university of groningen

faculty of science and engineering

Particle fluid simulation: visual design for effective user engagement

Bachelor's Project Computing Science

July 2024

Author: Djairo Hougee
Student Number: S4764153
First supervisor: Dr. mat. nat. Christian Kehl
Second supervisor: Dr. Julian Köllermeier

Abstract

Mismanagement of marine waste is a threat to delicate ecosystems such as the Waddenzee. A key aspect of managing such marine waste are volunteer-based beach cleanup actions, yet these actions depend on public interest on this subject. Interactive flow visualisations can provide a way for the public to engage with this important environmental issue, encouraging and maintaining management of marine waste in the public consciousness. To maximise the effectiveness of this engagement, research into effective design for particle flow visualisation is required.

This project performs a small-scale user experience study, which qualifies aspects of the user experience as captured in visual design elements. The user study expands the insight into aspects of visual design for fluid flow visualisation, specifically for user engagement via a scientific demonstrator. Taking into account the results of this user study, the project then puts forth two non-competing designs for Eulerian-Lagrangian particle fluid-flow visualisation.

Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my gratitude towards Dr. Christian Kehl, the supervisor of this project. Not only am I grateful he gave me the opportunity to work on this project in the first place, he has also been incredibly helpful throughout; from pre-processing the dataset to answering any questions I had, thank you.

Consequently, I would also like to thank Dr. Julian Köllermeier, the secondary supervisor of this project, for his supervision and advice throughout the project.

Lastly, I would like to thank Robin S. Ballantyne, for agreeing to work together on the base simulation program.

1 Introduction

1.1 Motivation

The increase of cargo shipping along the North West European Shelf has led to a corresponding increase in mismanaged plastic waste within the North Sea. Much of this waste ends up on coastlines, where it can often remain beached for months or years of time. To minimise harmful effects on the environment, organisations like 'Stichting de Noordzee' coordinate cleanup tours like the 'Boskalis Beach Cleanup tour' - a yearly event which aims to collect litter along large stretches of the Dutch coastline [5]. Cleanup actions like these depend on volunteers, and thus depend on public interest in beach cleaning and preservation.

To this end, public interest should be promoted and maintained to keep clean beaches in the public consciousness. An important aspect for awareness and engagement can be the visual design of scientific demonstrations. By designing fluid flow visualisations for engagement, the public may show more interest in important actions like the *Boskalis* tour.

1.2 Research question

This project addresses the relation of visualisation and public engagement by comparing and contrasting differences in visual designs for particle simulation. To this end, the question "How do different visualisations encourage retention of engagement for exploring the motion of floating marine debris" will be answered. In doing so, we gain insight into designing engaging visualisations for fluid simulation.

To compare the different visualisations, a real-time simulation program with different visualisation configurations was required. As existing programs either lack the ability to run interactively in real-time (for physics-derived models like Parcels [14], used for physical oceanography), or are incapable of simulating real oceanographic situations (as is the case for 3D graphics and video game engines), this program was developed as part of this project. The development of (parts of) this program were done in cooperation with my colleague Robin Ballantyne, who researched the impact of gamification on user engagement with regards to particle fluid simulation.

1.3 Contribution

This project puts forth two coherent visual designs making use of the same marks that simultaneously visualise both Eulerian fluid-flow fields and Lagrangian particle trajectories in a non-conflicting manner. These two visual designs are competing in the sense that they visualise the same data but each with a different focus.

Additionally, the program described above is designed for modularity: any visualisation design created in VTK can be easily adapted to work with the program. Furthermore, the simulation logic of the program is capable of running independently with only minimal changes required to the code; any latitude/longitude coordinate can be advected by the program with relative ease.

2 Related Work

2.1 Background: Oceanography

The field of computational fluid dynamics is mainly concentrated around solving the Navier-Stokes equations for a given context. The Navier-Stokes equations are a set of mathematical equations which describe the flow of an (incompressible) fluid over time. The solutions to these equations are not as solved as one may initially think; the inclusion of the nonlinear convective term introduces complications to the calculations such as turbulence and eddy flows [18].

2.1.1 Euler and Lagrange

The majority of fluid dynamics models can be categorized as either Eulerian or Lagrangian in nature. The main difference between these categories is the way in which they approach the discretization of the domain: Eulerian models split the domain into discrete chunks, and then observe one or more stationary chunks as the fluid moves through it. The Lagrangian approach instead treats the domain as a series of discrete particles, and will then follow one or more of these particles as it travels through the fluid [24]. This may be a bit hard to place in context, see Figure 1, which was sourced from [8], for another visualisation. In this figure, the cube represents the area of discrete measurement. Note how the cube is stationary in the Eulerian model, but moves along the line in the Lagrangian implementation.

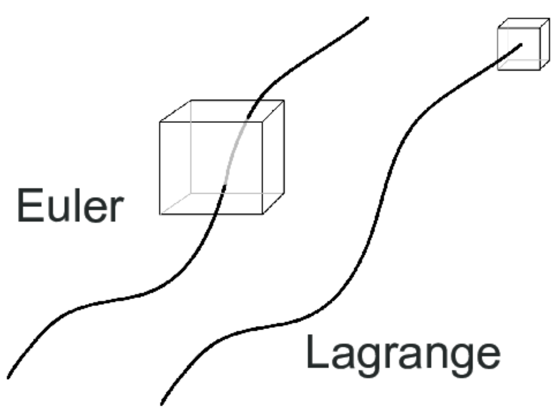


Figure 1: visualisation of Eulerian vs Lagrangian models.

Both of the approaches are used to resolve the Navier-Stokes equations - resulting in hydrodynamic velocities for the given part of the domain. These hydrodynamic velocities can then be applied to, for example, simulate the movement of particles in a fluid. Just as for the hydrodynamic velocities, calculating the movement of these particles can be done in Eulerian or Lagrangian frameworks. For this purpose, a Lagrangian framework is preferred, as these give more detailed information on individual particles than Eulerian models are capable of [9]. The main drawback of Lagrangian methods for interactive simulation is that Lagrangian simulations tend to be more computationally intensive compared to Eulerian methods, which is sub-optimal for real-time simulation [9].

2.1.2 Parcels

One such Lagrangian framework for ocean analysis is *Parcels*, which uses Ocean General Circulation Models (OGCMs) and employs Lagrangian methods to simulate Lagrangian particle-in-fluid transport. *Parcels* is highly configurable, making use of just-in-time compilation to convert (user-written) Ppython functions into much faster C code used "[...for] computing the Lagrangian transport kernels per particle."[16].

Computational requirements: Parcels is a versatile toolbox[1], but it is not well-suited to real-time simulations; its focus on modularity makes it extremely adaptable, but slower than is possible for a more fine-tuned implementation [13]. There are multiple ways in which a Parcels-like approach may be adapted for real-time simulation: as mentioned in [16], "...no assumptions about data layout or iteration protocol have been made in the high-level API of particle sets". By making these assumptions, which is possible for the more narrowed scope involved in simulating one type of dataset in one specific region of interest, the computational cost may be reduced, "...allowing more optimized implementations in the future." [16].

Another method of reducing computations is a reduction on the used hydrodynamic field data: "...extensive interaction with hydrodynamic field data constitutes a considerable cost of the overall computation...". However, doing so implies either a reduced accuracy or reduced resolution of the simulation - trade-offs which may be considered acceptable depending on the intended use of the simulation.

2.2 Background: Visual Design

Visual design is itself important in many fields of science, but especially so when dealing with large quantities of data. Jebeile characterizes this importance in three ways [10]:

- Visual representations can convey information on a lot of data simultaneously.
- Visual representations can reveal the minimal useful information contained within a simulation.

[1] introduced as "An increasingly versatile, open-source Lagrangian ocean simulation tool"[14]

- Visual representations make this information accessible by presenting it in an easy-to-understand manner.

The field of visual design has correspondingly been well-researched, with a large number of papers and books written about the subject. Munzner explains visual design as a composition of data, tasks, encoding, and idioms [19]. An example of encoding is the employment of glyphs, the design principles of which are analysed by Borgo et al. in their survey [4]. Post and van Walsum contribute a chapter to the IEEE Proceedings [21] on fluid flow visualisation, some principles of which remain practicable even now, over 30 years in the future. In this chapter they compare techniques of fluid flow visualisation, as applied to both 2D and 3D visualisations. The chapter treats 2D visualisations as 'default', noting that 3D equivalents of visualisation techniques often require extra attention to be visually palatable. The chapter focuses primarily on lines (i.e. contour lines, timelines, surfaces) and particles. Particles in the visualisation can be represented as a small object, a single point, or some mixture of both [21], while lines are either represented by a directly connect line or alternatively as a list of particles. The chapter briefly discusses the concepts of 'texture synthesis and mapping' as well - two techniques of visualisation which do not see widespread use in the field to this day. Further research on visualisation design is done by Munzner in [19], who characterizes eight rules of thumb to follow for visual design. These rules are:

- No Unjustified 3D
- No Unjustified 2D
- Eyes Beat Memory
- Resolution over Immersion
- Overview First, Zoom and Filter, Detail on Demand
- Responsiveness Is Required
- Get It Right in Black and White
- Function First, Form Next

With the above guidelines in mind, available visualisation techniques in the literature can be categorized under 'marks' and 'channels'; marks are the method of visualisation, while channels are changes to the visual appearance of a mark. These both have a number of possible expressions when used in physical oceanography, which are discussed below.

2.2.1 Marks

Marks are geometric primitives, through which information is conveyed to the user. Marks may require different dimensions to be displayed - a point needs zero dimensions, a line needs one, an area 2, and a volume 3 (although these are not used as often) [19]. In the field of physical oceanography, Kehl et al. classify the following marks in their paper 'Practices, Pitfalls and Guidelines in visualising Lagrangian Ocean analysis' [12]:

- coloured maps
  - scatter
  - continuous
  - categorized
  - categorical
- streamlines
- contours
- glyphs
  - circles
  - squares
  - triangles
  - arrows
  - blocks
- info visualisation
- cross sections
- noise maps
- texture synthesis

2.2.2 Channels

As for channels, these are ways to control the appearance of marks [19]. Each channel can be modulated individually, leading to vastly different results without changing which mark is used at all. The channels identified for physical oceanography by Kehl et al. [12] are:

- hue

- saturation/brightness
- transparency
- form (shape)
- boundary (countour)
- direction
- contrast
- structure

Not all channels convey the same (type of) information; Munzner considers two distinct classes of channels: ordered magnitude channels and categorical identity channels [19]. More interestingly, not all channels are expected to have the same impact: depending on whether one is interested in accuracy, discriminability, separability, or popout, different channels are going to be more or less effective. Considering only accuracy - the measure of how close a perceptive difference in the visualisation maps to an actual change in data, Munzner organises channels in figure 2.

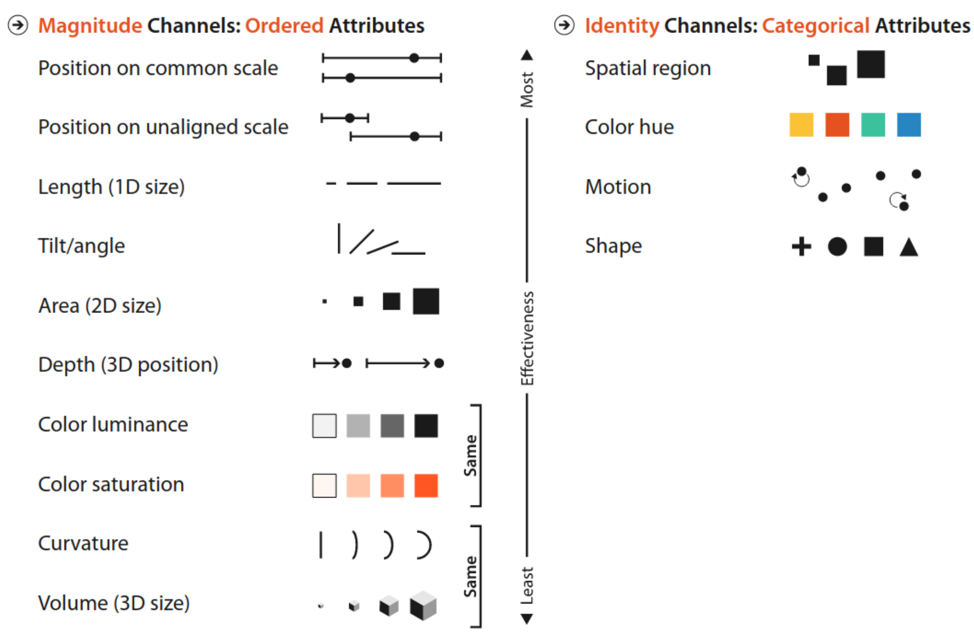


Figure 2: Accuracy of different channels, according to Munzner [19]

2.3 State of the Art

In this section we take a closer look at a few techniques for Eulerian/Lagrangian visualisations. To do so, each technique is introduced, placed within context of its' original paper, and analysed as either an Eulerian or Lagrangian visualisation. Then, the used marks and channels are discussed, of which an overview can also be found in table 1. Lastly, the research gap this proposal is interested in is examined.

Each technique is sorted underneath a header. These follow the format 'Eulerian/Lagrangian: used marks - used channels'. For example, 2 has the header 'Eulerian: glyph - direction, form', which indicates this technique is used for Eulerian visualisation, and makes use of the *glyph* mark - modulated by the *direction* and *form* channels.

Eulerian: glyph - direction, form In their 2021 paper 'Using machine learning and beach cleanup data to explain litter quantities along the Dutch North Sea coast' [11], Kaandorp et al. employ two visualisations: in figure 3, the intent is to show the midway points of Dutch beach cleanup tours - but the figure also conveys information on surface currents, which are represented using arrow-shaped glyphs. This figure hence makes use of the *glyph* mark, and further uses the *direction* and *form* channels to encode information about (Eulerian) surface currents.

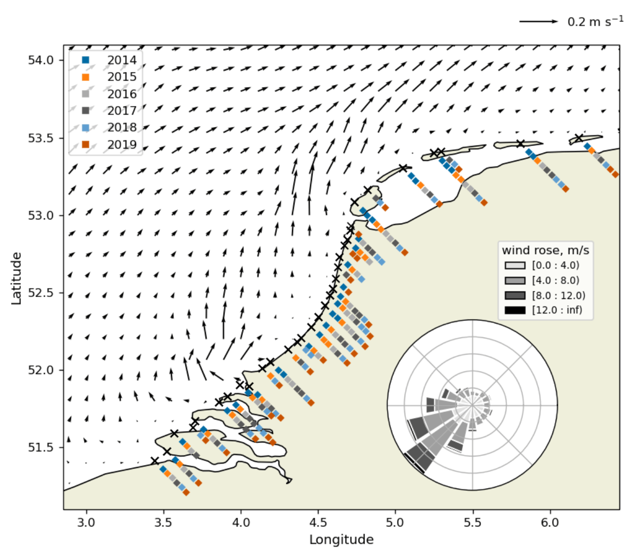


Figure 3: Figure 1 from the paper on litter quantities by Kaandorp et al. [11]

Lagrangian: continuous colour map - transparency The second figure of interest is shown in 4. This figure shows, in blue, the intensity of fishing in the North West European Shelf. This is represented using a continuous *colour map*, where the modulation of the *transparency* channel (with regards to the background - white) encodes the density of fishing operations. This figure furthermore conveys information on coastline waste production using the same mark/channel combination - albeit with a different *hue* - and the output of riverine plastic along the coastline in the form of a *glyph* / *hue* visualisation. These last two visualisations are of course not Lagrangian (nor Eulerian) in nature, but we note them here for completeness' sake.

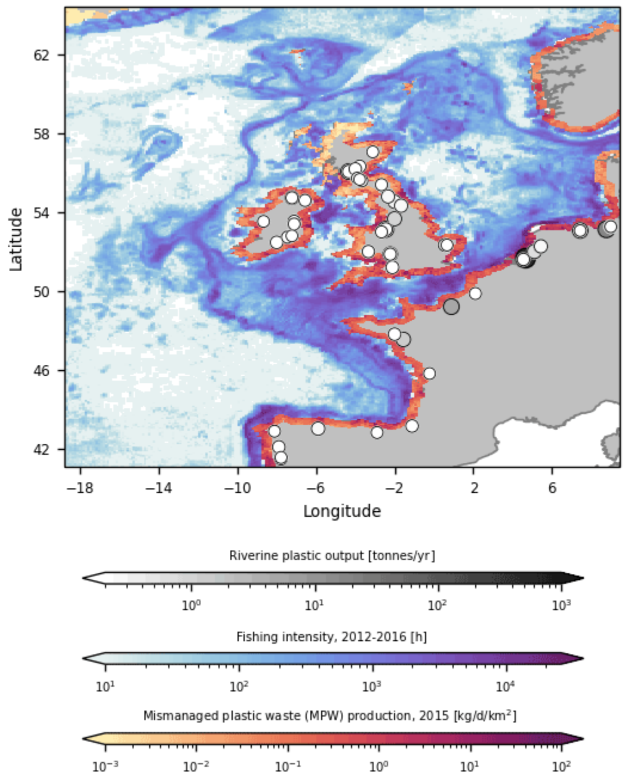


Figure 4: Figure 2 from the paper on litter quantities by Kaandorp et al. [11]

Lagrangian: Glyph - hue Another paper which deals with flow visualisation is 'Detecting flow features in scarce trajectory data using networks derived from symbolic itineraries: an application to surface drifters in the North Atlantic', by Wichmann et al. in 2020 [27]. Figure 5 shows basin connectivity, and uses *glyphs*, with variance in *hues* to indicate different initial drifter locations. This Lagrangian visualisation uses the *glyphs* to represent particle dispersion from initial basin locations.

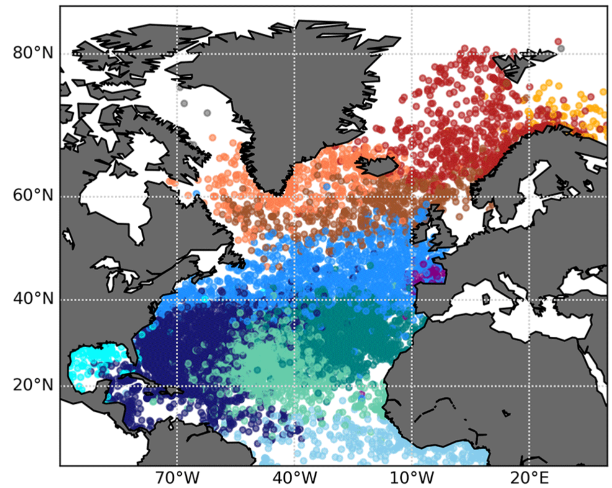


Figure 5: Figure 6 (adapted) from the surface drifter paper by Wichmann et al. [27]

Lagrangian: continuous colour map - saturation From the same paper, figure 6 displays the initial locations and final lifetimes of drifters from the dataset by Wichmann et al. - a Lagrangian visualisation. Subfigures (a) and (b) both use the same mark/channel combination of a continuous *colour map* where the *saturation* indicates density of the drifters.

Lagrangian: scatter colour map - hue Subfigure (c) in the same figure instead employs a *scatter plot*, where the *hue* of each element indicates the lifetime of a drifter. Note how for subfigure (c) specifically, the density of the markers occludes information conveyed by the hue - the density of the glyphs obscures their hue in the densest areas of the figure, decreasing readability.

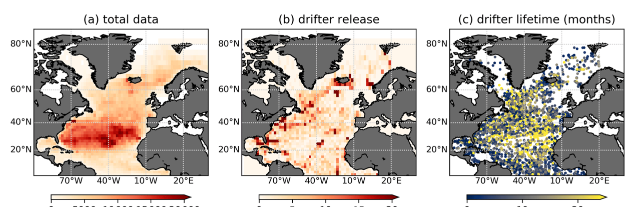


Figure 6: Figure 1 from the surface drifter paper by Wichmann et al. [27]

Eulerian/Lagrangian: texture synthesis A less thoroughly explored visualisation technique is analysed by Khlebnikov et al. who, in their 2012 paper 'Procedural Texture Synthesis for Zoom-Independent visualisation of Multivariate Data' explore the possibility of *texture synthesis* for data visualisation [15]. Figure 7 contains two distinct data sets: the first uses a continuous *colour map* with variations in *hue* to indicate average sea temperatures. Simultaneously, precipitation frequency is encoded using a *noise map*. The *noise map* alters the *saturation* of the displayed points. The *texture synthesis* in this figure comes from the simultaneous showing of the two datasets; by applying the precipitation *noise map* over the average temperature *colour map*, Khlebnikov et al. synthesise a new texture. The figure unfortunately displays neither Lagrangian particle nor Eulerian flow field visualisation, but it is easy to see how the techniques used here may be applied to a combination of Eulerian flow fields and Lagrangian particles - for example by encoding the flow fields in a *colour map*, and including the concentrations of particle densities by applying *noise*.

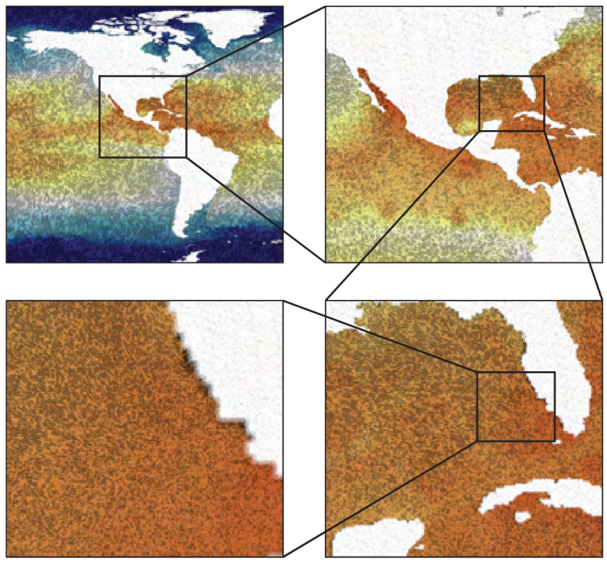


Figure 7: Figure 4 from the texture synthesis paper by Khlebnikov et al. [15]

2.3.1 Overview

Table 1 summarises the analysis of the techniques described in sections 2.3.x. This table contains each technique on a separate column, denoting which of the marks/channels are used with a mark ('X'). Each technique is referred to by its' (sub)section header number.

Marks & Channels	Colour map	Streamline	Contour	Glyph	Info visualisation	Cross section	Noise map	Texture synthesis	Hue	Saturation	Transparency	Form	Boundary	Direction	Contrast	Structure
1				E								E		E		
2	L			L					L	L						
3				L					L							
4	L									L						
5	L			L					L							
6	E						L	E/L	E	L						

Table 1: Analysed mark/channel combinations for each technique discussed in sections 2.3.X

There are a few mark/channel combinations which have not been analysed. For a number of these the reason is that there is little to no reason to use these combinations: a *hue*-modulated *colour map* Eulerian flow field visualisation with *hue*-modulated *contour*s for Lagrangian displacement would be very hard to make look appealing - not to mention having to convey the desired information as well. Other combinations are hard to find in published articles: there is a limited number of papers which use *texture synthesis* for their visualisations. Of these, even less apply this technique in the field of physical oceanography.

Of the analysed techniques, some combination of *colour map*ping, *glyph*s, and *hue* variation occur most frequently. Intuitively, this makes sense. *Colour map* and *glyph* marks can be used for both Eulerian and Lagrangian visualisation, with *hue* being a default channel editable in many visualisation toolkits. One way to put these together in a visually pleasing manner could be the following: *glyphs* to represent Lagrangian particles, with an underlying *hue* modulated *colour map* to indicate strength of Eulerian flow fields. This technique would encode only one dimension of the Eulerian flow fields however - if not only strength but direction of these fields is to be conveyed as well, another technique may be more appropriate.

2.3.2 Research Gap

In "True Colors of Oceanography", Thyng et al. provide guidelines intended to aid oceanographers in choosing appropriate colour maps to display data [26]. These guidelines account for the lightness and perceived perceptual changes of used colour schemes, while providing examples of sequential, diverging, and cyclic colour maps. While these guidelines are specifically designed for use in the field of oceanography - and thus exist on the intersection of physical oceanography and visual design - they only focus on two specific channels (hue and saturation/brightness) of visual design.

Similarly, in "Multivariate visualisation of Oceanography Data Using Decals", Rocha et al. investigate the visualisation of multiple oceanographic data attributes in a singular view [28]. Here, they incorporate information about the density, salinity levels, flow strength, and flow direction into one visualisation technique. Their methodology involves decomposing each information channel into one 'layer', which is then independently visualized. Note the similarities between this paper and this report's decomposition into Eulerian and Lagrangian layers. Again, however, Rocha et al. investigate only one visualisation of their data; there is little focus on comparative analysis between multiple visualisations of the same data.

The research gap, then, is as follows: there is only a small amount of research into the intersection of physical oceanography and visual design. Specifically, there are little to no papers which which qualitatively compare different visualisation techniques incorporating multiple information channels.

3 Methodology

In order to perform a qualitative comparison between different visualisation designs, a user study is conducted. This user study makes use of the simulation program developed for this project to facilitate user interaction. Details of both are described below.

3.1 Simulation Program

The simulation program built for this project simulates surface-level ocean currents (and particles thereon) in the North West European shelf using pre-existing flow data (see subsection 3). The program is interactive in nature, containing the following functionality:

- (un)pausing the simulation of the program
- placing particles, which are then simulated according to the flow data
- Zooming in/out, and panning of the camera over the simulated area
- switching between visualisation designs
- Modulating channels of a visualisation design

The frontend of the simulation program consists of two components: there is the simulation window, which shows the actual simulation state and allows for spawning particles and manipulating the camera. This is the area containing the map on the right of the program in figure 8.

The other component contains elements for modifying the visualisation, namely changing the active design and modulation of channels within a design. This is the area labelled 'Settings' in figure 8.

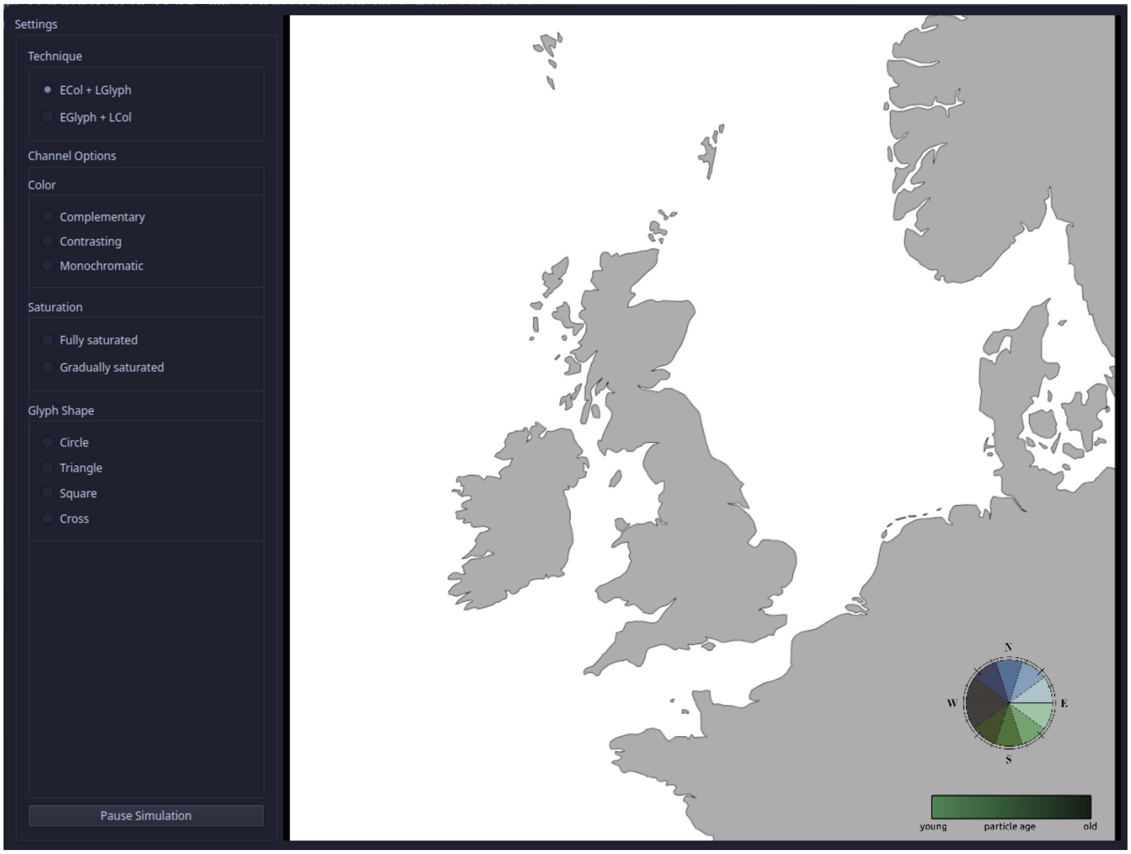
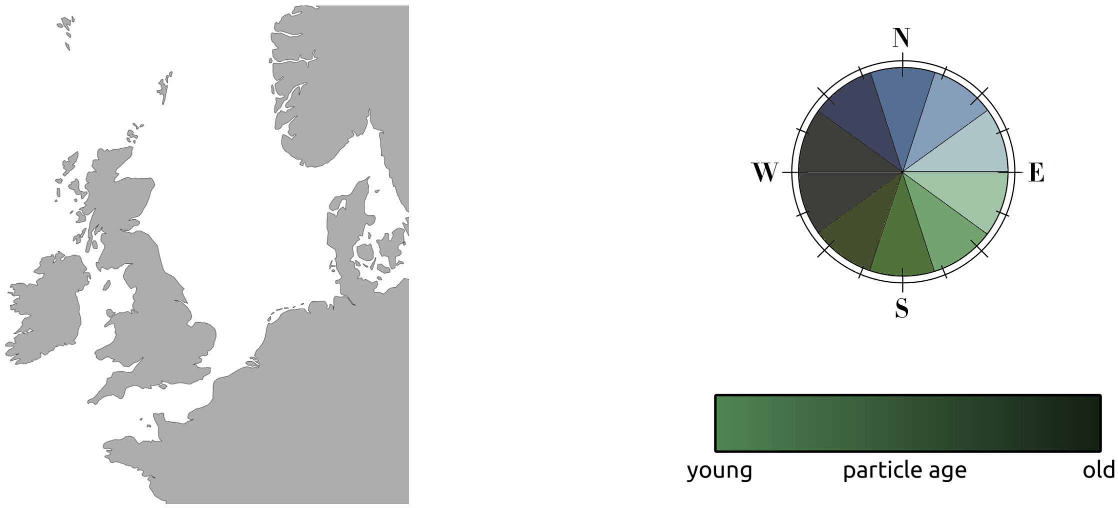


Figure 8: The simulation program

3.2 Visual Designs & Techniques

In this project, we compared two visualisation designs against each other. Each design has two separate layers - one for Eulerian velocity field data and one for Lagrangian particle data. There are a few corollary layers common between designs as well, namely a background layer (presented in figure 9a) and a legend layer (example shown in figure 9b).



(a) The background layer

(b) One example of the legend layer

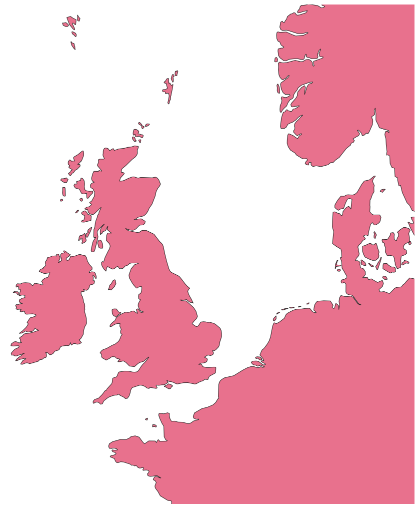
Figure 9: The two design-independent layers

Of these layers, the legend layer did not go through any drastic changes during the design process; any changes made to it were done only as a result of changes to the design of each channel configuration (different hues/saturation levels). The background layer, on the other hand, saw three different iterations: the first, very rudimentary implementation, used a screenshot from Openstreetmap [20] as background image. This suffered from multiple drawbacks, most prominent of which the low resolution of the image (as can be seen in figure 10a). The other issue with this implementation was the extraneous detail (city names and land colouration), which is not desired from a background image for this project.

The second iteration, seen in figure 10b, made use of a QGIS-based [22] render of the simulation domain. This render resolved the problems with resolution present in the initial design, but still did not have an ideal colouration: the use of hue for landmasses caused visual confusion when the Eulerian and Lagrangian layers were applied on top of it.

The final iteration then took the initial QGIS render, and modified the used colours; by making use of only gray-scale values, the hue channel is freed up from the background entirely. (This iteration is displayed in figure 9a)

(a) The original background image



(b) The initial QGIS-sourced background image

Figure 10: Iterations of the background layer throughout the design process (final design not included, see figure 9a)

3.2.1 Base design

The base design makes use of a *colour map* to visualise the Eulerian fluid-flow, and uses the *Glyph* mark for the Lagrangian particle component. The different configurations of this design are shown in figures 12a, 12b, 12c, and 13a.

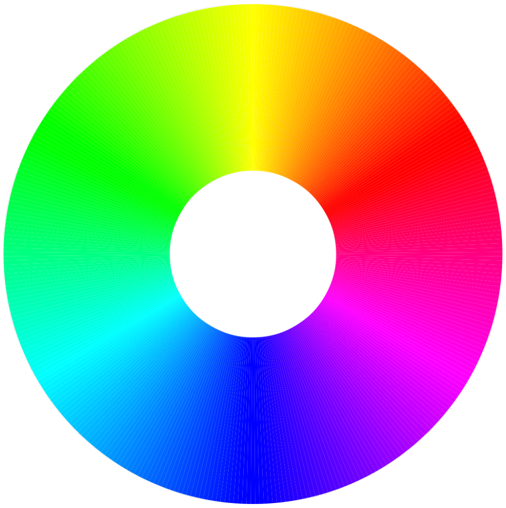
Lagrangian layer: the final implementation of the Lagrangian layer makes use of the *brightness* channel for particle age and *transparency* to indicate beached state of particles.

The initial design for this layer had these channels swapped, using *transparency* for particle age and *brightness* to indicate beached state. As the brightness channel led to more noticeable changes to the particles in the simulation - and the desired result was for age of particles to be more pronounced - these channels were swapped.

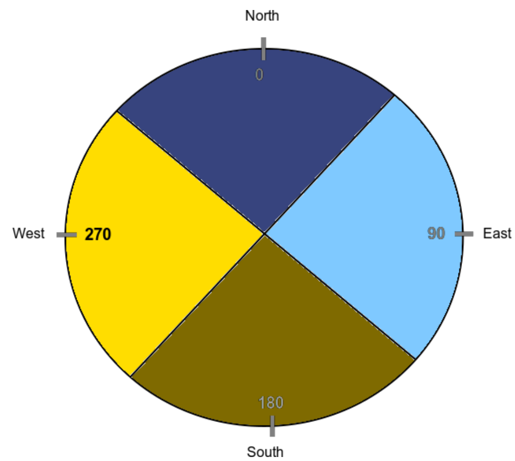
Eulerian layer: the Eulerian layer uses *transparency* to denote flow velocity magnitude and *hue* to indicate flow direction.

The transparency channel has been in place since the early stages of the design, but the hue channel went through several variations: the initial version used the full colour wheel to represent current directions (see figure 11a). This colour scheme resulted in two issues: not only is this colour map nonuniform in terms of visual perception, it also acyclic in nature; a cyclic colour map was required, so as to indicate similar directions with similar hues.

The second iteration of the colour map made use of a 4-tone cyclic mapping to indicate flow direction (displayed in figure 11b). While this solved the problem of cyclicity, the limited nature of four colours meant a desired level of detail was lost. The final version of the design makes use of the *corkO*, *romaO*, and a modified *9-class blues* colour scheme (see figures 12a, 12b, and 12c respectively). The first two of these maps make use of Fabio Crameria's scientific colour maps [25], while the last uses a modified version of a ColorBrewer [6] palette.



(a) A full colour wheel, like those used in the 'jet' and 'rainbow' colour maps



(b) A cyclic colour wheel

Figure 11: Iterations of colour wheels used throughout the design process (final designs not included, see figures 12a, 12b, 12c, and 13a)

3.2.2 Compared design

The compared design makes use of the opposite combination of marks, with *glyph*s for the Eulerian fluid-flow layer, and a *colour map* for the Lagrangian particles. The different configurations of this design are shown in figures 12d, 12e, 12f, and 13b.

Eulerian layer: the implementation of the Eulerian layer remained consistent throughout the design process, and uses the *direction* channel to indicate flow direction, further modulating the glyphs in *area* to indicate flow strength.

This layer makes use of arrow-shaped glyphs for their relative indexability, and the intuitive understanding of an arrow as indicating direction.

Lagrangian layer: the Lagrangian layer in the compared design uses the *hue* channel to indicate average particle age in a given cell, which is further modulated by *transparency* to indicate particle density.

The Lagrangian layer for this design was implemented after the Eulerian layer for the base design, meaning some of the issues we ran into while designing that layer could be avoided here entirely. As such, the initial designs implemented are the ones used in the final version of the program. These designs use reverse 'YlGn', (normal) 'Viridis' and reverse 'Greens' colour maps, all of which are based on the Matplotlib equivalents [7].

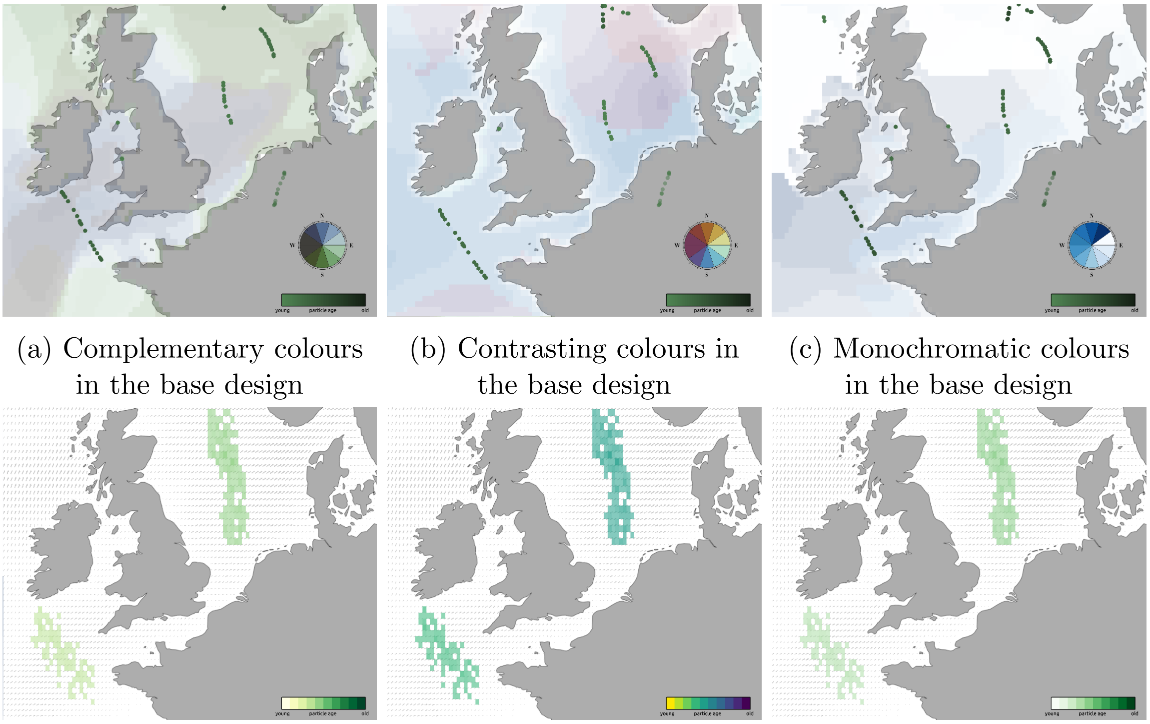
3.3 Modulation Channels

As mentioned above and visible in figure 8, the program supports a number of options for channel modulation. Specifically, the program supports modulation of *hue*, *saturation*, and *glyph shape*.

Hue: the first channel with different modulations is the hue channel. This channel supports three configurations, namely:

1. complementary colours
2. contrasting colours
3. monochromatic colours

Each design has a separate implementation of the above configurations. The three configurations for the base design can be found in figures 12a, 12b, and 12c, while the configurations for the compared design are shown in figures 12d, 12e, and 12f.



(a) Complementary colours in the base design
(b) Contrasting colours in the base design
(c) Monochromatic colours in the base design
(d) Complementary colours in the compared design
(e) Contrasting colours in the compared design
(f) Monochromatic colours in the compared design

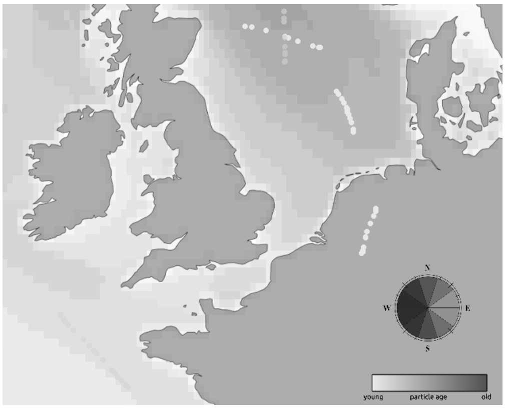
Figure 12: All available hue configurations for both designs

Saturation: next there is the saturation channel. Both techniques support two settings for this channel, those being:

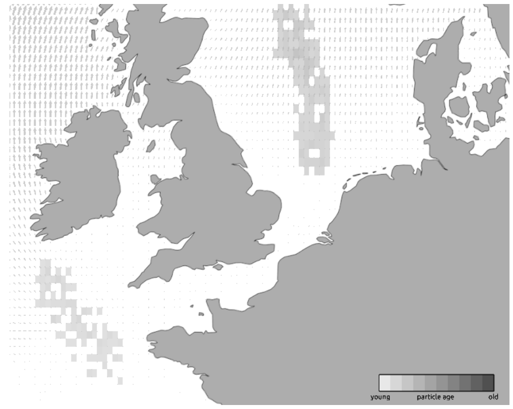
1. fully saturated
2. gradually saturated

The ‘gradually saturated’ option leads to a gray-scale result, as shown in figures 13a and 13b. This configuration makes use of a modified ‘GrayC’ colour map [25]. The ‘fully saturated’ selection applies the same saturation used by the different hue configurations. This results in the designs being coloured according to the active hue configuration (as shown in figure 12) while this setting is active.

Conversely, the ‘gradually saturated’ option takes precedence over any options for hue; changing the active hue setting has no effect so long as gradual saturation is selected.



(a) Gradual saturation in the base design



(b) Gradual saturation in the compared design

Figure 13: Gradual saturation in both designs

Glyph shape: lastly the glyph shape can be modulated. This setting only affects the base design, which uses glyphs to represent particles. There are four options for this channel:

1. circles
2. triangles
3. squares
4. crosses

All four options are shown in figure 14. An additional note: none of the indexed shapes (triangles, squares, crosses) are rotated during simulation; these shapes are always rendered facing North in the program.

(a) Circular glyphs

(b) Triangular glyphs

(c) Square glyphs

(d) Cross-shaped glyphs

Figure 14: Available glyph shapes in the base design

3.4 User Study

The user study was conducted in a white-box, supervised manner. Only one user participated at a time, ensuring any queries by the users could be answered without delay. The structure of the user study was as follows:

- First, each user was given a short introduction to the program, explaining what the program represented and which functionality it supported (particle spawning and visualisation switching).
- Then, the user was given free reign of the program and encouraged to explore the functionality at their leisure.
- Lastly, once the user indicated being done with the program (or after the user spent fifteen minutes interacting with the program, whichever happened first), the user was presented with a feedback form hosted through *Google Forms*.

The feedback form consisted of seven types of questions, specifically it contained questions about:

- Whether the participant is a student or not
- The time spent interacting with the program
- The information conveyed by a given design
- The visual appeal of a given design
- What the most appealing part of a given design was
- Which of the designs the user preferred
- Which of possible channel configurations the user preferred
- Direct feedback from the user

A full list of posed questions can be found in the appendix, under the section Feedback form user study.

3.5 Data

The simulation program can work with any HDF5 dataset, with a few restrictions: the dataset must be separated out into three files: 'grid.h5', 'hydrodynamic_V.h5' and 'hydrodynamic_U.h5'. The grid file must contain at least the three variables:

- latitude: a list of latitude points
- longitude: a list of longitude points
- times: a list of timestamps, starting at 0

The hydrodynamic files should contain at least the variables 'vo' and 'uo' respectively. These variables should then contain the latitudinal and longitudinal velocities of the regular grid described by the Cartesian product of the 'latitude' and 'longitude' variables in the grid file.

The dataset used during this project is sourced from *NEMO* and the *Copernicus Marine Service* [2], both of which provide datasets on the Atlantic-European North West Shelf. This data has been pre-processed by incorporating the effects of wind on surface velocities, for which data from the *WaveWatch* model has been used.

After pre-processing, the used dataset has a temporal resolution of 1 hour over the span of a complete year. Spatially, it contains 79 latitudinal and 116 longitudinal points, which span the area between (46.125, -15.875) and (62.125, 12.875) (latitude, longitude) respectively.

3.6 Approach

As outlined above, the approach to the experiment is to conduct a user study which results in actionable data we can analyse to gain an insight into visual design for interactive particle simulation.

The questions posed in the user study specifically attempted to investigate the following:

- Impact of different mark combinations on user engagement; the comparison of the base design and compared design. This comparison is done on two aspects:
  - visual appeal of each design.
  - adeptness of each design at conveying information about the simulation state.
- Impact of the modulation of different channels on user engagement; the comparison between different channel options. The different options each influence different channels:
  - Colour : modulates the *hue* channel
  - Saturation : modulates the *saturation/brightness* channel
  - Glyph shape : modulates the *form (shape)* channel

4 Implementation

4.1 Used technologies

The simulation is programmed in *C++*, and makes use of the *QT* software for GUI components. For visualisation, the library *VTK* is used. Lastly, the program makes use of the *netcdf-cxx* library to deal with loading and parsing data.

4.1.1 Programming language

C++ was chosen as programming language for its extensive ecosystem (unlike *Rust*), low-level access to memory management (unlike interpreted languages, i.e. *Python* and *Type/Javascript*), and object-oriented nature (unlike imperative languages like *C*).

4.1.2 GUI

Of the two most well-supported GUI libraries (in *C++*), *GTK* and *QT*, it is *QT* which is natively developed in *C++*. The *QT* framework also features a number of configurable widgets which streamlined the design process and made handling user-inputs for the simulation much easier.

As for a visualisation toolkit, this project needed a balance between ease of use and customizability. This means using bare *OpenGL* or a fully featured engine like *Magnus* was unfeasible. The Visualization Toolkit (*VTK*) provides advanced features in scientific visualisation (while managing the minutiae of platform-specific configurations under the hood), is cross-platform, and fully open-source[1].

4.1.3 Git

The code for the simulation program made use of the versioning control system Git, specifically in a *GitHub* repository[2].

[2] https://github.com/djairoh/interactive-track-and-trace

4.2 Program layout

The program structure consists of a few components: there is the GUI, the simulation logic, and the visualisation. These components map roughly to the libraries of *QT*, *netcdf-cxx*, and *VTK* respectively.

4.2.1 GUI

The GUI consists almost entirely of *QT* classes. All of these, save the VTK component, were designed using *Qt-Creator* [23].

To give the user a clear and immediate overview of available functionality, the settings panel consists mostly of radio buttons, which toggle the various options of the program. Underneath the settings panel is located a simple button, which toggles the simulation's progression. The settings panel is shown on the left-hand-side in figure 8.

The one component not modeled using QT-Creator is the VTK component. This component instead subclasses a builtin VTK class - the `QVTKOpenGLNativeWidget` class - specifically intended for integration with QT-enabled programs. By subclassing this widget, any arbitrary VTK component can be rendered inside a QT program. As implemented here, the entire simulation runs inside this component.

4.2.2 Simulation logic

The simulation part of the program consists of a self-contained package containing the functionality to read data (as described in section 3), as well as an interface and a few implementations for different advection kernels.

Each advection kernel implements an `advect` function which, given a triple of `time`, `latitude`, and `longitude`, will advect the given position for the given timestep and return a pair of `latitude` and `longitude`.

By using an abstract `kernel` class, the program allows for more modularity when it comes to actual simulating of particles - for example, one is free to implement any level of Runge-Kutta, linear, or other desired integration. During this project, the class `RK4Advectionkernel` was used. This kernel implements the fourth-order Runge Kutta method, minimising the drift of advected particles.

4.2.3 visualisation

The visualisation aspect of the program mainly deals in different VTK widgets and classes.

The primary component is the `Program` class, which manages the latter stages of the VTK pipeline. This class contains an attribute for a vector of `Technique`s, which can be modified through the `addTechnique`, `removeTechnique`, and `setActiveTechnique` functions.

The program class is also responsible for setting up a vtkTimer, which is used to drive the simulation logic of the whole program. The program will update each `Technique` every time the timer goes off, thus propagating the updating of particles down the pipeline.

Each `Technique` models one complete visualisation. The number of layers is variable; each one is completely independent. For this project, both designs contain a background, flow, particle, time, and legend layer. Each Technique also implements an `updateData` function, through which the Program can update the actual data contained in each Layer.

The `Layer` class contains everything associated with one instance of a vtkRenderer - meaning it renders one 'layer' of the total VTK pipeline. Each Layer manages its own layer in the VTK pipeline (meaning multiple Layer classes might render to the same VTK layer if required). Currently implemented are:

- BackgroundImage : reads and renders an image - layer 0.
- EColLayer : renders the flow field as a colour map - layer 1.
- EGlyphLayer : renders the flow field as glyphs - layer 1.
- LColLayer : renders the particles as a colour map - layer 2.
- LGlyphLayer : renders the particles as glyphs - layer 2.
- LegendLayer : renders the legends - layer 3.
- TimeLayer : renders a string showing how much time has passed since the simulation started - layer 4.

Note how the two E and L layers render to the same VTK layer respectively. This is because only one of each is intended to be used per Technique. Specifically, EColLayer and LGlyphLayer together form the core of the base design, whereas EGlyphLayer and LColLayer from the core of the compared design.

Each Layer implements a few additional functions, most of which are related to modulating the used channels. These functions are `setColourMode`, `setSaturationMode`, and `setGlyphStyle`. The actual effect of these functions differs per layer. Another function implemented by each Layer is an `updateData` function - the lowest in the hierarchy. There are also the `addObservers` and `removeObservers` functions, which respectively add and remove any observers from the vtkRenderWindowInteractor.

5 Results

We recorded responses of 13 participants in the user study. Of these 13, 12 indicated to be full-time students. As students are the primary target audience of this evaluation, only these 12 responses are included in further analysis.

Participants indicated that the interactive nature of the simulation was mostly clear to them, with 9 out of 12 responding with either 'very clear' or 'somewhat clear'. Of the considered participants, only 2 indicated the interactive nature to be unclear ('somewhat unclear', or 'very unclear'). A further 2 thought the interactivity to be 'neither unclear nor clear'. These results are also shown in figure 15.

Overall this indicates a majority of participants (75%) thought it was clear the program supports interactivity.

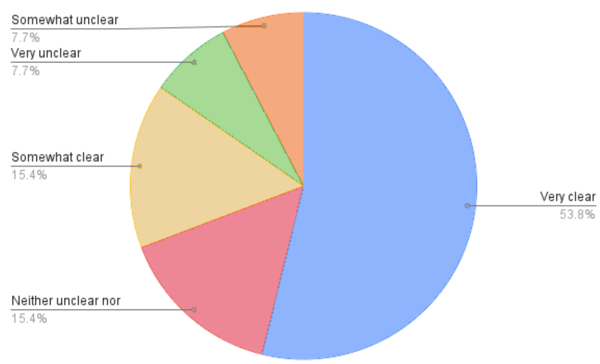


Figure 15: Responses to the question "How clear was it the simulation supports interaction?"

5.1 Interaction time

The survey contained two questions regarding interaction time; the first is more a measured data point than actual question, and asked the user how long they interacted with the program. A histogram of responses is included in figure 16.

The shortest interaction time recorded is 2 minutes, whereas the longest is 10. The mean interaction time is 5.8 minutes, with a standard deviation of 2.4 minutes. There are also two outliers, both of which exceed the upper bound at 9 and 10 minutes of interaction time respectively.

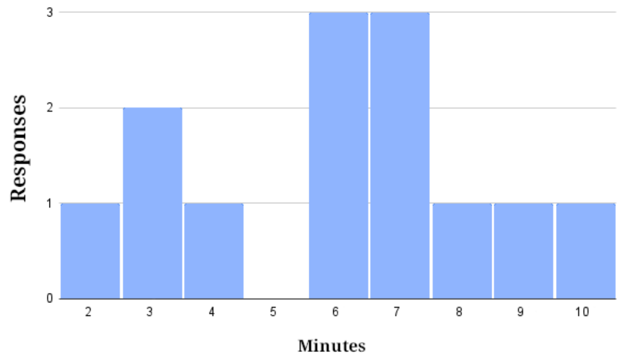


Figure 16: Interaction times in minutes

The second question of this nature asked users how long they would have continued interacting with the program for, if given unlimited time. For this question, 4 users entered the same number as for the previous question, indicating they would not have interacted with the program beyond the testing period (The other 8 participants entered values for this question higher than in the previous one).

Overall, the average time participants would have maximally engaged with the program is 10.3 minutes. There is a standard deviation of 6.0 minutes. The shortest maximum interaction time was 2 minutes (which is the actual time this participant interacted with the program as well), while the longest maximum interaction time is 25 minutes - which is also the one outlier for this question.

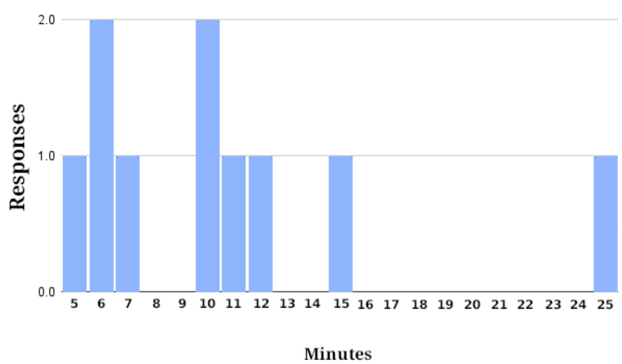


Figure 17: Maximum interaction time in minutes

5.2 Conveyed information

Each participant was asked to rate each design on how well a number of information channels was conveyed to the user. The information channels asked were:

- Speed of background current
- Direction of background current
- Movement of placed particles
- Age of placed particles
- Beached state of placed particles

The average rating of each channel for each design (as well as the average rating and standard deviation) is provided in table 2. There were exactly four channels of information which on average rated lower than a 4:

- Both background current channels for the base design (colour map)
- Speed of background current for the compared design (arrow glyphs)
- Beached state of particles for the compared design (colour map)

Similarly, there are two channels across both designs with an average rating higher than a 4.5:

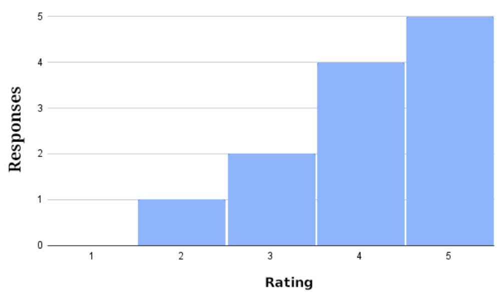
- Movement of placed particles for the base design (glyphs)
- Direction of background current for the compared design (arrow glyphs)

		Speed of background current	Direction of background current	Movement of placed particles	Age of placed particles	Beached state of placed particles
Base design	1	0	0	0	0	1
	2	1	1	0	1	0
	3	4	4	0	0	3
	4	3	3	3	5	0
	5	4	4	9	6	8
	Mean	3.8	3.6	4.8	4.3	4.2
	SD	1.0	1.2	0.5	0.9	1.3
Compared design	1	0	0	0	1	2
	2	2	1	0	1	1
	3	2	0	3	1	2
	4	4	2	5	2	3
	5	4	9	4	7	4
	Mean	3.8	4.6	4.1	4.1	3.5
	SD	1.1	0.9	0.8	1.4	1.5

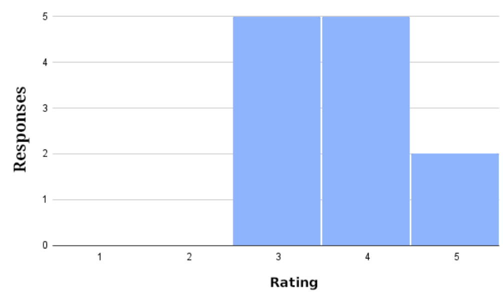
Table 2: Responses to the conveyed information questions

5.3 Visual appeal

Participants were asked to rate both visualisation designs on visual appeal, rated on a scale of 1-5. The base design was rated an average of 4.1 with a standard deviation of 1.0, while the compared design received an average rating of 3.8 with a standard deviation of 0.8 - see also figures 18a, 18b.



(a) Answers on the base design



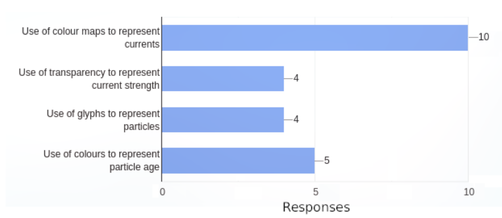
(b) Answers on the compared design

Figure 18: Responses on visual appeal of both designs

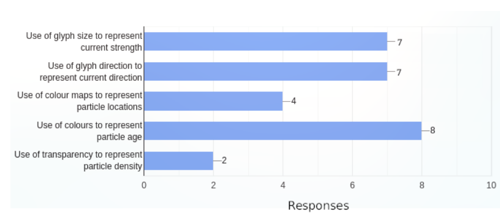
Next to these direct rating questions, the survey asked participants to identify which aspect(s) of each design they found most appealing. These questions were multi-select in nature, allowing participants to pick up to three elements they felt were most appealing.

For the base design, a majority of participants (10 responses) identified ‘the use of colour maps to represent currents’ as visually appealing, with an additional 5 indicating to like ’the use of colour to indicate particle age’. Additionally, the options ‘transparency to indicate current strength’, and ‘glyphs to present particles’ each were marked four times as visually appealing. These responses are summarized in figure 19a

The answers for the compared design had no clear favourite: while 8 responses identified ‘use of colours to represent particle age’ as visually appealing, the options ‘glyph size to represent current strength’ and ‘glyph direction to represent current direction’ each were chosen 7 times as well. A further 4 response indicated to like ‘colour maps to present particle locations’, with only 2 responses identifying ‘transparency to represent particle density’ as most visually appealing aspect. The responses to this question are also summarized, and can be seen in figure 19b.



(a) Answers on the base design



(b) Answers on the compared design

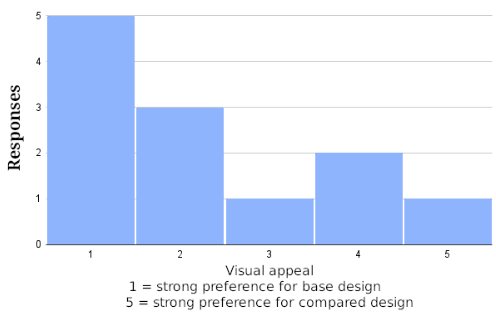
Figure 19: The most visually appealing aspects of each design, according to participants

5.4 Direct comparison

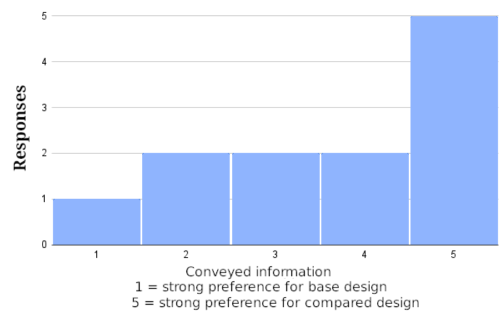
The survey contained two questions asking users to directly compare both designs, both on a scale of 1 to 5 (1 indicating a strong preference for the base design, whereas 5 indicates a strong preference for the compared design).

The first of these questions asked about visual appeal, and got an average value of 2.25, indicating a preference for the base design over the compared design. This is also supported by the modal value of 1: five participants indicated a strong preference for the base design, while only one participant indicated a strong preference for the compared design. These answers are also shown in the histogram in figure 20a.

The second question asked participants which of the designs they preferred in terms of *conveyed information*. The responses to this question indicate a preference for the compared design, with an average response of 3.7. The modal value here similarly supports the mean, with a value of 5. Interestingly, the number of responses for the values 1 and 5 exactly mirror that of the previous question - five participants indicated to strongly prefer the compared design, while only one participant strongly preferred the base design. The responses to this question are displayed in a histogram as well, and can be found in figure 20b.



(a) Answers on visual appeal



(b) Answers on conveyed information

Figure 20: Participant responses to direct comparison questions

5.5 Modulation Channels

For each available modulation channel (colour, saturation, glyph shape), participants were asked to rank the options in terms of preference.

5.5.1 Colour

For the colour option, participants indicated to prefer the options 'Complementary' and 'Contrasting' exactly equally, with each 6, 5, and 1 votes for 'Highest preference', 'Medium preference', and 'Lowest preference' respectively. The option 'Monochromatic' was clearly least favoured, receiving only 2 and 10 responses for 'Medium preference' and 'Least preference' respectively. These responses are included in figure 21.

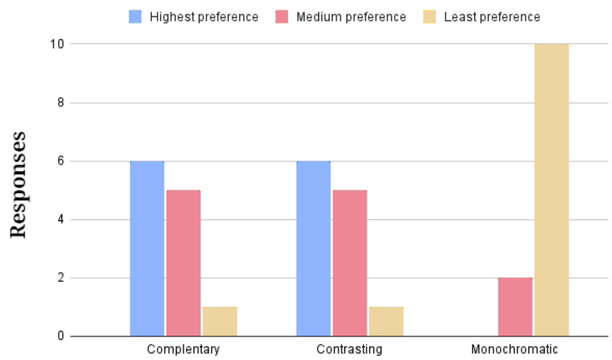


Figure 21: participant preference between colour options

After mapping each answer to a numerical point, according to the following:

- 'Lowest preference' = 1
- 'Medium preference' = 2
- 'Highest preference' = 3

The average rating of each option can provide an insight into preference; in this case, the complementary and contrasting options each have an average rating of 2.4, whereas the monochromatic option has an average rating of 1.2 - further cementing the user preference for complementary/contrasting colour schemes over a monochromatic one.

5.5.2 Saturation

For saturation, the majority of participants preferred full saturation (11 responses), with only one participant indicating to prefer the gradually saturated configuration. These responses are shown in figure 22.

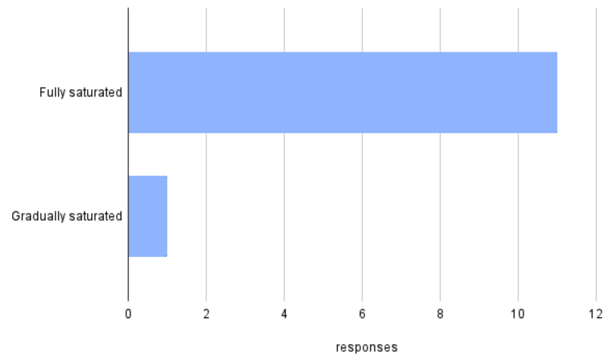


Figure 22: participant preference between saturation options

5.5.3 Glyph Shape

In terms of glyph shape, participants preferred circular glyphs the most, cross-shaped glyphs the least, and triangular and square glyphs in the middle. Individual counts of preference are included in figure 23.

Notably, circular glyphs not once received a rating of 'Lowest preference'. Conversely, cross-shaped glyphs were not once rated rated as 'Highest preference'.

Figure 23: participant preference between glyph shape options

Applying a similar mapping to numerical values as above, namely:

- 'Lowest preference' = 1
- 'Second lowest preference' = 2
- 'Second highest preference' = 3
- 'Highest preference' = 4

Once more the average values of each option may provide a more concrete insight into user preference. In this case, we observe the highest average value for circular glyphs (average value of 3.3). The lowest average value occurred for cross-shaped glyphs (average value of 1.8), while triangular glyphs have an average value of 2.8, and square-shaped glyphs get an average rating of 2.

5.6 Open Feedback

The last questions on the feedback form asked participants if there was anything in particular they wanted to share about the visualisation designs, the simulation program, or the user study itself.

5.6.1 visualisation designs

Of eight responses, one was simply ‘no’ (no feedback on visualisation designs). Two responses mentioned to prefer the base design in terms of visual appeal, but that the compared design was preferable in terms of conveyed information. One response mentioned to prefer the base design, while another mentioned to prefer the compared design - specifically because of the use of arrow-shaped glyphs. One response indicated a different glyph-shape to indicate beached state would have been helpful.

The last two responses included specific feedback about parts of the visualisation scheme: one noted the non-cyclic nature of the monochromatic colour option for flow-direction and that the participant did not like the use of transparency in visualisations. The second similarly noted occlusion on the transparency channel, indicating that ‘young particles [are] difficult to see’ when the complementary colour scheme is active in the compared design.

5.6.2 Simulation program

Two participants noted that the supported camera movement was not immediately clear. Another response included a suggestion to change the ‘Pause simulation’ button (see figure 24) to ‘Toggle simulation’. Furthermore, one participant noted that a different channel to indicate particle beaching would have been preferable - specifically the response mentions using a ‘specific mark’ to indicate as such.

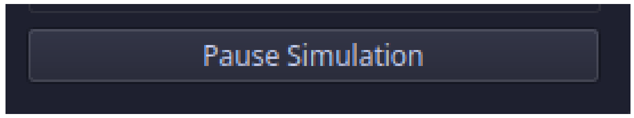


Figure 24: The ‘pause simulation’ button

5.6.3 User study

There were eight responses for the ‘User study’ category of open feedback, of which two simply responded ‘no’ (no feedback on this subject). Of the other six responses, one provided feedback on the survey tool used; the response suggests that the layout of Google forms’ buttons should be changed. The remaining five responses amount to saying either the participant enjoyed the user study or enjoyed playing with the simulation program.

6 Discussion

In this section we interpret the results observed in section 5. The first two questions - those about student status of the participant and clarity of interactivity - are not discussed further.

Interaction time

A majority of participants indicated a longer maximum interaction time than their actual interaction time, indicating a majority of participants would have continued interacting with the program longer than they did during the experiment. This implies the program might serve well at a scientific demonstrator; the fact that a majority of participants indicated they would have continued interacting with the program past the user study implies at least a curiosity on the subject.

Conveyed information

The use of glyphs was received especially well across both designs (average score of 4.8 and 4.6 respectively). This can be in part attributed to the discrete nature of the glyph mark: glyphs (as used in this project) uniquely provide a specific one-to-one mapping between data point and visualisation, resulting in higher clarity of conveyed information.

The least well conveyed information, on the other hand, both concern the colour map mark: The ‘beached state of placed particles’ for the compared design rated an average of 3.6. This rating may be explained by the non-discrete nature of the colour map mark: the colour map, by its’ nature, aggregates data on individual particles together into one cell (the size of which depends on the resolution of the colour map). In this aggregation, data on individual particles is necessarily lost, thus making it harder to observe when one particle is beached. The ‘direction of background current’ for the base design rated an average of 3.6. The reason for which is less immediately clear; the addition of a legend highlighting the cyclic nature of the used colour schemes should have cleared up any confusion on this point. Possibly, the fast changes in flow direction in the simulation (thus resulting in rapid changes in displayed hue for this information point) may have confused participants. No feedback indicating as such was given however, and as supplemental data (i.e. eye-tracking data, of which low resting times and high flicking rates of focus would have indicated the point of contention involved in rapidly changing colour maps) of participants is not available. This provides one avenue of possible further research, as is discussed in section 7.

Participants slightly preferred the base design over the compared design when asked about visual appeal, which conforms with the results from the direct comparison question (see below).

Direct comparison

The questions directly comparing the two designs results in a dichotomy between answers, wherein participants preferred the base design in terms of visual appeal, but liked the compared design better in terms of conveyed information. The reasoning behind this split is composed of the following:

The colour map used to represent currents in the base design is comparatively more appealing specifically in a zoomed-out, big-picture overview of the simulation domain. The continuous and gap-free nature of the colour map lending itself particularly well to this configuration - the colour map covers the entire background image, and contains no occlusion in the layer itself.

The arrow glyphs of the compared design, on the other hand, are less appealing - especially in the zoomed-out view of the program, which is the default. This design contains a fixed number of glyphs to render the current layer, which, combined with the varying strength of the flow-fields, leads to occlusion of glyphs. As can be seen in figure 25, this results in a 'busier' appearance of the compared design - which was identified by participants as being visually unappealing.

This is one area where the compared design can be definitively improved: by implementing (regular) subsampling of arrow glyphs, visual occlusion may be prevented, thus improving the visual appeal of the compared design. As a matter of fact, this is one of the bigger limitations of this design/the program as a whole (see subsection Limitations).

Figure 25: Arrow glyphs occluding due to high flow strength

Another reason participants preferred the base design is possibly the chosen colour palettes. Specifically, the compared design's 'complementary' colour scheme uses a lighter colouration for freshly-spawned particles. This colouration does not contrast well against the white background layer (as demonstrated in figure 26), resulting in an indistinctive, less appealing visualisation.

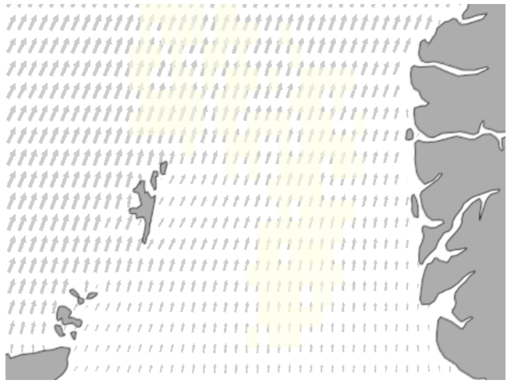
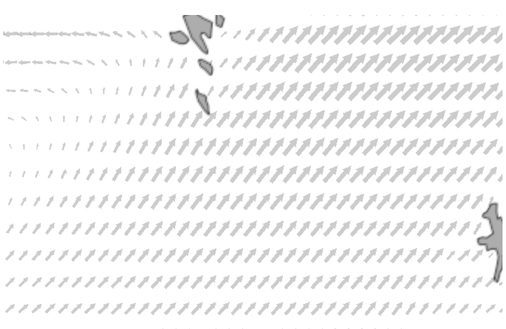


Figure 26: Low contrast of the complementary colour scheme in the compared design

The reason participants preferred the compared design in terms of conveyed information is similarly twofold: first of all, a rotating glyph provides a more accurate indication of the flow direction than a 10-part wind rose - resulting in the compared design providing more detailed information about the background fluid. Similarly, the way flow strength is encoded results in a more pronounced visual difference for the compared design over the base one. This is called the effectiveness of a channel, in Munzner's words [19]. Specifically, refer to figure 27a, where the modulating size of the arrow glyphs provides a more readily apparent difference in flow strength than the subtle modulation of transparency displayed in figure 27b.

The above does however come with a caveat: the effectiveness of the glyph size modulation (figure 27a) depends on the differently sized glyphs remaining distinct from each other. Local differences between glyphs are hard to discern or even invisible (remain unrendered) when observing the larger picture. This again, is where the strengths of a colour map may be more appropriate.



(a) Glyph size modulation as a result of differences in flow strength

(b) Transparency modulation as a result of differences in flow strength

Figure 27: Differences in flow strength visualisation between the base and compared designs

Modulation channels

Colour

Participants showed no preference between the ‘complementary’ and ‘contrasting’ colour schemes, but both of these configurations were clearly preferred over the ‘monochromatic’ colouration. The likely reason for this is that a monochromatic colour scale appears more uniform to the viewer; to borrow from Munzner again, there is less ‘popout’ [19] because there is less discrimination possible between extreme values - in this case, Eastern and Western or Northern and Southern currents.

Under this explanation the fact there is no preference between complementary and contrasting colour schemes lines up as well; both of these provide the desired popout between extreme values of the visualized data.

Saturation

A majority of participants preferred a fully saturated visualisation over a gradually saturated one. A likely reason for this is the visual distinctiveness, or lack thereof, of the gradual saturation option: in both designs, but especially in the base design, it can be quite hard to separate the current and particle layers from each other (see figures 28). This issue is compounded by the use of a white-gray background image, further reducing the ease with which participants could identify specific points of interest in the gradually saturated configuration.

By incorporating an additional channel (the hue channel), the fully saturated configuration has more options available to prevent the above issues of distinction, thus increasing visual saliency and boosting user engagement as compared to the gradually saturated configuration.

(a) in the base design

(b) in the compared design

Figure 28: Indistinctness in the gradual saturation configuration in both designs

Glyph shape

For glyph shape, participants preferred the order of circular > triangular > square > cross. We identify two possible reasons for this ordering:

First, there is the distinction between indexed and non-indexed shapes: of the tested glyph shapes, only circular glyphs are not indexable. The other three shapes can be modified to indicated direction (by rotating the shapes as they are displayed), but this was not done in this project. Indexed shapes are generally expected to rotate to indicate direction - especially when a time-dimension is involved as it is in this project. That the indexed shapes are instead static may have clashed with the way participants expected the glyphs to behave, thus resulting in a lower preference for these shapes.

The second reason for the observed ranking involves the ink coverage-ratio of the used glyph-shapes, as demonstrated in figure 29: the two least popular glyph shapes each have extreme ink coverage-ratios (close to 15% for cross-shaped glyphs and 100% for square glyphs), which results in a diminished or unwieldy appearance of these shapes respectively. The two more popular shapes meanwhile take up a medium amount of their bounding box (roughly 55% for triangular glyphs and about 80% for circular glyphs), which avoids the issues of diminutive appearance and unwieldiness inherent to the other two shapes and leads to a more pleasant appearance.

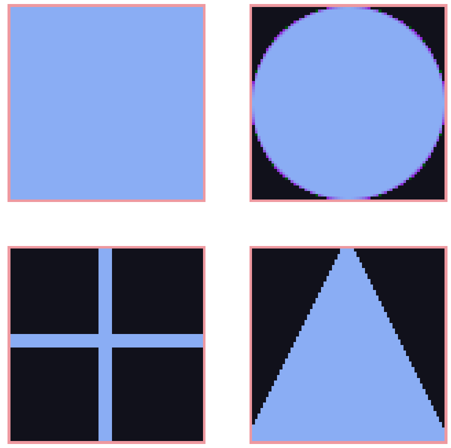


Figure 29: Glyph shapes and their bounding boxes

Circular glyphs specifically are the most popular of available shapes; aside from the reasons outlined above, humans are predisposed towards thinking of circles as ‘elegant’ or ‘beautiful’ [17], thus placing this shape clearly as the most appealing option of the tested glyph shapes.

6.1 User Engagement

After having discussed the results of each question separately, they can now be related back to the research question: how do different visualisations encourage retention of engagement for an interactive particle simulation?

We can see that interactivity is a key feature when it comes to user engagement; 75% of participants would have continued exploring the program. These participants had no deviating preference in term of visual design as compared to the other 25% of participants, indicating no clear deviation in visual design preference between those who are more interested in the subject of interactive particle simulation and those who are not.

Overall preference of participants leans towards the base design. This goes even more so for visual appeal specifically, which might indicate visual appeal to have a stronger impact on participants’ willingness to engage with the program than accuracy of conveyed information.

If the above reasoning holds, we can then say that maximising for visual appeal is preferred for purposes of user engagement - in this case, meaning the base design of a colour map for Eulerian flow-fields and a glyph-based mark for Lagrangian particle visualisation maximises user engagement retention. Of course, this maximisation should be done within reason: designing solely for visual appeal while neglecting realism in simulations would defeat the point of designing for simulations entirely.

In the above design, the Lagrangian particle glyphs should be either circular (non-indexed) or triangular (indexed) in nature, when designing for user engagement.

There is also an argument to be made that the compared design shines in high-fidelity situations - being as it is better at conveying information to the user - and a hybrid solution in which the program switches between the two designs might be ideal. This would require additional work in determining the ideal point at which to do so, i.e. at certain zoom levels or after a certain amount of time interacting with the program has elapsed. Depending on implementation this approach might however confuse the user - specifically if the transition between designs is not done smoothly (or, 'seamlessly' as Blascheck et al. recommend [3]), which might lead to the adverse effect of disruption of user engagement, 'breaking the flow' a user is in while interacting with the program.

Finally, to place both designs put forth in this project in the same framework under which the related literature in section 2 was analysed, we provide an overview of the used mark/channel combinations in both designs in table 3 (here, 1 = base design, and 2 = compared design). As can be seen here, the two designs are non-conflicting: both designs use exclusively different marks and channels for their respective Eulerian and Lagrangian components.

Marks & Channels	Colour map	Streamline	Contour	Glyph	Info visualisation	Cross section	Noise map	Texture synthesis	Hue	Saturation	Transparency	Form	Boundary	Direction	Contrast	Structure
1	E			L					E	L	E					
2	L			E					L		L			E		

Table 3: Used mark/channel combinations in both designs

6.2 Limitations

However, both the user study and the program itself have a few limitations, or areas which could benefit from further attention.

User study: the questionnaire used for the user study purposefully refrained from asking about any personally identifiable information, as this would result in complications with the GDPR. This was necessary to conduct the user study in the available time for this project, but did limit the types of questions that could be asked in the questionnaire.

Program: the program itself was developed to accept an arbitrary number of visualisation designs; however, each of these is limited both by the VTK framework and reliance on the simulation logic in the program, which expects a singular Latitude/Longitude pair per particle (removing the possibility of simulating anything that isn't a single particle, i.e larger objects).

The program also contains no functionality for subsampling of any data, as a result of complications with the VTK pipeline. Undoubtedly this technique can be implemented with the used technologies, but due to time constraints this was not done for this project. As discussed earlier, this is a technique from which the compared design specifically might have benefited in terms of visual appeal.

7 Future Work

While this project provides an exploratory insight into the subject of visual design for particle simulation, multiple avenues of further investigation remain open.

First off, the scope of this project was limited to a direct comparison of two visualisation designs - both of which largely conveyed the same information. By adding different channels of information to the designs, further insight may be gained on the best way to visualise this additional information when aiming for user engagement retention[3].

Additionally, visualisation channel modulation presents another research opportunity. Per Munzner's 'visualisation analysis and design', modulation of different channels will have more or less impact on the overall appearance of a visualisation (see also figure 2) [19]. More research into the influence of different channel modulations has the potential to offer deeper insights into user engagement as it relates to particle-flow visualisation.

Another possible avenue of investigation is the method of data collection; while a user study gives direct, actionable results, it also introduces the possibility for response bias. One way to circumvent this is by making use of indirect forms of data collection, like eye-tracking software.

Blascheck et al. [3] use eye-tracking data to investigate how participants discover functionality of an interactive visualisation. By incorporating similar markers such as 'fixtures', 'resting time', or 'Areas of Interest' the visually more salient aspects of a design can be identified, which may then give an insight into the most appealing aspects of the design, and accordingly, the effectiveness of those aspects at retaining user engagement.

Within the scope of a user study further improvements are possible as well: the questionnaire contained questions about interaction time which aimed to gauge the willingness of participants to keep interacting with the program. While useful for determining the overall interactiveness of the program, these questions did not give any insight into either of the visualisation designs. The user study could be improved by using a slightly modified setup: instead of allowing users free exploration of the program as a whole (and then noting the interaction times thereof), each design can be shown and interacted with separately. This would give more detailed data about the willingness of users to interact with each design individually.

Another change to the questionnaire: users were asked about the adeptness of each design at conveying certain channels of information. Modifying these questions to instead ask users to indicate which of the designs they preferred for each information channel - rated between 'strong preference for design 1' and 'strong preference for design 2' - could create a clearer juxtaposition between the designs, leading to further insight into which design is preferred by users. This approach was already used in the 'direct comparison' questions on

[3]For example, the incorporation of 'history lines' of particles, which was considered in the early stages of this project

the questionnaire.

Similarly, the questionnaire purposely avoided collecting personally identifiable information - so as to prevent complications with the GDPR - but questions of this nature could have provided context for user responses to other questions; i.e. visual impairment or colour vision deficiency of participants could impact the way participants experience modulation of the hue and saturation channels.

Lastly, the interactivity supported by the simulation software can be enhanced in multiple ways. Obviously the incorporation of additional channel modulations might require changes to the software - as would be the case for i.e. different levels of subsampling - but entirely separate methods of interaction might similarly have an impact on user engagement.

Another dimension of the simulation which went almost wholly unexplored is that of time-axis controls: the program as-is provides an exclusively constant time-delta ('playback speed') of the simulation. While the program supports (un)pausing the simulation, further options remain open for interacting with the time-dimension - such as an adjustable playback speed, or 'seeking' of a point in the simulation by fast-forwarding and fast-backwarding.

8 Conclusion

The primary research question this project aimed to investigate is the effect of different visualisation designs of particle-flow visualisation on retaining user engagement. To answer this question, we developed a simulation program and performed a user study comparing two visualisation designs against each other. The user study also included channel modulation of the tested visualisation designs in its scope.

After analyzing the results of the user study, we conclude that participants have an inclination towards visualisations containing high levels of saliency; this is mostly expressed through variance in the *hue* and *saturation* channels. Furthermore, when planning a visualisation design for user engagement, a clean visualisation is preferred over one that provides more accurate information. That is to say, discarding the accuracy of glyph-based flow visualisation in favour of a colour-map with a cyclic encoding of flow direction in the *hue* channel will lead to an increased favourable perception of the design.

However, if conveying information about the simulation state is considered paramount, it is advisable to make use of the more detailed glyph mark for flow visualisation. This similarly allows one to use the size of the glyphs to modulate for flow strength, which conveys more accurate information than the transparency modulation used in this project.

Furthermore, obvious contrast between the background and any visualisation is essential; in this project, one configuration led to a subpar contrast between represented particles and the background layer, significantly reducing visual appeal.

Lastly, unless there is a specific reason to do so, using the 'elegant' circular glyph shape is preferable over the more extreme ink coverage-ratios of squares or cross-shaped glyphs. If orienting of glyphs is considered important, the triangular glyph may be used instead; there is only a slight difference in preference for the circular glyphs over triangular ones.

References

[1] *About VTK*. URL: https://docs.vtk.org/en/latest/about.html.

[2] *Atlantic- European North West Shelf- Wave Physics Reanalysis*. URL: https://data.marine.copernicus.eu/product/NWSHELF_REANALYSIS_WAV_004_015/description.

[3] Tanja Blascheck et al. "Exploration strategies for discovery of interactivity in visualizations". In: *IEEE transactions on visualization and computer graphics* 25.2 (2018), pp. 1407–1420.

[4] Rita Borgo et al. "Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications." In: *Eurographics (state of the art reports)*. 2013, pp. 39–63.

[5] *Boskalis Beach Cleanup Tour*. URL: https://www.beachcleanuptour.nl/.

[6] *Colorbrewer: Color Advice for Maps*. URL: https://colorbrewer2.org/#type=sequential&scheme=Blues&n=9 (visited on 07/10/2024).

[7] *Colormap reference*. URL: https://matplotlib.org/stable/gallery/color/colormap_reference.html (visited on 07/10/2024).

[8] Hanna Corell. "Applications of ocean transport modelling". PhD thesis. Department of Meteorology, Stockholm University, 2012.

[9] Diana De Padova et al. "Comparison between the Lagrangian and Eulerian Approach for Simulating Regular and Solitary Waves Propagation, Breaking and Run-Up". In: *Applied Sciences* 11.20 (2021). ISSN: 2076-3417. DOI: 10.3390/app11209421. URL: https://www.mdpi.com/2076-3417/11/20/9421.

[10] Julie Jebeile. "Explaining with simulations: Why visual representations matter". In: *Perspectives on Science* 26.2 (2018), pp. 213–238.

[11] Mikael LA Kaandorp et al. "Using machine learning and beach cleanup data to explain litter quantities along the Dutch North Sea coast". In: *Ocean Science Discussions* 2021 (2021), pp. 1–33.

[12] C Kehl, RPB Fischer, and E van Sebille. "Practices, pitfalls and guidelines in visualising Lagrangian ocean analyses". In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 4 (2021), pp. 217–224.

[13] Christian Kehl et al. "Efficiently simulating Lagrangian particles in large-scale ocean flows—Data structures and their impact on geophysical applications". In: *Computers & Geosciences* 175 (2023), p. 105322.

[14] Christian Kehl et al. "Parcels 2.2-an increasingly versatile, open-source Lagrangian ocean simulation tool". In: *EGU general assembly conference abstracts*. 2021, EGU21–1033.

[15] Rostislav Khlebnikov et al. "Procedural Texture Synthesis for Zoom-Independent Visualization of Multivariate Data". In: *Computer Graphics Forum*. Vol. 31. 3pt4. Wiley Online Library. 2012, pp. 1355–1364.

[16] Michael Lange and Erik van Sebille. "Parcels v0. 9: prototyping a Lagrangian ocean analysis framework for the petascale age". In: *Geoscientific Model Development* 10.11 (2017), pp. 4175–4186.

[17] Manuel Lima. *Why do we find circles so beautiful?* URL: https://www.sciencefocus.com/science/why-do-we-find-circles-so-beautiful (visited on 06/24/2024).

[18] Grzegorz Łukaszewicz and Piotr Kalita. *Navier–Stokes Equations: An Introduction with Applications*. Springer, 2016.

[19] Tamara Munzner. *Visualization analysis and design*. CRC press, 2014.

[20] *Openstreetmap*. URL: https://www.openstreetmap.org/about.

[21] Frits H. Post and Theo van Walsum. "Fluid Flow Visualization". In: *Focus on Scientific Visualization*. Ed. by Hans Hagen, Heinrich Müller, and Gregory M. Nielson. Berlin, Heidelberg: Springer Berlin Heidelberg, 1993, pp. 1–40. ISBN: 978-3-642-77165-1. DOI: 10.1007/978-3-642-77165-1_1. URL: https://doi.org/10.1007/978-3-642-77165-1_1.

[22] *QGIS: A Free and Open Source Geographic Information System*. URL: https://qgis.org/en/site/index.html.

[23] *QT-Creator*. URL: https://github.com/qt-creator/qt-creator (visited on 06/20/2024).

[24] Milad Rakhsha, Christopher E. Kees, and Dan Negrut. "Lagrangian vs. Eulerian: An Analysis of Two Solution Methods for Free-Surface Flows and Fluid Solid Interaction Problems". In: *Fluids* 6.12 (2021). ISSN: 2311-5521. DOI: 10.3390/fluids6120460. URL: https://www.mdpi.com/2311-5521/6/12/460.

[25] *Scientific colour maps*. URL: https://www.fabiocrameri.ch/colourmaps/ (visited on 07/10/2024).

[26] Kristen M Thyng et al. "True colors of oceanography: Guidelines for effective and accurate colormap selection". In: *Oceanography* 29.3 (2016), pp. 9–13.

[27] David Wichmann et al. "Detecting flow features in scarce trajectory data using networks derived from symbolic itineraries: an application to surface drifters in the North Atlantic". In: *Nonlinear Processes in Geophysics Discussions* 2020 (2020), pp. 1–20.

[28] Cui Xie et al. "A survey on visual analysis of ocean data". In: *Visual Informatics* 3.3 (2019), pp. 113–128. ISSN: 2468-502X. DOI: https://doi.org/10.1016/j.visinf.2019.08.001. URL: https://www.sciencedirect.com/science/article/pii/S2468502X19300440.

Feedback form user study

* Indicates required question

1. Are you a student? *

Mark only one oval.

- Yes, full-time
- Yes, part-time
- No

2. How clear was it that the simulation supports interaction (camera movement, particle placement)? *

Mark only one oval.

- Very unclear
- Somewhat unclear
- Neither unclear nor clear
- Somewhat clear
- Very clear

3. What did you think of the introduction to the simulation? *

Mark only one oval.

- Not at all helpful
- Not very helpful
- Somewhat helpful
- Reasonably helpful
- Very helpful

Conveyed information

This section asks about the information technique 1 is supposed to convey.

6. How would you rate technique 1 (ECol + LGlyph) on conveying information about the speed of the background current? *

Mark only one oval.

	1	2	3	4	5	
Not	○	○	○	○	○	Very good

7. How would you rate technique 1 (ECol + LGlyph) on conveying information about the direction of the background current? *

Mark only one oval.

	1	2	3	4	5	
Not	○	○	○	○	○	Very good

8. How would you rate technique 1 (ECol + LGlyph) on conveying information about the movement of placed particles? *

Mark only one oval.

	1	2	3	4	5	
Not	○	○	○	○	○	Very good

9. How would you rate technique 1 (ECol + LGlyph) on conveying information about the age of placed particles? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very good

10. How would you rate technique 1 (ECol + LGlyph) on conveying information about the beached state of placed particles? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very good

Conveyed information 2

This section asks about the information technique 2 is supposed to convey.

11. How would you rate technique 2 (EGlyph + LCol) on conveying information about the speed of the background current? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very good

12. How would you rate technique 2 (EGlyph + LCol) on conveying information about the direction of the background current? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very good

13. How would you rate technique 2 (EGlyph + LCol) on conveying information about the movement of placed particles? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very good

14. How would you rate technique 2 (EGlyph + LCol) on conveying information about the age of placed particles? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very good

15. How would you rate technique 2 (EGlyph + LCol) on conveying information about the beached state of placed particles? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very good

Visual appeal

This section asks about the appeal of both techniques.

16. How appealing (visually speaking) would you say technique 1 is? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very appealing

17. What aspects of technique 1 did you find most appealing? Please select at most 3 options. *

Check all that apply.

- [ ] Use of colour maps to represent currents
- [ ] Use of transparency to represent current strength
- [ ] Use of glyphs to represent particles
- [ ] Use of colours to represent particle age
- [ ] Other: ____________________

18. How appealing (visually speaking) would you say technique 2 is? *

Mark only one oval.

	1	2	3	4	5	
Not	◯	◯	◯	◯	◯	Very appealing

19. What aspects of technique 2 did you find most appealing? Please select at most 3 options. *

Check all that apply.

- [ ] Use of glyph size to represent current strength
- [ ] Use of glyph direction to represent current direction
- [ ] Use of colour maps to represent particle locations
- [ ] Use of colours to represent particle age
- [ ] Use of transparency to represent particle density
- [ ] Other: ____________________

20. Which visualisation technique did you prefer in terms of visual appeal? *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strong preference for technique 2

21. Which visualisation technique did you prefer in terms of conveyed information? *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strong preference for technique 2

Modulation Channels

This sections asks about the available channel modulations.

22. Please rank your preference for the different colour options. *

Mark only one oval per row.

	Highest preference	Medium preference	Lowest preference
Complementary colours	⬭	⬭	⬭
Contrasting colours	⬭	⬭	⬭
Monochromatic	⬭	⬭	⬭

23. Please rank your preference for the different saturation options *

Mark only one oval per row.

	Highest preference	Lowest preference
Fully saturated	⬭	⬭
Gradually saturated	⬭	⬭

24. Please rank your preference for the glyph-shape options *

Mark only one oval per row.

	Highest preference	Second highest preference	Second lowest preference	Lowest preference
Circle	◯	◯	◯	◯
Triangle	◯	◯	◯	◯
Square	◯	◯	◯	◯
Cross	◯	◯	◯	◯

Closing questions

25. Do you have any other feedback about the visualisation techniques used?

26. Do you have any other feedback about the simulation program?

27. Do you have any feedback about the user study itself?