This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3209474

A Comparison of Spatiotemporal Visualizations for 3D Urban Analytics

Roberta Mota, Nivan Ferreira, Julio Daniel Silva, Marius Horga,
Marcos Lage, Luis Ceferino, Usman Alim, Ehud Sharlin, Fabio Miranda

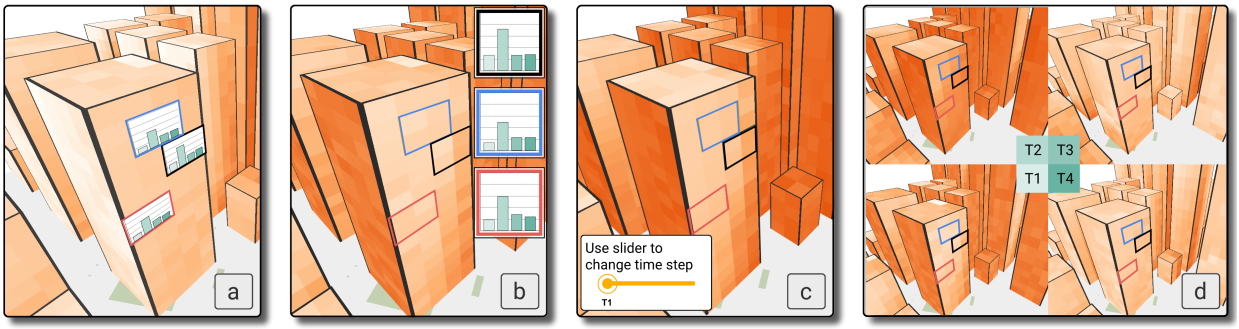

Fig. 1. The four spatiotemporal visualizations compared in our study: (a) `Embedded View`, (b) `Linked View`, (c) `Temporal Juxtaposition`, and (d) `Spatial Juxtaposition`. In this example, the visualizations encode the values of an artificially-created temporal attribute of the building facade with four time steps. In each visualization, three rectangular regions with colored outlines are placed along the building surface. When comparing their attribute values, the blue region has the minimum average attribute value between $[t_1, t_2]$ time steps, whereas the red and black regions have the maximum average values in t_3 and t_4, respectively.

Abstract—Recent technological innovations have led to an increase in the availability of 3D urban data, such as shadow, noise, solar potential, and earthquake simulations. These spatiotemporal datasets create opportunities for new visualizations to engage experts from different domains to study the dynamic behavior of urban spaces in this under explored dimension. However, designing 3D spatiotemporal urban visualizations is challenging, as it requires visual strategies to support analysis of time-varying data referent to the city geometry. Although different visual strategies have been used in 3D urban visual analytics, the question of how effective these visual designs are at supporting spatiotemporal analysis on building surfaces remains open. To investigate this, in this paper we first contribute a series of analytical tasks elicited after interviews with practitioners from three urban domains. We also contribute a quantitative user study comparing the effectiveness of four representative visual designs used to visualize 3D spatiotemporal urban data: spatial juxtaposition, temporal juxtaposition, linked view, and embedded view. Participants performed a series of tasks that required them to identify extreme values on building surfaces over time. Tasks varied in granularity for both space and time dimensions. Our results demonstrate that participants were more accurate using plot-based visualizations (linked view, embedded view) but faster using color-coded visualizations (spatial juxtaposition, temporal juxtaposition). Our results also show that, with increasing task complexity, plot-based visualizations perform better in preserving efficiency (time, accuracy) compared to color-coded visualizations. Based on our findings, we present a set of takeaways with design recommendations for 3D spatiotemporal urban visualizations for researchers and practitioners. Lastly, we report on a series of interviews with four practitioners, and their feedback and suggestions for further work on the visualizations to support 3D spatiotemporal urban data analysis.

Index Terms—Visualization, urban analytics, urban data, spatiotemporal data, empirical evaluation.

1 Introduction

Cities are the loci of economic activity and attract people looking for the myriad of services only offered by urban centers. With the growth of urbanization, the process and problems that shape and form cities become far more intertwined—transportation, housing, street layouts, and land use are all becoming more complex and interconnected by the day. In their attempt to meet the demands of an increasing number of urban dwellers, cities resort to planning initiatives that lead to the verticalization of their landscape. Far from being just another element of cities, their verticality is perhaps the *key* feature that defines cities. While serving as a potential solution to many of the problems faced by cities (*e.g.*, housing [8, 9]) verticalization can create new problems or exacerbate existing ones, *e.g.*, their impact on overall temperature [38]. Cognizant of this phenomenon and the fact that cities are characterized by their verticality, a growing number of domains (*e.g.*, civil engineering, urban planning, architecture) have gone beyond the usual flatland that defines a spatial region, and incorporated the third dimension (and its properties) in many of their analysis tasks in order to study the dynamic behavior of urban spaces. These tasks often rely on data that is intrinsic to the surface of buildings. Some examples include sound propagation simulation for studying noise in the urban environment [74], solar energy potential on facades for the installation of renewable energy equipment [41], outdoor heat exposure [49], sunlight access for the design of living spaces [37, 53, 55], and cellphone signal propagation

- *Roberta Mota, Julio Daniel Silva, Usman Alim, and Ehud Sharlin are with the University of Calgary, Canada. E-mails: {roberta.cabralmota, jd.silva, ualim, ehud} @ucalgary.ca.*
- *Nivan Ferreira is with the Universidade Federal de Pernambuco, Brazil. E-mail: nivan@cin.ufpe.br.*
- *Marius Horga and Fabio Miranda are with the University of Illinois at Chicago, United States. E-mails: {marius, fabiom} @uic.edu.*
- *Marcos Lage is with the Universidade Federal Fluminense, Brazil. E-mail: mlage@ic.uff.br.*
- *Luis Ceferino is with the New York University, United States. E-mail: ceferino@nyu.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.xxxxxxx

through urban canyons for the deployment of cellular towers [1]. The transitions to more sustainable environments, energy sources, housing, and technologies have underscored the importance of leveraging the geometry of cities in its entirety. As important in understanding cities as it is, 3D urban analytics has remained fairly unexplored. Unlike visual analysis on 2D maps, which has been thoroughly explored in the past, visualizations and visual analytics tools designed with 3D capabilities (and more importantly, that enable 3D-oriented tasks) are few and far between.

Designing visualizations for domain experts with 3D urban scenarios in mind is challenging, as it requires visual strategies to support analysis of the data referent to the city geometry. Tackling this is fundamental to uncovering features that are valuable for decision-making and problem solving in the aforementioned domains. Our goal in this work is, therefore, twofold. Recognizing the fact that 3D urban analytics is an unexplored yet growing area, our first goal is to have a deeper understanding of the requirements and analysis workflow usually employed by domain experts and practitioners in 3D urban analytics. To achieve this goal, we conducted a series of semi-structured interviews with experts from different domains in order to elicit a list of requirements that should be satisfied by visual analytics in the urban environment.

Our second goal is to better understand the effectiveness of different design strategies in the visualization of 3D spatiotemporal urban data: spatial juxtaposition, temporal juxtaposition, linked view and embedded view. To achieve this goal, based on the previously elicited requirements, we extended these designs to incorporate the geometries of cities and allow the visualization of data on building surfaces. We then performed a quantitative user study where participants performed a series of tasks that required them to identify extreme values on building surfaces over time. Our results point to the most performant designs to solve common 3D-oriented analytical tasks in the urban environment. Our contributions can be summarized as follows:

- A **task characterization** to inform visualization researchers and practitioners new to the 3D urban domain. We report on a series of semi-structured interviews with domain experts that routinely perform analytical tasks taking into account the 3D urban environment, and summarize visualization tasks encountered in this domain.
- A controlled **quantitative study** comparing user performance of four representative spatiotemporal visualizations, in which 32 participants were asked to perform tasks that required the identification of regions with extreme attribute values on building surfaces over time. We report and interpret study results, leading to a set of design recommendations on visualization choices for performant spatiotemporal analysis in 3D urban data.
- A series of **expert interviews** with four practitioners from the urban domain to get their perceptions and ideas for further work on the visualizations to support analysis of 3D time-varying urban data.

2 Related Work

In this section, we review previous visualization designs that combine space, time, and thematic data dimensions. The review is not restricted to urban contributions and includes general visualizations, to allow us to establish a set of common designs. We also survey and classify 3D urban visualizations according to the previously identified designs.

2.1 Visualization of 2D/3D spatiotemporal data

A survey conducted by Andrienko *et al*. on spatiotemporal visualizations for 2D geographic data mentions that existing visual strategies can be grouped into two categories [5]. *Linked views* segregate and display spatial and temporal dimensions in multiple coordinated views. *Embedded views* merge spatial and temporal aspects into a single view.

Linked views is a standard approach to display time-dependent geospatial data. It benefits from presenting data separately, which may prevent the occlusion of useful map information due to interference between the spatial and temporal data. However, a significant limitation of this approach is the additional screen space required to show side-by-side views. Additionally, the spatial decoupling between the 2D and 3D views may lead to mental burden related to the continuous context switching while relating the spatial and temporal data [21, 29, 88]. Turkay *et al*. proposed a visualization system to study how attributes vary over geographical spaces [78], where users create a linear selection sequence across a 2D map using a semi-automated brushing. After the brush sequence is computed, multiple linked sparklines are arranged in a 2D table layout, with sparkline depicting single attribute changes.

Embedded views typically use the map to convey the spatial referencing and superimpose time series plots as glyph overlays to present the changes in the data over time. Liu *et al*. embedded a circular time axis enclosing road trajectories on a map, and time-dependent attributes such as speed and duration of those trajectories are displayed as color-coded circular bar charts [50]. Andrienko *et al*. directly embedded ThemeRiver plots onto geographical maps, but this method led to severe clutter and occlusion of other map information [6]. To mitigate occlusion issues, Sun *et al*. applied a non-linear zooming algorithm to broaden map roads and overlaid temporal displays on the roads [75]. Abstraction and aggregation can also serve the purpose of de-cluttering and integrating space and time. Andrienko *et al*. discussed aggregation methods used for movement data [4]. Kim *et al.* group spatial 4D scientific visualizations into four categories: *juxtaposition*, *superimposition*, *interchangeable*, and *explicit encoding* [44]. Gleicher *et al.* follow a similar categorization in their taxonomy for infovis techniques [33]. These typologies are relevant as they discriminate visual strategies based on how data attributes are presented to the user, and as a result *which visual and motor perceptual mechanisms are used while comparing multiple data instances*.

Juxtaposition displays multiple data instances at the same time, but in different coordinate spaces, *e.g.*, side-by-side views [15, 42, 51]. The benefit is that all instances are simultaneously visible for comparison so that the user does not need to rely on their memory of past visualizations. However, juxtaposition leads to spatial decoupling, which hinders comparative judgments of slight variations between data instances. Additionally, scalability of juxtaposition poses two main challenges: available screen space and computational resources. When the number of data instances is high, the data encoding displayed in multiple views may become too small and hard to read. To deal with limited screen space, strategies include filtering to hide data instances that are of no interest to the user [46], clustering [51], or abstracting the visual representation of the data to take less screen area [42]. In order to maintain interactive frame rates as several model instances are being concurrently rendered—progressive computation and sampling can be used to alleviate computational costs [12, 39].

Johnson *et al*. propose multiple volumetric views juxtaposed in a grid arrangement for exploratory analysis of moderate-sized 4D cardiac data ensembles on the order of 10 data instances [39]. Each grid column refers to a single data instance, *i.e.*, one simulation run from the ensemble at a given time step. An entire grid row of volumetric views is created whenever users specify a box-shaped selection of a subvolume of interest on a given view (grid cell). In doing so, the technique follows a "global-to-local" exploratory approach: users progressively select sub-volumes of interest to interactively build up a comparative spatially-juxtaposed view layout; users can make visual comparisons of key features of interest across data instances and/or time steps.

Interchangeable displays only one data instance at a time and gradually transitions over time through either an automated or user-controlled animation; the data instances are co-registered in the same coordinate space. This has the advantage that each data instance is displayed at a full scale, without occlusion from other instances. However, since all data instances to compare are not simultaneously visible on the screen, users must rely on remembering past instances. This cognitive load increases with an increasing number of instances and may hinder comparison, especially for more detailed analyses [63, 82, 86]. Akiba *et al.* proposed an animation tool for volume visualization where users create a sequence by combining various individual effects that are set using transfer functions [2].

Superimposition displays two or more data instances at the same time and coordinate space. This is the only approach that displays multiple data instances in a way that is spatially co-registered, simultaneously visible, and preserves the original data. However, clutter and occlusion are major challenges that superimposed visual designs must

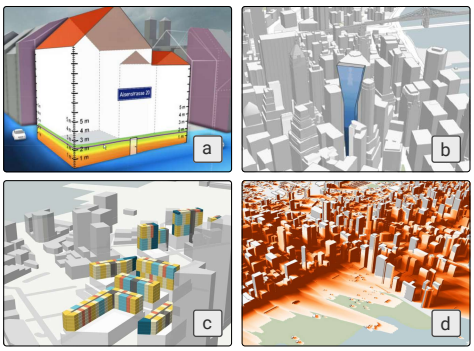

Fig. 2. Examples of data on urban surfaces from different domains: (a) flooding simulation analysis in the environmental domain [80], (b) view impact analysis in urban planning [23], (c) building design performance analysis in urban design [73], and (d) solar potential and sunlight access analysis in urban planning and architecture [53].

overcome to allow the comparison between multiple data instances. Glyphs are commonly used for overlaying various data instances in a 3D space [65, 68, 90]. Van Pelt *et al.* proposed a technique that combines established scientific visualizations with novel glyphs whose designs are based on concepts from information visualization [79]. The glyph-based visualizations support interactive filtering and details-on-demand zooming to set different zoom levels and provide progressively richer information within the glyph.

Explicit encoding does not display multiple data instances, but rather derives a composite data from all the original instances and only displays the composite result (*e.g.*, difference). The advantage of this approach is that the derived composite is displayed in full scale. In addition, for well-defined tasks, it is possible to set tailored encodings and visualize the composite result directly. On the other hand, explicit encoding does not preserve the original data; Therefore it is challenging to make comparisons between individual instances, even when the encoding is known [3, 30, 36]. Melanie *et al.* proposed a glyph-based visualization to study the progression of multiple sclerosis lesions from MRI scans over multiple time steps [76].

2.2 Visualization of 3D spatiotemporal urban data

Visualization systems are important tools for the analysis of urban data [24, 92]. The majority of these applications use a flat city metaphor—*i.e.*, a 2D map, to represent the city environment [25, 29, 40, 54, 92]. However, cities have inherent geometric features, leading to an increase in the number of 3D urban visualization systems in recent years (see Figure 2). These systems justify the need for 3D since specialists evaluate and communicate to stakeholders using 3D urban models (*e.g.*, discussing building designs with varied impacts on shadows, landmark visibility, and sky exposure [23, 28, 53, 58]).

After learning about the various visual designs for spatiotemporal data, we wanted to obtain an overview of the designs commonly used for 3D urban data. For this purpose, we reviewed recent 3D urban visualization systems that deal with multi-dimensional 3D urban data and classified them according to the visual designs discussed in the prior section—see classifications in Table 1. It is worth noting that, at a high level, the elicited spatiotemporal visual designs are essentially comparative designs that can be used for any multi-dimensional data type if reinterpreted slightly for the situation where multiple data dimensions are compared. Hence, our review includes all multi-dimensional data types such as temporal, multivariate, and uncertain.

From the literature, we searched papers on publication channels that publish quality work on visual analytics. These venues were either from the visualization or urban domains—*e.g.*, IEEE Transactions on Visualization and Computer Graphics; Computers & Graphics; Computer Graphics Forum; and Symposium on Simulation for Architecture and Urban Design. For these venues, although not exhaustively, we surveyed the archive published over the last ten years and selected 34 papers presenting 3D urban visualization systems. Table 1 displays a classification of recent 3D urban visualization systems. We observe that, concerning visual designs coming from 3D sci-vis, 3D urban vis systems heavily rely on *interchangeable* (27), followed by *juxtaposition* (7), and very few systems seem to use *explicit encoding* (2) and *superimposition* (1). Regarding designs drawn from 2D info-vis, we observe that most 3D urban vis systems do include 2D abstract visualizations (19). Among these, *linked views* is considerably more commonly used than *embedded view* (12 and 7, respectively).

Table 1. Classification of 3D urban visualization systems. The rows group systems according to Kim *et al.*'s taxonomy for 3D spatiotemporal scientific visualizations [44]. The first two columns group according to Andrienko *et al.*'s topology for 2D spatiotemporal information visualizations [5]. The third column refers to 3D urban vis systems that do not include 2D visualizations.

	LinkedV	EmbeddedV	No 2D
Juxtaposition	[34,46,48, 58,83,84]		[57]
Interchangeable	[28,31,67, 82,83,89]	[17,18,63,72, 77,80,91]	[13,16,22,43, 45,47,52,53,59, 62,64,69,71,87]
Superimposition	[58]		
Explicit encoding	[63]		[47,53,85]

3 Study Rationale

Unlike well-established sub-areas of visualization, such as scientific visualization, 3D urban visualization relies on ad-hoc solutions, primarily driven by industry standards or one-off collaborations between visualization researchers and domain experts. Our goal in this work is then to offer a more comprehensive understanding and characterization of visualization tasks and visualization designs that support exploration of 3D time-varying urban data. In order to achieve this, we structure our efforts in two different phases. First, we review the literature and perform a series of interviews with domain experts to elicit representative tasks and visualizations used in the urban domain. Second, guided by our previous findings, we select a subset of the elicited tasks and design a quantitative user study to measure the effectiveness of four techniques for the visualization of 3D spatiotemporal urban data.

This section reports on the first phase of our work. In Section 3.1, we detail interviews with experts and outline common tasks they perform for time-varying urban data analysis in 3D city models. In Section 3.2, we discuss representative approaches to visualizing 3D time-dependent urban data, with the second phase of our work detailed in Section 4.

3.1 Task characterization

We conducted a series of one-hour semi-structured interviews with domain experts: a civil engineer interested in urban resilience, disaster management, and earthquake simulations; an urban planner actively involved in the analysis of the solar potential of buildings; and an architect interested in the design of public and living spaces with more equitable access to sunlight and views to landmarks. Our first aim during these interviews was to better understand the types of temporal analyses they commonly perform with the purpose of deriving a *task characterization* that covers a range of representative tasks that 3D spatiotemporal urban visualizations must be able to support. Since these visualizations combine spatial, temporal, and thematic data dimensions, we frame our characterization after these data facets as four guiding **questions** asked during interviews (Q1-4), followed by a **discussion** of common use cases in their domains (D1-4), and condensed **responses** (R1-4).

Q1: What actions are relevant for temporal data analysis in 3D city models?

D1: In 3D city models, the visual analyses of spatiotemporal attributes are driven mainly by the spatial rather than the temporal dimension. Actions include *spatial search* (*e.g.*, search for buildings with increasing solar incidence throughout the day), *spatial comparison* (*e.g.*, compare shadow behaviors between different facades; which one is most shadowed during the winter?), and *temporal comparison* (*e.g.*, compare shadow behavior between summer and winter months on a given street; in which season does the shadow distribution have a sharper increase? In which month does it receive less shadow?).

R1: **Spatial search**, **Spatial comparison**, **Temporal comparison**

Q2: What spatial granularities are relevant for temporal data analysis in 3D city models?
D2: Domain specialists may be interested in different granularity levels. Architects run performance-driven analysis of candidate building designs (*e.g.*, daylight performance due to its influence on aesthetics and perception of space, energy conservation, and occupants' health and satisfaction). The evaluation of candidate designs typically operates on the *facade* level, and this includes both the *entire* and custom *region* of the facade. Whereas urban planners commonly perform multi-scale spatial analysis: from the *street* to *facade* to *building* to the *neighborhood* level (*e.g.*, how does the shadow behavior on a given *street/facade/building* vary due to a new building development?).
R2: **Street**, **Facade region**, **Facade**, **Building**, **Neighborhood**
Q3: What temporal granularities are relevant for temporal data analysis in 3D city models?
D3: Typical temporal resolutions include individual *time steps* and custom *time intervals*, and this may vary on a case-by-case basis according to the analysts' needs. For instance, when examining wind circulation behavior on a given neighborhood, they can focus their analyses on peak hours, or *monthly* and *seasonal* (*e.g.*, summer, winter) trends.
R3: **Time step**, **Time interval**, **Monthly**, **Seasonal**
Q4: What attribute behaviors are relevant for temporal data analysis in 3D city models?
D4: Domain experts commonly examine *extreme* attribute values (*e.g.*, which street receives most shadow during summer?), *average* values (*e.g.*, on a given facade region, what is the average solar incidence during winter months?), and *trends* (*e.g.*, on a given facade, is there small variation in wind circulation throughout the day?).
R4: **Extremes**, **Averages**, **Trends**

In our task characterization, the combination of the previous dimensions allow us to derive visualization tasks. For example, a common visualization task reported in urban planning involves the evaluation of sunlight access in the buildings of a given neighborhood, which can be specified as a **spatial search** for **facades** with **extreme** attribute values at certain **months** of the year. We build on this characterization to elaborate representative tasks for our user study (Section 4).

3.2 Visualization analysis

During our interviews, we learned that our domain collaborators worked with time-varying scalar functions defined on grids on top of the building surfaces. This data type therefore is set as the focus of study in this paper. We also learned that the experts' approach to visually examining time-dependent 3D urban data consists of color-coding attributes on the building surfaces, and using manual/automatic animations to display the data over time. They mentioned that due to the 3D nature and complexity of urban surfaces, relevant information may be missed or occluded and therefore requires users to restart the animation several times to inspect patterns in the visualization.

Following the interviews, we reviewed the literature in order to find alternative visual designs used to integrate temporal information into their 3D spatial references. As described in Section 2, our literature review identified a number of possible designs from which we selected four for comparison in our user study.

3.2.1 Selection of visualizations

To make our study tractable, our criteria for selecting representative visualization designs was twofold. First, we chose designs based on familiar 2D information visualizations in order to assess the efficiency of two ways of combining the physical data (building geometry) and thematic (temporal) data: *linked* and *embedded views*. Second, concerning designs based on 3D scientific visualizations, we chose only those derived from *juxtaposition* and *interchangeable* approaches. Not only are these familiar to the domain specialists, they are also the most popular designs used for 3D spatiotemporal data visualizations [44].

Our decision to discard explicit encoding designs is due to the fact that they do not present the entire original data values but rather only composite values derived from the original instances. Visualizing a meaningful derived relation between two or more data instances may provide a more focused encoding than would be possible by visualizing all instances separately. However, we are interested in designs that allow users to visualize individual data instances and to judge variations between them to maintain accordance with experts' analysis workflow.

Moreover, we also discard techniques based on superimposition since they often suffer from overplotting and strongly rely on carefully-tailored encodings [44]. The superimposed visual encodings need to not only communicate data semantics, but also be perceptually separable from one another so that users can perceive each encoding individually (*e.g.*, understand and discern encodings referent to a specific time interval). Therefore, we discard superimposition as we feel they would become a design challenge that could possibly bias study results.

Based on our choices, we identify four representative visual designs used to combine thematic, spatial, and temporal data into one visual representation. The visualizations can be classified into two major groups according to the graphical encoding of the spatiotemporal thematic attribute: *plot-based* and *color-coded* designs. Plot-based designs draw from 2D information visualizations and visually encode the attribute as the *position on a common scale* of an abstract time series plot. Color-coded designs build on 3D scientific visualizations and map the attribute to the *color* visual channel of the building geometry itself. The four visual designs are rationalized below, taking into account the fact that the data type of our time-varying thematic attribute is quantitative in nature. We refer the reader to the supplementary video for dynamic frames of each visualization.

Embedded view (`EmbeddedV`) displays all data instances on a bar chart embedded in the building surface geometry—Figure 1(a). The quantitative attribute values are encoded as the bar charts' ordinates, while time stamps are encoded as abscissas. The choice for encoding an attribute value as the vertical position on a common scale (bar chart) is because this visual channel is known to be one of the most effective for accurately decoding information [56]. Furthermore, adding a coordinate system together with reference lines can facilitate the judgment and comparison of different bars. This may be useful to mitigate perceptual issues due to perspective distortion and discrepant aspect ratios of different buildings—Figure 3(a,b).

To embed the bar chart into the building surface, we use a decal-based approach similar to Rocha *et al.* [65]. A decal map is built from the intersection of a sphere with a building surface; that is, the part of the surface that lies inside the sphere. This leads to an area that follows the surface geometry, like a decal. Afterwards, a local parametrization is computed to map the bar chart texture to the building's geometry.

Linked view (`LinkedV`) displays all data instances on a bar chart similar to embedded view; however, the chart is arranged on the right side of the screen, as depicted in Figure 1(b).

Temporal juxtaposition (`TemporalJX`) displays a single data instance at a time in full scale, and a user-controlled slider widget allows switching to other instances—see Figure 1(c). A sequential colormap encodes the temporal attribute, and this same coloring is used throughout all time steps.

Spatial juxtaposition (`SpatialJX`) arranges all data instances in an evenly-spaced 2×2 grid layout—see Figure 1(d). Since data instances are not registered in the same coordinate space, we extend the views with linked cameras and linked region highlights in order to provide a common context between the multiple views. A sequential colormap encodes the time-dependent quantitative attribute (*e.g.*, shadow).

These visualization designs, while satisfying our previously mentioned criteria, also posses certain drawbacks: `EmbeddedV` can distort visual information [81]; `LinkedV` requires additional screen space for side-by-side views and the decoupling between 2D and 3D views may lead to extraneous cognitive load related to the continuous context switching [21, 29, 88]; `TemporalJX` does not show all data instances simultaneously, users must rely on memory of past instances which also increases cognitive load [63, 82, 86]; and `SpatialJX` requires available screen space (see Section 2.1 for more details).

3.2.2 Interactions

The selected visualization designs support interactivity. The user can control the camera with the mouse—rotation via the right mouse button, zoom using the mouse wheel, and panning by pressing and holding

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3209474

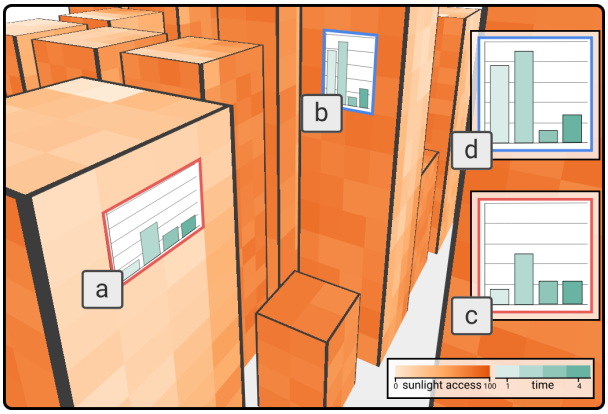

Fig. 3. Reference lines may assist comparative data analysis in at least two ways for `EmbeddedV`. (a, b) First, when plots are embedded in building geometries with discrepant resolutions. Second, as perspective may distort visual assets. A line, as well as the white area formed between the top of a bar and that line, provides common baseline scales for data comparison–*e.g.*, when comparing values in time $t4$. (c, d) Direct view of the same plots displayed in (a), (b) as a comparative baseline of perceptual artifacts: the same plots without perspective distortion and with equal resolution.

the mouse wheel while moving the mouse. Furthermore, in each visualization, users can also interact with regions. We define a *region* as a compound area composed of 2×2 grid cells along the building surface. The *region's attribute value* refers to the average values of all cells composing that region. In this sense, the color-coded visualizations (`SpatialJX`, `TemporalJX`) require users to mentally compute the mean value within a region based on its cells' colors. In the plot-based visualizations (`LinkedV`, `EmbeddedV`), users judge a region's mean value based on the vertical position along the bar chart.

Participants create a region by clicking the left mouse button over the building surface; drag a region by pressing and holding the middle mouse button over it while moving the mouse; and delete a region by clicking on it with the right mouse button. Users can define up to four regions simultaneously to compare attribute values on different parts of the building surface; regions are discriminated by color—Figure 1 depicts three regions in each visualization design.

4 Study Design

We now present our user study, whose goal was to compare the effectiveness of the four visualization techniques discussed in Section 3.2. The study followed a between-subjects design—participants were only exposed to one of the visualization techniques—and simulated an analysis scenario in which users were asked to perform some of the visualization tasks described in Section 3.1. Participants performed a total of 12 tasks. The task presentation order was counterbalanced using Latin squares, and the visualization assignment for the participants was randomized. In total, the study consisted of 32 participants $\times$ 12 tasks = 384 trials.

Datasets. We synthetically generated two 3D city models—one model for training and another for the main trials. The models incorporated common urban patterns (*e.g.*, medium-density downtown), and had, on average, 100 buildings with regular shapes and varying sizes. For each model, the time-varying attribute defined on the facades was created by discretizing each facade into $5\text{m} \times 5\text{m}$ cells. For each cell, we computed the accumulated sunlight access throughout 8 hours of a day for each one of the four yearly seasons, each season referring to one timestamp. Lastly, we added noise to increase data variability.

Apparatus. The visualizations were implemented using Unity. The experiments were conducted remotely on an Intel Core i3 and GeForce GTX Titan and used Zoom and Parsec for remote access.

Participants. We recruited a total of 35 participants but had to exclude one participant who declared during the session that s/he had misunderstood how to perform the tasks. We also excluded participants who had internet connection issues during the remote session. From the final

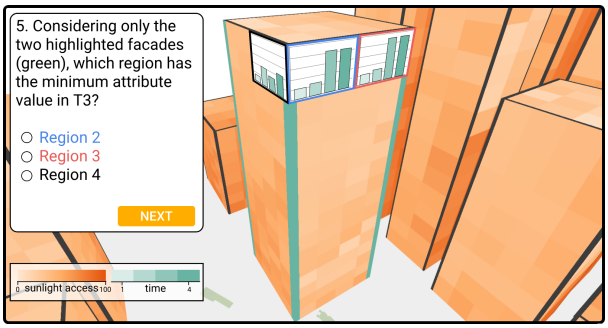

Fig. 4. `EmbeddedV` trial (Task #5, **multi-orientation 1 location** in **1 timestep**), with three user-created regions: blue (region 2), red (3), and black (4).

32 participants (9 female and 23 male), none reported any color deficiency, and they had normal or corrected-to-normal vision. Age ranged from 21 to 47 (avg. = 29.18, std. dev. = 6.38), and most of them were students (26 out of 32) from Computer Science graduate programs. Their backgrounds were mainly in visualization and HCI. They were all volunteers, and did not receive any monetary compensation.

4.1 Procedure

Each session consisted of three parts: introduction and tutorial, main trials, and post-study questionnaire (60–90 minutes total). In the introduction and tutorial, we explained the purpose of the experiment and the visualization to be used in the session. Then, the participants answered five training trials. After each trial, we would indicate if the answer was correct or not. They were encouraged to ask questions throughout the tutorial. If they made no errors and declared that they had no further questions, they would start the main trials. In the main trials, participants answered 12 questions and received no feedback regarding the correctness of their answers. They were asked to answer as quickly as possible without compromising accuracy.

Figure 4 shows a trial screenshot. For each trial, the building facade(s) pertaining to the trial was highlighted in green. Users could create, delete, or move regions along these target highlighted facade(s). On the top-left corner, the text box depicts the task number and question for the trial. It also lists all existing regions at any given time. By selecting a region on the list and clicking the 'next' button, the user submits that region as the task answer. On the right is the 3D urban visualization; the same visualization was used across the training and main trials—`SpatialJX`, `TemporalJX`, `LinkedV`, or `EmbeddedV`.

4.2 Measures

For each trial, we defined two objective metrics: 1) **task completion time** is measured from the moment participants saw the trial screen until they submitted an answer; and 2) **relative error** is computed as the absolute difference between the chosen attribute value and the true value, normalized by the task's maximum possible absolute error. During each trial, we also logged qualitative observations about participants' behaviors, interaction strategies, and aloud comments. After completing the trial, participants also filled out a post-study questionnaire to share positive and negative opinions, as well as ideas to improve the visualization used in their session.

4.3 Tasks

Guided by our task characterization, our study aims to compare the four visualizations in regards to a relevant task: to spatially search for the facade region with extreme attribute value over a certain time period. We had no hypothesis about which action would be more difficult: to characterize a region's attribute value (by averaging its cell values using position or color encoding), or to discriminate a region's attribute variation over time. Hence, we treated them as an integrated task that requires users to assess both the mean attribute value within a region and its changes over time, and varied the combinations of these two spatial and temporal dimensions during the construction of our tasks.

Figure 5 illustrates our task construction. To define our tasks, we used Peuquet's framework for geo-temporal data that describes the

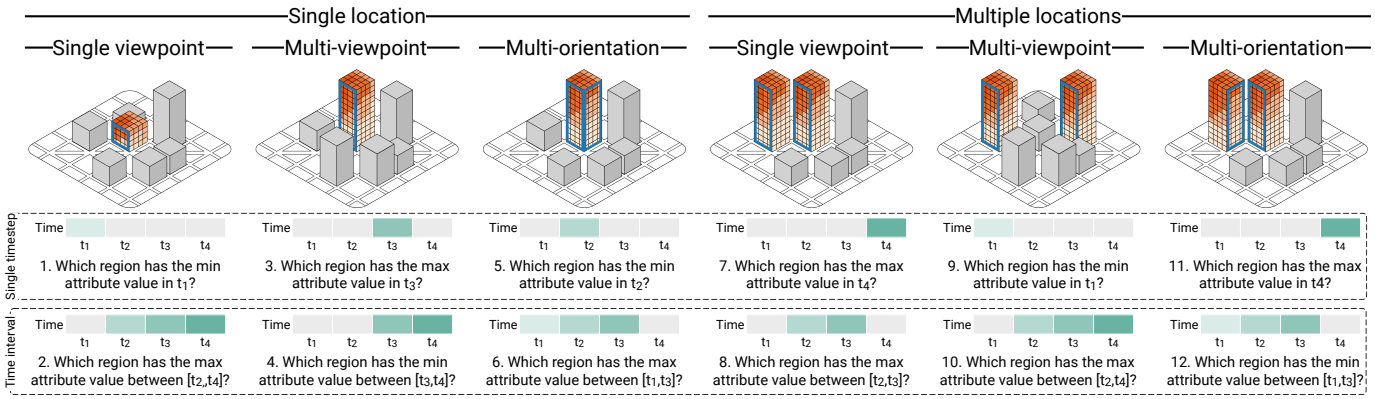

Fig. 5. Illustrative scheme of the twelve tasks of our study, with varying levels of spatial (*where*) and temporal (*when*) granularities. In each cell, the image illustrates the task scenario with highlighted facade(s), together with a sample question (*what*).

linked triad of *what*, *where*, and *when* [61]. Each task refers to a question in the format: *when* + *where* → *what*, where *what* is the participant's characterization of extreme attribute values. We varied the *when* and *where* dimensions in a way similar to other research [35, 60, 70], using varying spatial and temporal granularity levels. The temporal granularity *when* was divided into **1 timestep** and **time interval** (two or three consecutive time steps). The spatial granularity *where* was segmented into **1 location** (one building), and **N locations** (two buildings). The spatial granularity was further broken down into:

1. **Single** and **multi-viewpoint**: **Single-viewpoint** tasks target a single facade in **1 location**, and two facades in nearby buildings in **N locations**. **Multi-viewpoint** tasks target a facade partially occluded by another building in **1 location**, and two facades in fairly distant buildings in **N locations**. For **single-viewpoint** tasks, users can look at the target highlighted facade(s) from a single point of view. For **multi-viewpoint** tasks, users cannot capture the entire target facade(s) from a single viewpoint; thus, users must navigate the camera and continuously switch spatial context while searching for a facade region with extreme temporal value.

2. **Single** and **multi-orientation**: **Single orientation** tasks target two **single-viewpoint** facades facing the same direction. **Multi-orientation** tasks target two **single-viewpoint** facades whose surfaces are pairwise orthogonal. Note that these conditions occur in both **1 location** and **N locations**.

Crossing the spatial and temporal dimensions resulted in a matrix of 12 possible spatiotemporal tasks illustrated in Figure 5, along with a sample task question for each one of them.

4.4 Task conditions and complexity

Although complexity may come in many forms, in accordance with prior research [32], four conditions that drive the difficulty of our visualization tasks are: **number of time instances**, **number of spatial instances**, **camera navigation**, and **view angle**. The idea is that an increase in complexity will increase the number of spatial and temporal searches or camera movements that users must perform in order to find the answer.

While being unequivocal that increasing either the **number of time instances** or the **number of spatial instances** leads to an increase in complexity due to the extra searches an expert will be required to perform, **camera navigation** and the selection of a proper **view angle**, in particular, also increase complexity and lead to loss of performance. These depend on implicitly solving the inverse problem of where to direct our view, given a mental image of the rendering we *expect* to see. This is an inherently ill-posed problem as the expectation inferred from the rendering might not properly correlate with the 3D geometry because of common rendering artifacts (such as distortions caused by the camera projection, occlusion, etc.). This is typically, and intuitively, solved by trial and error: camera movements are tuned iteratively until adequate choices are found. As a concrete example, consider the case where data is presented in multiple locations. Occlusion prevents users from setting a single camera view that would work for all the data, which requires frequent manipulations of the viewpoint each time a different location is inspected. It is in this way that **camera navigation** and **view angle** conditions impact user performance: due to continuous camera view adjustments and context switching, time and attention are not used towards the user's utmost goal that is to understand the data.

Table 2. Description of task conditions (*rows*). Each task condition represents the difference in time and error between the two groups: lower and higher complexity (*columns*).

	Lower	Higher
Number of time instances	1 timestep	Time interval
Number of spatial instances	1 location	N locations
Camera navigation	Single viewpoint	Multi-viewpoint
View angle	Single orientation	Multi-orientation

To assess the complexity impact on user performance, we derive four task conditions from the elementary tasks listed in Figure 5. For each condition, we select a subset of tasks and divide them into two groups: lower and higher complexity. Next, we compute the differences in time and error between the two groups (lower and higher complexity) to obtain the impact in efficiency for that task condition. Table 2 illustrates the selection and grouping of elementary tasks for each task condition.

5 Results

We now report and interpret findings for task **completion time** and **relative error** from a total of 384 trials. We refer the reader to the supplementary material for the data files. Our reporting methodology uses estimation techniques and reports sample means with confidence intervals (CI) rather than *p*-value statistics [20, 26], following recommended practices [7, 19]. However, a *p*-value approach can always be obtained from the data by following similar techniques. For our inferential analysis, we use pairwise differences between means and their 95% CIs [1], indicating the range of plausible values for the population mean. We use BCa bootstrapping to construct the confidence intervals.

We first provide an overview of the results across the tasks to evaluate the efficiency of all four visualizations (Section 5.1). Afterwards, we present results for varied task conditions to investigate efficiency with increasing task complexity (Section 5.2).

5.1 Tasks

In this section, we investigate the question: "is there a difference in efficiency (time, error) between the four visual designs?". Figure 6 (top) displays mean times per visualization and Figure 7 (top) shows pairwise time comparisons for all tasks collectively. Participants spent 45.64 seconds on average (CI = [42.93, 48.44]) using `SpatialJX`, and slightly longer using `TemporalJX` (50.06s, [47.33, 52.85]). There is a pronounced increase in task duration in the results of both `LinkedV` (104.76s, [96.36, 113.28]) and `EmbeddedV` (104.97s, [95.98, 114.05]). Pairwise comparisons show strong evidence that plot-based visualizations were slower than color-coded visualizations (by 57.03s on average,

[1] A CI of *differences* that does not cross zero provides evidence of differences. The further away from zero and the smaller the CI, the stronger the evidence.

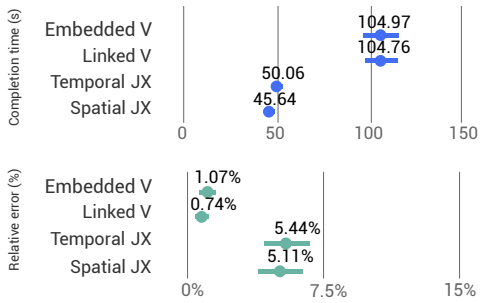

Fig. 6. Mean completion time (s) and relative error (%) for each visualization, for all tasks. Error bars represent 95% Bootstrap CIs.

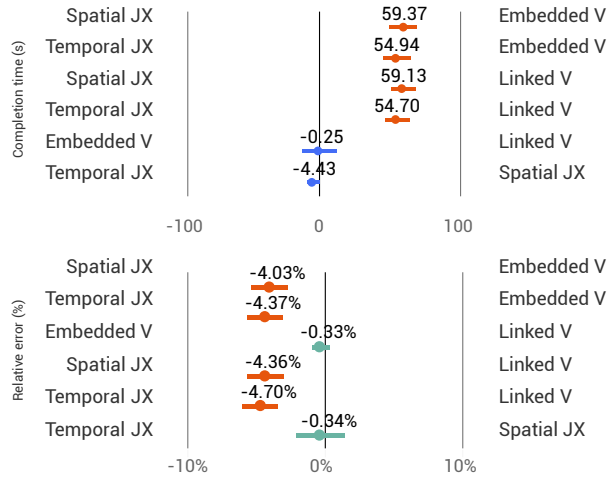

Fig. 7. Pairwise comparisons between visualizations for completion time (s) and relative error (%), for all tasks. Error bars represent 95% Bootstrap CIs, with evidence of differences in orange.[1]

[47.78, 66.15]). However, there is no clear evidence that completion times were different within each visualization group—*i.e.*, `LinkedV` *vs.* `EmbeddedV`, and `SpatialJX` *vs.* `TemporalJX`.

Figure 6 (bottom) displays mean errors and Figure 7 (bottom) shows mean error differences for all tasks collectively. Participants' average error was 1.07% ([0.62%, 1.57%]) when using `EmbeddedV`, and 0.74% ([0.38%, 1.15%]) with `LinkedV`. When using color-coded juxtaposed visualizations, the average error increased to 5.11% ([3.89%, 6.37%]) with `SpatialJX` and to 5.44% ([4.21%, 6.72%]) with `TemporalJX`. In this case, there are significant differences between plot-based and color-coded visualizations: the former is 4.36% on average ([3.06%, 5.69%]) more accurate than the latter. However, there is no clear difference within each visualization group.

5.2 Task conditions and complexity

In addition to differences in efficiency between the visualizations, we were also interested in investigating if their efficiency is affected with increasing task complexity, *i.e.*, "for each visual design, is there a difference in efficiency (time, error) with an increasing task complexity?". This sections reports results of efficiency impact based on the four task complexity conditions discussed in Section 4.4—namely, **number of time instances**, **number of spatial instances**, **camera navigation**, and **view angle**.

5.2.1 Completion time

Figure 8 (top) shows results across the task conditions and design. For the `TemporalJX` visualization, the **number of time instances** had the most impact on increasing completion time (15s, [10.4, 19.77]), followed by **camera navigation** (6.31s, [2.28, 15.3]) and **number of spatial instances** (6.16s, [0.81 , 11.62]). However, comparison within the **view angle** condition shows no clear increase in completion time. For `SpatialJX`, the **number of spatial instances** and **camera navigation** had the most impact on task completion time: 12.65s ([7.65, 17.64]) and 11.54s ([4.64, 18.76]), respectively; this dropped for **view angle** (7.29s, [2.22, 12.35]) and **number of time instances** (5.86s, [0.47, 11.39]). For `LinkedV`, the **number of spatial instances** and **camera navigation** had a significant impact on increasing task duration: 48.09s ([33.96, 61.59]) and 49.83s ([30.87, 68.35]), respectively. This dropped to 33.12s ([13.85, 51.68]) in the **view angle** condition, but no significant time difference was found with an increasing **number of time instances**. For `EmbeddedV`, similar to `LinkedV`, the **number of spatial instances** and **camera navigation** had a strong impact on increasing task duration: 51.9s ([36.88, 66.6]) and 61.16s ([42.87, 78.96]). This dropped to 43.2s ([23.85, 62.06]) in the **view angle** case. There is no evidence that increasing the **number of time instances** led to an increase in completion time.

5.2.2 Relative error

Figure 8 (bottom) shows the results on task error increase across the four task conditions.For `TemporalJX`, **number of time instances** had the most significant impact on task error increase (4.52%, [2.2%, 6.82%]), followed by **number of spatial instances** (3.25%, [0.82%, 5.66%]) and **camera navigation** (3.15%, [0.12%, 6.14%]). For `SpatialJX`, **camera navigation** had the most significant impact on error (4.61% increase, [1.44%, 7.79%]), and dropped to 2.55% ([0.06%, 4.96%]) with an increasing **number of spatial instances**. When considering `LinkedV`, none of the four task conditions had a significant impact on task error increase; and for `EmbeddedV`, **camera navigation** was the only condition that significantly impacted task error (1.28%, [0.28%, 2.4%]).

6 Discussion

In addition to examining the quantitative differences in performance between the four visualization designs, we also observed participants' behaviors and strategies when using each visualization. These observations, along with insights from participants' questionnaires, allowed us to more comprehensively characterize how they used each visualization.

6.1 Tasks

For the *elementary* tasks, we saw clear differences between plot-based and color-coded visualizations, in terms of both accuracy and task completion time—see Figure 7. Participants were considerably more accurate on average when using plot-based visualizations than when using color-coded juxtaposed visualizations. This tendency is consistent with participants' perceived difficulty in conducting the tasks. It was stated often in the post-study questionnaire that it is hard to perceive minor variability in the data through color encoding.

Takeaway #1: *Plots produce the most accurate results, and there is no clear winner between linked and embedded visualizations.*

However, the accuracy gain seen in the plot-based designs may have come at the expense of a substantial decrease in overall speed. Participants were considerably slower, possibly due to three factors: 1) plot-based visualizations required users to identify the target attribute values by scanning regions across the entire spatial domain, 2) since plots allowed participants to extract more subtle differences in the data, they likely spent more time on the task verifying their choices, and 3) occlusion—while most highlighted facades on the 3D city model were directly accessible, we observed that partially-occluded areas could make interactions that require direct access difficult (*e.g.*, mouse clicking and dragging regions). As a result, participants in our study often needed to adjust viewing angles (sometimes repeteadly) to reach a certain part of the building, drag the region, and see the visual plot.

Takeaway #2: *Plot-based visualizations require spatial scanning, which makes them slower to identify, verify, and access target data values compared to color-coded visualizations.*

6.2 Task conditions and complexity

6.2.1 Completion time

For the task conditions, we observed that task completion time increases substantially given an increasing **number of spatial instances**, **camera navigation**, and diverging **view angles** when using plot-based visualizations, possibly because of the three factors previously stated. However, we found no evidence that completion time increases with a higher **number of time instances**. This is likely because plot-based visualizations make it easy to see trends over time: collates all time steps across the individual bar chart; therefore, the cognitive load (time) did not significantly

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3209474

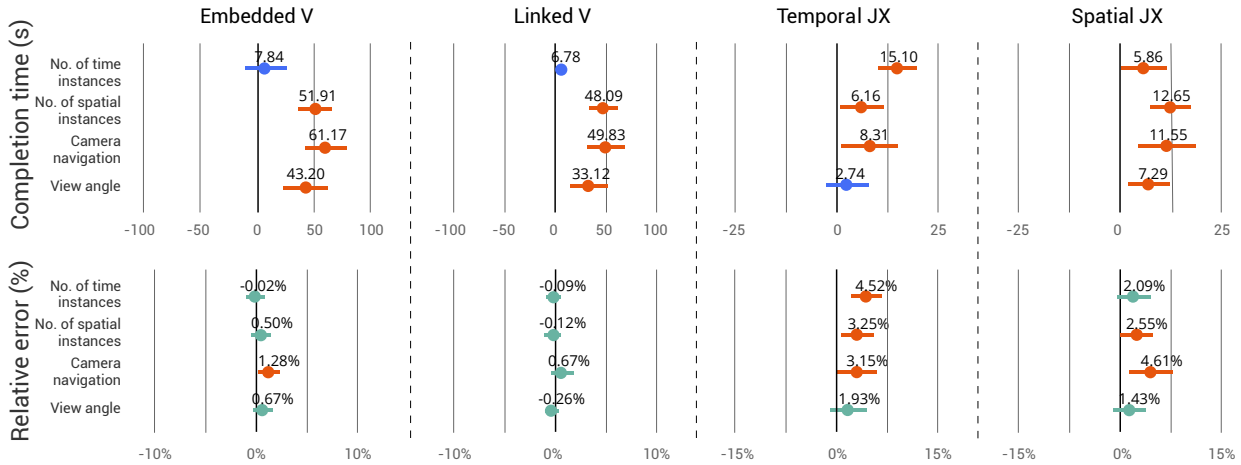

Fig. 8. Increase in completion time (*s*) and relative error (%). Rows: task conditions. Columns: visualization designs. Note that plots may present different time/error scales. Error bars represent 95% Bootstrap CIs, with evidence of differences in orange.[1]

Number of time instances	→	LinkedV, EmbeddedV
Number of time instances	→	LinkedV, EmbeddedV, SpatialJX
Number of spatial instances	→	LinkedV, EmbeddedV
Camera navigation	→	LinkedV
View angle	→	LinkedV, EmbeddedV, SpatialJX, TemporalJX

Fig. 9. The designs that scale best in completion time and error with increasing condition complexity (*i.e.*, no evidence of increase in time/error).

add up. The same was not true for spatially and temporally juxtaposed visualizations, where participants spent longer with an increasing **number of time instances** likely because participants had a fair amount of difficulty to collate information across different spatial and temporal contexts, as we observed. Although spatially-juxtaposed visualization displays all data instances at the same time, collating information across somewhat distant screen areas seems to affect cognitive load (time).

Takeaway #3: *Plot-based visualizations scale better in completion time given an increasing number of time instances as they collate all information needed at a specific location.*

We also observed that the time increase for plot-based visualizations tends to be more than twice the increase for the color-coded ones for task conditions that cover a wider or more convoluted spatial domain—*e.g.*, the **number of spatial instances** condition considers a higher number of building facades, and the **camera navigation** condition deals with partially-occluded facades. To illustrate, with an increasing **number of spatial instances**, the time increase for `SpatialJX` was 12s against 48s for `LinkedV`. This tendency may reflect our prior finding that spatial scanning is a bottleneck for task duration in plot-based visualizations: it requires not only sequential scan but also physical accessibility to partially-occluded surface areas (***Takeaway #2***).

The results of our study indicate that participants spent substantially longer when using plot-based visualizations; however, these designs scale well when analysis grows difficult with an increasing **number of time instances**, likely because the plots collate all the information needed to identify temporal tendencies. Given these findings, Figure 9 (top) shows recommendations of visual designs that were most effective in task duration with increasing task condition complexity.

6.2.2 Relative error

`LinkedV` did not have significant error increase for *any* of the four task conditions. A similar pattern was observed in `EmbeddedV` as little evidence of error increase was found only in the **camera navigation** condition—and we believe that the required context switches between the different building facades was the dominant causal factor. Hence, the plot representation seems to sustain accuracy irrespective of the number of locations and timestamps taken into account.

Takeaway #4: *Plot-based visualizations scale better in accuracy given an increasing number of spatial instances, temporal instances, or camera navigation. Yet, continuous context switching may be a concern for embedded visualizations.*

One interesting finding from this study is the inconclusive evidence of task error increase with varying **view angle** for all four visualizations. This is particularly intriguing as `EmbeddedV` raises concerns due to distortion of visual information partly caused by perspective. The lack of observed differences may be due to low statistical power; but we believe it is more likely due to the referencing lines added to the plot design. Participants' comments indicate that they often used the lines, as well as the empty space between the top of the bar and the line, to make comparisons within and between individual plots—see Figure 3. Yet, since aliasing artifacts appear on the reference lines with increasing camera distance, we observed that participants had to move close enough to the building model in order to make them legible.

Takeaway #5: *Reference lines can serve as a visual aid in plot-based visualizations to facilitate comparison, at least from a legible distance.*

The results of our study indicate that plot-based visualizations (`EmbeddedV` and `LinkedV`) can not only enhance the legibility of spatiotemporal data on 3D city models, but also scale well when analysis grows more complex. Based on our observations, Figure 9 (bottom) suggests the visual designs that were most effective in task error with increasing task condition complexity. Lastly, our design recommendations depicted in Figure 9 demonstrate that, considering increases in complexity in task conditions, plot-based visualizations more often preserve efficiency (time, accuracy) compared to color-coded ones.

7 Practitioners' Perspective

In addition to examining quantitative differences in performance between the visual designs, we conducted interviews with four practitioners from the urban domain; two were the same who advised us on task requirements for spatiotemporal analysis in 3D urban visualizations.

Procedure. Each session was divided into three parts. First, we started with a description of the motivation and goal of our study. Second, we performed four tutorials; these were the same tutorials previously used in our study with the purpose of introducing the analysis tasks and each of the four visualizations. Throughout the tutorials, participants were encouraged to pose questions and asked to think aloud to express reasoning. The researcher guided the trials, and adopted an inquiry-based learning approach throughout the tutorial by posing questions to the participants—*e.g.*, "*can you tell me a location to place a region?*" and "*can you indicate which region has the highest average attribute value in t_1 and t_2?*", and relying on them to assess the underlying task analysis and recommend courses of action to complete each trial. The third part of the session consisted of a semi-structured interview to collect participants' impressions as well as suggestions for further work on the visual designs. Sessions lasted 45 to 60 minutes and were video recorded. Comments were transcribed, and an open coding approach was applied. Quotes have been lightly edited to remove extraneous or repeated words, without changing the semantics of the sentence.

Results. Overall, all participants agreed on the usefulness of the visualizations to support analysis of 3D time-varying urban data. Although

color encoding was acknowledged to be most familiar, three of the four participants reported a preference for the plot representation. They highlighted the fact that plots can inform various features of the temporal function (*e.g.*, local/global extrema, increasing/decreasing intervals) that would otherwise be difficult to extract from colors (P1, P2). They also argued that they were more confident when comparing variations between values using plots than they were with colors (P1, P2, P4). One participant reported a preference for a hybrid visualization as "*colors could be used for rapid screening, and plots could be used to get and compare regions based on more precise data*" (P3). Some participants remarked that embedded views minimize the seam between the plots and their spatial references in a way that reduces mental workload as "*you no longer need to associate spatial regions with plots (...) you're (spatially) manipulating the plots themselves*" (P4). Others remarked that linked views allow plots to be always "accessible" (P2) and "adjacent" (P1) to one another, even when comparing distant regions.

We also asked participants to provide suggestions for further work to build on the current designs. Participants suggested investigating plot designs. An example are designs to aid comparison between regions: "stacked" (P1) or "multiseries" (P3) bar charts, or even charts that encode operations between data points from two or more regions—*e.g.*, "*the difference in the solar irradiance from candidate areas to install solar panels*" (P3). The comparative designs could "*be particularly useful for comparing data from dispersed regions using e.g., embedded views*" (P1). Another example is multivariate designs to aid the co-analysis of damages incurred to both outer and inner layers of the building—"*it could be very interesting to see, e.g., in a story, not only the damage (e.g., stress, displacement, fatalities, cost of repair) of the facade but also the damage profile of internal areas (e.g., partition walls) (...) maybe by projecting it to the facade itself*" (P1).

Participants also suggested considering plot interactions. An example is filtering as "*time responses of structural damages tend to vary significantly along the building facade due to vertical earthquake motions (...) apply filters on the plots to specify separate time periods makes sense for assessing damage severity e.g., at a story by story level*" (P3). Other examples include hovering, reordering, and even orienting plots since "*for vertical earthquake motions, comparing regions' responses at different story levels (..) 'rotating' the bar charts and (reference) lines would likely make it easier comparing them*" (P3). Lastly, participants suggested other region shapes to cover areas considered during damage analysis: story, multi-story, and corners (P1, P2, P3).

8 Conclusion, Limitations, and Future Work

In this paper, we first presented a domain task characterization as a series of abstract tasks that benefit from 3D spatiotemporal urban visualization by conducting interviews with experts from urban planning, architecture, and civil engineering; we hope that this serves as a resource for the visualization community to better understand the growing area of 3D urban analytics. Afterwards, we selected a subset of these tasks and presented a quantitative study comparing user performance of four representative visual designs used for the analysis of time-varying 3D urban data. Our findings indicate that participants were more accurate using plot-based visualizations but faster using color-coded visualizations. The plot-based visualizations improve accuracy for analytical tasks as participants tend to perceive higher discriminability in the data with the *position on a common scale* encoding compared to *color*—interestingly, this result is aligned with perceptual accuracy rankings previously found in controlled experiments for 2D information visualizations [14, 56]. Our findings also reveal that plot-based visualizations more often preserve user performance compared to color-coded visualizations with increasing task complexity. Based on our findings, we were able to derive a series of takeaways that offer practical design guidelines for visualization researchers and practitioners. Lastly, we presented a qualitative evaluation with four practitioners from the 3D urban domain. The interviews confirmed the overall usefulness of the visualizations for 3D spatiotemporal urban data analysis, and revealed practitioners' impressions and suggestions for further work.

Limitations and future work. We acknowledge some limitations of our study. Although our quantitative study included participants with considerable experience with 3D visualizations, none of them were practitioners working with 3D urban visualizations in a professional setting. We also acknowledge that the findings of our study may be limited by a relatively small sample of participants—mostly graduate students. We feel that general trends found in our study would still apply across domain practitioners as they are familiar with 3D visualizations, similarly to the recruited participants. However, further work is needed to verify how our study results alter with a large sample and how well the visualizations fare when used by domain practitioners. Additionally, we recruited participants who might be more familiar with some of the techniques (*e.g.*, temporal juxtaposition) than others (*e.g.*, embedded view). While we believe the general trends still apply, it is possible that this effect could bias the results in favor of some techniques.

We also acknowledge the limitations of our visual designs in tackling a key challenge encountered in 3D urban spaces: occlusion. We carefully controlled the amount of occlusion between buildings in our trials as we primarily aimed at evaluating the influence of space and time at different granularities on users' ability to interpret attribute values with different visualizations. Hence, we refrained from adding too many factors to an already complex experimental design. Our discussion section provides thoughts on how (partial) occlusion impacts performance, but further work is needed and could consider a combination of our designs with de-occlusion designs such as the ones presented by Elmqvist *et al.* [27]. Another limitation is that we used a single 3D city model in the study, which reflects only a subset of possible geographical configurations. Additional studies may be necessary to measure participants' performance in a more diverse range of model types. While plot-based visualizations performed well for a medium-density model with partial occlusion, they may provide less of a benefit for a high-density model with strong occlusion between buildings. Our city model also did not include buildings with facade textures that portray structural information. Further work could consider blending thematic and structural data, and study to what degree users perceive each information individually. Furthermore, our city model only included buildings with regular geometry. It is possible that irregular geometries could affect the perception of the representation of the thematic variable in our visualizations. Although we believe general trends of our study findings would persist in building models with less pronounced deformations (*e.g.*, facades with smooth, wide curves), further work is needed to consider cases with acute geometric irregularity. In such cases, irregularities may cause undesired distortions on the plot representation of an embedded visualization that conforms to the surface geometry like ours; therefore, further work could investigate how well the visualization fares with increasing levels of irregularity and consider redesigning it for legibility purposes, if needed. Some redesign ideas could build on prior work from GIS research on varying the level-of-detail of building models [10], or consider using proxy surfaces as simplified, smoother approximations of a given polygonal model [11, 66].

Finally, although we covered a spectrum of analytical tasks elicited in our task characterization with domain experts, our study was also limited to a few tasks—**spatial search** for a **facade region** with **extreme** attribute value over a time period (a **time step** or a **time interval**). It is possible that other task settings could have yielded different study results and may be worth further investigating—*e.g.*, in **temporal comparison** tasks, where target locations are already known by users, it is very likely that time performance differences between plot-based and color-coded visualizations would be less pronounced. Additionally, due to the study feasibility purposes, we limited our tasks to deal with only a few spatial and temporal instances. Although our results provide insights on how well visualizations scale in such cases, future work could concentrate on perceptual scalability and investigate finer granularity levels.

Acknowledgements

We would like to thank the anonymous reviewers for their constructive comments and feedback. We acknowledge the support of the CGS-D Program [CGSD3-534866-2019] and Discovery grants [RGPIN/05320-2019, RGPIN/05303-2019] from NSERC, CNPq [316963/2021-6], and FAPERJ [E-26/202.915/2019, E-26/211.134/2019].

REFERENCES

[1] O. Ahmadien, H. F. Ates, T. Baykas, and B. K. Gunturk. Predicting path loss distribution of an area from satellite images using deep learning. *IEEE Access*, 8:64982–64991, 2020.

[2] H. Akiba, C. Wang, and K.-L. Ma. Aniviz: A template-based animation tool for volume visualization. *IEEE Computer Graphics and Applications*, 30(5):61–71, 2009.

[3] O. S. Alabi, X. Wu, J. M. Harter, M. Phadke, L. Pinto, H. Petersen, S. Bass, M. Keifer, S. Zhong, C. Healey, et al. Comparative visualization of ensembles using ensemble surface slicing. 8294:318–329, 2012.

[4] G. Andrienko and N. Andrienko. Spatio-temporal aggregation for visual analysis of movements. In *IEEE Symposium on Visual Analytics Science and Technology*, pp. 51–58, 2008.

[5] 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.

[6] N. Andrienko and G. Andrienko. Interactive visual tools to explore spatio-temporal variation. In *International Conference on Advanced Visual Interfaces*, pp. 417–420, 2004.

[7] A. P. Association et al. *Publication Manual of the American Psychological Association*. American Psychological Association, 2019.

[8] J. Barr and J. Johnson. Skyscrapers and the happiness of cities. *Eastern economic journal*, 46(2):344–377, 2020.

[9] J. Barr, B. Mizrach, and K. Mundra. Skyscraper height and the business cycle: Separating myth from reality. *Applied Economics*, 47(2):148–160, 2015.

[10] F. Biljecki, H. Ledoux, J. Stoter, and J. Zhao. Formalisation of the level of detail in 3D city modelling. *Computers, Environment and Urban Systems*, 48:1–15, 2014.

[11] S. Calderon and T. Boubekeur. Bounding proxies for shape approximation. *ACM Transactions on Graphics*, 36(4):57–1, 2017.

[12] J. Choo and H. Park. Customizing computational methods for visual analytics with big data. *IEEE Computer Graphics and Applications*, 33(4):22–28, 2013.

[13] A. Christodoulou, M. Vola, and G. Rikken. Case study for the application of multidisciplinary computational design assessment and constructability optimisation tools. In *Symposium on Simulation for Architecture and Urban Design*, pp. 1–8, 2018.

[14] W. S. Cleveland and R. McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387):531–554, 1984.

[15] D. Coffey, C.-L. Lin, A. G. Erdman, and D. F. Keefe. Design by dragging: An interface for creative forward and inverse design with simulation ensembles. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2783–2791, 2013.

[16] D. Cornel, A. Buttinger-Kreuzhuber, A. Konev, Z. Horváth, M. Wimmer, R. Heidrich, and J. Waser. Interactive visualization of flood and heavy rain simulations. *Computer Graphics Forum*, 38(3):25–39, 2019.

[17] D. Cornel, A. Konev, B. Sadransky, Z. Horváth, A. Brambilla, I. Viola, and J. Waser. Composite flow maps. *Computer Graphics Forum*, 35(3):461–470, 2016.

[18] D. Cornel, A. Konev, B. Sadransky, Z. Horváth, E. Gröller, and J. Waser. Visualization of object-centered vulnerability to possible flood hazards. *Computer Graphics Forum*, 34(3):331–340, 2015.

[19] G. Cumming. The new statistics: Why and how. *Psychological Science*, 25(1):7–29, 2014.

[20] G. Cumming and S. Finch. Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist*, 60(2):170, 2005.

[21] Z. Deng, D. Weng, J. Chen, R. Liu, Z. Wang, J. Bao, Y. Zheng, and Y. Wu. Airvis: Visual analytics of air pollution propagation. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):800–810, 2019.

[22] T. Dogan, S. Samaranayake, and N. Saraf. Urbano: A new tool to promote mobility-aware urban design, active transportation modeling and access analysis for amenities and public transport. In *Symposium on Simulation for Architecture and Urban Design*, pp. 1–8, 2018.

[23] H. Doraiswamy, N. Ferreira, M. Lage, H. Vo, L. Wilson, H. Werner, M. Park, and C. Silva. Topology-based catalogue exploration framework for identifying view-enhanced tower designs. *ACM Transactions on Graphics*, 34(6):1–13, 2015.

[24] H. Doraiswamy, J. Freire, M. Lage, F. Miranda, and C. Silva. Spatio-temporal urban data analysis: A visual analytics perspective. *IEEE Computer Graphics and Applications*, 38(5):26–35, 2018.

[25] H. Doraiswamy, E. Tzirita Zacharatou, F. Miranda, M. Lage, A. Ailamaki, C. T. Silva, and J. Freire. Interactive visual exploration of spatio-temporal urban data sets using urbane. In *International Conference on Management of Data*, pp. 1693–1696, 2018.

[26] P. Dragicevic. Fair statistical communication in HCI. In *Modern Statistical Methods for HCI*, pp. 291–330. 2016.

[27] N. Elmqvist and P. Tsigas. A taxonomy of 3D occlusion management for visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(5):1095–1109, 2008.

[28] N. Ferreira, M. Lage, H. Doraiswamy, H. Vo, L. Wilson, H. Werner, M. Park, and C. Silva. Urbane: A 3D framework to support data driven decision making in urban development. In *IEEE Conference on Visual Analytics Science and Technology*, pp. 97–104, 2015.

[29] N. Ferreira, J. Poco, H. T. Vo, J. Freire, and C. T. Silva. Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2149–2158, 2013.

[30] S. Frey, F. Sadlo, and T. Ertl. Visualization of temporal similarity in field data. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2023–2032, 2012.

[31] J. Gautier, M. Brédif, and S. Christophe. Co-visualization of air temperature and urban data for visual exploration. In *2020 IEEE Visualization Conference*, pp. 71–75, 2020.

[32] M. Gleicher. Considerations for visualizing comparison. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):413–423, 2017.

[33] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts. Visual comparison for information visualization. *Information Visualization*, 10(4):289–309, 2011.

[34] R. Goldstein, S. Breslav, K. Walmsley, and A. Khan. Spaceanalysis: A tool for pathfinding, visibility, and acoustics analyses in generative design workflows. In *Symposium on Simulation for Architecture and Urban Design*, pp. 1–8, 2020.

[35] S. Goodwin, J. Dykes, A. Slingsby, and C. Turkay. Visualizing multiple variables across scale and geography. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):599–608, 2015.

[36] L. Hao, C. G. Healey, and S. A. Bass. Effective visualization of temporal ensembles. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):787–796, 2015.

[37] J. Hraška. Approaches, methods and tools of rights of access to sunlight around the world. *Slovak Journal of Civil Engineering*, 27:45–52, 2019.

[38] IPCC. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2021. In press.

[39] S. Johnson, D. Orban, H. B. Runesha, L. Meng, B. Juhnke, A. Erdman, F. Samsel, and D. F. Keefe. Bento box: An interactive and zoomable small multiples technique for visualizing 4D simulation ensembles in virtual reality. *Frontiers in Robotics and AI*, 6:61, 2019.

[40] A. Karduni, I. Cho, G. Wessel, W. Ribarsky, E. Sauda, and W. Dou. Urban space explorer: A visual analytics system for urban planning. *IEEE Computer Graphics and Applications*, 37(5):50–60, 2017.

[41] S. Kaynak, B. Kaynak, and A. Özmen. A software tool development study for solar energy potential analysis. *Energy and Buildings*, 162:134–143, 2018.

[42] D. Keefe, M. Ewert, W. Ribarsky, and R. Chang. Interactive coordinated multiple-view visualization of biomechanical motion data. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1383–1390, 2009.

[43] C. E. Kilsedar and M. A. Brovelli. Multidimensional visualization and processing of big open urban geospatial data on the web. *ISPRS International Journal of Geo-Information*, 9(7):434, 2020.

[44] K. Kim, J. V. Carlis, and D. F. Keefe. Comparison techniques utilized in spatial 3D and 4D data visualizations: A survey and future directions. *Computers & Graphics*, 67:138–147, 2017.

[45] A. Konev, M. Matusich, I. Viola, H. Schulze, D. Cornel, and J. Waser. Fast cutaway visualization of sub-terrain tubular networks. *Computers & Graphics*, 75:25–35, 2018.

[46] A. Konev, J. Waser, B. Sadransky, D. Cornel, R. A. Perdigao, Z. Horváth, and M. E. Gröller. Run watchers: Automatic simulation-based decision support in flood management. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1873–1882, 2014.

[47] B. König, D. Limberger, J. Klimke, B. Hagedorn, and J. Döllner. RoomCanvas: A visualization system for spatiotemporal temperature data in

smart homes. In *Eurographics Conference on Visualization*, 2021.

[48] B. Krietemeyer and R. El Kontar. A method for integrating an UBEM with GIS for spatiotemporal visualization and analysis. In *Symposium on Simulation for Architecture and Urban Design*, pp. 87–94, 2019.

[49] X. Li and G. Wang. GPU parallel computing for mapping urban outdoor heat exposure. *Theoretical and Applied Climatology*, 145(3):1101–1111, 2021.

[50] H. Liu, Y. Gao, L. Lu, S. Liu, H. Qu, and L. M. Ni. Visual analysis of route diversity. In *IEEE Conference on Visual Analytics Science and Technology*, pp. 171–180, 2011.

[51] A. Lu and H.-W. Shen. Interactive storyboard for overall time-varying data visualization. In *IEEE Pacific Visualization Symposium*, pp. 143–150, 2008.

[52] H. Malm and P. Jenning. Cityfiction: Scenarios for densification. In *Symposium on Simulation for Architecture and Urban Design*, pp. 1–8, 2019.

[53] F. Miranda, H. Doraiswamy, M. Lage, L. Wilson, M. Hsieh, and C. T. Silva. Shadow accrual maps: Efficient accumulation of city-scale shadows over time. *IEEE Transactions on Visualization and Computer Graphics*, 25(3):1559–1574, 2019.

[54] F. Miranda, M. Hosseini, M. Lage, H. Doraiswamy, G. Dove, and C. T. Silva. Urban mosaic: Visual exploration of streetscapes using large-scale image data. In *ACM CHI Conference on Human Factors in Computing Systems*, p. 1–15, 2020.

[55] E. Morello and C. Ratti. Sunscapes: 'solar envelopes' and the analysis of urban dems. *Computers, Environment and Urban Systems*, 33(1):26–34, 2009.

[56] T. Munzner. *Visualization analysis and design*. CRC press, 2014.

[57] A. Noyman and K. Larson. A deep image of the city: Generative urban-design visualization. In *Symposium on Simulation for Architecture and Urban Design*, pp. 1–8, 2020.

[58] T. Ortner, J. Sorger, H. Steinlechner, G. Hesina, H. Piringer, and E. Gröller. Vis-A-Ware: Integrating spatial and non-spatial visualization for visibility-aware urban planning. *IEEE Transactions on Visualization and Computer Graphics*, 23(2):1139–1151, 2017.

[59] S. Pasewaldt, M. Trapp, and J. Döllner. Multi-perspective detail+ overview visualization for 3D building exploration. In *Theory and Practice of Computer Graphics*, pp. 57–64, 2013.

[60] V. Peña-Araya, E. Pietriga, and A. Bezerianos. A comparison of visualizations for identifying correlation over space and time. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):375–385, 2019.

[61] D. J. Peuquet. It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems. *Association of American Geographers*, 84(3):441–461, 1994.

[62] R. Ramyar, E. Zarghami, and M. Bryant. Spatio-temporal planning of urban neighborhoods in the context of global climate change: Lessons for urban form design in tehran, iran. *Sustainable Cities and Society*, 51:101554, 2019.

[63] H. Ribičić, J. Waser, R. Fuchs, G. Blöschl, and E. Gröller. Visual analysis and steering of flooding simulations. *IEEE Transactions on Visualization and Computer Graphics*, 19(6):1062–1075, 2012.

[64] H. Ribičić, J. Waser, R. Gurbat, B. Sadransky, and M. E. Gröller. Sketching uncertainty into simulations. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2255–2264, 2012.

[65] A. Rocha, U. Alim, J. D. Silva, and M. C. Sousa. Decal-maps: Real-time layering of decals on surfaces for multivariate visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):821–830, 2016.

[66] A. Rocha, J. D. Silva, U. R. Alim, S. Carpendale, and M. C. Sousa. Decal-lenses: Interactive lenses on surfaces for multivariate visualization. *IEEE Transactions on Visualization and Computer Graphics*, 25(8):2568–2582, 2018.

[67] S. Rockcastle, M. L. Ámundadóttir, and M. Andersen. Ocuvis: a web-based visualizer for simulated daylight performance. In *Symposium on Simulation for Architecture and Urban Design*, pp. 1–8, 2018.

[68] J. Sanyal, S. Zhang, J. Dyer, A. Mercer, P. Amburn, and R. Moorhead. Noodles: A tool for visualization of numerical weather model ensemble uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1421–1430, 2010.

[69] D. Schaumann, S. Moon, M. Usman, R. Goldstein, S. Breslav, A. Khan, P. Faloutsos, and M. Kapadia. Toward a multi-level and multi-paradigm platform for building occupant simulation. In *Symposium on Simulation for Architecture and Urban Design*, pp. 169–176, 2019.

[70] J. Schiewe. Task-oriented visualization approaches for landscape and urban change analysis. *International Journal of Geo-Information*, 7(8):288, 2018.

[71] A. Semmo and J. Döllner. Interactive image filtering for level-of-abstraction texturing of virtual 3D scenes. *Computers & Graphics*, 52:181–198, 2015.

[72] R. Sicat, J. Li, J. Choi, M. Cordeil, W.-K. Jeong, B. Bach, and H. Pfister. DXR: A toolkit for building immersive data visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):715–725, 2018.

[73] Side Walks Lab. Delve — A Real Estate Generative Design Product. `https://www.sidewalklabs.com/products/delve`. [Online; accessed 2022].

[74] J. Stoter, H. de Kluijver, and V. Kurakula. 3D noise mapping in urban areas. *International Journal of Geographical Information Science*, 22(8):907–924, 2008.

[75] G. Sun, R. Liang, H. Qu, and Y. Wu. Embedding spatio-temporal information into maps by route-zooming. *IEEE Transactions on Visualization and Computer Graphics*, 23(5):1506–1519, 2016.

[76] M. K. Tory, T. Möller, and M. S. Atkins. Visualization of time-varying MRI data for MS lesion analysis. 4319:590–598, 2001.

[77] M. Trapp, F. Schlegel, S. Pasewaldt, and J. Döllner. Rendering procedural textures for visualization of thematic data in 3D geovirtual environments. In *International Conference on Information Visualization Theory and Applications*, pp. 282–288, 2019.

[78] C. Turkay, A. Slingsby, H. Hauser, J. Wood, and J. Dykes. Attribute signatures: Dynamic visual summaries for analyzing multivariate geographical data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2033–2042, 2014.

[79] R. Van Pelt, R. Gasteiger, K. Lawonn, M. Meuschke, and B. Preim. Comparative blood flow visualization for cerebral aneurysm treatment assessment. *Computer Graphics Forum*, 33(3):131–140, 2014.

[80] M. Vuckovic, J. Schmidt, T. Ortner, and D. Cornel. Combining 2D and 3D visualization with visual analytics in the environmental domain. *Information*, 13(1):7, 2021.

[81] C. Ware. *Information visualization: perception for design*. Morgan Kaufmann, 2019.

[82] J. Waser, A. Konev, B. Sadransky, Z. Horváth, H. Ribičić, R. Carnecky, P. Kluding, and B. Schindler. Many plans: Multidimensional ensembles for visual decision support in flood management. *Computer Graphics Forum*, 33(3):281–290, 2014.

[83] J. Waser, H. Ribičić, R. Fuchs, C. Hirsch, B. Schindler, G. Blöschl, and E. Gröller. Nodes on Ropes: A Comprehensive Data and Control Flow for Steering Ensemble Simulations. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1872–1881, 2011.

[84] L. Wilson, J. Danforth, C. C. Davila, and D. Harvey. How to generate a thousand master plans: A framework for computational urban design. In *Symposium on Simulation for Architecture and Urban Design*, 2019.

[85] D. Wolosiuk and A. Mahdavi. Application of ontologically structured data for building performance analysis. In *Symposium on Simulation for Architecture and Urban Design*, pp. 1–6, 2020.

[86] J. Woodring and H.-W. Shen. Multi-variate, time varying, and comparative visualization with contextual cues. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):909–916, 2006.

[87] L. Xu, A. Francisco, J. E. Taylor, and N. Mohammadi. Urban energy data visualization and management: Evaluating community-scale eco-feedback approaches. *Journal of Management in Engineering*, 37(2):04020111, 2021.

[88] Y. Yang, T. Dwyer, S. Goodwin, and K. Marriott. Many-to-many geographically-embedded flow visualisation: An evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):411–420, 2016.

[89] W. Zeng and Y. Ye. Vitalvizor: A visual analytics system for studying urban vitality. *IEEE Computer Graphics and Applications*, 38(5):38–53, 2018.

[90] C. Zhang, T. Schultz, K. Lawonn, E. Eisemann, and A. Vilanova. Glyph-based comparative visualization for diffusion tensor fields. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):797–806, 2015.

[91] C. Zhang, W. Zeng, and L. Liu. UrbanVR: An immersive analytics system for context-aware urban design. *Computers & Graphics*, 99:128–138, 2021.

[92] Y. Zheng, W. Wu, Y. Chen, H. Qu, and L. M. Ni. Visual analytics in urban computing: An overview. *IEEE Transactions on Big Data*, 2(3):276–296, 2016.