumann, and Christian Tominski. 2014. *Visualization of Trajectory Attributes in Space–Time Cube and Trajectory Wall*. Springer Berlin Heidelberg, Berlin, Heidelberg, 157–163. DOI : [http://dx.doi.org/10.1007/978-3-642-32618-9\\_11](http://dx.doi.org/10.1007/978-3-642-32618-9_11)
- [4] Natalia Andrienko and Gennady Andrienko. 2013. Visual analytics of movement: An overview of methods, tools and procedures. *Information Visualization* 12, 1 (2013), 3–24. DOI : <http://dx.doi.org/10.1177/1473871612457601>
- [5] Daniel Archambault, Helen Purchase, and Bruno Pinaud. 2010. Animation, small multiples, and the effect of mental map preservation in dynamic graphs. *IEEE Transactions on Visualization and Computer Graphics* 17, 4 (2010), 539–552. DOI : <http://dx.doi.org/10.1109/TVCG.2010.78>
- [6] Benjamin Bach, Emmanuel Pietriga, and Jean-Daniel Fekete. 2014. GraphDiaries: Animated Transitions and Temporal Navigation for Dynamic Networks. *IEEE Transactions on Visualization and Computer Graphics* 20, 5 (May 2014), 740–754. DOI : <http://dx.doi.org/10.1109/TVCG.2013.254>
- [7] Duygu Balcan, Vittoria Colizza, Bruno Gonçalves, Hao Hu, José J. Ramasco, and Alessandro Vespignani. 2009. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences* 106, 51 (2009), 21484–21489. DOI : <http://dx.doi.org/10.1073/pnas.0906910106>
- [8] Roger Beecham, Jason Dykes, Wouter Meulemans, Aidan Slingsby, Cagatay Turkay, and Jo Wood. 2017. Map LineUps: Effects of spatial structure on graphical inference. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan. 2017), 391–400. DOI : <http://dx.doi.org/10.1109/TVCG.2016.2598862>
- [9] Julio A. Benavides, William Valderrama, and Daniel G. Streicker. 2016. Spatial expansions and travelling waves of rabies in vampire bats. *Proceedings of the Royal Society B: Biological Sciences* 283, 1832 (2016), 20160328. DOI : <http://dx.doi.org/10.1098/rspb.2016.0328>
- [10] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D3 Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (Dec. 2011), 2301–2309. DOI : <http://dx.doi.org/10.1109/TVCG.2011.185>
- [11] Ilya Boyandin, Enrico Bertini, and Denis Lalanne. 2012. A Qualitative Study on the Exploration of Temporal Changes in Flow Maps with Animation and Small-Multiples. *Computer Graphics Forum* 31, 3pt2 (2012), 1005–1014. DOI : <http://dx.doi.org/10.1111/j.1467-8659.2012.03093.x>
- [12] M. Brehmer, B. Lee, P. Isenberg, and E. K. Choe. 2020. A Comparative Evaluation of Animation and Small Multiples for Trend Visualization on Mobile Phones. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (Jan 2020), 364–374. DOI : <http://dx.doi.org/10.1109/TVCG.2019.2934397>
- [13] J. Buchmüller, D. Jäckle, E. Cakmak, U. Brandes, and D. A. Keim. 2019. MotionRugs: Visualizing Collective Trends in Space and Time. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (Jan 2019), 76–86. DOI : <http://dx.doi.org/10.1109/TVCG.2018.2865049>
- [14] N. Cao, Y. Lin, X. Sun, D. Lazer, S. Liu, and H. Qu. 2012. Whisper: Tracing the Spatiotemporal Process of Information Diffusion in Real Time. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (Dec 2012), 2649–2658. DOI : <http://dx.doi.org/10.1109/TVCG.2012.291>
- [15] Segolene Charaudeau, Khashayar Pakdaman, and Pierre-Yves Boëlle. 2014. Commuter Mobility and the Spread of Infectious Diseases: Application to Influenza in France. *PLOS ONE* 9, 1 (01 2014), 1–9. DOI : <http://dx.doi.org/10.1371/journal.pone.0083002>
- [16] S. Chen, X. Yuan, Z. Wang, C. Guo, J. Liang, Z. Wang, X. Zhang, and J. Zhang. 2016. Interactive Visual Discovering of Movement Patterns from Sparsely Sampled Geo-tagged Social Media Data. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (Jan 2016), 270–279. DOI : <http://dx.doi.org/10.1109/TVCG.2015.2467619>
- [17] G. Christakos, R.A. Olea, and H.-L. Yu. 2007. Recent results on the spatiotemporal modelling and comparative analysis of Black Death and bubonic plague epidemics. *Public Health* 121, 9 (2007), 700 – 720. DOI : <http://dx.doi.org/10.1016/j.puhe.2006.12.011>
- [18] Andy Cockburn, Carl Gutwin, and Alan Dix. 2018. HARK No More: On the Preregistration of CHI Experiments. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, Article 141, 12 pages. DOI : <http://dx.doi.org/10.1145/3173574.3173715>
- [19] P. Craig, N. R. Seiler, and A. D. O. Cervantes. 2014. Animated Geo-temporal Clusters for Exploratory Search in Event Data Document Collections. In *2014 18th International Conference on Information Visualisation*. 157–163. DOI : <http://dx.doi.org/10.1109/IV.2014.69>

- [20] Geoff Cumming. 2014. The New Statistics: Why and How. *Psychological Science* 25, 1 (2014), 7–29. DOI: <http://dx.doi.org/10.1177/0956797613504966>
- [21] Z. Deng, D. Weng, J. Chen, R. Liu, Z. Wang, J. Bao, Y. Zheng, and Y. Wu. 2020. AirVis: Visual Analytics of Air Pollution Propagation. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (Jan 2020), 800–810. DOI: <http://dx.doi.org/10.1109/TVCG.2019.2934670>
- [22] Doantam Phan, Ling Xiao, R. Yeh, and P. Hanrahan. 2005. Flow map layout. In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005*. 219–224. DOI: <http://dx.doi.org/10.1109/INFVIS.2005.1532150>
- [23] Simon Dobson. 2017–2018. epydemic: Epidemic simulations on networks in Python. <https://github.com/simoninireland/epydemic>. (2017–2018).
- [24] Somayeh Dodge, Robert Weibel, and Anna-Katharina Lautenschütz. 2008. Towards a taxonomy of movement patterns. *Information visualization* 7, 3-4 (2008), 240–252. DOI: <http://dx.doi.org/10.1057/PALGRAVE.IVS.9500182>
- [25] Pierre Dragicevic. 2016. Fair statistical communication in HCI. In *Modern Statistical Methods for HCI*, Judy Robertson and Maurits Kaptein (Eds.). Springer International Publishing, Chapter 13, 291–330. DOI: [http://dx.doi.org/10.1007/978-3-319-26633-6\\_13](http://dx.doi.org/10.1007/978-3-319-26633-6_13)
- [26] Gytis Dudas, Luiz Max Carvalho, ..., and Andrew Rambaut. 2017. Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature* 544, 7650 (2017), 309. DOI: <http://dx.doi.org/10.1038/nature22040>
- [27] Alexander Dunkel, Gennady Andrienko, Natalia Andrienko, Dirk Burghardt, Eva Hauthal, and Ross Purves. 2019. A conceptual framework for studying collective reactions to events in location-based social media. *International Journal of Geographical Information Science* 33, 4 (2019), 780–804. DOI: <http://dx.doi.org/10.1080/13658816.2018.1546390>
- [28] Jacob Eisenstein, Brendan T. O’Connor, Noah A. Smith, and Eric P. Xing. 2014. Diffusion of Lexical Change in Social Media. In *PloS one*. DOI: <http://dx.doi.org/10.1371/journal.pone.0113114>
- [29] Daniel T. Gillespie. 1977. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* 81, 25 (1977), 2340–2361. DOI: <http://dx.doi.org/10.1021/j100540a008>
- [30] Graells-Garrido, Eduardo, Caro, Diego, and Parra, Denis. 2018. Inferring modes of transportation using mobile phone data. *EPJ Data Sci.* 7, 1 (2018), 49. DOI: <http://dx.doi.org/10.1140/epjds/s13688-018-0177-1>
- [31] Amy L Griffin, Alan M MacEachren, Frank Hardisty, Erik Steiner, and Bonan Li. 2006. 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 (2006), 740–753. DOI: <http://dx.doi.org/10.1111/j.1467-8306.2006.00514.x>
- [32] Joachim Gudmundsson, Patrick Laube, and Thomas Wölle. 2008. Movement Patterns in Spatio-Temporal Data. In *Encyclopedia of GIS*. DOI: [http://dx.doi.org/10.1007/978-3-319-17885-1\\_823](http://dx.doi.org/10.1007/978-3-319-17885-1_823)
- [33] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. 2008. Exploring Network Structure, Dynamics, and Function using NetworkX. [http://conference.scipy.org/proceedings/SciPy2008/paper\\_2/](http://conference.scipy.org/proceedings/SciPy2008/paper_2/). In *Proceedings of the 7th Python in Science Conference*. 11 – 15.
- [34] Mark Harrower and Sara Fabrikant. 2008. *The Role of Map Animation for Geographic Visualization*. John Wiley & Sons, Ltd, 49–65. DOI: <http://dx.doi.org/10.1002/9780470987643.ch4>
- [35] Krishna Y Kamath, James Caverlee, Kyumin Lee, and Zhiyuan Cheng. 2013. Spatio-temporal dynamics of online memes: a study of geo-tagged tweets. In *Proceedings of the 22nd international conference on World Wide Web*. ACM, 667–678. DOI: <http://dx.doi.org/10.1145/2488388.2488447>
- [36] William Ogilvy Kermack and Anderson G McKendrick. 1927. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* 115, 772 (1927), 700–721. DOI: <http://dx.doi.org/10.1098/rspa.1927.0118>
- [37] Sungye Kim, Ross Maciejewski, Abish Malik, Yun Jang, David S Ebert, and Tobias Isenberg. 2013. Bristle Maps: a multivariate abstraction technique for geovisualization. *IEEE Transactions on Visualization and Computer Graphics* 19, 9 (Sept. 2013), 1438–1454. DOI: <http://dx.doi.org/10.1109/TVCG.2013.66>
- [38] Andreas Kjellin, Lars Winkler Pettersson, Stefan Seipel, and Mats Lind. 2010. Different Levels of 3D: An Evaluation of Visualized Discrete Spatiotemporal Data in Space-Time Cubes. *Information Visualization* 9, 2 (2010), 152–164. DOI: <http://dx.doi.org/10.1057/ivs.2009.8>
- [39] A. Koussoulakou and M.J. Kraak. 1992. Spatia-temporal maps and cartographic communication. *The Cartographic Journal* 29, 2 (1992), 101–108. DOI: <http://dx.doi.org/10.1179/000870492787859745>
- [40] Menno-Jan Kraak. 2003. The space-time cube revisited from a geovisualization perspective. In *Proceedings of the 21st International Cartographic Conference (ICC)*. 1988–1996.
- [41] John C Lang, Hans De Sterck, Jamieson L Kaiser, and Joel C Miller. 2018. Analytic models for SIR disease spread on random spatial networks. *Journal of Complex Networks* 6, 6 (03 2018), 948–970. DOI: <http://dx.doi.org/10.1093/comnet/cny004>

- [42] X. Liang, R. Aggarwal, A. Cherif, A. Gumel, G. Mascaro, and R. Maciejewski. 2016. Visualizing Malaria Spread Under Climate Variability. In *Proceedings of the Workshop on Visualisation in Environmental Sciences (EnvirVis '16)*. Eurographics Association, Goslar Germany, Germany, 29–33. DOI: <http://dx.doi.org/10.2312/envirvis.20161104>
- [43] Mark A Livingston and Jonathan W Decker. 2011. Evaluation of trend localization with multi-variate visualizations. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (Dec. 2011), 2053–2062. DOI: <http://dx.doi.org/10.1109/TVCG.2011.194>
- [44] Amy Lobben. 2008. Influence of Data Properties on Animated Maps. *Annals of the Association of American Geographers* 98, 3 (2008), 583–603. DOI: <http://dx.doi.org/10.1080/00045600802046577>
- [45] María-Jesús Lobo, Caroline Appert, and Emmanuel Pietriga. 2019. Animation Plans for Before-and-After Satellite Images. *IEEE Transactions on Visualization and Computer Graphics (TVCG)* 25, 2 (Feb 2019), 1347–1360. DOI: <http://dx.doi.org/10.1109/TVCG.2018.2796557>
- [46] Ross Maciejewski, Philip Livengood, Stephen Rudolph, Timothy F. Collins, David S. Ebert, Robert T. Brigantic, Courtney D. Corley, George A. Muller, and Stephen W. Sanders. 2011. A pandemic influenza modeling and visualization tool. *Journal of Visual Languages & Computing* 22, 4 (2011), 268 – 278. DOI: <http://dx.doi.org/10.1016/j.jvlc.2011.04.002>
- [47] Seth A. Marvel, Travis Martin, Charles R. Doering, David Lusseau, and M. E. J. Newman. 2013. The small-world effect is a modern phenomenon. <https://arxiv.org/abs/1310.2636>, *arXiv e-prints* (Oct 2013), arXiv:1310.2636.
- [48] Naoki Masuda, Hiroyoshi Miwa, and Norio Konno. 2005. Geographical threshold graphs with small-world and scale-free properties. *Phys. Rev. E* 71 (Mar 2005), 036108. Issue 3. DOI: <http://dx.doi.org/10.1103/PhysRevE.71.036108>
- [49] Wouter Meulemans, Jason Dykes, Aidan Slingsby, Cagatay Turkay, and Jo Wood. 2017. Small Multiples with Gaps. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan. 2017), 381–390. DOI: <http://dx.doi.org/10.1109/TVCG.2016.2598542>
- [50] Mark EJ Newman. 2002. Spread of epidemic disease on networks. *Physical review E* 66, 1 (2002), 016128. DOI: <http://dx.doi.org/10.1103/PhysRevE.66.016128>
- [51] S. Nusrat, M. J. Alam, and S. Kobourov. 2018a. Evaluating Cartogram Effectiveness. *IEEE Transactions on Visualization and Computer Graphics* 24, 2 (Feb 2018), 1077–1090. DOI: <http://dx.doi.org/10.1109/TVCG.2016.2642109>
- [52] Sabrina Nusrat, Muhammad Jawaherul Alam, Carlos Scheidegger, and Stephen Kobourov. 2018b. Cartogram Visualization for Bivariate Geo-Statistical Data. *IEEE Transactions on Visualization and Computer Graphics* 24, 10 (Oct. 2018), 2675–2688. DOI: <http://dx.doi.org/10.1109/TVCG.2017.2765330>
- [53] WorldHealth Organization. 2019. Ebola Virus Disease. <https://www.who.int/en/news-room/fact-sheets/detail/ebola-virus-disease>. (2019). Last accessed: 2019-09-04.
- [54] Christine Parent, Stefano Spaccapietra, Chiara Renso, Gennady Andrienko, Natalia Andrienko, Vania Bogorny, Maria Luisa Damiani, Aris Gkoulalas-Divanis, Jose Macedo, Nikos Pelekis, Yannis Theodoridis, and Zhixian Yan. 2013. Semantic Trajectories Modeling and Analysis. *ACM Comput. Surv.* 45, 4, Article 42 (Aug. 2013), 32 pages. DOI: <http://dx.doi.org/10.1145/2501654.2501656>
- [55] V. Peña-Araya, E. Pietriga, and A. Bezerianos. 2019. A Comparison of Visualizations for Identifying Correlation over Space and Time. *IEEE Transactions on Visualization and Computer Graphics* (2019), 1–1. DOI: <http://dx.doi.org/10.1109/TVCG.2019.2934807>
- [56] Jonathan C. Roberts. 2007. State of the Art: Coordinated & Multiple Views in Exploratory Visualization. In *Proceedings of the Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV '07)*. IEEE Computer Society, 61–71. DOI: <http://dx.doi.org/10.1109/CMV.2007.20>
- [57] George Robertson, Roland Fernandez, Danyel Fisher, Bongshin Lee, and John Stasko. 2008. Effectiveness of animation in trend visualization. *IEEE transactions on visualization and computer graphics* 14, 6 (2008), 1325–1332. DOI: <http://dx.doi.org/10.1109/TVCG.2008.125>
- [58] Hugo Romat, Dylan Lebout, Emmanuel Pietriga, and Caroline Appert. 2019. Influence of Color and Size of Particles on Their Perceived Speed in Node-Link Diagrams. In *Human-Computer Interaction – INTERACT 2019*. Springer International Publishing, Cham, 619–637. DOI: [http://dx.doi.org/10.1007/978-3-030-29384-0\\_37](http://dx.doi.org/10.1007/978-3-030-29384-0_37)
- [59] Terry A. Slocum, Robert S. Sluter, Fritz Connor Kessler, and Stephen C. Yoder. 2004. A qualitative evaluation of maptime, a program for exploring spatiotemporal point data. *Cartographica* 39, 3 (1 1 2004), 43–68. DOI: <http://dx.doi.org/10.3138/92T3-T928-8105-88X7>
- [60] Loes Soetens, Susan Hahné, and Jacco Wallinga. 2017. Dot map cartograms for detection of infectious disease outbreaks: an application to Q fever, the Netherlands and pertussis, Germany. *Eurosurveillance* 22, 26 (2017). DOI: <http://dx.doi.org/10.2807/1560-7917.ES.2017.22.26.30562>
- [61] D. A. Szafir. 2018. Modeling Color Difference for Visualization Design. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (Jan 2018), 392–401. DOI: <http://dx.doi.org/10.1109/TVCG.2017.2744359>

- [62] Barbara Tversky, Julie Bauer Morrison, and Mireille Betrancourt. 2002. Animation: Can It Facilitate? *Int. J. Hum.-Comput. Stud.* 57, 4 (Oct. 2002), 247–262. DOI: <http://dx.doi.org/10.1006/ijhc.2002.1017>
- [63] Cécile Viboud, Ottar N. Bjørnstad, David L. Smith, Lone Simonsen, Mark A. Miller, and Bryan T. Grenfell. 2006. Synchrony, Waves, and Spatial Hierarchies in the Spread of Influenza. *Science* 312, 5772 (2006), 447–451. DOI: <http://dx.doi.org/10.1126/science.1125237>
- [64] Javier Vidal-Mateo, Ugo Mellone, Pascual López-López, Javier De La Puente, Clara García-Ripollés, Ana Bermejo, and Vicente Urios. 2016. Wind effects on the migration routes of trans-Saharan soaring raptors: geographical, seasonal, and interspecific variation. *Current Zoology* 62, 2 (03 2016), 89–97. DOI: <http://dx.doi.org/10.1093/cz/zow008>
- [65] T. von Landesberger, F. Brodkorb, P. Roskosch, N. Andrienko, G. Andrienko, and A. Kerren. 2016. MobilityGraphs: Visual Analysis of Mass Mobility Dynamics via Spatio-Temporal Graphs and Clustering. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (Jan 2016), 11–20. DOI: <http://dx.doi.org/10.1109/TVCG.2015.2468111>
- [66] Z. Wang, M. Lu, X. Yuan, J. Zhang, and H. v. d. Wetering. 2013. Visual Traffic Jam Analysis Based on Trajectory Data. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Dec 2013), 2159–2168. DOI: <http://dx.doi.org/10.1109/TVCG.2013.228>
- [67] Jo Wood, Jason Dykes, and Aidan Slingsby. 2010. Visualisation of Origins, Destinations and Flows with OD Maps. *The Cartographic Journal* 47, 2 (2010), 117–129. DOI: <http://dx.doi.org/10.1179/000870410X12658023467367>
- [68] Y. Yang, T. Dwyer, S. Goodwin, and K. Marriott. 2017. Many-to-Many Geographically-Embedded Flow Visualisation: An Evaluation. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan 2017), 411–420. DOI: <http://dx.doi.org/10.1109/TVCG.2016.2598885>