Master Thesis in Geographical Information Science nr 139

Perception accuracy and user acceptance of legend designs for opacity data mapping in GIS

Jonas Tillman

2022
Department of
Physical Geography and Ecosystem Science
Centre for Geographical Information Systems
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Jonas Tillman (2022). **Perception accuracy and user acceptance of legend designs for opacity data mapping in GIS**

Master degree thesis, 30 credits in Master in Geographical Information Systems (GIS)
Department of Physical Geography and Ecosystem Science, Lund University

Perception accuracy and user acceptance of legend designs for opacity data mapping in GIS

Jonas Tillman
Master thesis, 30 credits, in Geographical Information Sciences

Supervisor: Micael Runnström PhD
GIS Centre
Dept of Physical Geography and Ecosystem Science
Lund University

Abstract

In a GIS system, the need to encode geospatial data without standardized cartographic representations, such as population data, weather information, etc., on top of a map is common. This can be done with a layer on top of a base map, with the effect that the base map is partially or fully hidden. Reducing the opacity (or its equivalent - increasing the transparency) of this overlay data layer is a frequently seen solution to show the geographic context. With the colours of the base map being combined with the overlay layer's colours, the resulting visualization can become difficult to interpret for the end-user.

To help the user decode the values encoded in maps, legends are a common tool for non-interactive maps and data visualizations. This user study investigates the decoding accuracy using a map without a legend, as well as 4 different legend designs for opacity-mapped data overlayed on a static base map. A secondary objective is to measure how helpful the users considered the different legend designs were to decode values.

Baseline categories for comparison were (i) no legend - only having the range of the data values being expressed in text and (ii) a legend design imitating the ArcGIS legend for opacity data mapping.

Three different legend designs were produced to introduce more contextualisation from the map background to the background of the legend and reducing the distance from the legend to the data. This was done using (iii) a sample of the map as background for the legend (iv) having the most common colours of the map base layer as legend background (v) attaching a legend directly to the edge of the overlay data area.

Using a web interface, the users were requested to visually estimate the value at the location of a marker within the overlay data area. In statistical analysis of the results, there was clear statistical effect in reduced errors when having a legend compared to when no legend was included. There was, however, no statistically significant difference in estimation/perception errors between the legend designs tested.

The acceptance of respondents - defined as how useful they considered the legend types were to help estimate the value - did have statistically higher estimates when sampling the map background (marked as iii when introduced in text above) and when attaching the legend to the data area (v) compared to the default ArcGIS design (i).

Keywords: GIS, Data Visualisation, Legends, Opacity, Transparency, Visual Data Encoding, Perception

Contents

List of Figures

List of Tables

List of Abbreviations

API Application Programming Interface

CSS Cascading Style Sheets

ESRI Environmental Systems Research Institute, Inc. **GIS** Geographic Information System

HTML5 Hypertext Markup Language, 5th major version **OLS** Ordinary Least Squares

RGBA Red, Green, Blue, Alpha

SVG Scalable Vector Graphics

UI User

1 Introduction

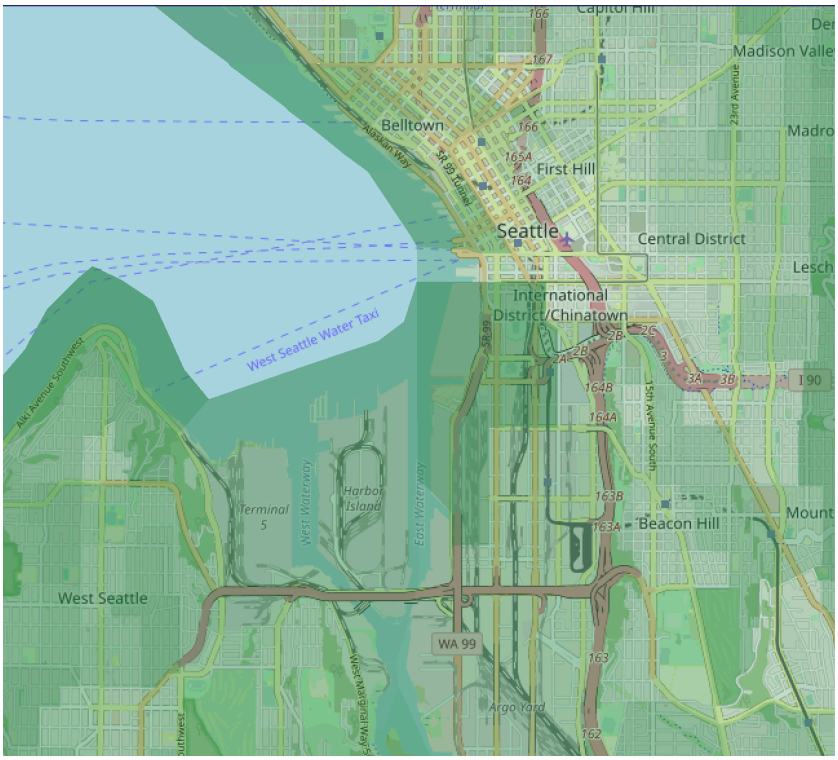
1.1 Opacity/transparency in GIS visualizations

Having a base map for geographic context and the need to visualize other geospatial data as an overlay layer on top of that geographic context results in the background being partly or fully concealed. To show part of the geographical context, a common method is to reduce the opacity of the overlay data layer. This is also sometimes referred to as increasing the inverse of opacity; transparency.

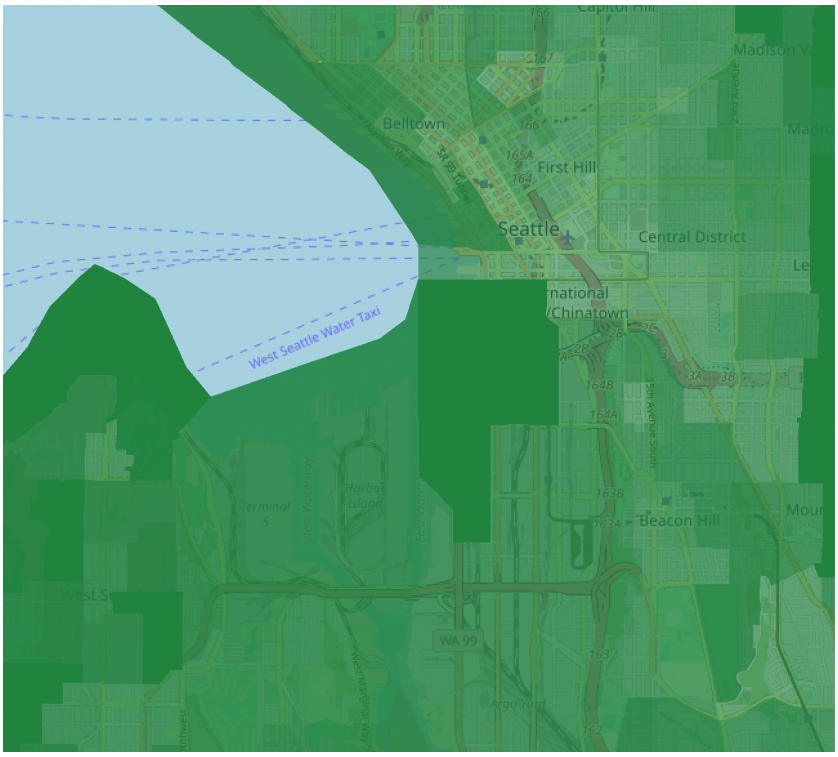
Ware (2020) highlights some issues of using low opacity layers in his seminal work on information visualization. In particular, he describes that colours and objects that the user perceives become composites of the merged layers:

> *In many visualization problems, it is desirable to present data in a layered form. This is especially common in geographic information systems (GISs). So that the contents of different layers are simultaneously visible, a useful technique is to present one layer of data transparently over another; however, there are many perceptual pitfalls in doing this. The contents of the different layers will always interfere with each other to some extent, and sometimes the two layers will fuse perceptually so that it is impossible to determine to which layer a given object belongs (Ware, 2020, p. 217).*

An illustrative example of how choosing lower opacity can show more of the base map, while higher opacity can make the geographic context difficult to see can be found in Figure 1. In this figure, the same data is mapped to 2 different ranges of opacity values. Figure 1a has the data mapped between opacity values of 0.2-0.6, while Figure 1b starts at 0.6 and ends at fully opaque (0.6-1.0). With the low opacity, there is more of the base layer colour influencing the colour that is displayed to the user, which in turn could make decoding the data more difficult.



(a) Opacity mapping 0.2-0.6



(b) Opacity mapping 0.6-1.0 (1.0=fully opaque)

Figure 1: Low and high opacity overlay examples

This difficult decoding task can be made easier by interaction in web-based GIS systems. In those systems, a user can interact with the visualization to get values from e.g. tooltips on hover of the mouse over parts of the map. For static visualizations and on touch-screen devices (such as smartphones or tablets) mouse-hovering is not possible. A legend is therefore often used both in interactive and static web-GIS visualizations.

This study aims to look at the efficiency of legend designs in visual opacity-to-value decoding tasks. The study also investigates which of 4 legend designs the participants of the study perceive as most helpful in the decoding task.

This is done using a web-based graphical user interface. With this interface as a data-gathering tool, a group of people with medium to high previous experience using map services were asked to estimate the value of a marker when shown examples that were using different legend designs.

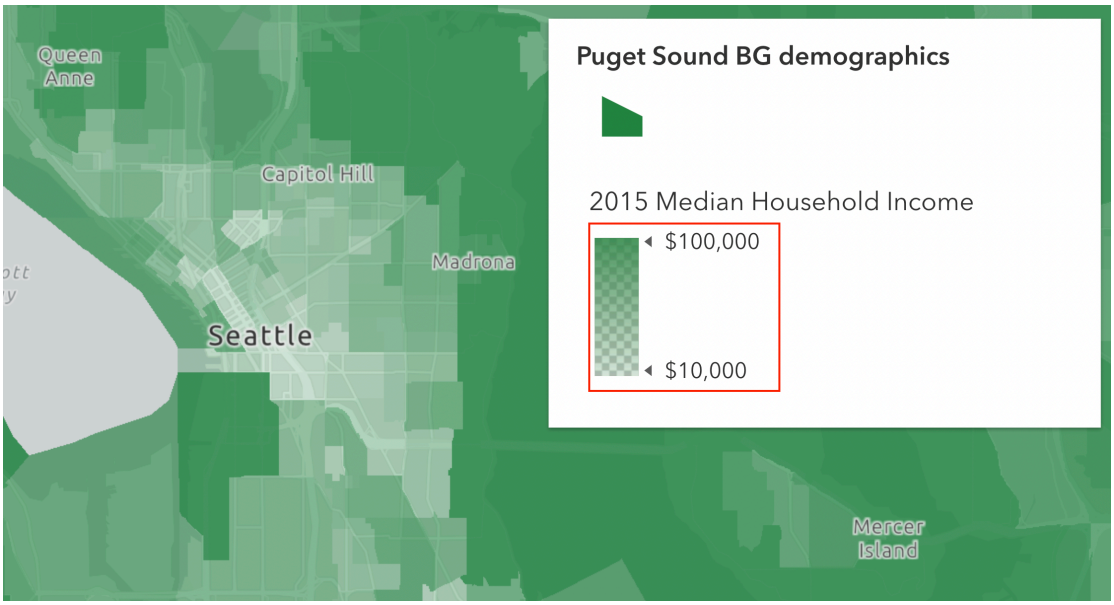
A user study was chosen due to being able to capture the responses of participants in a controlled environment. The user study is an umbrella term for the process of understanding the impact of design choices on an audience. This may have qualitative (attempts to measure the results through collecting quantitative data) and quantitative parts (e.g. interviews), but for this study both the estimation tasks and acceptance questions are handled through quantitative means. It would be possible to get more responses by publishing it on the Internet than to observe and measure user responses while having the author of the study present in the room. However, to collect clean data it was it was considered useful to ask the participants if the task was fully understood before starting the data decoding and user acceptance survey tasks. These decisions were made during a testing phase, where some testing subjects did not bother to read the instructions, and did not make use of the legends during the decoding task until after going through a few of the examples.

Additionally, with a smaller group of participants (34 in this study, in the age range between 25-50), you can also easily collect responses on the impressions of the legend designs. In comparison to more qualitative methodologies, such as in-depth interviews, you can get standardized and measurable results. The choice of a user study does come with some caveats in what population the chosen sample of users is expected to be representative of. For this study, having users with some familiarity with reading maps

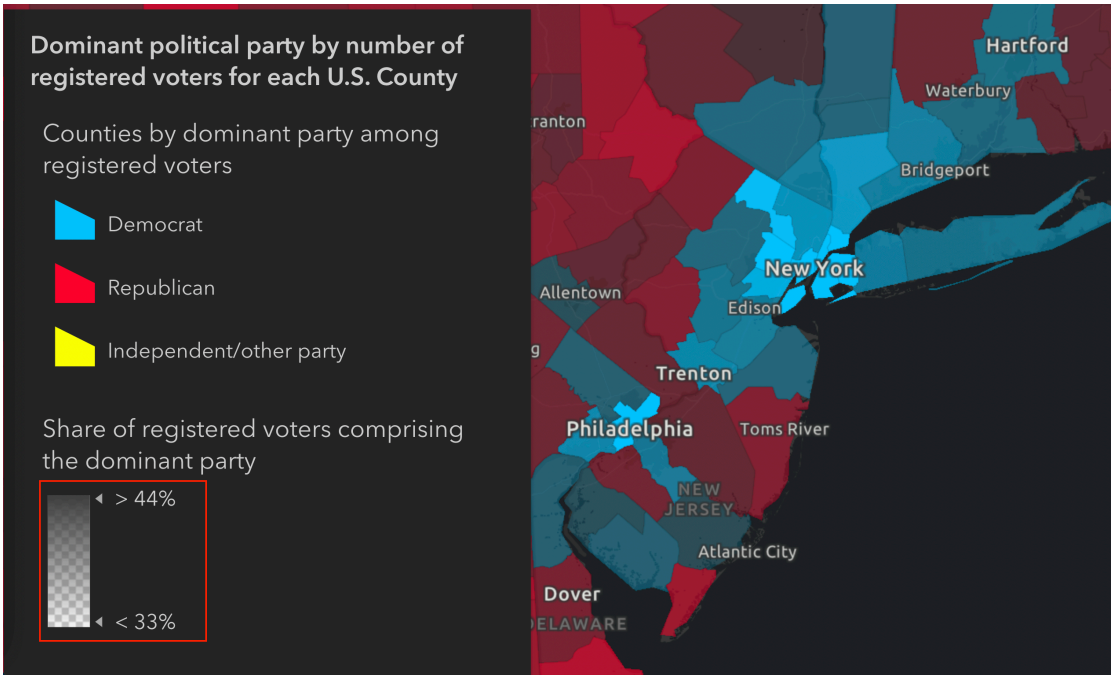
or at least being heavily exposed to maps through navigation applications were chosen to ensure that common map elements and representations were familiar.

The legend design of the enterprise GIS company Esri's "ArcGIS online" tool (ESRI, n.d.a, n.d.b) was together with the absence of a legend used as baselines to evaluate alternative legend designs against.

ESRI's "ArcGIS Online" examples (ArcGIS, n.d.a, n.d.b) using transparency are showing a legend that has a checkered background with white and mid-grey values (Figures 2a and 2b). This legend representation of transparency lacks any context of the visualization's base layer/map and is always displayed as a checkered mid-grey background under the opacity mapping.



(a) Opacity data mapping and legend with colour scale, highlighted with a red bounding box



(b) Opacity data mapping and legend, highlighted with a red bounding box

Figure 2: ArcGIS opacity legend and data mapping (ArcGIS, n.d.a., n.d.b)

1.2 Research questions

- **RQ1** - Do legends increase the accuracy of decoding opacity-mapped values over knowing only the range of the data being mapped (i.e. when no legend is included)?
- **RQ2** - How is decoding accuracy affected by legend designs that introduce more context from the base map or reduce the distance from legend to opacity-encoded data?
- **RQ3** - Do subjects perceive legend designs with more context from the base map as more helpful in decoding opacity mapped values?
- **RQ4** - How are decision times and user selection behaviour affected by legend designs?

1.3 Limitations

The study looks at static 2D maps with not chance for interaction to ensure only the legend could be used to decode the values. Most web-based maps these days have some mechanism for interaction.

The base map was kept constant and using only one colour scheme that was very close to the commonly used map cartographic representation of e.g. Google maps. The results may differ if having an opacity layer on top of simpler base map with less colours and details, or more details as when having satellite images as the base layer.

The opacity data layer was continuous, where each pixel has a slightly different opacity data value compared to the neighboring pixels. It is common to have larger polygons with the same data value, which may be easier for the user to decode.

The sample of the study focused on people with some previous experience with web-based maps, which may limit the external validity for the general population that may be less used to using map elements such as legends.

Lastly, the tested designs were created by the author of the study, and there may be alternative designs which are more optimal.

1.4 Disposition

Section *2. Background* provides contextualization and a walk-through of related literature. Section *3. Methodology* describes the legend designs used in the study as well as how the collected data was analyzed. *4. Results* includes summary statistics, key results and robustness tests, which are elaborated upon in section *5. Discussion*. Section 6 concludes the thesis.

2 Background

The ubiquity of web mapping and the increased ease of creating customized maps for distribution on the web makes online maps a part of most people's daily life. Panko (n.d.) estimates that 77% of smartphone owners regularly use navigation apps, where map visualizations are a key part of every interface. Google maps had over one billion monthly users in 2014, representing 41% of Internet users worldwide (Veenendaal et al., 2017). According to a site measuring the presence of different web technologies online, maps were present in around 25% of the top 100,000 sites in July 2021, of which Google maps accounted for 74% (BuiltWith, n.d.). In the list of frequently used map technologies, there are also other alternatives, such as MapBox, ArcGis, and Leaflet.js. It is not easy to customize the base map that comes from centrally controlled servers, resulting in the use of colour scales and semi-transparent heat map visualizations as a way to add data to map visualizations.

This frequent exposure to maps creates a multiplying effect in the exposure of even less frequently seen map types in terms of impressions (how many people see even less common visualizations). Lower opacity visualizations[1] are suitable for visualizations of e.g. high-impact events such as typhoons/hurricanes. There are very few previous studies on the perception accuracy of such map designs as will be outlined in section 2.1.

2.1 Related studies

Literature that intersects with the scope of this study can be found in a few different research fields.

In the area of perception and data visualization studies, there is a long history of low-level studies measuring the ability of humans to correctly perceive visual data encodings, such as position, length, angle, circle area, hue, etc. (Cleveland and McGill, 1985).

[1] mapping of data directly to opacity or mapping of data to other channels, such as a colour scale in combination with lowered opacity to reduce occlusion.

There are also studies on perception effects from context, such as luminosity perception of a grey area depending on surrounding areas (Adelson, 1993). For higher-level tasks, Tufte (2001) exemplifies attempts to bring these smaller encodings together to describe good data visualization practices. This corpus of literature usually tries to create design guidelines and best practices for data visualizations.

These efforts to quantify design choices and data encoding methods have been criticized, where e.g. Muzner (2013) describes a gap between these low-level tasks and high-level tasks in the data visualization literature. In recent years, researchers have commonly attempted to bridge the gap towards other scientific areas (Kim et al., 2019, 2017), incorporating mental models, uncertainty, and Bayesian prior beliefs of subjects into the perception task. The goal in this literature is to bridge the combination of the lower-level tasks and relate them to efficient user experiences through e.g. preattentive processing - meaning differentiating objects of interest to the viewer by giving it different encoding (e.g. giving data points of interest emphasis by encoding them in different colours than the other data points) (Findlay and Gilchrist, 1998; Treisman, 1985; Wolfe and Horowitz, 2004).

A similar discourse is also present in the field of map visualizations. Cognitive cartography blends the fields of cognitive methods with how maps are used. This cover perception of e.g. colours, the need for users to keep information in their long term or working memory. The way this has commonly been studied through low-level identification and decoding tasks and creating a controlled environment, where subjects may not interact with maps in a similar way they would outside of the experimental setting has been criticized, as for the larger literature on data visualization outline above (Montello, 2002).

Studies that cover transparency directly or indirectly for 2D data visualization are rare. Jo et al. (2019) briefly mentions opacity:

> *To scale scatterplots, several approaches have been proposed, such as adaptive opacity [15, 30, 32] and aggregation [13, 53]. However, adaptive opacity does not scale well with the number of categories since multiple categorical colors become ambiguous when blended[..].*

For map and perception studies, colour choices in a cartographic context is a large corpus, with Brewer (2006) having provided colour palette tools from studying the use of colour schemes from a cartographic/GIS perspective. These online tools help in creating suitable colour scales taking into consideration factors such as contrast, usability for people with colour blindness, etc. (Harrower and Brewer, 2003). The colour schemes/palettes provided by the tool have been extensively used also in other data visualization fields outside of map-making.

Brewer (1997) have also studied the use of colour scales, which is an alternative to using only a single colour as in this study. Brewer (1997) states that "*Cartographers have long discouraged the use of spectral, or rainbow, color schemes on thematic maps of quantitative geographic data, though such color use is common in GIS and scientific visualization*," but goes into detail on how spectral schemes can be useful if designed with care.

This level of care is always required when choosing multiple colours to encode data, as the distance between different colours is perceptually ordered only for short sections, but not for the full colour spectrum. Ware (2020) exemplifies the issue:

> *This can be demonstrated by the following test. Give someone a series of gray paint chips and ask them to place them in order. They will happily comply with either a dark-to-light ordering or a light-to-dark ordering. Give the same person paint chips with the colors red, green, yellow, and blue and ask them to place them in order, and the result will be varied. For most people, the request will not seem particularly meaningful. (Ware, 2020, p. 128).*

Nevertheless, colour is one of the most common ways to encode and communicate information in maps and other visualization types, and the design of colour schemes for different use cases is a field where many variations and use cases are studied. In the greater literature on how to represent uncertainty in maps (e.g. Cheong et al., 2016; Leitner and Buttenfield, 2000), there is also attempts at creating dedicated colour schemes, as in Seipel and Lim (2017)'s study on the effectiveness of colour schemes when assessing flood risk using maps.

Colours and transparency interact by definition, as the output per pixel is a calculated composite of the colour of the transparent object and the colour of the items behind

(Porter and Duff, 1984). In the context of 3D models, and the ability to perceive objects clearly this is a major and well-studied problem. Seipel et al. (2020) have shown that for 3D models, distinct colours (evaluated using not only hue, but also perceptual properties) when fully opaque becomes difficult to distinguish when transparency is increased due to colour blending.

Other studies have looked at transparency/opacity more generally in distinguishing objects in 3D scenes, finding that reducing the transparency number of levels in the scene can help users in identification tasks (Wang et al., 2017).

Targeted literature on legends is often highly tied to different visualization types. In a walk-through of proper data visualization practices, Evergreen and Metzner suggest that there is clear importance of data-proximity of legend usage and placement:

> *Because human eyesight has only a narrow range of focus, graphics should be placed very near their associated text (Malamed, 2009; Ware, 2012). […] if a legend is required in the visualization, it should be so near the corresponding data points that no eye movements are needed to relate the two.* (Evergreen and Metzner, 2013).

A large part of the literature studying legends in the context of maps has a clear design focus. Peterson (1999) introduced an interactive legend that allowed users to make changes to the map through the legend interactions, finding that users can get a better understanding of maps through the interactions. Other design studies have looked at fairly narrow questions related to legend designs for maps. Examples of such design studies have been using additional diagrams in combination with classic legends as a supplement to choropleth maps (Chien et al., 2019; Cromley and Cromley, 2009; Cromley and Ye, 2006), and in other cases substituting the legend for a frequency histogram altogether (Kumar, 2004).

In a more holistic view on legend usage/design within web GIS, Cybulski (2016) analyzed if animated maps that are published on the internet have designs that follow long-established guidelines within cartography. Cybulski (2016) finds that these maps do follow the guidelines in general, but also that they are adapting to new technologies and user circumstances.

No intersection between legends and transparency, either in GIS-context or elsewhere, has been found in the literature review. This is largely supported in the literature study "Grouping Rules for Effective Legend Design" by Qin and Li (2017) that aimed to find whether there were cartographic rules for effective legend designs *"[…]it is found that only one study was dedicated to the building of cartographic rules for effective legend design"*.

Kiik et al. (2017) evaluates some visualization alternatives in the cartographic design of polygons, transparency being one of the alternatives, together with outlines/boundaries, hatches (texture within the polygons) and icons. The results with examples of multiple polygons the hatches design was more effective along many of the metrics gathered using eye-tracking technology, but the transparency design was preferred by most of the subjects. Kiik et al. (2017)'s study used multiple overlayed polygons with low opacity in their examples, creating a colour blending effect between the polygons.

3 Methodology

Four different map legend designs, as well as a version without legend, were generated and then shown to respondents through a web user interface. Using this interface, the task for the respondents was to estimate the value of the opacity-mapped overlay at the location of a marker.

The generation of the examples allowed all parts of the designs apart from the legends to be kept constant. This was done to control the exposure to the treatment (different legend designs, or absence of legends).

3.1 Legend design and generation

3.1.1 Legend design types

The map designs were all static 2D maps shown using web technologies. The subjects were shown different legend designs and were asked to estimate the value at the location of a marker. The data was represented as a spatially distributed phenomenon mapped to opacity in a polygon overlaid on a base map, as in the ArcGIS online maps described in section 1.1.

The 4 legend designs and the version without legend, are presented below, together with some justifications – where all legends and data mappings are linear in the opacity/alpha channel. This ought to introduce fewer complications than data-encoding using colour channels due to a linear encoding of the alpha/opacity-channel in the RGBA-model[2]:

> *It's worth pointing out that unlike the color components which are often encoded using a non-linear transformation, alpha is stored linearly – encoded value of 0.5 corresponds to alpha value of 0.5 (Ciechanowski, 2019)*

Baseline 1 No legend ["Headline" in tables], (Figure 3a) - Only having a title indicating the range of values. Used to test the correctness of visual decoding by

[2] RGBA = Red, Green, Blue, Alpha

subjects without legend. Used as the first and last example in the progression presented to the subjects.

Baseline 2 – ArcGIS Legend Imitation ["Checkered" in tables], (Figure 3b) - The design is used as a baseline for models that explore if alternative legend choices are helping in decoding values compared to the current "industry standard."

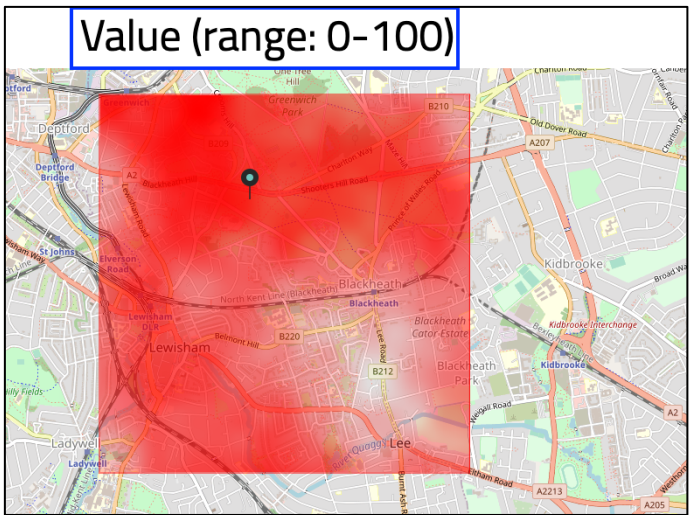
One likely downside making it more difficult for subjects to accurately decode the data values is the lack of background context in the legend, i.e. the colours in the legend do not interact with a coloured background as it does in the visualization. The remaining legend design types are all designed to combine the map base-layer background information into the legend.

Legend with sampled context ["Sampled" in tables], (Figure 3c) - Sampling a rectangular area of the base layer as background to the opacity mapping in the legend.

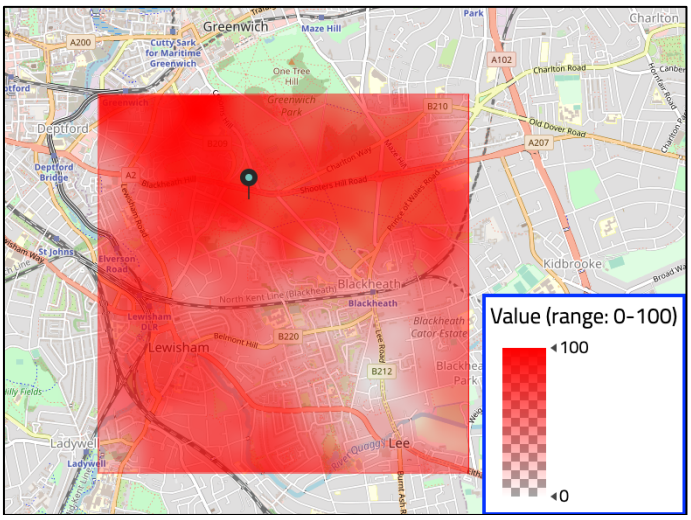
Legend with clustered colour bands ["Clustered" in tables], (Figure 3d) - Using the most common colours in the base layer as a background for the legend. In the Clustered legend design, the 10 most common colours of the background map[3] were displayed as strips/columns in a vertical manner behind the opacity data mapping. This way the most prevalent colours in the map provide a context for data-encoding in the legend.

Annotated Outline ["Annotated" in tables], (Figure 3e) - The legend placement may heavily affect the users' ability to keep the information in near memory while moving the eyes back and forth between visualized data and the legend. In an attempt to move the information closer to the data an outline legend was placed next to the data polygon. The legend is contextualized by having the legend opacity viewed directly on top of the base map layer.

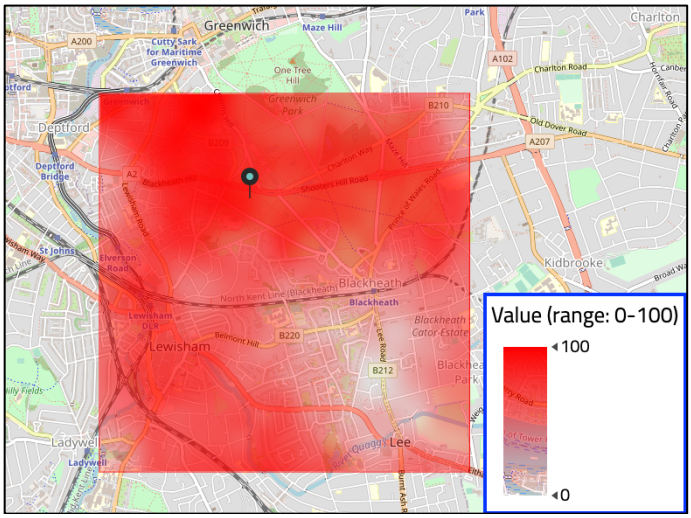
[3] Using kMeans clustering algorithm implementation in the base R language, was used to extract 10 "colour centers" representing common colours in the base map.



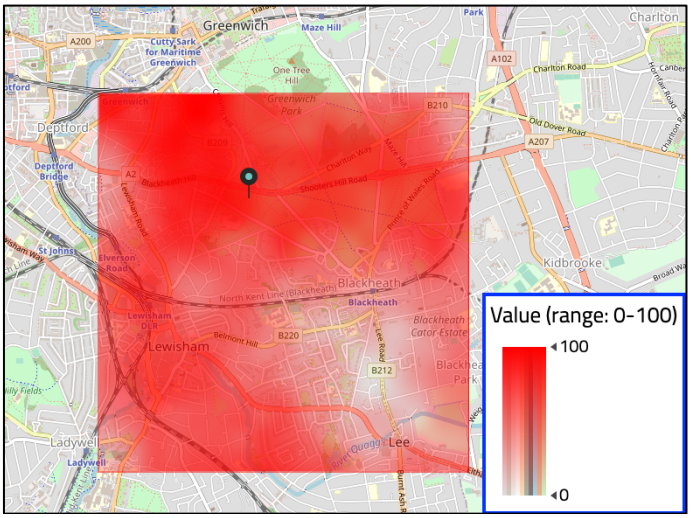
(a) Baseline 1 - No legend [”Headline”]



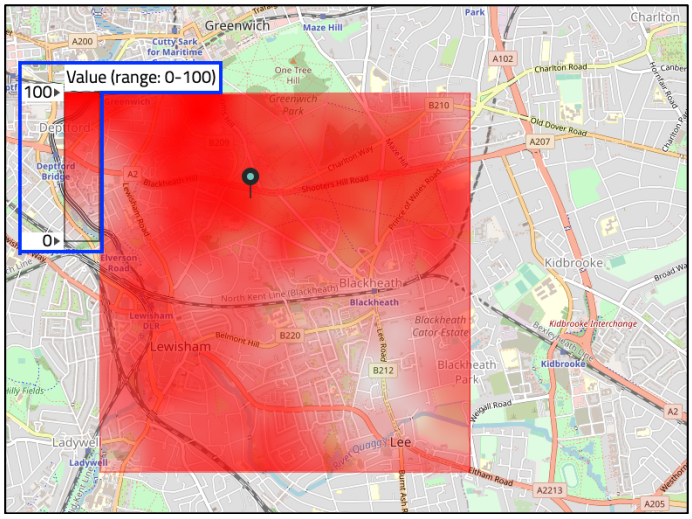
(b) Baseline 2 – ArcGIS Imitation [”Checkered”]



(c) sampled context [”Sampled”]



(d) clustered colour bands [”Clustered”]



(e) Annotated Outline [”Annotated”]

Figure 3: Visualization legend types and design with no legend (enlarged legend for clarity, highlighted in blue)

3.1.2 Map example design and generation

Several layers and elements were produced for the sample map images[4]. These building blocks of the example maps are described below.

Base Map - The Javascript library OpenLayers[5] was used to download a background map to a large[6] HTML5 Canvas element (web-browser raster-API). The area was chosen arbitrarily as Greenwich, London. The zoom level for the map was chosen manually to get an area with a diverse set of background colours representing water, buildings, and other infrastructure.

Opacity Data Layer - In order not to have respondents making use of pre-existing knowledge or assumptions, continuous data was generated using simulated Perlin-noise (Perlin, 1985)[7]. The Perlin-noise family of algorithms are an alternative to other ways to generate data, such as using random noise, pure trigonometrical functions or statistical distributions. Due to the unpredictable and highly natural look of Perlin-noise, it is commonly used in computer-generated imagery (e.g. in computer-generated art and in video games).

All generated data were normalized to have the lowest value to be 0, and the highest being 100.

Two different types were generated: one using pure Perlin noise within the area (see Figure 4a), and the other one using "fall-off": having higher data values in the center of the data area, with decreasing values towards the edges (Figure 4b).

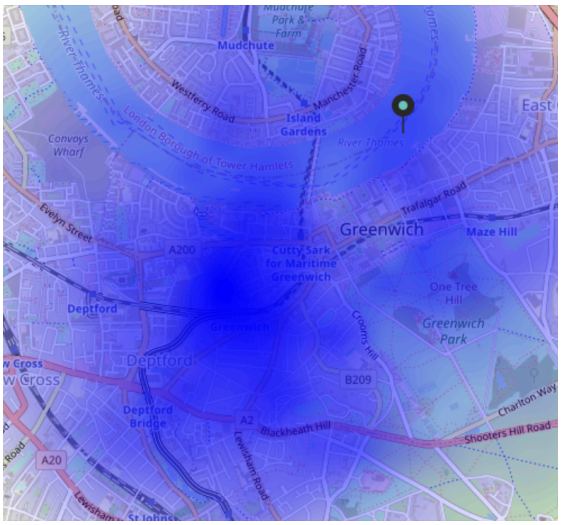
---

[4] All code available at https://github.com/Tille88/thesis-map-generation

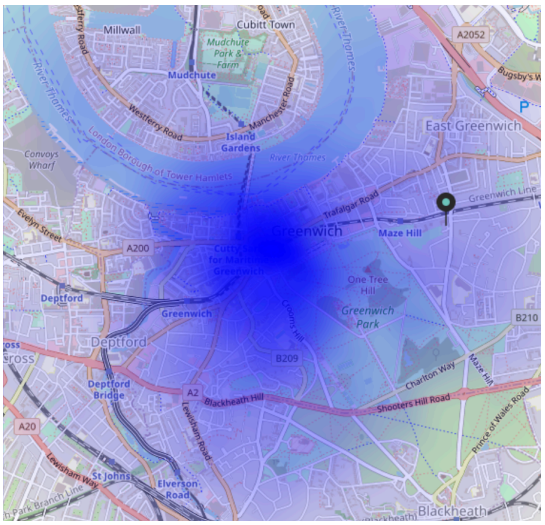
[5] ttps://openlayers.org/

[6] larger than 1,600*1,600 pixels, and later down-sampled due to not knowing usage at the data generation stage.

[7] The implementation used to generate the noise was https://github.com/p5py/p5/blob/master/p5/pmath/rand.py



(a) No falloff



(b) With falloff

(c) Marker

Figure 4: Map elements

The Perlin noise data was put into the opacity value for a canvas image element, with 1 of the 3 RGB-colour channels set to the max value, e.g. for red, canvas per that pixel was rgba(255, 0, 0, opacityValue). The same data overlay mapped to the different colours are shown in Figure 5.

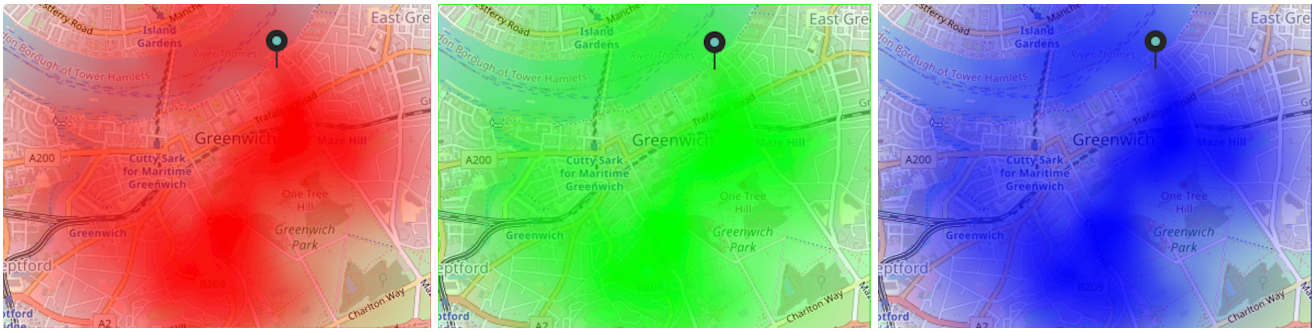


Figure 5: Identical data and marker positions for all colors (red, green, blue)

Marker - A marker was designed using the vector-based SVG-API in the browser and overlaid at a randomized location within the data-extent (see Figure 4c). This was

merged into the Canvas raster representation, and the value of the data layer at the location was stored.

Legends - All legends were created using the Canvas API, similarly as was done for the base map and data layer. The different legend types have been described in detail in section 3.1.1.

3.1.3 External validity and limitations of example designs

To control the setting for the data collection, the examples were somewhat contrived.

- Today, rather few complex data visualizations are lacking annotation or possibilities for interaction.
- All data area mappings were strictly rectangular to allow for easy automated placement of the Annotated legend types.
- Only one opacity data layer was included. However, this seems like a good design choiceas multiple and possibly overlapping opacity layers would result in compounding opaqueness and more complex colour blending.
- To keep the designs similar for the decoding task, the size of the legends were chosen to imitate the ArcGIS design. Verbal feedback was given by one respondent that it would be useful to have larger legends.
- The base map layer was kept constant, and may not be representative of other areas. It is not clear how opacity mapping and legend are working for less complex (e.g. maps/backgrounds with less/no colours) or more complex backgrounds (e.g. aerial/satellite imagery).

None of these factors, except the last, ought to have strong implications for the external validity of the study.

3.2 Data collection

3.2.1 Procedure and user interface (UI)

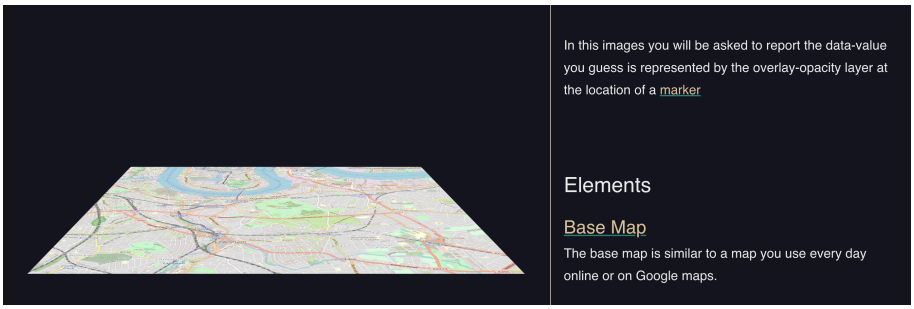
A front-end UI was developed using HTML5, CSS, and Javascript so that it could be displayed using any web browser[8]. In this system, the respondents were taken through three distinct stages:

View 1 - Introduction of task (Part shown in figure 6)

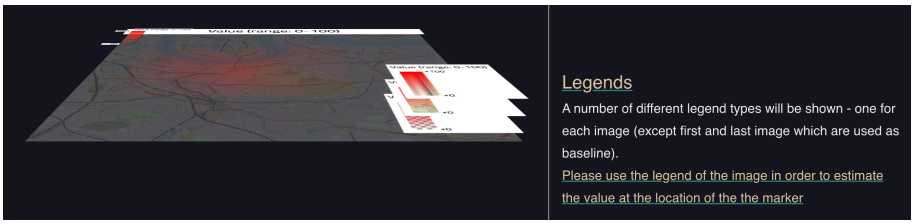
Introduction with instructions of the decoding task and the UI. The elements of the map (base map, data layer, marker, and legend) were shown visually together with the UI elements used to submit responses (slider, submit button). The respondents were explicitly asked to use the legends to estimate the value of the data layer at the location of a randomly generated marker.

Following the introduction, and a verbal check on understanding the task, the session was initialized by clicking a button in View 1.

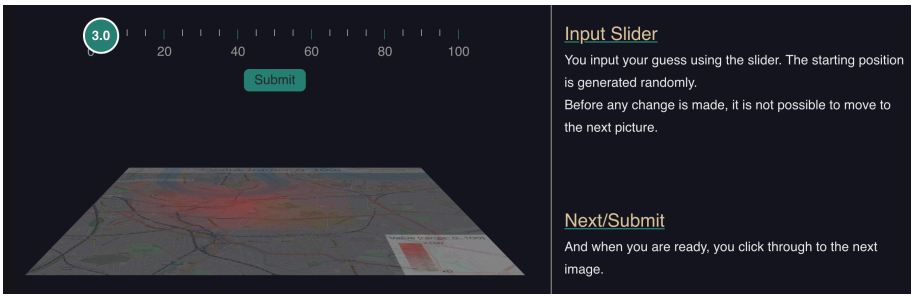
[8] All frontend code can be found at https://github.com/Tille88/thesis-front-end



(a) Elements - base map introduction



(b) Elements - legends introduction



(c) Elements - response element introduction

Figure 6: Introduction of task

View 2 - Visual decoding of 10 examples, exposing the legend designs in semi-random order

Using the web interface of *View 2*, the users visually estimated the value at the location of a marker within the overlay data area.

Ten images picked by the back-end server from a large number of generated example combinations[9] were presented to the respondent one example at a time.

[9] All possible combinations of colours (3, being red/green/blue), random data and marker location variations (21 examples that were generated through automated scripts) and designs (5, made up from 4 legend designs + 1 no legend), resulting in 3 * 21 * 5 = 315 different examples

The progression of the images always started and ended with the baseline of Headline/no-legend type, and had the order of the 4 legend variations picked uniformly random for the example progression images 2-5 and 6-9 respectively.

The colours of the data layer were available in pure red, green, and blue for all examples. For the first image in the progression, a combination of these three colours (e.g. [“Blue”, “Red”, "Green”]) was randomly generated and repeated for examples 1-10. The reason for changing the colours between the examples was to reduce the risk of the respondents remembering the colour-to-data mapping between examples.

The progression for each respondent was generated using a combination of colour, data example (data layer and marker location), and legend types using a simple algorithm.

As an example to how a different progression would be generated for each user, for the progression of images 1-10, three lists were generated:

1. A random list of numbers from the number of examples (21), i.e. between 1-21, where each number could only be picked once per progression.
   Example list of 10 elements (i): [3, 14, 4, 10, 16, 2, 7, 9, 12, 20]

2. A colour progression, e.g. [“green,” “blue,” “red”] that was circulated: green, blue, red, then starting with green again.
   Example list of 10 elements (ii): [“green,” “blue,” “red,” “green,” “blue,” “red,” “green,” “blue,” “red,” “green”]

3. A selection of legend types always starting and ending with Headline/no-legend: [**first 5 elements:** headline, random order of 4 legend types, **last 5 elements:** random order of 4 legend types, headline].
   Example list of 10 elements (iii): [“no-legend,” “sampled,” “clustered,” “annotated,” “checkered,” “clustered,” “checkered,” “sampled,” “annotated,” “no-legend”]

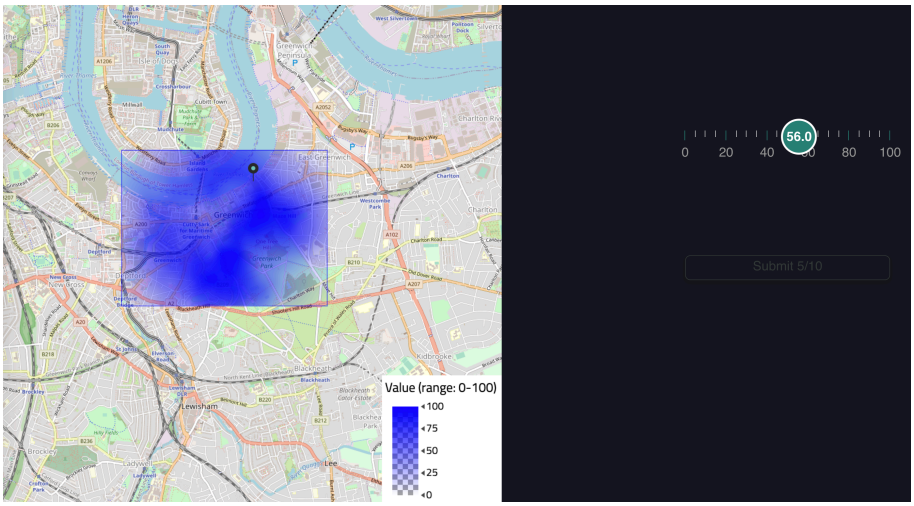
4. Then these lists (1-3) were combined into a progression, corresponding to the file names for the generated example images
   Example combined lists (i, ii, and iii): [“3-green-no-legend,” “14-blue-sampled,” “4-red-clustered,” “10-green-annotated,” “16-blue-checkered,” “2-

red-clustered," "7-green-checkered," "9-blue-sampled," "12-red-annotated," "20-green-no-legend"]

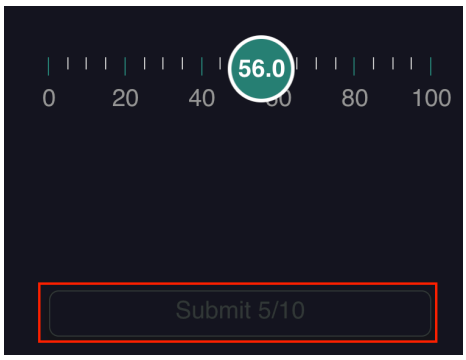
The response (estimate of the value at the marker) and some interactions by the respondents were persisted as the respondents went through the generated progression of 10 examples. The response for each example was picked using a slider in the UI covering the range of the data (0-100). The initial value of the slider was random, and the respondents were required to change the response before being able to progress to the next example image (UI example shown in Figure 7a).

At the load of the image, a timer was started, and some events (mouse hovering over image, response changes of the slider, time of event relative to the example loading, etc.) were stored in the browser's memory. At the time of submission of an example, all event data, the time from load until submission, and the final response value was persisted server-side. This was repeated 10 times.

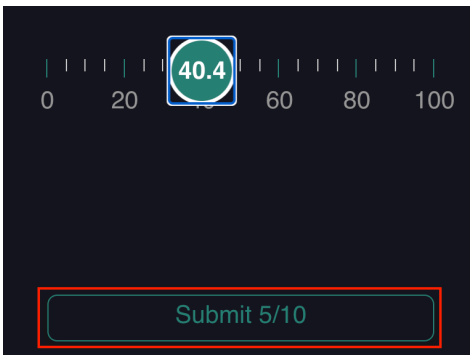
The user was not able to go back once an answer had been submitted, and the "submit button" was not active until the user had updated the value from the initially random value in the range of 0-100 (see Figure 7b and 7c).



(a) Progression UI



(b) Submit inactive – highlight in red



(c) Submit active – highlight in red

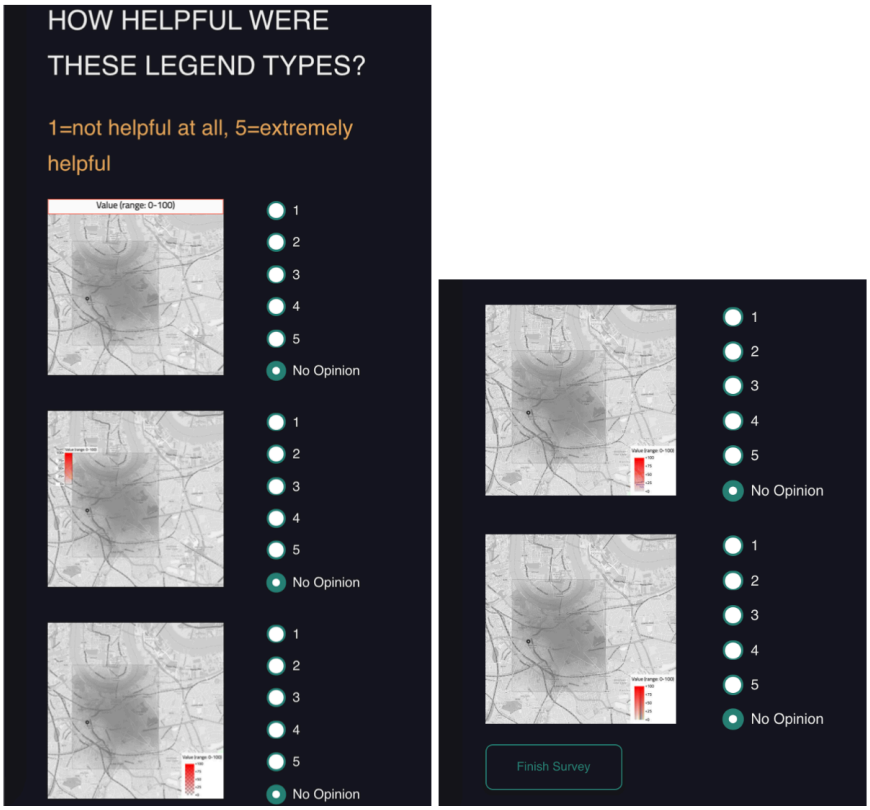
Figure 7: Progression UI

The reason for choosing 10 examples per respondent was to ensure a large amount of data could be collected with a limited number of respondents. During testing prior to data collection, this amount of examples was established to be possible to complete within less than ten minutes, and not be very tiring for the respondent.

Having each respondent being able to see each legend design twice, and have Headline/no-legend designs as first and last example, also allows easier testing of learning effects and being able to check if there are any major differences between not having a legend after being exposed to many examples.

View 3 - User acceptance

After progressing through the 10 example images, the respondents were asked about how helpful they considered the different legend types were for their task. This information was also sent back and persisted to the database (Figure 8).



(a) Acceptance UI top (b) Acceptance UI bottom

Figure 8: Acceptance UI

After having to choose very precise numbers of a continuous scale between 0-100 for the prior step, a common web-form format and only five choices were chosen for user acceptance. These values can not be seen to be fully comparable, as they are more subjective and less precise than estimation errors, which can be measured to the decimal level. In other words, for each respondent, you get an ordinal scale, but it is not necessarily comparable between respondents. That is also the reason why each respondent had to give a value for each legend design to have a balanced data set for analysis.

3.2.2 Participants in the study

The survey data was collected between Jan 9th and Feb 6th, 2021, using convenience sampling. A sample of participants with moderate to good experience of map reading was considered ideal for a relatively small sample size, and to increase the chance of understanding the task quickly. This sample also ought to be homogeneous in relation to the scope and research questions of the study, increasing the chance of clear and easily interpretable data. Participants with a background in software development, UI/UX designers and employees of companies with map services were requested to participate. As the data was collected during a time with COVID-restrictions, most participants were either working or visiting the workplace of the author of the study, resulting in most of the 34 respondents being employees at the Volvo Car Corporation Asia Pacific Head Quarter in Shanghai, China, as well as vendors visiting the location.

All participants had at least a bachelor degree. Most held Master degrees in technical disciplines, and one participant a doctorate in mathematics.

The sample consisted of 8 women and 26 men. Approximately one-third of the participants work with geospatial data on a weekly basis, while most of the remaining participants spend significant time in vehicle testing and planning field tests routes using maps similar to the base map of the study. All of the participants were in the age range of 25-50, with most below the age of 35.

The participants are on average expected to have better eyesight than the general population due to their relatively young age but are more likely to have some variation of colour blindness due to the high proportion of men participating in the study. Colour blindness or eyesight was not asked or tested during the data collection.

The initial plan was to put the survey online using cloud technologies for wider distribution. During the testing stage, it became clear that there was a need to check that the instructions were understood to reduce the variation in the data. Hence, a more qualitative methodology was chosen where the author was present in the room and asked the respondent if the task was understood after reading the instructions described in section 3.2.1.

This resulted in a more consistent study, where the same laptop and Chrome web browser was used by all data collection, and the lighting conditions were kept

consistent. All 34 respondents did therefore see the examples of the same size on an identical screen and using the same rendering software.

Due to each respondent submitting responses to 10 example images, the resulting analyzed data sample was quite large. During the collection phase, some base models were estimated repeatedly to evaluate the results of a larger sample. At around 20 participants the results stabilized, but data collection was continued until the pool of suitable participants at the Volvo office was exhausted.

3.3 Data analysis

Multiple models are estimated to evaluate the effect of the independent categorical variable of legend design. The dependent variables are estimation errors and acceptance scores. A full list of variables is given in section 3.3.1.

Outlined in section 3.3.2, Bartlett's statistical test is conducted to evaluate how the variance of the error are differing between the legends designs. This is followed by Ordinary Least Squares (OLS) estimation with dependent variables that are not naturally centered around 0 - absolute estimation error (|*PERCEPTION ERROR*| in the following text) and acceptance scores (*ACCEPTANCE*). Each legend design is encoded as separate dummy variables and estimated relative to a reference category that is chosen to isolate the effects of interest to answer the research questions.

Section 3.3.3 describes the use of control variables and subsamples to be able to test the robustness of the results in section 3.3.2.

3.3.1 Data and variable description

Dependent Variables

PERCEPTION ERROR - distance between the response value to the actual value of the location of the marker per visualization.

|*PERCEPTION ERROR*| - absolute perception error. For linear modelling, there is a need to look at estimated mean differences between categories. Simply using *PERCEPTION ERROR* would result in all categorical differences being statistically insignificant unless the variance around 0 is heavily skewed.

ACCEPTANCE - objective valuation by participants on the usefulness of each legend type on a scale of 1-5. Opt-out responses of "No opinion" were allowed and treated as missing values in the statistical modelling.

Independent Variables

LEGEND - categorical variable for each of the 4 legend types and no legend. The only part of visualization that is not randomized, and designed to be the treatment variable for which effects are to be estimated. Having values of Headline (no legend), Checkered, Clustered, Sampled, and Annotated, corresponding to section 3.1.

Three variables were collected with the aim to find a good proxy variable for uncertainty:

SUBMIT TIME - time in seconds from the rendering of the example until the subject submitted their response.

INPUT CHANGES - number of times the respondent changes their answer with the input-slider before submitting.

HOVER EVENTS - events of hovering the mouse cursor over the example image.

Two variables were collected as potential control variables:

PROGRESSION - for each respondent, the numbering of the examples in the progression from 1-10. Used to check for learning and/or fatigue effects.

COLOUR - visualization presented having opacity data layer mapped colour channel as pure red, green, or blue.

3.3.2 Hypothesis testing models

For *PERCEPTION ERROR* the average value is likely centered around 0. Most statistical models are constructed to compare mean values between categories, which would not be able to give any indication of differences between the *PERCEPTION ERROR* for different legend designs. However, the distribution variance is of interest to estimate.

The parametric Bartlett's test (Snedecor and Cochran, 1989) is used to test if multiple categories all have equal variances. In the study, Bartlett's test is used to evaluate if

variances differ in a statistically significant way between the different legend types. This is done for both the full data including the Headline/no-legend type, as well as a subset excluding the Headline type. Significant results for the tests would indicate that the distribution of errors differs between any of the legend categories included in the sample.

Due to that Bartlett's test assumes that data is normally distributed, Shapiro-Wilk tests (Shapiro and Wilk, 1965) are first conducted to assert if the normality assumption holds. This will be done for all legend types pooled, with Headline removed from the pooled sample, as well as for the subsample of each legend type separately.

For variables that are not centered around 0, |*PERCEPTION ERROR*|, *ACCEPTANCE*, *SUBMIT TIME*, *INPUT CHANGES* and *HOVER EVENTS*, are examined using regular OLS techniques. As the treatment variable *LEGEND* is categorical, each category is transformed into a separate dummy variable for each legend type.

To avoid multicollinearity, the baseline legend categories are interpreted as the intercept α-estimate, against which the β-estimates and corresponding test statistics are compared against.

The baseline OLS models estimated are of the form

$$|PERCEPTIONERROR|_i = \alpha + \beta_i LEGEND_i + \varepsilon_i \quad (1)$$

and

$$ACCEPTANCE_i = \alpha + \beta_i LEGEND_i + \varepsilon_i \quad (2)$$

3.3.3 Robustness checks

OLS models are estimated to see if there are any indications of time effects as the respondent went through the progression of 10 examples. This could be in either direction due to learning and/or fatigue.

Learning effects are likely to be negligible. This is due to that the subjects were not given any feedback on the correctness of their responses during the progression, and that different legend designs and data overlay colours were rotated.

$$|PERCEPTIONERROR|_i = \alpha + \beta_1 PROGRESSION_i + \varepsilon_i \quad (3)$$

An alternative model is run on the subset of only the first and last example (Headline/no-legend) encoded as a dummy variable, to see if the response accuracy differed between the first and the last examples shown to respondents.

$$|PERCEPTIONERROR|_i = \alpha + \beta_1 PROGRESSIONLAST_i + \varepsilon_i \quad (4)$$

The results from these tests also feed into the possible need to do subsample analysis using only e.g. the first 5 responses of each respondent.

Subset analysis using only the respondents that are likely to be the most engaged or skilled, as defined as the respondents with below-median average |*PERCEPTION ERROR*| are conducted to see if the results differ in a significant way.

4 Results

4.1 Summary statistics and exploratory graphs

Treatment

Each of the 34 respondents completed the full survey, resulting in 340 responses to the data-progression, and as seen they are distributed equally among the legend types by design (Table 1).

Legend	N
Annotated	68
Checkered	68
Clustered	68
Headline	68
Sampled	68
Total	340

Table 1: Responses by legend type

Dependent Variables

PERCEPTION ERROR - As seen in Figure 9, the mean is centered around 0 for all categories of *LEGEND*. There are some visually distinguishable differences in distribution between the Headline (no legend) category having the largest variance and more outliers compared to the examples with legends. The figure has the mean indicated by a black dot, with 1 standard deviation in each direction shown by a range-whisker for each category.

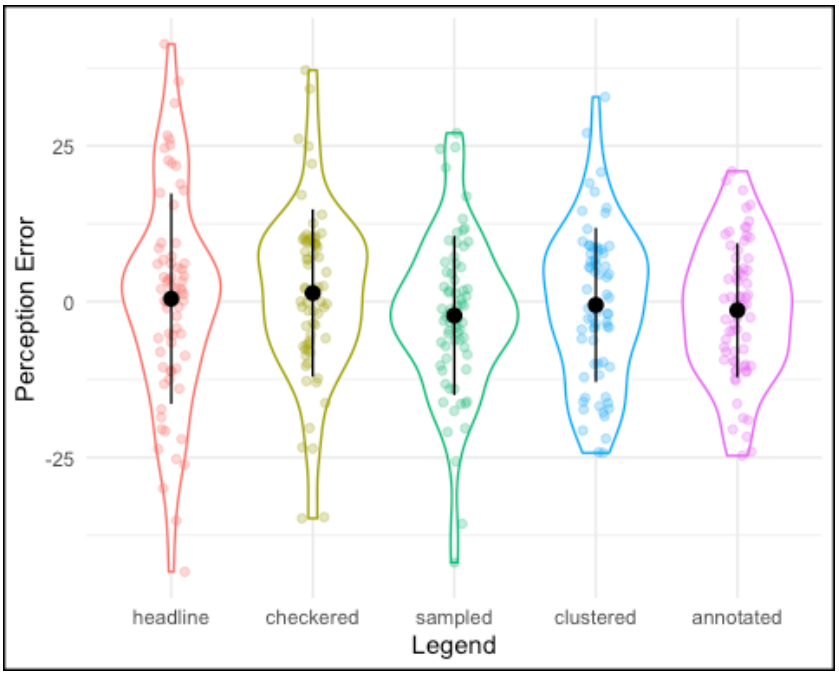


Figure 9: Perception error by visualization category

|*PERCEPTION ERROR*| - Figure 10 shows the same data as in Figure 9 with an absolute transformation. The mean differs between the categories, being highest for Headline type, and lowest for Annotated legend type. The variance of |*PERCEPTION ERROR*| is quite large throughout. The vast majority of categories have absolute errors heavily clustered below 20, but there are outliers above 30 for all categories except Annotated.

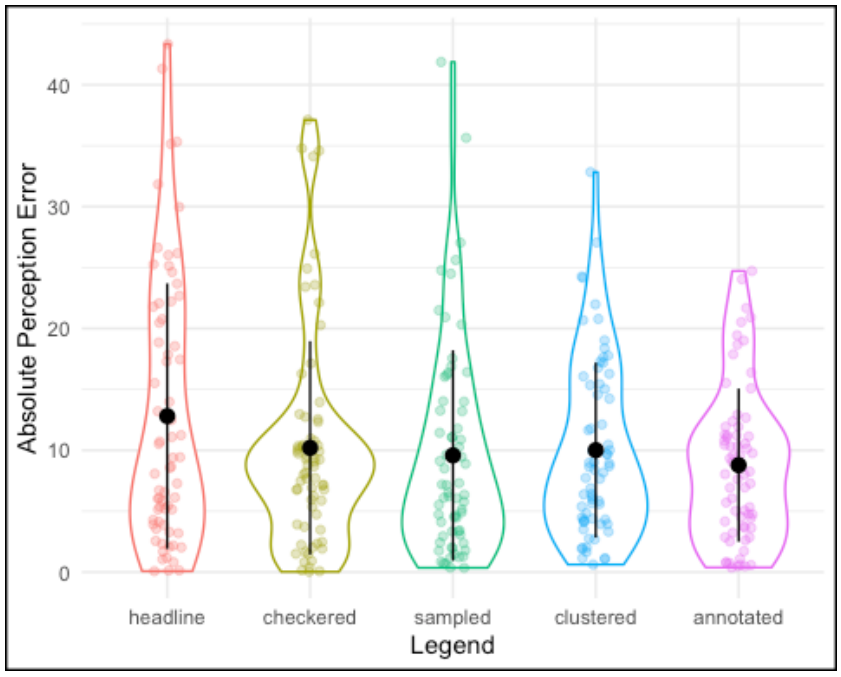


Figure 10: Absolute error by visualization category

ACCEPTANCE - Table 2 indicates that the highest *ACCEPTANCE* scores are for Annotated and Sampled legend types. Headline/no-legend has the lowest acceptance scores.

Acceptance	Headline	Checkered	Sampled	Clustered	Annotated
1	11	0	0	1	0
2	9	5	1	10	2
3	12	13	5	11	4
4	1	8	14	7	17
5	0	6	13	4	10
No Opinion	1	2	1	1	1

Table 2: Acceptance response count cross-tabulation

Independent/Control Variables Candidates

In the data collection phase, it became clear that all the uncertainty-proxy variables would be unlikely to be valid proxies for respondent uncertainty.

SUBMIT TIME varied because of incoming phone calls, the need to ask questions about the design types and many other external factors, *INPUT CHANGES* varied due to how the respondents used the slider[10]. *HOVER EVENTS* did not vary much within-subjects as the mouse cursor wasn't consistently used as a visual guide (some users did not use the mouse cursor, but kept it still while using fingers on the screen to make comparisons).

Plots for *SUBMIT TIME*, *INPUT CHANGES* and *HOVER EVENTS* (Figures 12, 13 and 14) can be found in Appendix A.

COLOUR - From visual inspection, no strong difference between the colours red/green/blue examples were perceptible (see Figure 15 in Appendix A). Statistical estimates also had no statistical differences in |*PERCEPTION ERROR*| between the colours (Table 14 in Appendix A), reducing the need to include colour as a control variable in any models.

4.2 Dependent variable: PERCEPTION ERROR

4.2.1 Normality tests

As shown in Table 3, all the p-values for the Shapiro-Wilk test are above 0.05, meaning the null of normality is not rejected for any of the tests run. In other words, we can assume that the data is normally distributed, no matter how it is separated by legend types, or if the statistical test is run on the full combined sample.

[10] Some respondents picked their response in one click, while other respondents clicked the slider to put it in focus, and then use the arrow keys to change the input value. In the former case, this would result in one *INPUT CHANGE*-event, and in the latter, there would be a separate event logged for each arrow key-press incrementing/decrementing the value by 0.1 each click.

	Sample	Shapiro Wilk Test Statistics	P-Value
1	Combined	0.992	0.061
2	Headline Removed	0.992	0.169
3	Headline	0.988	0.738
4	Sampled	0.976	0.211
5	Clustered	0.979	0.313
6	Checkered	0.967	0.071
7	Annotated	0.987	0.705

Table 3: Shapiro-Wilk normality tests

4.2.2 Bartlett's test of homogeneity of variances

Table 4 shows that the Headline types show the largest variance/standard deviation for the error, followed by Checkered, Sampled, Clustered, and Annotated legend types.

	Legend	Standard Dev.
1	Headline	16.9
2	Checkered	13.4
3	Sampled	12.8
4	Clustered	12.4
5	Annotated	10.8

Table 4: Error standard deviation by legend type

In the statistical tests, the Headline type was used first as a baseline, to determine if these differences were statistically significant, then a subset removing all the Headline data points from the sample was used.

As seen from Table 5, there is a statistically significant difference in the variance for the sample including the Headline type (p-value < 0.05). This indicates that when including the Headline category, the variance of the *PERCEPTION ERROR* is not the same for all categories.

	Sample	Test.Statistic	P.Value
1	All, including Headline	15.018	0.005
2	Excluding Headline	3.433	0.330

Table 5: Bartlett's test results

The significance does not remain when excluding the Headline category from the sample (p=0.33). The null of the variance for the legend designs are the same can not be rejected, i.e. using Bartlett's test, the legend designs does not show differences in error variance.

4.3 Dependent variable: |PERCEPTION ERROR|

The way of reading Table 6, as well as Tables 7, 10, 11, 12 and 13 are identical. The dependent variable is displayed at the top corresponding to the variable on the left side of equation (1) in section 3.3.2. Two different model columns are listing which variable are used as reference estimate α in equation (1). For each column the estimate for that category can be found under the "Constant/Reference" column. So for Model (1), the average error for the Headline design is 12.8. This also means that for model (2) where Checkered is the reference category, there is no estimate for the Checkered column.

The estimates for the other rows (Annotated, Checkered, Clustered and Sampled), corresponds the β-estimates in equation (1), and are relative to the reference category. In Model (1) in Table 6, Annotated has an average error relative to the headline category of -4.02, meaning that the actual estimated error is 12.805–4.024=8.781.

The regression results in Table 6 are consistent with the variance tests, where model (1) estimates are relative to the Headline reference category.

All other legend types display lower errors and are all statistically significantly lower than the errors for Headline. The Annotated type had the lowest average error of -4.02 lower errors than the Headline design.

In model (2), where Checkered is used as a reference category (chosen due to both the research question formulation and the data distribution) the other designs are showing comparatively lower estimated errors, but none of those results was statistically significant.

	Dependent variable:	
	|PERCEPTION ERROR|	
	Headline Reference	Checkered Reference
	(1)	(2)
Annotated	−4.024***	−1.420
	(1.458)	(1.336)
Checkered	−2.604*	
	(1.458)	
Clustered	−2.793*	−0.189
	(1.458)	(1.336)
Sampled	−3.226**	−0.622
	(1.458)	(1.336)
Constant/Reference	12.805***	10.201***
	(1.031)	(0.945)
Observations	340	272
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 6: Absolute perception error models

4.4 Dependent variable: ACCEPTANCE

Table 7 shows the average *ACCEPTANCE*-score for each legend type. As expected, having no legend at all (Headline) have the lowest acceptance scores, while the Sampled and Annotated legend types have average acceptance scores above 4 out of a maximum of 5.

	Legend	Mean
1	Sampled	4.18
2	Annotated	4.06
3	Checkered	3.47
4	Clustered	3.09
5	Headline	2.09

Table 7: Average acceptance scores by legend type

As seen from the regression results in Table 8, all acceptance means are statistically higher than the Headline category, and both Sampled and Annotated legend types are statistically significantly considered more helpful than the Clustered legend type by the respondents. The Clustered legend type has a lower acceptance score compared to the Checkered legend, but not statistically significantly so.

	Dependent variable:	
	ACCEPTANCE	
	Headline Reference	Checkered Reference
	(1)	(2)
Annotated	1.970***	0.592**
	(0.228)	(0.230)
Checkered	1.378***	
	(0.230)	
Clustered	1.000***	−0.378
	(0.228)	(0.230)
Sampled	2.091***	0.713***
	(0.228)	(0.230)
Constant/Reference	2.091***	3.469***
	(0.161)	(0.164)
Observations	164	131
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 8: Acceptance-score models

4.5 Robustness tests

The variables that were collected as potential proxy variables for uncertainty (*SUBMIT TIME*, *INPUT CHANGES*, and *HOVER EVENTS*) exhibited high variability and large outliers, likely due to external factors. Separate regression models against the legend types are shown in Tables 11, 12 and 13 in Appendix A. *SUBMIT TIME* and *INPUT CHANGES* are significantly lower for the other legend types compared to Headline/no-legend. However, that was likely due to the first interaction with the UI was for the

Headline type by design. At this point, many respondents did ask questions and changed their input continuously before submitting the answer. Comparison between other categories did not give any statistically significant differences.

To check for clear learning or fatigue effects that would merit subsample analysis, an OLS model was run on |*PERCEPTION ERROR*| against the linear progression variable (discrete values expressed as integers 1-10). Another model using only the subsample of the first and last images, i.e. all Headline type responses was estimated to check if the respondents learned to memorize the opacity data mapping without the use of legends.

The results can be found in Table 9. As seen, neither the linear trend variable nor the categorical variable for the last image in the progression was statistically significant. The estimate is even showing a slightly higher error for the last example on average compared to the first for the gathered sample of 68 (34 respondents, 2 examples per respondent).

	Dependent variable:	
	|PERCEPTION ERROR|	
	Linear trend	Subsample fist/last, Reference=First
	(1)	(2)
ProgressionLinear	0.110	
	(0.162)	
Progression Last		1.834
		(2.655)
Constant/Reference	9.672***	11.888***
	(1.004)	(1.878)
Observations	340	68
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 9: Progression effects regression results

These results indicate that conducting further analysis on e.g. a subsample of the first 5 examples in the progression for each respondent is of little value, as it would most likely only result in halving the sample size.

A subsample of the 50th percentile on average more accurate respondents was created to evaluate if this would remove large error outliers and give more significant estimates. Identical models as in section 4.3, using |*PERCEPTION ERROR*| as dependent variable were estimated.

Visual inspection based on Figure 11 shows that there is still a large variance in the distribution. The Headline legend type still has the highest average absolute error. However, the Clustered, and not the Checkered legend design, has the second-highest average absolute error for the subsample.

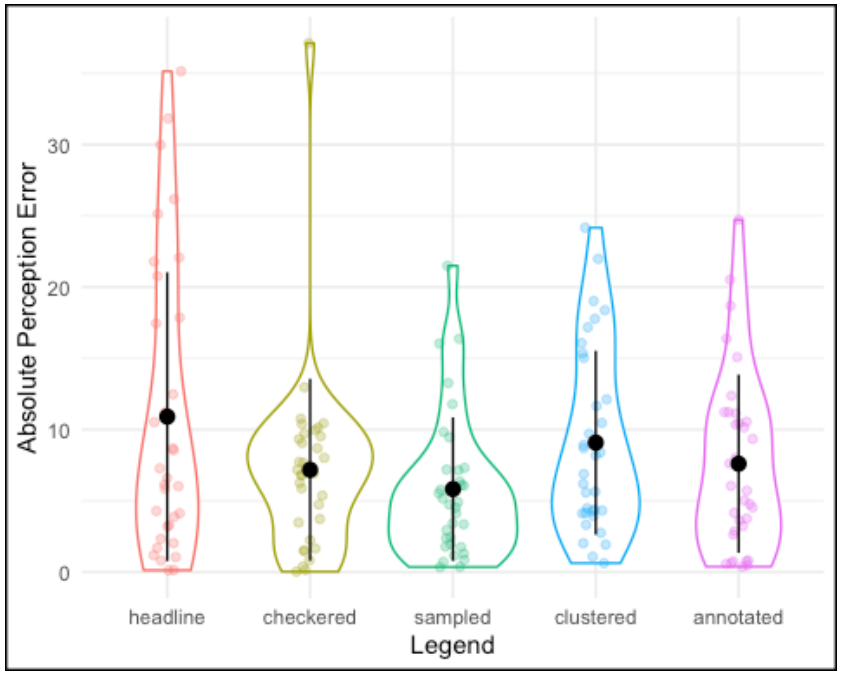


Figure 11: Absolute error - 50th percentile most accurate respondent subset

As seen in Table 10, which corresponds to Table 6 for the full sample, all estimates except for Clustered still have lower absolute errors compared to the Headline category that are statistically significant. For comparison, Checkered is still kept as a reference category in the models where Headline types have been excluded from the sample. No statistically significant results were estimated for model (2), as was the case when using the full sample.

	Dependent variable:	
	|PERCEPTION ERROR|	
	Headline Reference	Checkered Reference
	(1)	(2)
Annotated	−3.301* (1.715)	0.426 (1.471)
Checkered	−3.727** (1.715)	
Clustered	−1.824 (1.715)	1.904 (1.471)
Sampled	−5.083*** (1.715)	−1.356 (1.471)
Constant/Reference	10.903*** (1.212)	7.176*** (1.040)
Observations	170	136
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 10: Absolute perception error models - above 50th percentile accuracy subsample

5 Discussion

Both the Bartlett's test results and the dummy-variable OLS models are consistent in establishing that for visualizations using transparency mapped overlays on top of a static base map, there is an influence of legend choices - as to if legends are included or not. It is evident that there is both a lower variance of errors in estimations from the respondents and more accurate results when there is a legend compared to when no legend (Headline type) is present. In other words, a legend improves the ability to make visual comparisons and reduces estimation errors (answering **RQ1**).

There are no clear results indicating differences in response accuracy depending on the type/design of legend, and there is quite a large variation in response accuracy in the sample overall. These results indicate that the decoding support from a legend helps create a reference point for the user, but that it is not very precise or sensitive to the design of the legend used (**RQ2**). These results for **RQ1** and **RQ2** were robust to subsamples of respondents with average lower errors. There are no indications that there are learning effects where the users are able to memorize the opacity to value mappings without making use of the legends, meaning that the results from using the full sample can be taken at face value, and there is no obvious way to selectively remove observations.

The most robust results were from the user preferences/acceptance of the different legend types (**RQ3**), where both Sampled and Annotated (more contextualized and higher placement proximity to data) were preferred to the Checkered legend type ("industry standard" imitation).

For the Annotated legend design type, this may be due to the reduced need for "attention switching" (Kern et al., 2010) following higher proximity to the data and marker. Another factor that is relevant for both Sampled and Annotated designs is that there may be higher familiarity for the user, as they have the background of the map superimposed behind the legend bar for both these legend types.

Notably as well was that the number of times the users changed their responses before submitting by clicking a button, or the duration until the users submitted their responses from the visualization rendering was not well-suited to act as proxy variables for

respondent confidence or uncertainty (not being possible to draw any valid conclusions to in the study to **RQ4**).

During the data collection this may have been due to different respondents making use of the computer and UI-input elements in vastly different ways, and sometimes being distracted from the task or asking clarifying questions.

This resulted in that it was not possible to draw conclusions from the decision times and user selection behaviour for different legend designs. For future studies, this could likely be improved by using eye-tracking technology, where you could track that the duration the user is actively looking at the map, instead of trying to make use of proxy variables. This method has been used extensively in map design research (Brychtova and Coltekin, 2016; Çöltekin et al., 2009; Dong et al., 2014).

Regarding the general design of the study, it would be of interest to see how the results would differ when using different map backgrounds. Another variation that is common in use and has not been tested is having discrete (non-continuous) opacity-mapped data - i.e. where larger regions have the same opacity value. Users ability to perceive the opacity mapping where they have larger areas with the same value may differ from when there is pixel-level variation. This would also result in legends where the opacity values are binned, not showing the whole linear scale.

The sampling was performed using non-random convenience sampling. As outlined in section 3.2.2, this was chosen to keep external factors, such as screen size and lighting, as consistent as possible. Due to access to a suitable group of technically adept map users and practitioners, the results ought to be representative of how experienced map users can utilize different legend designs. Future studies could investigate legends for users with less experience of mapping products, as no claims can be made that certain legend designs are helpful for users that don't have a lot of prior exposure to common map elements.

In using only one colour for the data opacity layer for each example, the effects of colour scales being somewhat arbitrary, as described by Ware (2020) did not need to be accounted for. It is, however, common in visualizations to have a constant opacity for a polygon, and use a colour scale to encode values within the polygon, and it is also possible to have double encoding (have the same variable displayed using both opacity

variations and colour encodings simultaneously). For this study, it would have been difficult to isolate and measure the effect of transparency variations if using such encodings.

The different colours used for the examples (red, green and blue) did not result in differing decoding accuracy. One or more of the participants were likely to have some kind of colour blindness, as most of the participants were men, and about 1 in 12 men sees colour differently from most of the population (the same number for women is 1 in 200) (Lee et al., 2020). In this study, it is unclear if that would affect the difficulty of the decoding task, or increase issues with blending for specific colours. The respondents were not asked to distinguish between elements of different colours, but to evaluate opacity on top of a base map, where the combined colour blending results in general trends of contrast. If any systematic biases for colour blind decoding accuracy would exist for colour blending with the red, green or blue overlay and the base map, such effects may have become obvious from accuracy differences for different colours. However, that would likely require a much larger sample, or using a dedicated sample of users of with types of colour blindness that are heavily represented in any chosen examples.

Relating to Seipel et al. (2020), where the task was to distinguish objects there ought to be less complication from colour blending in this study, due to only one lower opacity data layer on top of a fully opaque base layer. With the data encoded layers, it is possible to distinguish trends over a larger area of the base map. The opacity layer is still blended with the base map. Hence, in lower opacity areas, you would get more blending with the base layer, which may affect precision in the estimation from the respondents. This may have created less precision in some estimations, but is a natural effect of using opacity/transparency as an encoding mechanism. There could be certain colour combinations from the base map that makes the decoding task more difficult, which would be of interest to evaluate in future studies. In this study, those effects ought to have been smoothed out by randomization, and not have biased the results in any particular direction.

In comparison to Kiik et al. (2017), the usage of only a single opacity layer did circumvent the issue of having blending from multiple low opacity layers.

In the larger literature on data visualization, there have been discussions on the usefulness of legends in general. Alternatives such as simplifying the visualizations, directly labelling the data, altering visual weight and making use of motion to remove the need for legends have been suggested (Evergreen and Metzner, 2013).

In the field of cartography, legends containing map symbols are commonly listed as one of the essential elements of a map together with title, scale bar, north arrow and a few other elements (Peterson, 2009, p. 17). This study showed that legends can effectively increase decoding accuracy for complex visualizations (**RQ1**), where the techniques described by Evergreen and Metzner (2013) may not be easy to implement.

The legend design variations created for this study to investigate a reduced need for "attention switching" through data-proximity from the legend (Evergreen and Metzner, 2013; Kern et al., 2010) did not result in significantly higher decoding accuracy (**RQ2**), but was one of the designs preferred by the respondents (**RQ3**).

6 Conclusions

When using transparency mapping, there were strong and statistically significant results that map designs using legends provide reduced visual estimation errors compared to when no legend is present (**RQ1**). This was concluded through variance analysis, dummy-variable OLS models and robustness tests, where all results provided clear reduced estimation errors when providing a legend for the user.

However, there were no statistically differing errors when comparing the different legend designs created for this study (**RQ2**). The inclusion of the legend for opacity mapped data seems to provide a reference point to the user, but none of the designs resulted in low enough variation for the differences in decoding errors to be statistically significant.

In terms of user preference/acceptance of the different legend designs Sampled and Annotated designs were preferred (**RQ3**). These were also the two designs that had the overall lowest estimation errors. These designs had in common that they put the data in context of the base map through either being directly overlayed on the map, or having a sampled rectangle from the map as background. In the case of the Annotated legend design, the distance between the legend and the data/marker was also reduced compared to the other legend designs.

The duration until the users submitted responses or the number of times they changed their response before submitting were influenced by too many external factors to provide useful inference to respondent confidence or uncertainty (**RQ4**). In future studies alternative methodology could alleviate these issues, as outlined in Section 5.

References

Adelson, E., 1993. Perceptual organization and the judgment of brightness. Science 262, 2042–2044. https://doi.org/10.1126/science.8266102

ArcGIS, n.d.a. Data-driven opacity [online] [WWW Document]. URL https://developers.arcgis.com/javascript/latest/sample-code/visualization-vv-opacity/ (accessed 2021-03-20).

ArcGIS, n.d.b. Create a custom visualization using arcade [online] [WWW Document]. URL https://developers.arcgis.com/javascript/latest/sample-code/visualization-arcade/index.html (accessed 2021-03-20).

Brehmer, M., Munzner, T., 2013. A multi-level typology of abstract visualization tasks. IEEE Transactions on Visualization and Computer Graphics 19, 2376–85. https://doi.org/10.1109/TVCG.2013.124

Brewer, C.A., 2006. Basic mapping principles for visualizing cancer data using geographic information systems (GIS). American Journal of Preventive Medicine 30, S25–S36. https://doi.org/10.1016/j.amepre.2005.09.007

Brewer, C.A., 1997. Spectral schemes: Controversial color use on maps. Cartography & Geographic Information Systems 24, 203.

Brychtova, A., Coltekin, A., 2016. An empirical user study for measuring the influence of colour distance and font size in map reading using eye tracking. Cartographic Journal 53, 202–212.

BuiltWith, n.d. Maps usage distribution in the top 100k sites [online] [WWW Document]. URL https://trends.builtwith.com/mapping/maps/traffic/Top-100k (accessed 2021-07-16).

Cheong, L., Bleisch, S., Kealy, A., Tolhurst, K., Wilkening, T., Duckham, M., 2016. Evaluating the impact of visualization of wildfire hazard upon decision-making under uncertainty. International Journal of Geographical Information Science 30, 1377–1404.

Chien, T., Wang, H., Hsu, C., Kuo, S., 2019. Choropleth map legend design for visualizing the most influential areas in article citation disparities: A bibliometric study. Medicine (Baltimore) 98. https://doi.org/10.1097/MD.0000000000017527

Ciechanowski, B., 2019. Alpha compositing [online] [WWW Document]. URL https://ciechanow.ski/alpha-compositing/ (accessed 2021-02-20).

Cleveland, W.S., McGill, R., 1985. Graphical perception and graphical methods for analyzing scientific data. Science 229, 828–833. https://doi.org/10.1126/science.229.4716.828

Cromley, R.G., Cromley, E.K., 2009. Choropleth map legend design for visualizing community health disparities. International Journal of Health Geographics 8. https://doi.org/10.1186/1476-072X-8-52

Cromley, R.G., Ye, Y., 2006. Ogive-based legends for choropleth mapping. Cartography and Geographic Information Science 33, 257–268. https://doi.org/10.1559/152304006779500650

Cybulski, P., 2016. Design rules and practices for animated maps online. Journal of Spatial Science 61, 461–471. https://doi.org/10.1080/14498596.2016.1147394

Çöltekin, A., Heil, B., Garlandini, S., Fabrikant, S.I., 2009. Evaluating the effectiveness of interactive map interface designs: A case study integrating usability metrics with eye-movement analysis. Cartography and Geographic Information Science 36, 5–17.

Dong, W., Liao, H., Roth, R.E., Wang, S., 2014. Eye tracking to explore the potential of enhanced imagery basemaps in web mapping. Cartographic Journal 51, 313–329.

ESRI, n.d.a. GIS software | ArcGIS products for the cloud, mobile apps & desktop [WWW Document]. URL https://www.esri.com/en-us/arcgis/products/index (accessed 2020-09-13).

ESRI, n.d.b. Enterprise GIS system | ArcGIS enterprise - geospatial platform [WWW Document]. URL https://www.esri.com/en-us/arcgis/products/arcgis-enterprise/overview (accessed 2020-09-13).

Evergreen, S., Metzner, C., 2013. Design principles for data visualization in evaluation.

Findlay, J.M., Gilchrist, I.D., 1998. Chapter 13 - eye guidance and visual search 295–312. https://doi.org/10.1016/B978-008043361-5/50014-6

Harrower, M., Brewer, C.A., 2003. ColorBrewer.org: An online tool for selecting colour schemes for maps. Cartographic Journal 40, 27–37.

Jo, J., Vernier, F., Dragicevic, P., Fekete, J., 2019. A declarative rendering model for multiclass density maps. IEEE Transactions on Visualization and Computer Graphics 25, 470–480. https://doi.org/10.1109/TVCG.2018.2865141

Kern, D., Marshall, P., Schmidt, A., 2010. Gazemarks: Gaze-based visual placeholders to ease attention switching, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '10. Association for Computing Machinery, New York, NY, USA, pp. 2093–2102. https://doi.org/10.1145/1753326.1753646

Kiik, A., Nyström, M., Harrie, L., 2017. Cartographic design matters : A comparison of thematic polygon design. Cartographic Journal 54, 24–35.

Kim, Y.-S., Reinecke, K., Hullman, J., 2017. Explaining the gap: Visualizing one's predictions improves recall and comprehension of data, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17. Association for Computing Machinery, New York, NY, USA, pp. 1375–1386. https://doi.org/10.1145/3025453.3025592

Kim, Y.-S., Walls, L.A., Krafft, P., Hullman, J., 2019. A bayesian cognition approach to improve data visualization, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19. Association for Computing Machinery, New York, NY, USA, pp. 1–14. https://doi.org/10.1145/3290605.3300912

Kumar, N., 2004. Frequency histogram legend in the choropleth map: A substitute to traditional legends. Cartography and Geographic Information Science 31, 217–236. https://doi.org/10.1559/1523040042742411

Lee, H., Lee, E., Choi, G., 2020. Wayfinding signage for people with color blindness. Journal of Interior Design 45, 35–54.

Leitner, M., Buttenfield, B.P., 2000. Guidelines for the display of attribute certainty. Cartography and Geographic Information Science 27, 3–14.

Montello, D.R., 2002. Cognitive map-design research in the twentieth century: Theoretical and empirical approaches. Cartography and Geographic Information Science 29, 283–304.

Panko, R., n.d. The popularity of google maps: Trends in navigation apps in 2018 [WWW Document]. URL https://themanifest.com/mobile-apps/popularity-google-maps-trends-navigation-apps-2018 (accessed 2021-07-16).

Perlin, K., 1985. An image synthesizer, in: Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '85. Association for Computing Machinery, New York, NY, USA, pp. 287–296. https://doi.org/10.1145/325334.325247

Peterson, G.N., 2009. GIS cartography : A guide to effective map design. CRC Press.

Peterson, M.P., 1999. Active legends for interactive cartographic animation. International Journal of Geographical Information Science 13, 375–383. https://doi.org/10.1080/136588199241256

Porter, T., Duff, T., 1984. Compositing digital images. Proceedings of the 11th Annual Conference: Computer Graphics & Interactive Techniques 253–259.

Qin, Z., Li, Z., 2017. Grouping rules for effective legend design. The Cartographic Journal 54, 36–47. https://doi.org/10.1080/00087041.2016.1148105

Seipel, S., Andrée, M., Larsson, K., Paasch, J.M., Paulsson, J., 2020. Visualization of 3D property data and assessment of the impact of rendering attributes. Journal of Geovisualization and Spatial Analysis 4, 1–17.

Seipel, S., Lim, N.J., 2017. Color map design for visualization in flood risk assessment. International Journal of Geographical Information Science 31, 2286–2309.

Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). Biometrika 52, 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Snedecor, G.W., Cochran, W.G., 1989. Statistical methods. Iowa State University Press.

Treisman, A., 1985. Preattentive processing in vision. Computer Vision, Graphics, and Image Processing 31, 156–177. https://doi.org/https://doi.org/10.1016/S0734-189X(85)80004-9

Tufte, E.R., 2001. The visual display of quantitative information, 2nd ed. Graphics Press, Cheshire, CT.

Veenendaal, B., Brovelli, M.A., Songnian, L., 2017. Review of Web Mapping: Eras, Trends and Directions. International Journal of Geo-Information 6, 317. https://doi.org/10.3390/ijgi6100317

Wang, C., Pouliot, J., Hubert, F., 2017. How users perceive transparency in the 3D visualization of cadastre: Testing its usability in an online questionnaire. GeoInformatica 21, 599–618.

Ware, C., 2020. Information visualization: Perception for design, 4th ed. Elsevier.

Wolfe, J.M., Horowitz, T.S., 2004. What attributes guide the deployment of visual attention and how do they do it? Nature Reviews Neuroscience 5, 495–501. https://doi.org/https://doi.org/10.1038/nrn1411

Appendices

Appendix A - plots and tables

Note: All analysis code and data for reproducing results can be found at https://github.com/Tille88/thesis-data-analysis

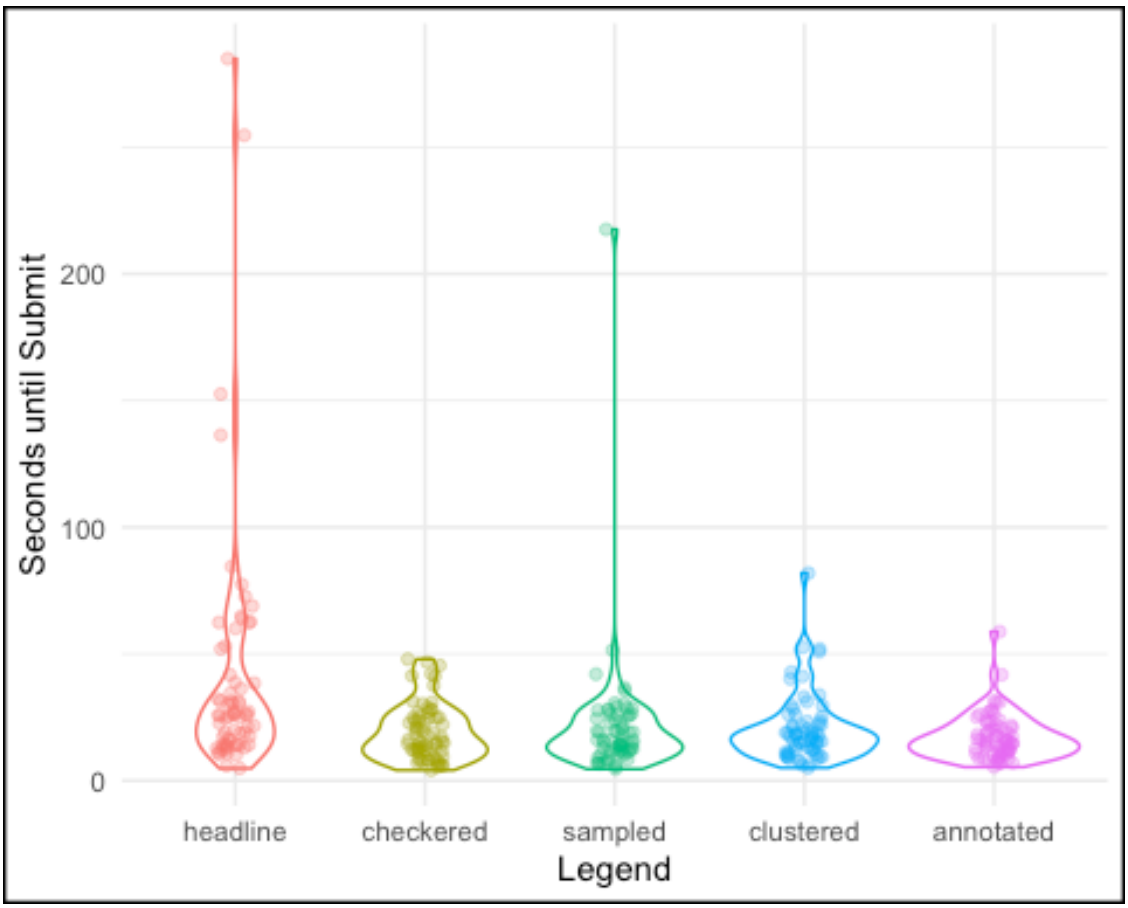


Figure 12: Submit time across legend types

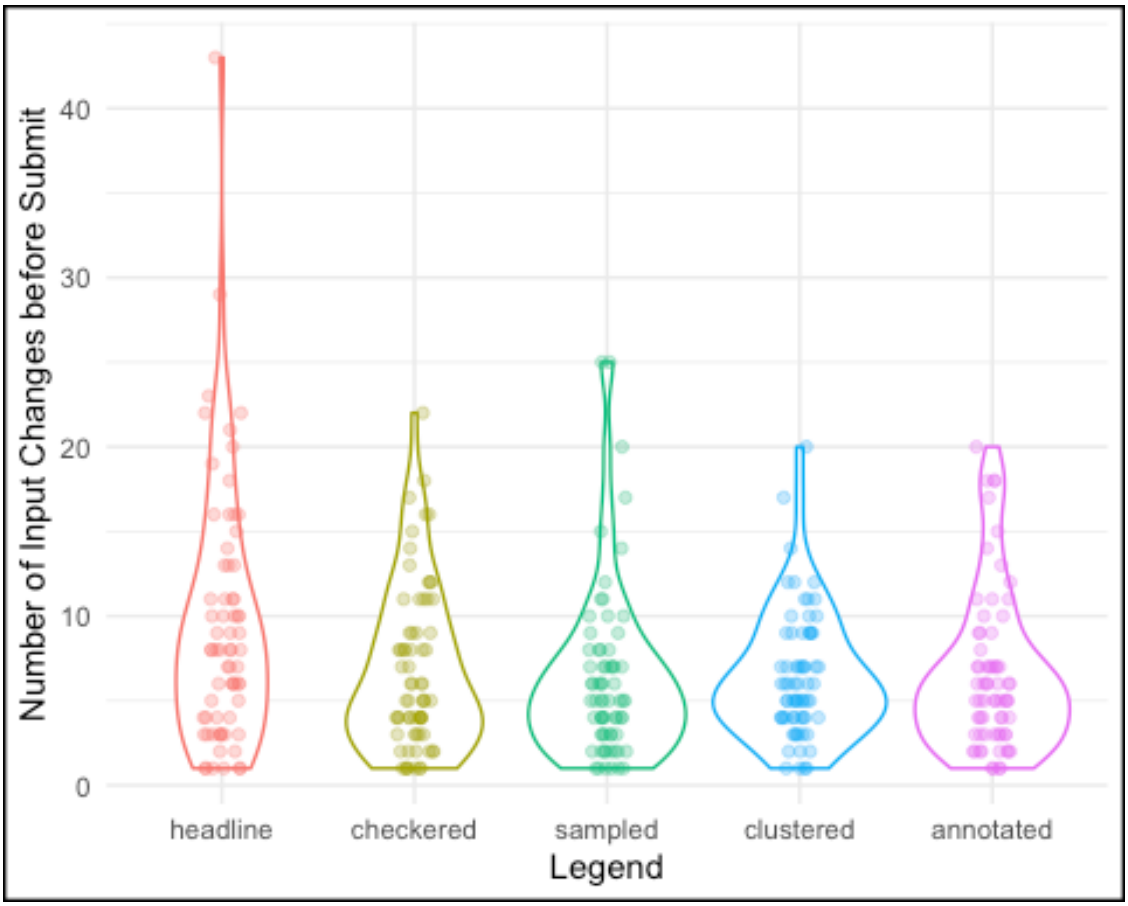


Figure 13: Input changes across legend types

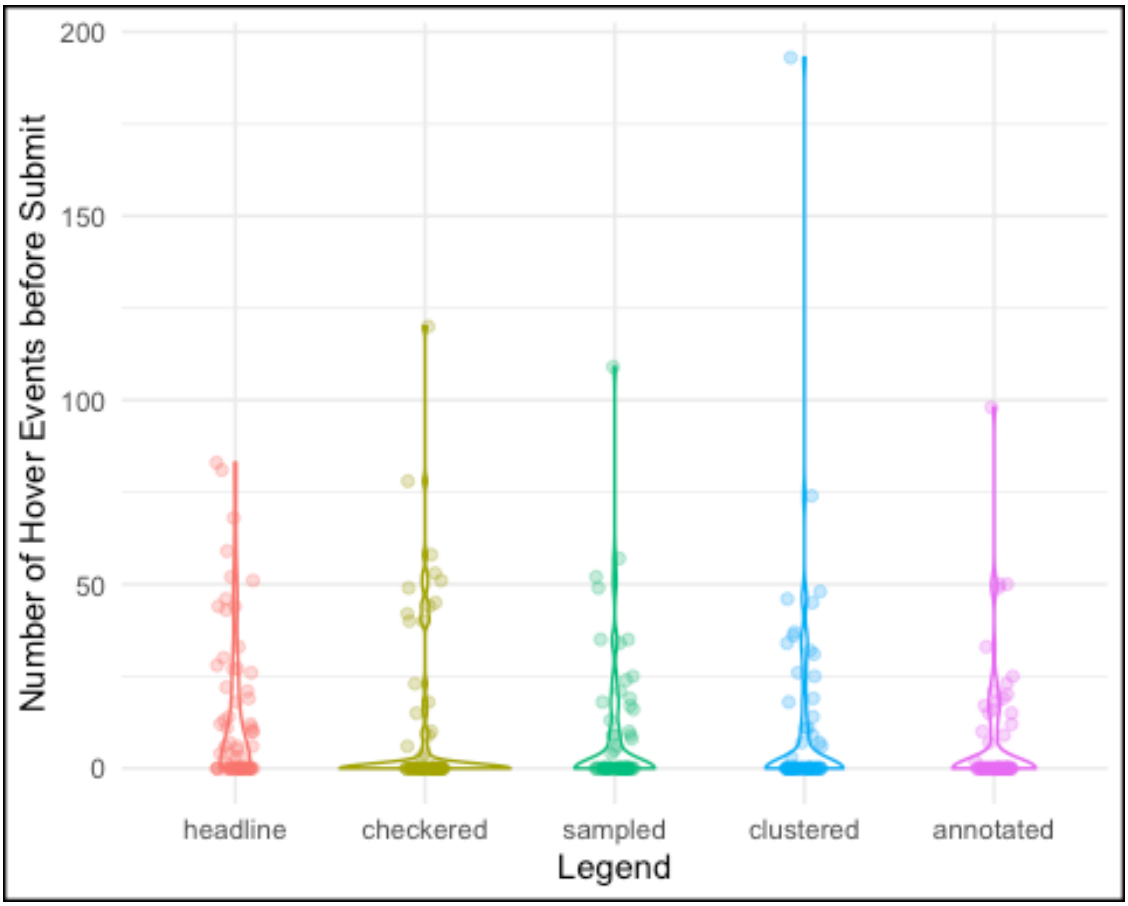


Figure 14: Hover events across legend types

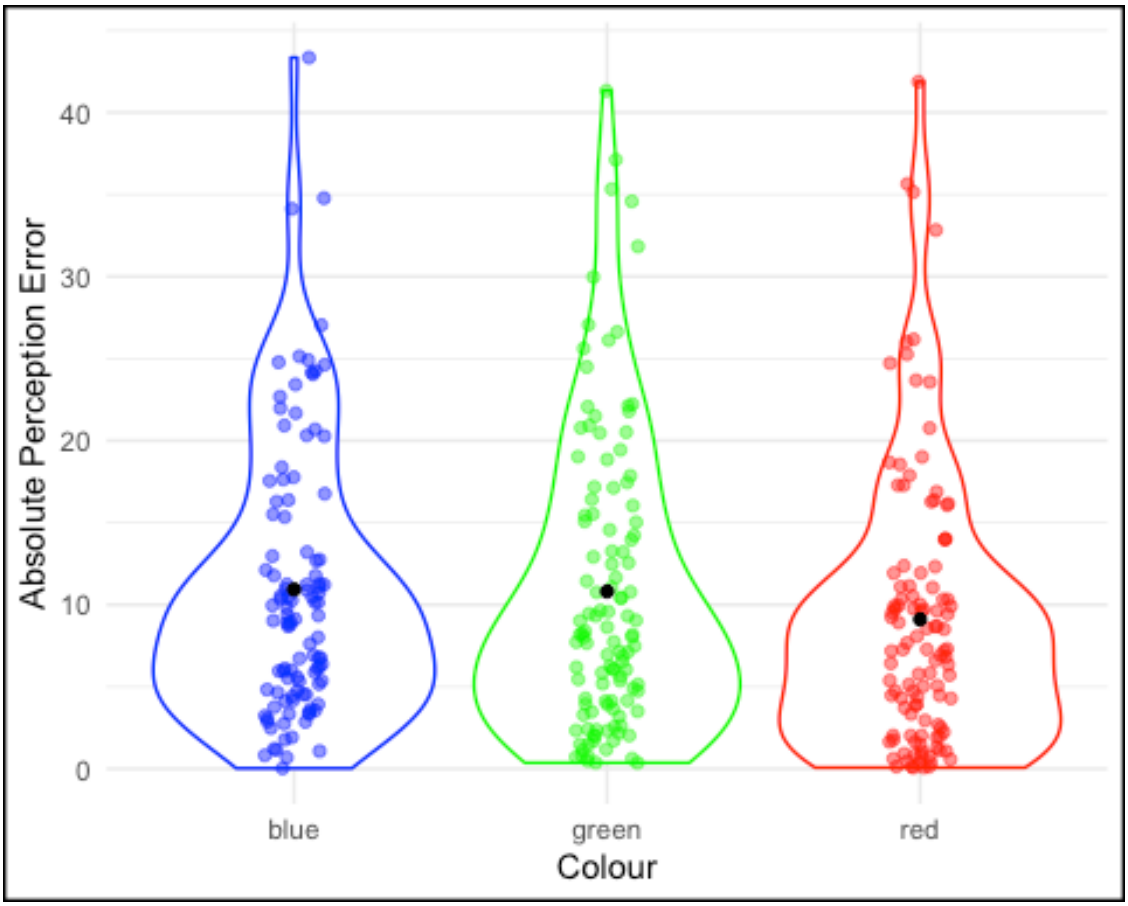


Figure 15: Absolute perception error by colour types

	Dependent variable:	
	SUBMIT TIME	
	Headline Reference	Checkered Reference
	(1)	(2)
Annotated	−22.908***	−1.913
	(4.480)	(2.771)
Checkered	−20.995***	
	(4.480)	
Clustered	−19.662***	1.333
	(4.480)	(2.771)
Sampled	−19.033***	1.961
	(4.480)	(2.771)
Constant/Reference	40.162***	19.167***
	(3.168)	(1.959)
Observations	340	272
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 11: Time to submit regression results

	Dependent variable:	
	INPUT CHANGES	
	Headline Reference	Checkered Reference
	(1)	(2)
Annotated	−2.926***	−0.235
	(0.911)	(0.786)
Checkered	−2.691***	
	(0.911)	
Clustered	−3.044***	−0.353
	(0.911)	(0.786)
Sampled	−3.118***	−0.426
	(0.911)	(0.786)
Constant/Reference	9.456***	6.765***
	(0.644)	(0.556)
Observations	340	272
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 12: Input changes regression results

	Dependent variable:	
	HOVER EVENTS	
	Headline Reference	Checkered Reference
	(1)	(2)
Annotated	−6.632*	−2.926
	(3.676)	(3.703)
Checkered	−3.706	
	(3.676)	
Clustered	−3.353	0.353
	(3.676)	(3.703)
Sampled	−5.529	−1.824
	(3.676)	(3.703)
Constant/Reference	14.118***	10.412***
	(2.600)	(2.619)
Observations	340	272
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 13: Hover events regression results

	Dependent variable:
	|PERCEPTION ERROR|
	Full sample
Green	−0.122
	(1.138)
Red	−1.821
	(1.148)
Constant/Reference = Blue	10.928***
	(0.822)
Observations	340
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 14: Absolute perception error by colour

Series from Lund University

Department of Physical Geography and Ecosystem Science

Master Thesis in Geographical Information Science

1. *Anthony Lawther:* The application of GIS-based binary logistic regression for slope failure susceptibility mapping in the Western Grampian Mountains, Scotland (2008).

2. *Rickard Hansen:* Daily mobility in Grenoble Metropolitan Region, France. Applied GIS methods in time geographical research (2008).

3. *Emil Bayramov:* Environmental monitoring of bio-restoration activities using GIS and Remote Sensing (2009).

4. *Rafael Villarreal Pacheco:* Applications of Geographic Information Systems as an analytical and visualization tool for mass real estate valuation: a case study of Fontibon District, Bogota, Columbia (2009).

5. *Siri Oestreich Waage:* a case study of route solving for oversized transport: The use of GIS functionalities in transport of transformers, as part of maintaining a reliable power infrastructure (2010).

6. *Edgar Pimiento:* Shallow landslide susceptibility – Modelling and validation (2010).

7. *Martina Schäfer:* Near real-time mapping of floodwater mosquito breeding sites using aerial photographs (2010).

8. *August Pieter van Waarden-Nagel:* Land use evaluation to assess the outcome of the programme of rehabilitation measures for the river Rhine in the Netherlands (2010).

9. *Samira Muhammad:* Development and implementation of air quality data mart for Ontario, Canada: A case study of air quality in Ontario using OLAP tool. (2010).

10. *Fredros Oketch Okumu*: Using remotely sensed data to explore spatial and temporal relationships between photosynthetic productivity of vegetation and malaria transmission intensities in selected parts of Africa (2011).

11. *Svajunas Plunge:* Advanced decision support methods for solving diffuse water pollution problems (2011).

12. *Jonathan Higgins:* Monitoring urban growth in greater Lagos: A case study using GIS to monitor the urban growth of Lagos 1990 - 2008 and produce future growth prospects for the city (2011).

13. *Mårten Karlberg:* Mobile Map Client API: Design and Implementation for Android (2011).

14. *Jeanette McBride:* Mapping Chicago area urban tree canopy using color infrared imagery (2011).

15. *Andrew Farina:* Exploring the relationship between land surface temperature and vegetation abundance for urban heat island mitigation in Seville, Spain (2011).

16. *David Kanyari*: Nairobi City Journey Planner: An online and a Mobile Application (2011).

17. *Laura V. Drews:* Multi-criteria GIS analysis for siting of small wind power plants - A case study from Berlin (2012).

18. *Qaisar Nadeem:* Best living neighborhood in the city - A GIS based multi criteria evaluation of ArRiyadh City (2012).

19. *Ahmed Mohamed El Saeid Mustafa:* Development of a photo voltaic building rooftop integration analysis tool for GIS for Dokki District, Cairo, Egypt (2012).

20. *Daniel Patrick Taylor*: Eastern Oyster Aquaculture: Estuarine Remediation via Site Suitability and Spatially Explicit Carrying Capacity Modeling in Virginia's Chesapeake Bay (2013).

21. *Angeleta Oveta Wilson:* A Participatory GIS approach to *unearthing* Manchester's Cultural Heritage '*gold mine'* (2013).

22. *Ola Svensson:* Visibility and Tholos Tombs in the Messenian Landscape: A Comparative Case Study of the Pylian Hinterlands and the Soulima Valley (2013).

23. *Monika Ogden:* Land use impact on water quality in two river systems in South Africa (2013).

24. *Stefan Rova:* A GIS based approach assessing phosphorus load impact on Lake Flaten in Salem, Sweden (2013).

25. *Yann Buhot:* Analysis of the history of landscape changes over a period of 200 years. How can we predict past landscape pattern scenario and the impact on habitat diversity? (2013).

26. *Christina Fotiou:* Evaluating habitat suitability and spectral heterogeneity models to predict weed species presence (2014).

27. *Inese Linuza:* Accuracy Assessment in Glacier Change Analysis (2014).

28. *Agnieszka Griffin:* Domestic energy consumption and social living standards: a GIS analysis within the Greater London Authority area (2014).

29. *Brynja Guðmundsdóttir:* Detection of potential arable land with remote sensing and GIS - A Case Study for Kjósarhreppur (2014).

30. *Oleksandr Nekrasov:* Processing of MODIS Vegetation Indices for analysis of agricultural droughts in the southern Ukraine between the years 2000-2012 (2014).

31. *Sarah Tressel:* Recommendations for a polar Earth science portal in the context of Arctic Spatial Data Infrastructure (2014).

32. *Caroline Gevaert:* Combining Hyperspectral UAV and Multispectral Formosat-2 Imagery for Precision Agriculture Applications (2014).

33. *Salem Jamal-Uddeen:* Using GeoTools to implement the multi-criteria evaluation analysis - weighted linear combination model (2014).

34. *Samanah Seyedi-Shandiz:* Schematic representation of geographical railway network at the Swedish Transport Administration (2014).

35. *Kazi Masel Ullah:* Urban Land-use planning using Geographical Information System and analytical hierarchy process: case study Dhaka City (2014).

36. *Alexia Chang-Wailing Spitteler:* Development of a web application based on MCDA and GIS for the decision support of river and floodplain rehabilitation projects (2014).

37. *Alessandro De Martino:* Geographic accessibility analysis and evaluation of potential changes to the public transportation system in the City of Milan (2014).

38. *Alireza Mollasalehi:* GIS Based Modelling for Fuel Reduction Using Controlled Burn in Australia. Case Study: Logan City, QLD (2015).

39. *Negin A. Sanati:* Chronic Kidney Disease Mortality in Costa Rica; Geographical Distribution, Spatial Analysis and Non-traditional Risk Factors (2015).

40. *Karen McIntyre:* Benthic mapping of the Bluefields Bay fish sanctuary, Jamaica (2015).

41. *Kees van Duijvendijk:* Feasibility of a low-cost weather sensor network for agricultural purposes: A preliminary assessment (2015).

42. *Sebastian Andersson Hylander:* Evaluation of cultural ecosystem services using GIS (2015).

43. *Deborah Bowyer:* Measuring Urban Growth, Urban Form and Accessibility as Indicators of Urban Sprawl in Hamilton, New Zealand (2015).

44. *Stefan Arvidsson:* Relationship between tree species composition and phenology extracted from satellite data in Swedish forests (2015).

45. *Damián Giménez Cruz*: GIS-based optimal localisation of beekeeping in rural Kenya (2016).

46. *Alejandra Narváez Vallejo:* Can the introduction of the topographic indices in LPJ-GUESS improve the spatial representation of environmental variables? (2016).

47. *Anna Lundgren:* Development of a method for mapping the highest coastline in Sweden using breaklines extracted from high resolution digital elevation models (2016).

48. *Oluwatomi Esther Adejoro:* Does location also matter? A spatial analysis of social achievements of young South Australians (2016).

49. *Hristo Dobrev Tomov:* Automated temporal NDVI analysis over the Middle East for the period 1982 - 2010 (2016).

50. *Vincent Muller:* Impact of Security Context on Mobile Clinic Activities A GIS Multi Criteria Evaluation based on an MSF Humanitarian Mission in Cameroon (2016).

51. *Gezahagn Negash Seboka:* Spatial Assessment of NDVI as an Indicator of Desertification in Ethiopia using Remote Sensing and GIS (2016).

52. *Holly Buhler:* Evaluation of Interfacility Medical Transport Journey Times in Southeastern British Columbia. (2016).

53. *Lars Ole Grottenberg*: Assessing the ability to share spatial data between emergency management organisations in the High North (2016).

54. *Sean Grant:* The Right Tree in the Right Place: Using GIS to Maximize the Net Benefits from Urban Forests (2016).

55. *Irshad Jamal:* Multi-Criteria GIS Analysis for School Site Selection in Gorno-Badakhshan Autonomous Oblast, Tajikistan (2016).

56. *Fulgencio Sanmartín:* Wisdom-volkano: A novel tool based on open GIS and time-series visualization to analyse and share volcanic data (2016).

57. *Nezha Acil:* Remote sensing-based monitoring of snow cover dynamics and its influence on vegetation growth in the Middle Atlas Mountains (2016).

58. *Julia Hjalmarsson:* A Weighty Issue: Estimation of Fire Size with Geographically Weighted Logistic Regression (2016).

59. *Mathewos Tamiru Amato:* Using multi-criteria evaluation and GIS for chronic food and nutrition insecurity indicators analysis in Ethiopia (2016).

60. *Karim Alaa El Din Mohamed Soliman El Attar:* Bicycling Suitability in Downtown, Cairo, Egypt (2016).

61. *Gilbert Akol Echelai:* Asset Management: Integrating GIS as a Decision Support Tool in Meter Management in National Water and Sewerage Corporation (2016).

62. *Terje Slinning:* Analytic comparison of multibeam echo soundings (2016).

63. *Gréta Hlín Sveinsdóttir:* GIS-based MCDA for decision support: A framework for wind farm siting in Iceland (2017).

64. *Jonas Sjögren:* Consequences of a flood in Kristianstad, Sweden: A GIS-based analysis of impacts on important societal functions (2017).

65. *Nadine Raska:* 3D geologic subsurface modelling within the Mackenzie Plain, Northwest Territories, Canada (2017).

66. *Panagiotis Symeonidis*: Study of spatial and temporal variation of atmospheric optical parameters and their relation with PM 2.5 concentration over Europe using GIS technologies (2017).

67. *Michaela Bobeck:* A GIS-based Multi-Criteria Decision Analysis of Wind Farm Site Suitability in New South Wales, Australia, from a Sustainable Development Perspective (2017).

68. *Raghdaa Eissa*: Developing a GIS Model for the Assessment of Outdoor Recreational Facilities in New Cities Case Study: Tenth of Ramadan City, Egypt (2017).

69. *Zahra Khais Shahid*: Biofuel plantations and isoprene emissions in Svea and Götaland (2017).

70. *Mirza Amir Liaquat Baig*: Using geographical information systems in epidemiology: Mapping and analyzing occurrence of diarrhea in urban - residential area of Islamabad, Pakistan (2017).

71. *Joakim Jörwall*: Quantitative model of Present and Future well-being in the EU-28: A spatial Multi-Criteria Evaluation of socioeconomic and climatic comfort factors (2017).

72. *Elin Haettner*: Energy Poverty in the Dublin Region: Modelling Geographies of Risk (2017).

73. *Harry Eriksson*: Geochemistry of stream plants and its statistical relations to soil- and bedrock geology, slope directions and till geochemistry. A GIS-analysis of small catchments in northern Sweden (2017).

74. *Daniel Gardevärn:* PPGIS and Public meetings – An evaluation of public participation methods for urban planning (2017).

75. *Kim Friberg:* Sensitivity Analysis and Calibration of Multi Energy Balance Land Surface Model Parameters (2017).

76. *Viktor Svanerud:* Taking the bus to the park? A study of accessibility to green areas in Gothenburg through different modes of transport (2017).

77. *Lisa-Gaye Greene*: Deadly Designs: The Impact of Road Design on Road Crash Patterns along Jamaica's North Coast Highway (2017).

78. *Katarina Jemec Parker*: Spatial and temporal analysis of fecal indicator bacteria concentrations in beach water in San Diego, California (2017).

79. *Angela Kabiru*: An Exploratory Study of Middle Stone Age and Later Stone Age Site Locations in Kenya's Central Rift Valley Using Landscape Analysis: A GIS Approach (2017).

80. *Kristean Björkmann*: Subjective Well-Being and Environment: A GIS-Based Analysis (2018).

81. *Williams Erhunmonmen Ojo*: Measuring spatial accessibility to healthcare for people living with HIV-AIDS in southern Nigeria (2018).

82. *Daniel Assefa*: Developing Data Extraction and Dynamic Data Visualization (Styling) Modules for Web GIS Risk Assessment System (WGRAS). (2018).

83. *Adela Nistora*: Inundation scenarios in a changing climate: assessing potential impacts of sea-level rise on the coast of South-East England (2018).

84. *Marc Seliger*: Thirsty landscapes - Investigating growing irrigation water consumption and potential conservation measures within Utah's largest master-planned community: Daybreak (2018).

85. *Luka Jovičić*: Spatial Data Harmonisation in Regional Context in Accordance with INSPIRE Implementing Rules (2018).

86. *Christina Kourdounouli*: Analysis of Urban Ecosystem Condition Indicators for the Large Urban Zones and City Cores in EU (2018).

87. *Jeremy Azzopardi*: Effect of distance measures and feature representations on distance-based accessibility measures (2018).

88. *Patrick Kabatha*: An open source web GIS tool for analysis and visualization of elephant GPS telemetry data, alongside environmental and anthropogenic variables (2018).

89. *Richard Alphonce Giliba*: Effects of Climate Change on Potential Geographical Distribution of Prunus africana (African cherry) in the Eastern Arc Mountain Forests of Tanzania (2018).

90. *Eiður Kristinn Eiðsson*: Transformation and linking of authoritative multi-scale geodata for the Semantic Web: A case study of Swedish national building data sets (2018).

91. *Niamh Harty*: HOP!: a PGIS and citizen science approach to monitoring the condition of upland paths (2018).

92. *José Estuardo Jara Alvear*: Solar photovoltaic potential to complement hydropower in Ecuador: A GIS-based framework of analysis (2018).

93. *Brendan O'Neill*: Multicriteria Site Suitability for Algal Biofuel Production Facilities (2018).

94. *Roman Spataru*: Spatial-temporal GIS analysis in public health – a case study of polio disease (2018).

95. *Alicja Miodońska*: Assessing evolution of ice caps in Suðurland, Iceland, in years 1986 - 2014, using multispectral satellite imagery (2019).

96. *Dennis Lindell Schettini*: A Spatial Analysis of Homicide Crime's Distribution and Association with Deprivation in Stockholm Between 2010-2017 (2019).

97. *Damiano Vesentini*: The Po Delta Biosphere Reserve: Management challenges and priorities deriving from anthropogenic pressure and sea level rise (2019).

98. *Emilie Arnesten*: Impacts of future sea level rise and high water on roads, railways and environmental objects: a GIS analysis of the potential effects of increasing sea levels and highest projected high water in Scania, Sweden (2019).

99. *Syed Muhammad Amir Raza*: Comparison of geospatial support in RDF stores: Evaluation for ICOS Carbon Portal metadata (2019).

100. *Hemin Tofiq*: Investigating the accuracy of Digital Elevation Models from UAV images in areas with low contrast: A sandy beach as a case study (2019).

101. *Evangelos Vafeiadis*: Exploring the distribution of accessibility by public transport using spatial analysis. A case study for retail concentrations and public hospitals in Athens (2019).

102. *Milan Sekulic*: Multi-Criteria GIS modelling for optimal alignment of roadway by-passes in the Tlokweng Planning Area, Botswana (2019).

103. *Ingrid Piirisaar*: A multi-criteria GIS analysis for siting of utility-scale photovoltaic solar plants in county Kilkenny, Ireland (2019).

104. *Nigel Fox*: Plant phenology and climate change: possible effect on the onset of various wild plant species' first flowering day in the UK (2019).

105. *Gunnar Hesch*: Linking conflict events and cropland development in Afghanistan, 2001 to 2011, using MODIS land cover data and Uppsala Conflict Data Programme (2019).

106. *Elijah Njoku*: Analysis of spatial-temporal pattern of Land Surface Temperature (LST) due to NDVI and elevation in Ilorin, Nigeria (2019).

107. *Katalin Bunyevácz*: Development of a GIS methodology to evaluate informal urban green areas for inclusion in a community governance program (2019).

108. *Paul dos Santos*: Automating synthetic trip data generation for an agent-based simulation of urban mobility (2019).

109. *Robert O' Dwyer*: Land cover changes in Southern Sweden from the mid-Holocene to present day: Insights for ecosystem service assessments (2019).

110. *Daniel Klingmyr*: Global scale patterns and trends in tropospheric NO2 concentrations (2019).

111. *Marwa Farouk Elkabbany*: Sea Level Rise Vulnerability Assessment for Abu Dhabi, United Arab Emirates (2019).

112. *Jip Jan van Zoonen*: Aspects of Error Quantification and Evaluation in Digital Elevation Models for Glacier Surfaces (2020).

113. *Georgios Efthymiou*: The use of bicycles in a mid-sized city – benefits and obstacles identified using a questionnaire and GIS (2020).

114. *Haruna Olayiwola Jimoh*: Assessment of Urban Sprawl in MOWE/IBAFO Axis of Ogun State using GIS Capabilities (2020).

115. *Nikolaos Barmpas Zachariadis*: Development of an iOS, Augmented Reality for disaster management (2020).

116. *Ida Storm*: ICOS Atmospheric Stations: Spatial Characterization of CO2 Footprint Areas and Evaluating the Uncertainties of Modelled CO2 Concentrations (2020).

117. *Alon Zuta*: Evaluation of water stress mapping methods in vineyards using airborne thermal imaging (2020).

118. *Marcus Eriksson*: Evaluating structural landscape development in the municipality Upplands-Bro, using landscape metrics indices (2020).

119. *Ane Rahbek Vierø*: Connectivity for Cyclists? A Network Analysis of Copenhagen's Bike Lanes (2020).

120. *Cecilia Baggini*: Changes in habitat suitability for three declining Anatidae species in saltmarshes on the Mersey estuary, North-West England (2020).

121. *Bakrad Balabanian*: Transportation and Its Effect on Student Performance (2020).

122. *Ali Al Farid*: Knowledge and Data Driven Approaches for Hydrocarbon Microseepage Characterizations: An Application of Satellite Remote Sensing (2020).

123. *Bartlomiej Kolodziejczyk*: Distribution Modelling of Gene Drive-Modified Mosquitoes and Their Effects on Wild Populations (2020).

124. *Alexis Cazorla*: Decreasing organic nitrogen concentrations in European water bodies - links to organic carbon trends and land cover (2020).

125. *Kharid Mwakoba*: Remote sensing analysis of land cover/use conditions of community-based wildlife conservation areas in Tanzania (2021).

126. *Chinatsu Endo*: Remote Sensing Based Pre-Season Yellow Rust Early Warning in Oromia, Ethiopia (2021).

127. *Berit Mohr*: Using remote sensing and land abandonment as a proxy for long-term human out-migration. A Case Study: Al-Hassakeh Governorate, Syria (2021).

128. *Kanchana Nirmali Bandaranayake*: Considering future precipitation in delineation locations for water storage systems - Case study Sri Lanka (2021).

129. *Emma Bylund*: Dynamics of net primary production and food availability in the aftermath of the 2004 and 2007 desert locust outbreaks in Niger and Yemen (2021).

130. *Shawn Pace*: Urban infrastructure inundation risk from permanent sea-level rise scenarios in London (UK), Bangkok (Thailand) and Mumbai (India): A comparative analysis (2021).

131. *Oskar Evert Johansson*: The hydrodynamic impacts of Estuarine Oyster reefs, and the application of drone technology to this study (2021).

132. *Pritam Kumarsingh*: A Case Study to develop and test GIS/SDSS methods to assess the production capacity of a Cocoa Site in Trinidad and Tobago (2021).

133. *Muhammad Imran Khan*: Property Tax Mapping and Assessment using GIS (2021).

134. *Domna Kanari*: Mining geosocial data from Flickr to explore tourism patterns: The case study of Athens (2021).

135. Mona Tykesson Klubien: Livestock-MRSA in Danish pig farms (2021).

136. *Ove Njøten*: Comparing radar satellites. Use of Sentinel-1 leads to an increase in oil spill alerts in Norwegian waters (2021).

137. *Panagiotis Patrinos*: Change of heating fuel consumption patterns produced by the economic crisis in Greece (2021).

138. *Lukasz Langowski*: Assessing the suitability of using Sentinel-1A SAR multi-temporal imagery to detect fallow periods between rice crops (2021).

139. *Jonas Tillman*: Perception accuracy and user acceptance of legend designs for opacity data mapping in GIS (2022).