Cartography M.Sc.

Master thesis

Combining novel visualizations of temporal changes with maps using satellite time series of vegetation moisture

Monika Krautschneider

Technical University of Munich

TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

TECHNISCHE UNIVERSITÄT DRESDEN

UNIVERSITY OF TWENTE.
ITC

2023

Statement of Authorship

Herewith I declare that I am the sole author of the submitted Master's thesis entitled:

"Combining novel visualizations of temporal changes with maps using satellite time series of vegetation moisture"

I have fully referenced the ideas and work of others, whether published or unpublished. Literal or analogous citations are clearly marked as such.

Vienna, 6.11.2023

Abstract

This thesis offers a practical and theoretical review on visualizing time series data for Life Fuel Moisture Content, Fire Weather Index, and Temperature, which all play significant roles in forecasting wildfire occurrences and spread. The study explores various techniques, including combining novel visualizations of Warming Stripes and the Climate Spiral with maps, to present the multivariate data of these fire risk factors. One visualization employs the Warming Stripes color scheme and small multiples to display the evolution and comparison of all three variables over time. Another visualization unites the Climate Spiral with choropleth maps to emphasize the seasonality and data accessibility of Life Fuel Moisture Content and Fire Weather Index. Furthermore, this thesis will provide an evaluation framework that demonstrates the superiority of the created visualizations compared to other selected spatio-temporal visualizations. Additionally, this research addresses a gap in the literature by evaluating the awareness and effectiveness of Warming Stripes. However, the survey results did not support the reputation of this visualization.

Key words: spatio-temporal visualizations, life fuel moisture content, wildfire risk, warming stripes, climate spiral

Contents

Figures

Tables

List of Abbreviations

EDA	Exploratory data analysis
FWI	Fire weather index
LFMC	Life fuel moisture content
NA	Not available (in terms of data)
TEMP	Temperature

1 Introduction

In this section the topic of this thesis will be introduced. First, key elements of the research will be given for an insight into the topics covered within this work. In order to contextualize this thesis, a review of the fields and methods from which this thesis draws inspiration from will follow. After defining the state of the art in the creation of spatio-temporal visualizations, the research objectives and research questions to be addressed in this study will be stated.

1.1 Key elements

Climate change has evolved from a topic projected to show its effects in the near future to a problem that is already demonstrating its consequences today. One of the outcomes that occurred faster than expected by the scientific community predicted is the intensification of wildfires (NASA's Jet Propulsion Laboratory, 2023). As global temperatures continue to rise, the frequency and intensity of fire weather will also increase further. Therefore, there is an urgent need to improve models used for the estimation of fires and to understand the variables that contribute to favorable conditions for wildfires (Jones et al., 2022).

Similar to climate change is the estimation and documentation of wildfires a multifaceted phenomenon. Several factors contribute to the spread and occurrence of wildfires such as fuel availability, fuel moisture, climatic conditions, and sources of ignition (Chuvieco et al., 2010; Forkel et al., 2023; Jones et al., 2022). Some of these factors like life fuel moisture content (LFMC) are still highly researched on. LFMC represents the moisture content of the living vegetation canopy and is a well-established concept in wildfire risk assessment. However, the large-scale derivation of LFMC by microwave remote sensing is a novel technique. This data acquisition method enables global and daily retrieval of LFMC data over the past few decades, resulting in a significant increase in available data compared to traditional site measurement methods (Forkel et al., 2023).

Especially in times of high computational power and massive amounts of data, accessible as never before but mostly unfiltered, it is important to find communication tools that help the audience outside of the scientific community understand the information being worked on. Scientific visualizations play an important role for connecting scientific research and their findings with people outside of the academic world. They are powerful tools to create a bridge between the scientific community and people which are interested in their findings. Therefore, the combination of filtered information as a visually pleasing representation has the advantage of communicating complex contents in a coherent way to an audience which is not familiar with the subject (Grainger et al. 2016).

In recent years, two novel visualizations depicting variables related to climate change in form of increasing global temperatures have gained widespread attention on social and public media. The Climate Spiral and Warming Stripes, both created by Ed Hawkins are claimed to be effective visualizations due to their popularity and familiarity (e.g., Dixon, 2022; National Centre for Atmospheric Science, 2018; O'Connor, 2023; Sengupta, 2022). Although these visualizations are widely used to convey the message of increasing global temperatures, their familiarity as well as their effectiveness due to simplicity have not yet been empirically studied. Nonetheless, these visualization concepts of conveying information in a simple manner should be utilized to communicate various variables that are affected by climate change and adapted to the needs of visualizing a multifaceted phenomenon.

This research will address the identified research gaps of the lack of empirical research on the popularity and effectiveness of Warming Stripes and as well as on the adaptation of the novel visualizations Climate Spiral and Warming Stripes to incorporate multivariate data within static visualizations. For this study, environmental fire risk factors of LFMC, fire weather index (a metric that measures the potential for fire spread based on weather conditions and fuel moisture) and temperature will be considered. Additionally, an incorporation of maps for this thesis is aimed to illustrate geographical patterns and regional variations of these variables which are not visible within the original visualizations.

1.2 Motivation

Visualizations are captivating and valuable tools that convey information effectively to people. On the one hand they can help clarify complex concepts but on the other hand can also be used to awake interest in people even if they are not familiar with the context. Warming Stripes, in particular, are said to be capable of accomplishing both, which is why the author found them intriguing. Combining the widely known visualization with cartography through the addition of cartographic methods was found to be an ideal topic.

Furthermore, the author's background in Environmental System Sciences during their bachelor's degree, gave a theoretical background on the effects of anthropogenic climate change. Nonetheless, working with large-scale data sources such as satellite time series was a rather unfamiliar topic but the challenge of combining knowledge and creating a visualization from the raw data to the final result was gladly accepted.

Making data clear and understandable for an audience that might not be familiar with can be seen as a big motivator. In times of excessive data availability, it is found important to being able to communicate the main messages in a way that also people outside of the scientific community can understand it.

1.3 Literature Review

Most of the time the methods described apply explicitly to spatio-temporal data, thus data which incorporates a temporal and a spatial reference. The notation (spatio-)temporal data is used whenever concepts are applicable to temporal as well as spatio-temporal data.

A literature review was selected as a method to examine essential information on the creation and evaluation of temporal and spatio-temporal visualizations. Correspondingly, literature from various fields including Cartography, Data and Information Visualization, Data Analysis and Statistics was studied. The chapter includes a theoretical visualization preparation workflow was created to understand the crucial elements in the creation of a spatio-temporal visualization. This includes identifying the target audience, characterizing the data, identifying the key messages through tasks and selecting an appropriate type of visualization.

1.3.1 Visualization process

The creation of temporal and spatio-temporal visualizations heavily relies on the target audience, the features of the selected data and the purpose for which the visualization is being made (Knaflic, 2015). It is challenging, if not unfeasible, to create a general process for (spatio-)temporal visualizations that encompasses different types of data and/or visualizations for a broad audience. However, some aspects to consider when creating (spatio-)temporal visualizations are applicable to all workflows. As such, a universal visualization preparation workflow is suggested on the following pages, with the intention of facilitating the creation of spatio-temporal visualizations and in our case of the modification of novel visualizations for a multivariate dataset. The workflow was created based on literature on (spatio-)temporal visualizations. For instance, Aigner et al. (2011) introduced a workflow that revolves around three fundamental questions regarding the so-called "visualization problem" of visualizing temporal data in general. Before creating a temporal visualization, the questions that should be asked are: "What?", "Why?" and "How?". These same three questions are also referenced by Munzner (2015) for the preparation of visualizations. The questions deal with 1. the data and temporal primitives used for the visualization ("What?"), 2. the tasks that define the purpose of the visualization ("Why?") and 3. the tools or methods for the presentation ("How?). Although Andrienko & Andrienko (2006) did not explicitly list these questions, they also divided their visualization workflow on spatial-temporal data into data, tasks, and tools. Moreover, literature on (spatio-)temporal visualizations frequently overlook a crucial aspect, which is a central element in the literature on data visualization and information visualization: the consideration and specification of the target audience.

The following topics need to be inspected carefully when creating a spatio-temporal visualization according to Aigner et al. (2011); Andrienko & Andrienko (2006); Knaflic (2015); Munzner (2015):

1) Target audience
2) Data classification
3) Task specification
4) Modelling of Time
5) Visualization types

1.3.1.1 Target audience

The initial step for creating a spatio-temporal visualization involves considering the target audience. According to Knaflic (2015), a more specific description of the audience facilitates determining the required depiction and visualization style. Moreover, tailoring the message to the audience's interests and knowledge enhances the visualization's impact. Based on Knaflic's work Deacon et al., (2020) formulated questions on how to define the target audience of data visualizations and the message the creator of the visualization wants to transmit. The questions concerning the audience itself include their background (e.g., educational, or professional experience) as well as their technical knowledge (familiarity with the subject). The producer of visualizations must always remember to take the general message they are attempting to communicate into account and the degree of detail necessary to deliver this message. The purpose for communicating content to a specific audience should be considered by examining why it is being communicated to that specific audience in the first place. It is necessary to define which actions or types of understanding the audience is expected to have after viewing the visualizations. This approach simplifies the content and reduces the complexity of the information presented.

1.3.1.2 Data Classification

Several studies in the field of data visualization (Aigner et al., 2011; Andrienko & Andrienko, 2006; Munzner, 2015, amongst others) put forward that data classification is an essential component of the visualization process. Different approaches on data classification of (spatio-)temporal data are available in literature and are employed to better comprehend the data and determine the type of visualization that best suits the selected data.

Andrienko & Andrienko (2006) stress the importance of distinguishing data elements into referential components (referrers) and characteristic components (attributes) to understand the data structure and identify dependent and independent variables. Referential components define the context of the data (e.g., where and when a phenomenon takes place), while characteristic components describe the actual measurement, or in other words the retrieved attributes. While referrers can be chosen arbitrarily (e.g., year of observation or chosen spatial extent like a country or a region), attributes are dependent based on the choice of referrers.

According to Peuquet's (1994) triad framework, spatio-temporal data can be divided into three elements: location, time and object. By classifying these elements, fundamental questions which each temporal visualization should be able to answer, can be identified:

- When + Where → What: By combining the information on time and location, the question on what object is depicted can be answered.
- What + When → Where: By combining information on the object and time, the question on the location can be answered:
- Where + What → When: By combining information on the location and time, the question on time can be answered.

Also, Aigner et al. (2011) provide a comprehensive overview on how to classify data to gain better understanding of its structure. They distinguish between four categories: scale, frame of reference, kind of data and number of variables (see Figure 1Figure 1). Data is either qualitative or quantitative based on their scale. Quantitative data is described by variables that are based on a metric range, allowing them to be compared, while qualitative data is not. Abstract data without spatial references and spatial data can be distinguished within the frame of reference. Furthermore, the kind of data is also important when creating a visualization, as noted by the authors. A discrete point in time and space characterizes an event, while a continuous condition describes states. Distinguishing between event and state data can be challenging, and while it is possible to combine both types within a single visualization, it is essential to clearly communicate the type of data to the audience. Aigner et al. (2011) also highlight the importance of differentiating between univariate and multivariate data, as not all visualization methods are suitable for the latter.

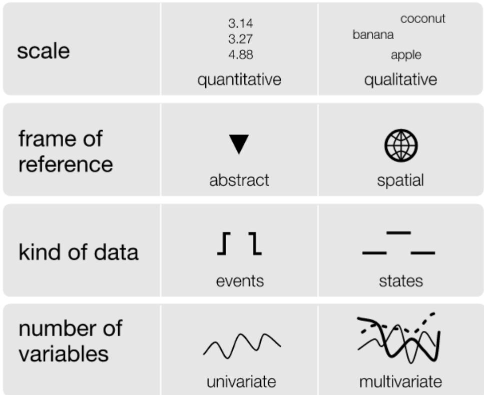


Figure 1: Data characterization by Aigner et. al. (2011)

1.3.1.3 Task specification

Identifying tasks or inquiries to understand the focus or the intended message of the visualization is the next step in creating a visualization. These tasks assist in discovering the fundamental messages that must be conveyed to the audience and offer direction when constructing the visualizations to avoid deviation from the messages (Aigner et al., 2011).

Tasks are specified during the creation of a visualization to identify messages, that should be depicted within the visualization, and to define questions that should be answered with the visualization. MacEachren (1995) defined seven temporal tasks descriptions (see Table 1) which incorporate questions to help structure the content of a visualization and effectively communicate complex temporal information. The aim of a spatio-temporal visualization has to address at least one of the inquiries defined in the tasks, with the possibility of incorporating information covering several tasks.

	Question	**Starting Point**	**Search for**	**User task type**	**Example**
Existence of data element	*Does* a data element exist at a specific time?	Time point or time interval	Data element at that time	identification	Was a measurement made in June, 1960?
Temporal location	*When* does a data element exist in time?	Data element	Time point or time interval	localization	When did the Olympic Games in Vancouver start?
Time interval	*How long* is the time span of beginning and end of the data element?	Data element	Duration (beginning and end of data element)	localization	How long was the processing time of dataset A?
Temporal pattern	*How often* does a data element occur?	Time point or time interval	Frequency of data elements at portion of time and based on this the detection of a pattern	identification	How often was Jane sick last year?
Rate of change	*How fast* is a data element changing?	Data element	Magnitude of change over time	localization	How did the price of gasoline vary last year?

	/ *How much difference* is there from data element from time to time?				
Sequence	In *what order* do data elements occur?	Data elements	Temporal order of different data elements	localization	Did the explosion happen before or after the car accident?
Synchronization	Do data elements exist *together*?	Data elements	Occurrence at the same point of time	localization	Is Jill's birthday on Easter Monday this year?

Table 1: Temporal tasks adapted from MacEachren (1995, as cited in Aigner et al., 2011)

Furthermore, the tasks can be classified into two main types: localization tasks which are specified by the knowledge of at least one or more data elements, together with finding out the time and space of this element, while identification tasks involve the search for data elements or objects with known time and space components (Aigner et al., 2011). This division of tasks shows similarity with Peuquet's (1994) triad framework of interconnections between location, time and object, which was discussed earlier.

Andrienko & Andrienko (2006) divided their task model a bit differently. By categorizing the message of a visualization into specific tasks, it becomes easier to understand how the visualization should be designed and at what scope the data needs to be handled. These insights into the data facilitates finding the most effective way to present complex spatio-temporal information to the specified target audience. The task distinction helps in determining whether the information presented by the visualization creator concerns the entire dataset or specific points in time. Their task model is based on the task definition of Bertin (1983, as cited in Andrienko & Andrienko, 2006), which differentiates tasks based on the scope of the data that is analyzed. Accordingly, the authors developed a task taxonomy, distinguishing at the first level between elementary and synoptic tasks (see Figure 2). Elementary tasks describe questions on single data values, while synoptic tasks investigate sets or even entire datasets.

Within synoptic tasks, a further differentiation can be made between descriptive tasks and connectional tasks. The former seeks patterns and significant behaviors, while the latter strives to investigate these behaviors further by identifying relationships within them or cause-and-effect-relationships with other variables.

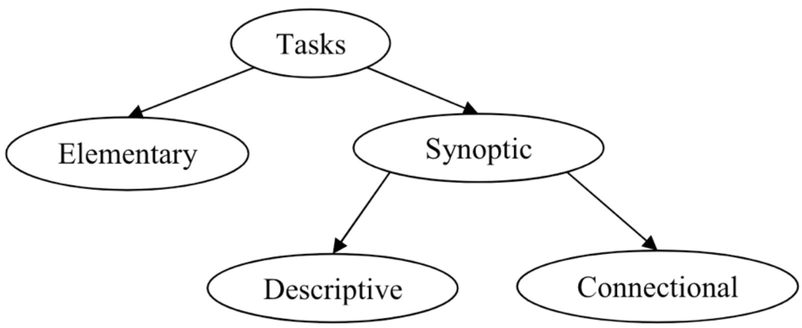


Figure 2: Task classification framework by Andrienko & Andrienko (2006)

Other task definition approaches can be found in literature which may draw on or incorporate elements of task definitions presented by MacEachren (1995) and Andrienko & Andrienko (2006). Nonetheless, these two approaches were selected based on their established effectiveness in the research of (spatio-)temporal visualizations.

1.3.1.4 Types of Visualization

Going back to the visualization problem defined by Aigner et al. (2011), the questions on "What?" described by data classification and the "Why?" described by task specification, can be answered with the approaches explained on the previous pages. With the answers of those two questions, also the answer to "How?" in terms of how the information can be visualized, can be answered.

This subchapter shows an excerpt of types of (spatio-)temporal visualizations that are suitable for the depiction of the selected remote sensing data on LFMC, FWI and TEMP. The visualizations have been selected based on the data classification and task specification which were theoretically described in the previous pages. It is important to note that these visualizations are not the only ones suitable; they present only a small fraction of the available options in depicting these datasets. Nonetheless, a set of visualizations for these datasets were carefully selected based on their appropriateness. A thorough comparison of the advantages and disadvantages of these (spatio-)temporal visualizations is provided.

Line Plot

The most common way of visualizing time series is by using line plots. Cartesian planes typically display data points placed from left (representing earlier points in time) to right (recent points) in a linear progression of values over time. Line graphs depict a continuous timeline by connecting these data points (Aigner et al., 2011). This leads to a clearly understood temporal evolution by the audience, but this illusion of continuity may lead to drawing false conclusions by the viewers. Although line graphs use simple lines, when several lines with different trends cross each other, it can create a visually complex depiction. Therefore, the audience needs to maintain a constant connection to the legend (Monmonier, 1989). It is noteworthy that line plots inherently include geographical references. Nonetheless, incorporating spatial features can be achieved by labeling the lines with the corresponding geographic names (see Figure 3) or by using line charts as symbols on choropleth maps (Aigner et al., 2011; Monmonier, 1990).

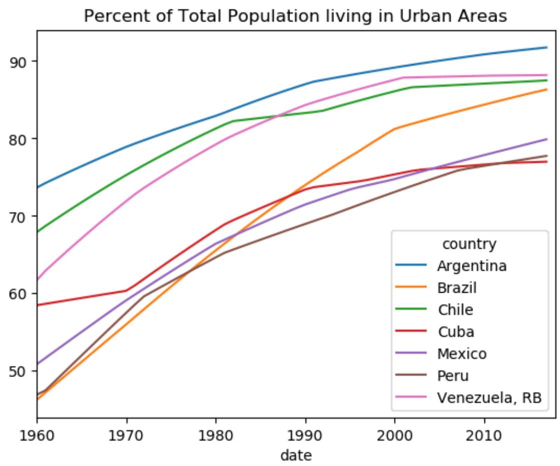


Figure 3: Example of a Line Graph with Geographical References from Mikecz (2019)

Horizon Graph

Horizon graphs were originally developed for analyzing financial stock data and are still commonly used in that field due to their capacity for comparing multiple time series within limited space (Reijner, 2008). The creation of a horizon plot is executed by modifying a simple line plot (see Figure 4). A divergent color scheme is employed with negative values depicted in red and positive values in blue. Moreover, the data range is segmented into various categories within the color spaces using uniformly distributed thresholds. As values increase, color saturation also intensifies. Additionally, negative values are mirrored horizontally at the baseline which is located at position 0 (Aigner et al., 2011; Reijner, 2008). This visualization type allows for the display of more data within a single visualization, which is highly beneficial for depictions in limited spaces. Nevertheless, the visualization could potentially distort patterns and mentally separating the graphs in negative and positive values may be cognitively demanding (Jabbari et al., 2018). Furthermore, the colors need to be adjusted to specific data; while using red for negative and blue for positive values might work in some cases, this color scheme leads to confusion with other variables such temperature.

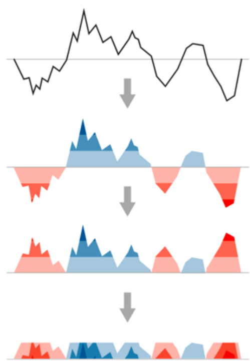


Figure 4: Process of Creation of a Horizon Graph from Aigner et al. (2011)

Small Multiples

Small multiples (see Figure 5) are a form of time series visualization that displays a sequence of maps with each map representing a single point in time. They are considered a general visualization concept rather than a specific visualization type according to Tufte (1983/1990, as cited in Aigner et al., 2011). The depiction of time is linear, and each map can be compared to a single data point within a line plot. Aigner et al. (2011) note that the concept can be applied to various cartographic visualization techniques such as those defined by Bertin (1983), which include changes in position, size, shape, value, hue, orientation and texture (see Figure 6). However, it is important to note that the use of these techniques on comparably small maps results in a loss of detail. Additionally, small multiples have the same disadvantage as line plots in only depicting time events and presenting snapshots, without revealing the time between the maps. Depicting a non-linear trend can lead to confusion when only specific points in time are displayed; the viewer cannot know how the variables behave between those points. Furthermore, the same interval between the individual maps must be chosen, otherwise this will lead to confusion.

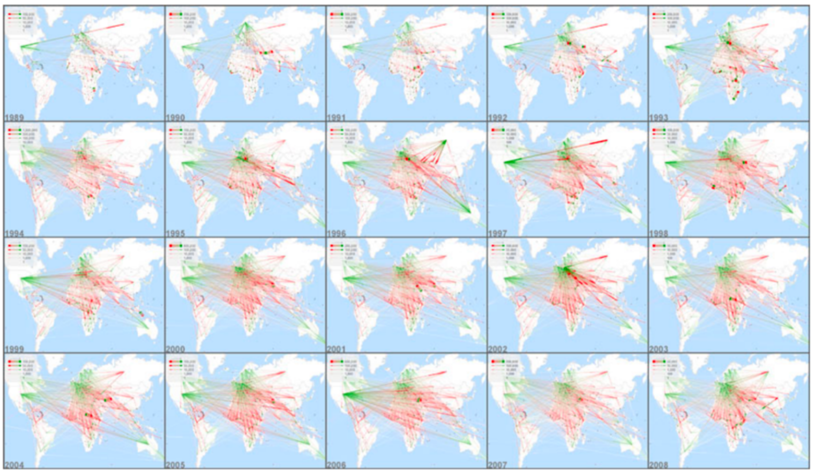


Figure 5: Small Multiples with Link Maps from Aigner et al. (2011)

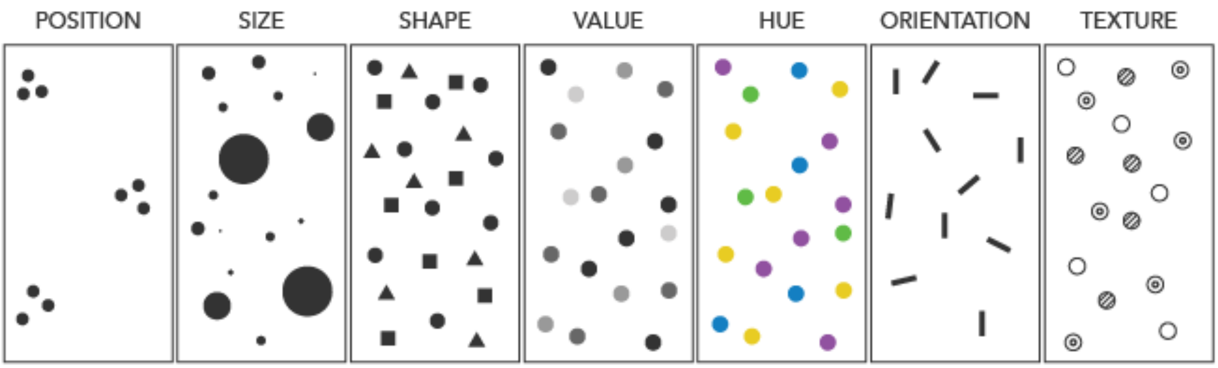


Figure 6: Visual Variables by Bertin (1983) from Axis Maps (n.d)

Climate Spiral

The visualization type that the Climate Spiral (see Figure 7) uses is based on a spiral graph which is a commonly used temporal depiction. Graphs like this are often used to reveal seasonal patterns, which are not visible within linear depictions, just the correct cycle needs to be chosen for revealing these patterns (Aigner et al., 2011). The animated visualization shows a circular graph with monthly temperature anomalies starting from 1850 to initially 2016. Each monthly's temperature value is depicted as a data point which is connected to the previous and the consecutive point in time. Even though the visualization type of spiral graphs existed already before, this specific visualization was well received. First published in 2016 on the twitter account of the creator Hawkins E. [@ed_hawkins] (2016) the Climate Spiral was widely shared on social media and other media platforms. As of 2019, the original tweet alone had been viewed by more than 3.7 million people, not including views on other media channels. The Climate Spiral also appeared during the opening ceremony of the Olympic Games in Rio. The people involved in the process of this animated visualization state several reasons why it was that successful. First, the familiarity of the audience to the variable temperature; people understand immediately what kind of data is depicted. Second, the Climate Spiral was created by scientists, who are people are usually conceived as trustworthy and last, the graph incorporates balance between being extraordinary and not overly complex (E. Hawkins et al., 2019).

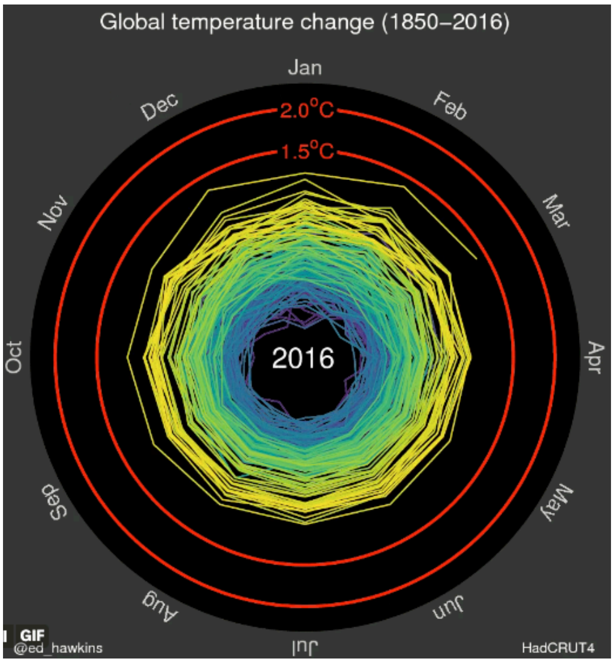


Figure 7: Screenshot of animated Climate Spiral by Hawkins E. [@ed_hawkins] (2016)

Helix Icons

Another spatio-temporal visualization is a map combined with helix icons (see Figure 8). These 3D map icons follow a depiction in a three-dimensional space in which the x- and y-axes demonstrate the location, and the z-axis shows the temporal development of the variable. A helix ribbon extends vertically upwards from the planar depiction of the map, with each part of the helix ribbon representing a point of time and displaying values with the use of different colors. The icons are designed to show multivariate data in an interactive environment. A disadvantage of this visualization is that a part of the information is always hidden due to the 3D-effect. Hence, helix icons are designed for interactive visualizations which allow users to navigate through the map with ease (Aigner et al., 2011; Tominski et al., 2005). This visualization is mentioned to illustrate that also interactive visualizations were considered for this thesis even though the final product aims for a static visualization. Given the objective of modifying visualizations, interactive visualizations were examined with the understanding that they needed adaptation for static uses.

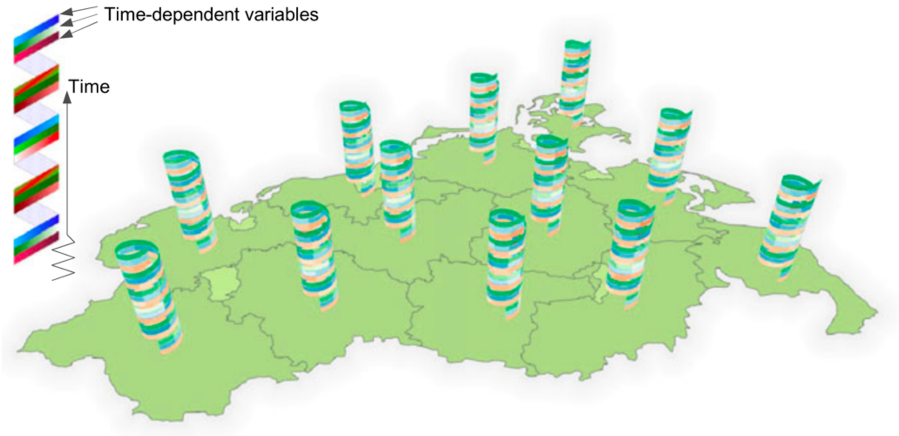


Figure 8: Helix Icons on a Choropleth Map from Aigner et al. (2011)

Warming Stripes

Little empirical evidence exists regarding the effectiveness of Warming Stripes (see Figure 9) since no studies have been published providing proof to their advantages and disadvantages. However, the visualization has become widespread on public and social media and is being credited with opening conversations about increasing global temperatures and, more broadly, climate change. Warming Stripes display temperature data through vertical bars that are color-coded. Blue shades denote temperatures below the average temperature, while red shades represent temperatures above the global average. The visualization shows a time series from 1850 to 2022, each year representing one stripe, and no further textual or numerical information. O'Connor (2023) states that the design of the visualization and explicitly the choice of color conveys the message to the audience in a simple and understandable way. Furthermore, the author points out that the audience can comprehend the message transmitted by the visualization, regardless of their age, cultural background, or knowledge. Warming Stripes' simplicity is noteworthy as it permits the depiction of lengthy time series with multiple data points due to the thinness of the stripes.

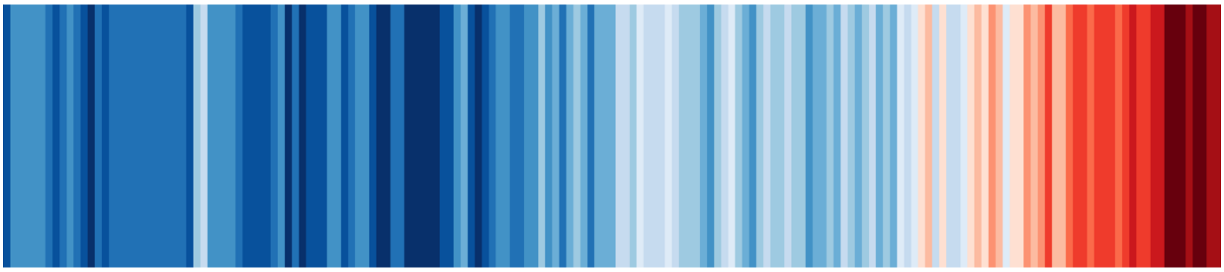


Figure 9: Warming Stripes in Original Form without Title, Labels nor Legend from E. Hawkins (n.d.)

Warming Stripes are a common way to visualize the increase of global temperature for a broader audience. Sengupta (2022) from The New York Times even categorizes the visualization as one of the most popular ways to depict increasing temperature. Also, Greta Thunberg, environmental activist and the leading figure of the global climate strike movement Fridays for Future, chose Warming Stripes as the cover of her book (Fridays For Future, 2023; Poole, 2022).

One of the research gaps concerning Warming Stripes is the lack of empirical evidence on the awareness and effectiveness of this visualization. This thesis aims to address this question, with further research objectives detailed in the following subchapter.

In summary this visualization process is a product of a combination of sources that focus on (spatio)temporal visualizations but it is important to note that as Nocke et al. (2004) correctly points out, the visualization of scientific research is not straight-forward due to the variety of parameters that can differ within the process. Due to the diversity of variables, tasks, and visualization types it is important to get an overview of the data used for each visualization and then the further process can be adjusted to the respective needs. Therefore, this visualization process will be applied and adjusted to the selected variables LFMC, FWI and TEMP within Chapter 2.2.3.

1.4 Research objective and research questions

This research is divided into three main objectives with corresponding research questions:

Within the literature review it became clear that the preparation of a visualization should follow specific steps to create appropriate visualizations according to the data that is used. For the specific use case of the selected multivariate data of life fuel moisture content, fire weather index and temperature, key messages and tasks need to be defined in order to evaluate the suitability of novel and historic visualizations for this data.

Research Objective 1	Creation of evaluation framework on novel and historic temporal and spatio-temporal visualizations for time series data visualizations of LFMC, FWI and TEMP
Research Question 1	What are the key messages that need to be incorporated within a spatio-temporal visualization for LFMC, FWI and TEMP when showing the overall development of these three variables from 1988 - 2016?

One of the research gaps concerning Warming Stripes is the lack of empirical evidence on the awareness and effectiveness of this visualization. No study could be found in which a user group determined their effectiveness, nor any evidence on the awareness can be given.

Research Objective 2	Evaluation of novel visualization Warming Stripes to test their awareness and effectiveness
Research Question 2	Can Warming Stripes be considered as a successful data visualization in terms of awareness and effectiveness?

The last objective this research will address is the combination of novel visualizations with maps for the specific data of LFMC, FWI and TEMP. This combination of these fire risk variables has not yet been visualized with novel visualization types and therefore, this research tries to fill this gap.

Research Objective 3	Combining two novel visualizations with maps for a static representation of time series of LFMC, FWI and TEMP
Research Question 3.1	Can novel visualizations of Warming Stripes and Climate Spiral be modified to a spatio-temporal visualization of multivariate data of LFMC, FWI, and TEMP in combination with a map?
Research Question 3.2	If the combination of Warming Stripes/Climate Spiral with data of LFMC, FWI, and TEMP is not possible, which other spatio-temporal visualizations can be created to visualize these time series?

1.5 Thesis structure

The outline of this thesis starts with an overview of the used datasets and their unique properties respectively. The different methods used within this work, including an adapted visualization process, a survey to evaluate a chosen novel visualization, an exploratory data analysis and the creation of alternative visualizations, will be described in the same chapter. The main part of this research can be found in chapter 3 in which the results of each method will be described as well as discussed. Furthermore, this chapter will also include a discussion on combined findings and evaluations on the the created visualizations. The last chapter involves the conclusion of this thesis and highlights the results.

2 Data and Methodology

Jones et al. (2022) identifies three factors that influence the occurrence and spread of wildfires: availability of fuel, climatic conditions, and sources of ignition. In addition to fuel availability, vegetation moisture also plays a crucial role in wildfire estimation, as components with lower water content are more prone to ignite. (Chuvieco et al., 2010; Forkel et al., 2023). Based on this, environmental fire risk variables were chosen for their visualization and to present their spatio-temporal development. In the next pages the selected variables LFMC, FWI and TEMP will be explained. Additionally, the explanation of LFMC will also address the novel data acquisition method utilized.

2.1 Data properties

The datasets provided by JProf. Matthias Forkel and Luisa Schmidt in form of netcdf files were pre-processed in terms of resampling with the nearest neighbor method to achieve the same spatial resolution and the calculation of the monthly average for all datasets. These operations resulted in a coherent set of datasets with the same temporal as well as spatial resolution for this research. The table below (see Table 2) provides an overview of the datasets and their respective properties after the supervisors' pre-processing.

Dataset	Variable	Spatial resolution and coverage	Temporal resolution and coverage	Reference
LFMC	LFMC (%)	0.25 x 0.25° - Global	Monthly Aug 1987 – Jul 2017	Forkel et al., 2023
FWI	MERRA-2 FWI	0.25 x 0.25° Global	Monthly Jan 1987 – Dec 2017	Field et al., 2015
TEMP	2 meter above ground temperature (K)	0.25 x 0.25° Global	Monthly Aug 1987 – Dec 2017	Hersbach et al., 2023

Table 2: Properties of used datasets

Due to different temporal coverages, the temporal extent was limited to January 1988 to December 2016 to solely encompass complete years. Furthermore, the entire TEMP dataset was converted from Kelvin to Celsius during all procedures.

2.1.1 Fire weather index

Fire weather index is a metric used to estimate the potential for fire ignition and spread. The index combines meteorological data such as wind speed, temperature, precipitation, and humidity as well as data on fuel moisture (Ellis et al., 2022; Natural Resources Canada, n.d.). Jones et al. (2022) examined the impact on anthropogenic climate change induced factors on FWI and found evidence on the significant increases in fire weather in most regions of the world and in recent decades due to climate change. Temperature, humidity, and wind patterns are significantly affected by anthropogenic climate change and are also important factors influencing fire weather.

2.1.2 Two meter above ground temperature

The variable 2 meter above ground temperature obtained from the Copernicus datastore comprises temperature measurements recorded at 2 meters across all global surfaces, including land, sea, and inland waters. The reanalysis data was derived through data assimilation, a method that combines data from various sources, including in-situ measurements, such as weather stations or aircraft, and remote sensing data from satellites (European Centre for Medium-Range Weather Forecasts 2020, 2023). Compared to the two other variables, temperature data offers additional values above marine water bodies which were excluded for the purpose of maintaining consistent spatial coverage. The dataset has a unique characteristic, in the sense that it contains no data gaps, setting it apart from the other datasets. This variable was chosen because it is an important factor for wildfire risks but also due to one of the reasons why it was selected for the Climate Spiral: the familiarity of the audience with this variable.

2.1.3 Life fuel moisture content

Fuel moisture represents the water content of the vegetation canopy and can be divided into life fuel moisture which is living vegetation and dead fuel moisture which considers vegetation that is already dried out with moisture contents below 30% (National Centers for Environmental Information, n.d.; Yebra et al., 2019). Life fuel moisture content (LFMC) is a vegetation property that provides information on the moisture content of the vegetation canopy and is frequently utilized in wildfire ignition and spread estimation (Forkel et al., 2023). To obtain LFMC data, vegetation components such as green grasses and leaves are traditionally weighed, dried, and reweighed. The ratio of water mass to dry biomass results as LFMC (Yebra et al., 2013):

$$LFMC = \frac{m_{fresh} - m_{dry}}{m_{dry}}$$

Higher LFMC values indicate higher water content, resulting in decreased ignition risk, longer ignition times and reduced wildfire spread. Therefore, LFMC serves as an essential predictor for wildfire prediction and modeling (Yebra et al., 2013).

Data acquisition of LFMC through field sampling is widely employed but it has some disadvantages. First, the retrieval of LFMC by field samples is labor-intensive and the process needs to be repeated over time to capture the temporal variation of LFMC. Another disadvantage is that on-site LFMC data retrieval is not coordinated by an organization or research group but rather by several researchers and agencies (Yebra et al., 2019). In 2019 Yebra et al., published an openly accessible dataset called *Globe-LFMC*. This dataset combines information from 1,383 sampling sites across 11 countries to validate and calibrate remote sensing algorithms for LFMC research via satellite observation. However, it should be noted that there is an uneven spatial distribution of the observation sites within the dataset. Therefore, extensive research on estimating LFMC from satellite observations, with the goal of achieving a globally accessible spatial retrieval of LFMC.

Satellite observations can be used to calculate LFMC through either visible and infrared or microwave domains of the electromagnetic spectrum. However, these variants do not directly retrieve LFMC but rather calculate it indirectly as a proxy variable. Retrieval of LFMC in the spectral and infrared region is heavily influenced by weather conditions, as wavelengths in that spectrum cannot penetrate through clouds or when the elevation angles of the sun are too low. Therefore, this method cannot utilize satellite imagery on days with high cloud coverage or low sun illumination. On the other hand, LFMC retrieved through microwaves is not affected by weather conditions as microwaves can penetrate cloud cover and are not influenced by sun angles (Forkel et al., 2023; Yebra et al., 2013).

The dataset utilized within this thesis was developed by the authors of Forkel et al. (2023) and its methodology of retrieving LFMC from microwaves follows the same workflow outlined in their publication earlier that year. The only disparity between the dataset described in the paper and the one employed in this study is the substitution of GLOBMAP data for the variable leaf area index instead of using MODIS data.

2.2 Visualization Process for Visualization of LFMC, FWI and TEMP

The workflow based on the literature review which was suggested in Chapter 1.3.1 is now inspected with consideration of the variables of LFMC, FWI and TEMP. Therefore, the target audience, data classification, task specification and visualization types will be defined specifically for the chosen variables. In the end of this subchapter also an evaluation framework will be presented which gives an overview of historic and novel visualization types which are suitable for the selected remote sensing time series of LFMC, FWI and TEMP. The framework gives insight into the intended visualizations and their characteristics.

2.2.1 Target audience

As mentioned by Knaflic (2015) in Section 1.3.1 the target audience should on the one hand be defined by the background and technical knowledge of the people the visualization is aimed at. On the other hand, also the message needs to be defined clearly because an understanding of why the target audience should be informed on this information and what understanding is expected from the audience after seeing the visualization is needed.

The main message on the targeted visualizations is to show the development of LFMC, FWI and TEMP within the time span of 1988 to 2016. People who might not have heard about these variables should be able to understand an overall development of these variables throughout the years. In particular, the visualization should address an audience that does not have a professional or educational experience on LFMC, FWI and TEMP but they should be interested in learning about topics they are unfamiliar with. Even though the audience should not be necessarily experts on fire risk factors, they should be able to comprehend the information of the visualization. Therefore, the age of the target group will be set to a minimum of 12 years. No actions are expected from the audience after seeing the visualizations but a rough understanding on what LFMC, FWI and TEMP is and their development throughout the years is aimed for. The technical knowledge should at least include that multivariate data sets do not need to be explained, nor general temporal visualizations in the form of linear time series or cycle graphs. The audience should be able to read a spatio-temporal visualization without any help other than the legend.

2.2.2 Data classification

The classification of data will give an overall insight into what exactly wants to be depicted within the visualization and helps forming the visualization's main messages. We will follow the introduced data classification models of Andrienko & Andrienko (2006), Peuquet (1994) and Aigner et al. (2011) from Section 1.3.1 to describe the variables of LFMC, FWI and TEMP further in a more explicit manner.

The classification model by Andrienko & Andrienko (2006) provides a good basis for an initial data inspection. The method distinguishes in referential components (referrers) and characteristic components (attributes). In the case of the variables of LFMC, FWI and TEMP which all cover the same temporal and spatial extent, two referrers and three attributes can be identified. The referrers are represented by the time and the location of the data. The spatial referrer is discretized because it is raster data and therefore divided into uniform areas. The temporal referrer is discretized as well but on the level of monthly data. The three attributes of the selected data are the measurements of LFMC, FWI and TEMP, which are defined by monthly aggregation due to calculation of the mean. This simple classification method shows us that location and time are independent components and LFMC, FWI and TEMP are the dependent components. The attributes are dependent on the chosen spatial extent. Andrienko & Andrienko (2006) explicitly point out that for climatic data, cyclic patterns should be considered as well.

By applying the data classification triad framework by Peuquet (1994) to our data the main components location, time and objects can be identified. Location describes the spatial position within the global datasets, time incorporates the monthly values from 1988 to 2016 and the attributes are LFMC, FWI and TEMP measurements.

According to Aigner et al. (2011) the data can be categorized as quantitative data because the measurements are based on a metric scale which allows numeric comparisons. Furthermore, the data is obviously spatial data because it incorporates geographic information. Even though the data only has monthly measurements and does not have data at any other point of time, the kind of data can still be categorized as states and not events, because of the continuity. Furthermore, the data we are handling in this thesis is multivariate data because it consists of three variables: LFMC, FWI and TEMP.

All these categorization methods give a simple insight into what kind of data needs to be visualized. This is part of the first stage of visualizing the variables to know exactly what is handled and what should be focused on.

2.2.3 Task specification

The next stage of the visualization process is understanding why exactly the visualization should be created because this also defines the content of the final visualization; when knowing what message one wants to transmit to the audience, also the content is shaped. Tasks should be considered which help determining the idea of the content. We will begin by defining main components required for the intended visualizations, utilizing temporal task descriptions from MacEachren (1995). Two visualizations are aimed for during this research that cover the whole time series from 1988 to 2016.

The respective key messages of the visualizations and the classification of these key messages as tasks based on MacEachren (1995) are defined in Table 3.

	Key messages	Task definition based on MacEachren (1995)
Visualization 1	Regional variability Comparison of all three variables Overall trend of all three variables	Existence of data element Temporal location Time interval Rate of change
Visualization 2	Periodic patterns of monthly data Depiction of missing data Nine biogeographical regions	Existence of data element Temporal location Time interval Rate of change

Table 3: Key Messages and Task Definitions for Visualization 1 and Visualization 2

Furthermore, the tasks can be considered as synoptic tasks and even further in descriptive tasks based on the task taxonomy of Andrienko & Andrienko (2006). Since the key messages concern the development of LFMC, FWI and TEMP throughout the whole time period, the tasks which are executed within data analysis are concerning the whole dataset or set of values when only one variable is considered; this type of tasks is considered as synoptic tasks. Additionally, all visualization's key messages can be categorized within descriptive tasks because the visualization aims for highlighting patterns and significant behaviors in space and time of LFMC, FWI and TEMP.

2.2.4 Types of visualizations

The final stage involves selecting appropriate visualizations for specific tasks and variables. This research aims to adapt novel visualizations to match the needs of the selected variables. Therefore, we will investigate these visualizations and determine whether they already meet the criteria in terms of key messages and showing a spatial representation in the form of a map.

The original Warming Stripes need several modifications to suit the key messages of the intended visualizations (see Table 4). First, a map needs to be incorporated. The map will be a choropleth map with global extent and with biomes (see Figure 10) as dividing geographical areas due to the regional variety of LFMC and the variety of this variable across different vegetation types. This also leads to the fulfilment of the second key message of this visualization which is regional variability. The last point which needs to be incorporated is the comparison of all three variables. This will be achieved by creating three different stripes: Warming Stripes for TEMP, Vegetation Stripes for LFMC and Fire Stripes for FWI. The overall trend is one of the strong points of the original Warming Stripes and therefore is already included.

	Warming Stripes	Modification
Spatial representation	No	Incorporation of a map
Regional variability	No	Division of map by biomes
Comparison of three variables	No	Creation of three distinctive stripes
Overall Trend	Yes	-

Table 4: Fulfilments on Warming Stripes with key messages of Visualization 1 and needed Modifications

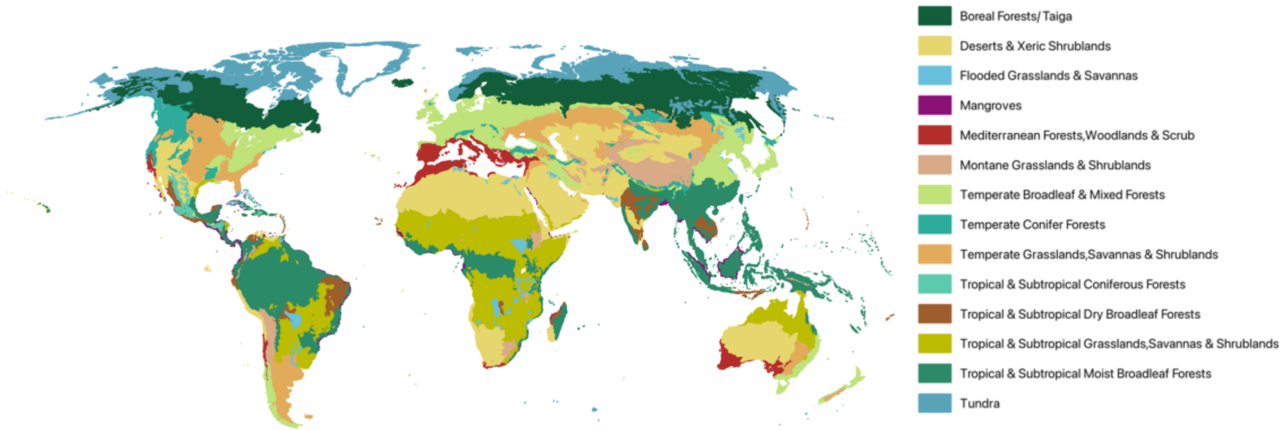


Figure 10: Global Map of Biomes

For the second visualization the Climate Spiral will be inspected for their suitability with the key messages of Visualization 2 (see Table 5). Climate Spirals already meet two out of the four criteria for Visualization 2. Therefore, only a map needs to be incorporated which will be a map of Europe showing biogeographical regions (see Figure 11). Also, the creation of three spirals per geographical unit will be necessary to show a comparison of all three variables.

	Climate Spiral	Modification
Spatial representation	No	Incorporation of a map
Periodic patterns of monthly data	Yes	-
Comparison of three variables	No	Creation of three distinctive spirals
Overall Trend	Yes	-

Table 5: Fulfilments on Climate Spiral with key messages of Visualization 2 and needed Modifications

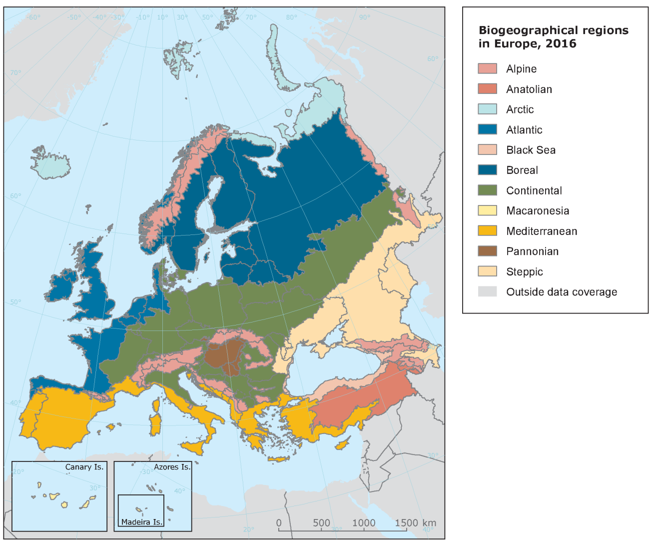


Figure 11: Map of Biogeographical regions in Europe from European Environment Agency (2023)

2.3 Survey to evaluate Warming Stripes

The primary aim of the survey was to empirically assess the awareness of Warming Stripes as well as to investigate potential variations in the familiarity across different demographic categories such as age groups and levels of education. The reputation of Warming Stripes is that the visualization functions well for starting the conversation about increasing global temperatures and subsequently on climate change. The visualization does not show any labels, legend nor title. Furthermore, since no research has been made so far to proof the effectiveness in consideration of understandability and intuitiveness of this novel visualization, one of the objectives of the survey was to test this. Therefore, participants were asked for a short explanation on what Warming Stripes (might) depict without being given any additional information. For the survey the Warming Stripes as they were published initially - without labels nor titles - were derived.

The survey was conducted at the "Dresdner Lange Nacht der Wissenschaften" – the long night of sciences of Dresden on the 30. June 2023 at the main auditorium of the Technical University of Dresden. The long night of sciences is a concept adapted by the long night of museums, which like the latter opens the gates for people interested in sciences. The target audience of long nights of sciences include people that are interested in broadening their knowledge within science and learn about concepts and methods which they are not familiar with in an interactive and experimenting way. The program is not only designed for German-speaking adults but also include program elements for children and international visitors (Landeshauptstadt Dresden, 2023).

To make the survey inclusive for most of the visitors of the long night of sciences, the questions and possible answers were created in German and English. Visitors of the long night of sciences were directly approached and asked to fill out the questionnaire by themselves and without any additional information on the questions. Therefore, the participants were not influenced and could fill out the form anonymously. Additionally, a QR-Code including the same questionnaire was created, printed, and placed opposite of the booth of the Environmental Remote Sensing group of TU Dresden.

A summary of the outline of the survey as well as the type of survey questions are illustrated in Figure 12. The complete survey is attached in the Appendix. Part 1 of the questionnaire includes questions on the age and education level of the participants and whether the visualization had been seen before. Depending on the answer to the latter question, participants continued with Part 2a if they have seen the visualization before or with Part 2b if they have not. For partakers which have seen the visualization before, questions followed on whether they have seen the visualization once or several times and in which context they did. Furthermore, they were asked to describe the visualization in an open-ended answer. The last two questions of Part 2a contain information on the visualization with other variables. Part 2b solely includes one question, in which participants are asked to provide a brief description on what they think this visualization might represent. In that case also guiding questions were provided to facilitate the interpretation (What is the significance of the stripes and the colors? Which data is presented here?). This question was asked to test for the intuitiveness of the visualization; to find out if the visualization can be interpreted without any additional information than the stripes themselves.

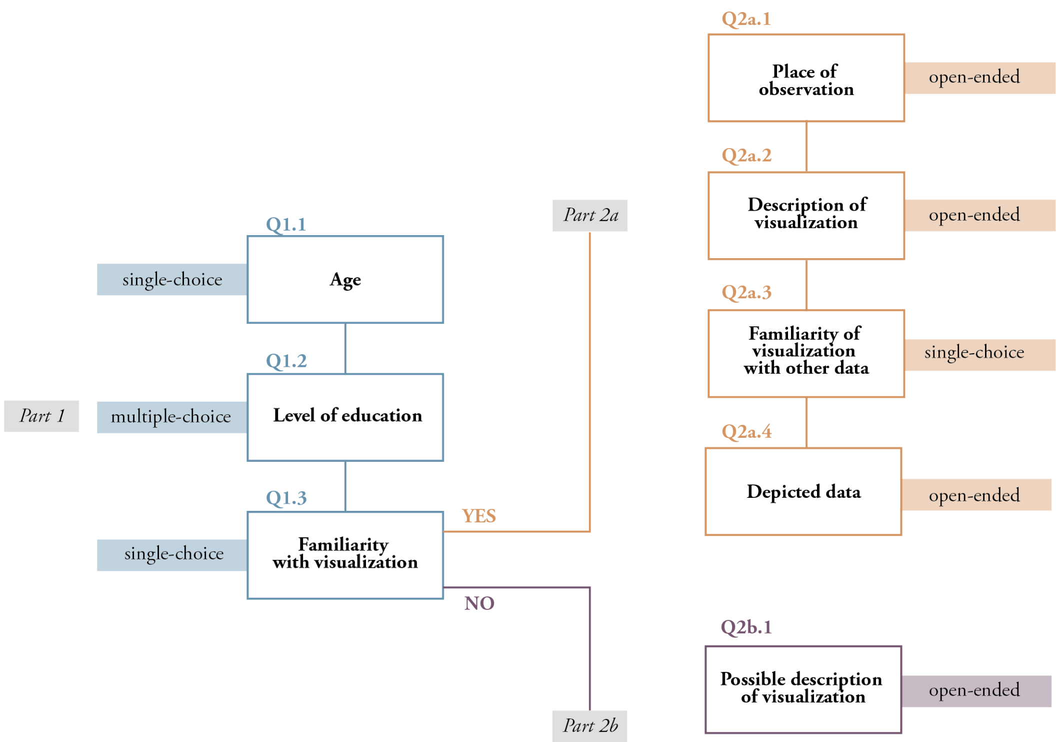


Figure 12: Outline of survey on evaluation of Warming Stripes

The manually filled-out answers, together with the answers from the online survey were transferred to Excel and translated to English. Due to open-ended questions, the transfer of the hand-written answers involved more difficulty than expected because some answers lacked legibility. The answers to single-, and multiple-choice questions were coded to allow easier processing. Open-ended questions were categorized into common categories by semantic similarity. An extract of the table consisting of all answers can be seen in **Fehler! Verweisquelle konnte nicht gefunden werden.**. The results of the survey can be found in Chapter 3.1.

Q1.1 (single-ch) age 1 = under 15 2 = 16-30 3 = 31-45 4 = 46-60 5 = over 60	**Q1.2 (multiple-ch) high. Level of education** 1= no compl. educ. 2= primary school 3=Realschule 4=secondary school 5=apprenticeship 6=bachelor 7=master 8=doctorate 9=other	**Q1.3 (single-ch) familiarity with visualization** 1=yes, several 2=yes, once 3=no	**Q2a.1 (open-ended) place of observation**	**Q2a.2 (open-ended) familiarity = yes; brief description**	**Q2b.1 (open-ended) familiarity = no; brief description**	**Q2a.3 familiarity visualization with other data** 1= yes 2 = no	**Q2a.4 depicted data**
2	4	2	Phone background	no idea			
2	4	1	climate data, student assistant job, news, studies	**Temperature data** of the last 100 years of a region		2	maybe prepiciation
5	7	3			**LQTB**, dark colors are more concise		
2	6	1	Street Art, poster, protests, literature, TV	Increase of **global average temperature** of the last 100 years		2	
2	4	3			**Temperature, Gerhard Richter**		
4	7	1	Internet, TV, banner, several occasions	**Warming Stripes** show the development of **average temperature** of the earth		2	
3	3	1	Paint box, thermal image verification of ENEV	**Cold and warm zones**		2	
1	2	3			Scan-**Code, Curtain**		
4	5	1	Laboratory	how **strong a chemcial is, Heat/Cold radiation**		2	

Figure 13: Transferred survey results translated to English

2.4 Adaptation of Warming Stripes with multivariate data and global spatial representation

The data for the 14 biomes were retrieved from Dempsey, 2021 in form of raster data which show the same resolution as the other variables with a global extent. Biomes are the largest geographic biotic unit and are defined by similar environmental conditions such as vegetation and similar life forms like animals (Augustyn, 2023). LFMC varies with vegetation types, hence a classification that includes a distinction between vegetation types was selected.

Due to the shorter depicted time period of only 29 years, our adaptation of Warming Stripes will resemble rather blocks than stripes. Therefore, we will from now on refer Vegetation (LFMC), Fire (FWI) and Warming Blocks (TEMP) whenever the adaptation of Warming Stripes to our data is mentioned.

The workflow for the creation of the distinctive blocks is described in the following section. A graphical overview of the main steps can be seen in Figure 14. To begin, the variables FWI and TEMP were respectively spatially intersected with LFMC. In this step only cells with values that exist within LFMC variable were considered. This was done to account for imbalances in missing data due to missing values in LFMC but not in TEMP. This step was the only one that differed between the variables, the following steps were consistent across all the data.

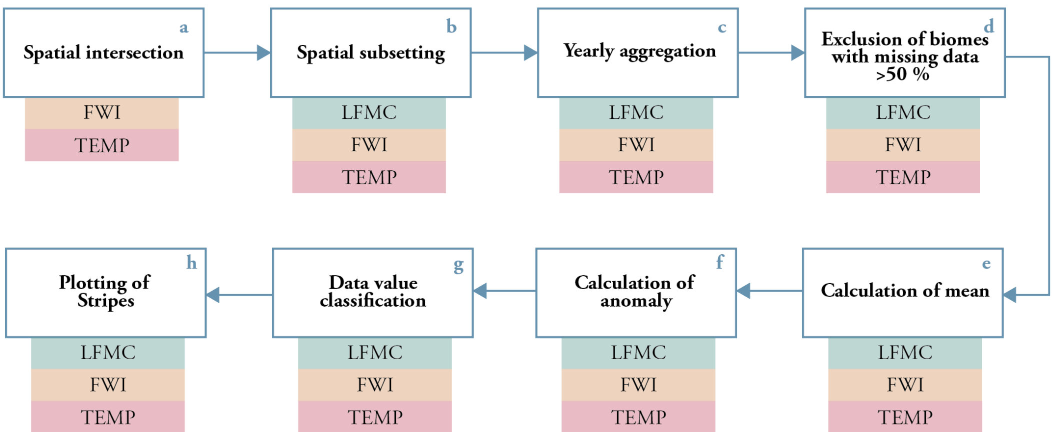


Figure 14: Workflow of creation of Vegetation Blocks, Fire Blocks and Warming Blocks

Subsequently, spatial subsetting of each variable with all 14 biomes was executed. This resulted in the creation of 14 distinctive raster files for each variable, only including the spatial extent of the respective biome. Following this, yearly aggregation was performed by calculating the yearly average using monthly values. Missing values were excluded in this operation and this step was repeated for all variables individually.

To exclude years with high rates of missing data, each layer of the LFMC biome variable was inspected and the percentage of missing data was calculated. There were no significant differences in terms of missing data across years within each biome, but LFMC data availability varied significantly by biome. Therefore, the approach was modified from initially excluding individual layers with high percentages of missing data to excluding biomes with high percentages of missing data. To achieve this, the mean percentage of missing data per biome for all years was computed. Table 6 indicates that most biomes had missing data rates ranging from 5 to 31% but some had relatively high missing data values. Therefore, a threshold of at least 60% of data availability (maximum of 40% of missing data) was established to exclude biomes with high levels of missing data. Montane grasslands & shrublands (47% of missing data), Tundra (54%), Desert & xeric shrublands (74%) and Mangroves (72%) were excluded from the subsequent analysis. As a result, only 10 biomes were analyzed in the following steps.

Biome	**Mean percentage of missing data in %**
Tropical & Subtropical Moist Broadleaf Forests	8
Tropical & Subtropical Dry Broadleaf Forests	8
Tropical & Subtropical Coniferous Forests	5
Temperate Broadleaf & Mixed Forests	13
Temperate Conifer Forests	11
Boreal Forests/Taiga	12
Tropical & Subtropical Grasslands, Savannas & Shrublands	21
Temperate Grasslands, Savannas & Shrublands	8
Flooded Grasslands & Savannas	21
Montane grasslands & Shrublands	**47**
Tundra	**54**
Mediterranean Forests, Woodlands & Scrub	31
Desert & Xeric Shrublands	**74**
Mangroves	**72**

Table 6: Mean percentage of missing data per biome

At this stage in the creation of the visualization, the variables per biome were still in the form of spatrasters, with each 0.25° x 0.25° cell containing LFMC, FWI and TEMP values respectively. To create blocks which follow the model of Warming Stripes, only one yearly value was necessary. Therefore, the mean for each year, biome and variable was calculated. Missing values were again excluded for the calculation of the mean. After this step, the data type of the individual variables changed from rasters to tables.

For each variable, a table was available that contained the mean annual value of LFMC, FWI and TEMP for all years from 1988 to 2016. To calculate the anomaly, a reference value was calculated for each biome and variable by calculating the mean value for the whole period. Next, the annual mean of LFMC, FWI, and TEMP was subtracted from the reference value.

For the classification of the data range, various approaches were tested, including using equal steps, different numbers of percentiles, and quantiles were tried. However, the same approach as with the original Warming Stripes was eventually selected with steps deviating from the baseline 0. For TEMP equal steps were chosen because the data distribution allowed it. Anyway, for each variable the histogram of all nine biomes were inspected together. Based on the data distribution, the classification for the legend and the Blocks were chosen.

The last step in the blocks creation process involved plotting the blocks. An online tutorial by Dumble (2021) was followed for this step. The colors for TEMP were set to a color scale from blue to red. A brown to green color scale was chosen for LFMC to represent dry (brown) and healthy, moist (green) vegetation canopy. Additionally, a color ramp from violet to orange was chosen. The association with fire led to the choice of orange. For low FWI anomaly values, violet was chosen as it sits opposite of orange on a color wheel.

The placement of all elements along with a choropleth map of the biomes and legends (see Figure 15) for each variable was accomplished in Adobe Illustrator. An example of Warming, Fire, and Vegetation Blocks of Mediterranean Forests, Woodlands and Scrubs can be seen in Figure 16.

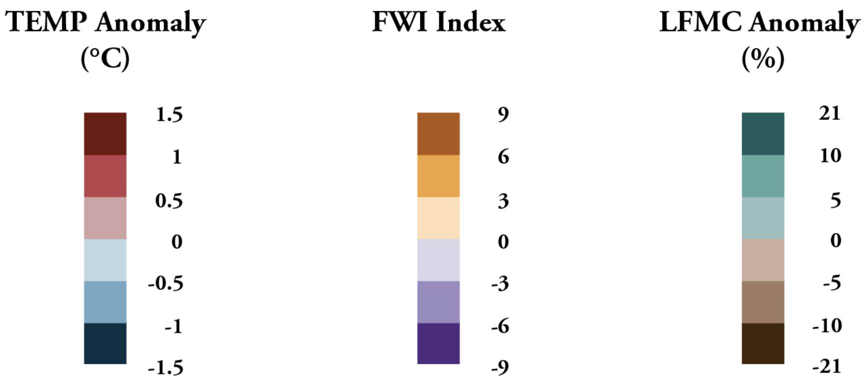


Figure 15: Legends of Warming, Fire and Vegetation Blocks based on data classifications

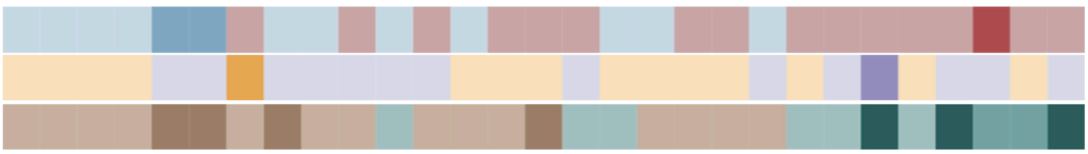


Figure 16: Warming, Fire and Vegetation Blocks of Mediterranean Forests, Woodlands & Scrubs; Blocks are Adaptations of Warming Stipes, each Block represents one Year and its respective Anomaly Value

The resulting visualization and a detailed discussion can be seen in Chapter 3.2. At this point it should be noted that the visualization was found to be unsuitable for the displaying of FWI data at this scale. Therefore, it was decided to take a thorough look at the data using the method of exploratory data analysis and to investigate other suitable (spatio-)temporal visualizations for the combination of these data in order to fulfil the goals of the visualization as defined by the key messages in Section 2.2.3.

2.5 Exploratory Data Analysis of Missing Data

For further proceedings exploratory data analysis was chosen. EDA is a method that was first introduced by Tukey (1977), who describes EDA as "graphical detective work". EDA is a method setting focus on the exploration of data and their properties with a combination of statistical summaries and graphical depictions of data. A key difference between EDA and the classical statistical way of data analysis is the explorative way of looking at data; instead of already having a hypothesis in mind, EDA supports a more descriptive way of unfolding data properties and finding a hypothesis that can be tested later on (Andrienko & Andrienko, 2006; Theus, 2005). There are no specific tools or operations that are involved in EDA since it highly depends on the nature of the individual data (Andrienko & Andrienko, 2006). It is important to note that this method shift in analyzing data was achieved with the improvement of computational power, which made the process of visualizing large datasets even possible in the first place (Theus, 2005). For a further inspection of the datasets this method was chosen, even though tasks and key messages were already defined.

The opinions on defining tasks before EDA is performed are diverse according to Andrienko & Andrienko (2006). While some researchers state that task definitions are not necessary and EDA should be carried out like described by Tukey (1977) without defining questions beforehand, others which also include Andrienko & Andrienko (2006) are convinced that some tasks are always defined beforehand. Looking at a dataset without any bias is according to the authors almost impossible because there is always some direction the analyst is aiming to explore. They claim that from the beginning an explorer has a general synoptic task in their mind, which includes describing the most important patterns of the dataset and after identifying these, further tasks and key messages are defined to break down more specific trends. This approach was also taken within this research. Broad synoptic tasks were defined in Chapter 2.2.3.

2.5.1 Inspection of Seasonals and Total Missing Data

Due to the unsuccessful visualization of LFMC, FWI and TEMP with the method of Warming Stripes, a smaller extend was chosen to proceed. Therefore, the same spatial extent as for Visualization 2 was chosen which is the biogeographical regions of Europe. Moreover, EDA was conducted to this spatial extent. The data for the biogeographical regions were derived from the European Environment Agency (2023) in the form of shapefiles. For the following steps the data on the biogeographical regions was prepared in QGIS before proceeding with EDA with R. The shapefile was filtered by removing the categories “not Europe” and “Macaronesia” from the attribute table. Macaronesia was removed because the islands were too small for the resolution of the raster and therefore had no data values. Afterwards the area was disaggregated, resulting in a layer only containing the outer borders of the area without any regional subunits. Before saving the file and proceeding in R, the borders were also simplified because the borders were overly detailed.

The following steps which include the inspection of the variables for seasonal patterns and missing data were all conducted in R. The first step was to clip the spatial extent of the LFMC, FWI and TEMP data to the borders of the European biogeographical regions. Then, a subset for each season (see Table 7) for each variable was created. The result of this operation were separate subsets for each variable (LFMC, FWI, and TEMP) including respective seasons. For each cell of each distinctive subset, the number of missing values was counted for the duration of the whole time period. Each inspected subset is composed of 87 layers due to 29 years within the time series and each season containing three months. The maximum value of missing values per cell can thus be 87.

Season	Months
Spring	March, April and May
Summer	June, July and August

Autumn	September, October and November
Winter	December, January and February

Table 7: Division of Months per Season for Exploratory Data Analysis

Figure 17 shows a compilation of maps which displaying the percentages of missing values per cell for FWI and LFMC across Europe. TEMP did not include any missing data and, thus, was not included in further analysis on missing data. Due to this graphical exploration of the missing values per seasons, patterns were discovered. Significant variations in data availability both seasonally and geographically were revealed. While inspecting temporal differences, FWI displays low rates of missing data during summer across Europe, but only in Southern European and Western European countries throughout the other seasons. Furthermore, in the spring and autumn seasons data availability decreases at higher latitudes.

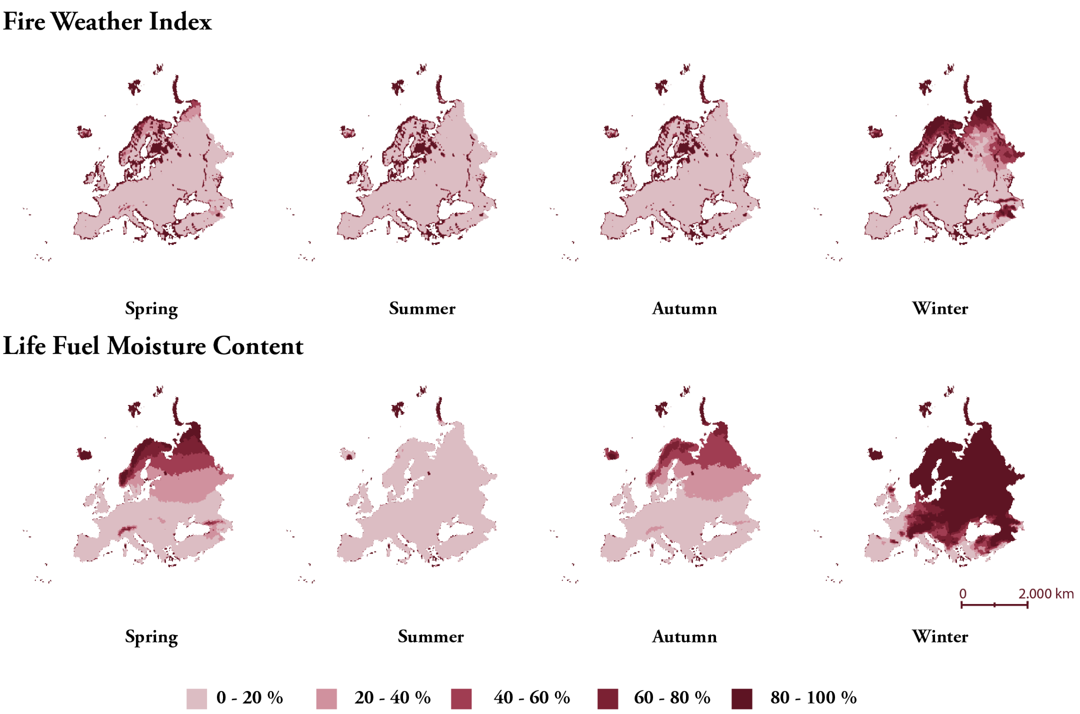


Figure 17: Percentage of Missing Values per Cell by Seasons for FWI and LFMC

Another spatial pattern, which was not visible before visualizing missing values, was the difference in the extent of the area between the datasets. While TEMP and FWI display the same extent, LFMC reveals a prominent line along maritime boundaries. This is because LFMC does not consider cells that contain a certain percentage of water, which is included in the other two variables (Forkel et al., 2023). Also, for LFMC higher latitudes show missing data in the winter.

This made it clear that it is advisable to investigate the values per season rather than the mean values of the entire time period. Averaging the values as illustrated in Chapter 2.4, would distort the mean value since there are more values available during the summer season, leading to a false message.

Consequently, to avoid distorting the averaged values later on, layers with a high number of missing data values were excluded from further analysis.

2.5.2 Inspection of Monthly Aggregated Missing Values

To exclude monthly layers that have excessive missing values, the corresponding datasets were inspected graphically and statistically. First, the LFMC time series was examined, and then the FWI time series. It is important to note that a common threshold of the exclusion was avoided, because the data structure was considered individually.

Basic statistical operations were performed to determine key measures of every variable, including the minimum, maximum, mean, and median of missing values per layer. The results of these operations can be seen in Table 8. Furthermore, these results were also visualized by creating a boxplot. Figure 18 depicts that the majority of the NA rates of the layers of LFMC lie in the range up to 49% while the layers with a higher percentage of NAs are considered as outliers. Therefore, 48.56% was set as a threshold for the exclusion of layers. All layers with a higher value of missing data were excluded. For FWI the data range of missing values had a wider range from 9% to 9% and as can be seen in Figure 19 the missing data rate was more evenly distributed across the layers. Therefore, there are no outliers visible, and another approach had to be implemented to exclude layers with high missing values. The decision was made to exclude all layers that were not within the third quartile. So, 25% of the data was excluded and these layers showed a minimum of 68.9% of missing values. The selected layers were not deleted but all their values were set to NA. Therefore, the temporal component in the form of the referrer is still available but the attributes are empty at that temporal position. For the variable LFMC, 29 out of 348 layers were set to NA. Similarly, for the variable FWI, 87 out of 349 layers were set to NA.

	Minimum	Maximum	Mean	Median
LFMC	26%	67%	33%	28%
FWI	9%	91%	38%	28%

Table 8: Statistical Operations (Minimum, Maximum, Mean, and Median) for Missing Data per Month for LFMC and FWI

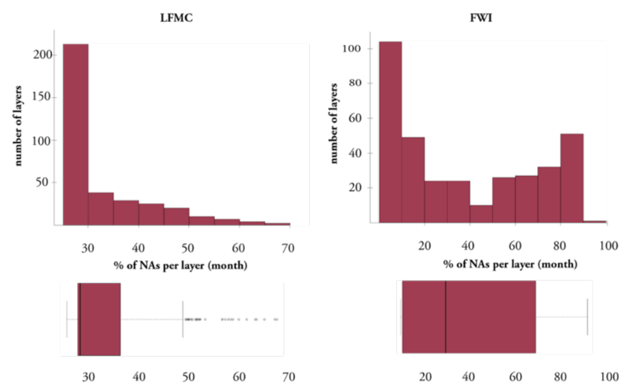


Figure 18: Histogram and Boxplot of Missing Data per Dataset

2.6 Alternative visualizations

In this chapter, the method of work for the visualizations will be presented that were successfully capable to incorporate all the key messages.

2.6.1 Visualization 1: Adaptation of Warming Stripes with Small Multiples

The key messages from Table 3 for Visualization 1 were re-evaluated together with the selected visualizations from Section 1.3.1.4 in order to develop a new visualization method as an adaptation of Warming Stripes. Additionally, the recent insights derived from computing EDA were considered. As a result, only summer seasons were considered for data analysis since they provide the most data availability, rather than taking values for the entire year. Calculating the mean using biased data leads to inaccurate results. Furthermore, the spatial extent was diminished, on the one hand because data availability shows better quality in that region, but also because Southern Europe is most exposed to wildfires. The visualization type chosen is small multiples because this visualization already meets criteria (see Table 9) and the colors which are the essential part of the interpretation of Warming Stripes, can also be incorporated within this visualization.

	Small Multiples with Pixel-based Colors	**Modification**
Spatial representation	Yes	-
Regional variability	Yes	-
Comparison of three variables	No	Creation of 3 distinctive map series
Overall Trend	Yes	-

Table 9: Fulfilments on Small Multiples with key messages of Visualization 1 and needed Modifications

For the creation of small multiples for Visualization 1, the same process was applied to all three variables. The entire data preparation process until the plotting of the individual maps was executed in R. Initially, a prepared shapefile of Southern Europe in QGIS was masked to the LFMC, FWI and TEMP time series, which can be considered as spatial subsetting. Afterwards the reference value for each cell was computed by taking the mean of the entire time series, which only contained values from the summer season. Yearly averages for each cell were calculated by aggregating the cell values within one year. Then the anomalies per cell were calculated by subtracting the reference value from each yearly value.

Each dataset was carefully inspected again for data classification. Following the classification of Warming Stripes which classifies the data in even steps with symmetric intervals, a classification deviating from 0 in even steps was aimed for. After inspecting the individual histograms of the anomaly values, a classification from 0 to symmetric intervals was chosen for all datasets but not even steps were considered because the data distribution differed from a symmetric even distribution. The decision not to use even steps was made because such a method would have led to classes with very few data values at the extreme ends of the distribution, namely at the classes with the highest and lowest values. The selected classification method allows for a more balanced data distribution within each class and leads to a clearer and more informative representation of the data's characteristics. Afterwards, the individual maps for each year and each variable were plotted in R and saved locally.

The arrangement of the plots was executed in Adobe Illustrator. The goal was to arrange the elements in a way that the whole time series with all three variables would fit on an A4 page. Not only the maps for all three variables of every year should be incorporated but also additional information on the variables which includes a few introducing sentences for LFMC, FWI and TEMP and on wildfires. Additionally, a line graph depicting the burned areas of Southern European countries within the time frame was visualized alongside. The data for the chart was derived from the European Environment Agency (2021) in form of a csv file. The data preparation took place in R which included filtering for countries within Southern Europe and cropping the temporal extent to 1988 to 2016. Afterwards, the result was transferred to Illustrator in order to visualize them. The result of this visualization as well as the discussion on the visualization can be seen in Chapter 3.3.

2.6.2 Visualization 2: Adaptation of Climate Spiral

For the second visualization the Climate Spiral was adapted to a static spatio-temporal visualization. The tasks that were defined for this visualization are biogeographical regions of Europe. The shapefile of the biogeographical regions of Europe were loaded into R, and also further data preparation as well as the plotting of the spirals took place in R. The data that was used for this visualization was the data after the exclusion of layers with high values of missing data.

First, the workflow in R will be described. The regions of the shapefile were separately saved, and each region was masked to a variable respectively. This step resulted in ten regions which had to be considered for each variable. In this visualization the focus is on monthly values and not on yearly values as in the visualizations before. Each region's reference value is represented by its monthly mean. Therefore, a loop was created that calculates the monthly average for each region and each variable. Furthermore, another loop was created that iterates through all 348 months, resulting in a single mean value for each region for each variable. The anomalies were computed by subtracting the reference value from the corresponding monthly aggregated values. These computed anomalies were then saved as dataframes.

For the visualization of the dataframes for each region to spiral plots, video tutorials by Riffomonas Project (2022a, 2022b) were followed and modified for the specific variables. Also, for this visualization each data classification was chosen based on the data distribution and range. Again, for reasons of comparability, the data range as well as the classifications were chosen based on the total anomaly values per variable. After plotting and saving the total of 30 spiral plots, the results were integrated into Illustrator.

The choropleth map of the biogeographical regions was modified in QGIS because the level of detail of the obtained shapefile from the European Environment Agency (2023) was too high to be used. Additionally, this level of detail was not needed for the visualization. Therefore, a simplification of the polygons with the Visvalingam algorithm was executed with a tolerance of 1.000 meters. Furthermore, smoothing was applied to the resulting polygons. Afterwards, this component was also transferred to Illustrator where the visualizations were prepared.

Just as for Visualization 1, also for this visualization a depiction of all three variables with a spatial representation in the form of a map was aimed for but unfortunately this could not be achieved and therefore three distinctive maps were created, each covering one A4 page and showing the development of LFMC, FWI and TEMP respectively in biogeographical regions of Europe. The results and the discussion of these spatio-temporal visualizations can be seen in Chapter 3.4.

3 Results and Discussion

3.1 Survey

The survey carried out during the long night of sciences event in Dresden, had a total of 42 participants. Among them, 39 responded by filling out the form manually in person, while the remaining 3 did so online. The results of the survey were both graphically and textually processed and can be found in the following section. The demographic information for the participants is presented in both Figure 19, illustrating age distribution, and Figure 20, displaying the level of education attained. Despite aiming for an even age representation, uneven willingness to participate across age groups lead to the majority of participants being between the age of 16 and 30.

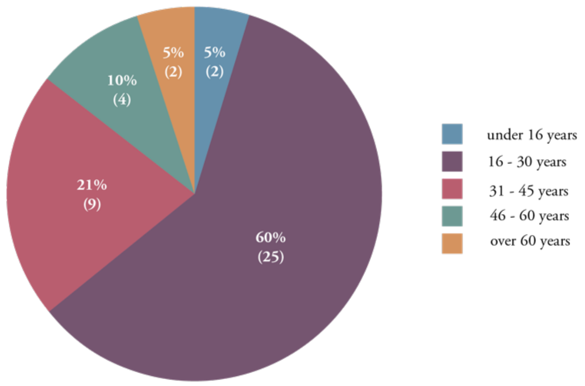


Figure 19: Age distribution of participants of survey

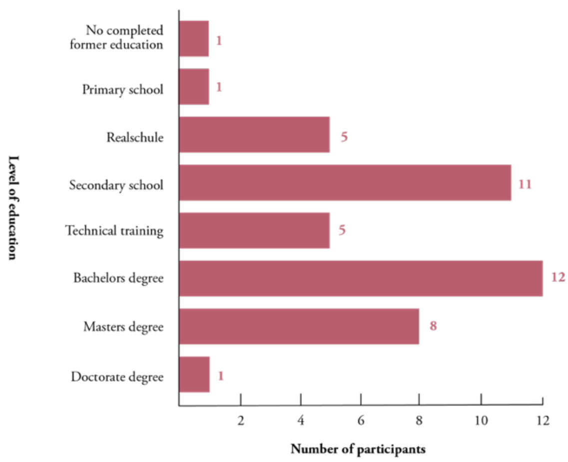


Figure 20: Distribution of highest accomplished educational level across survey participants

One of the primary goals of this survey was to examine the familiarity of Warming Stripes across various age groups and educational backgrounds. Figure 21 shows the result of the question regarding participants' prior exposure to this visualization. Despite the common perception of Warming Stripes being widely recognized, the survey revealed that 62% of participants were not familiar with this visualization. Of those surveyed, 21% reported seeing the visualization multiple times while 17% had seen it once, which results in more than a third of the participants. However, less than half of the individuals who stated they had seen Warming Stripes before were able to describe the visualization correctly. This indicates that Warming Stripes do seem familiar to about one-third of the people but more than half of those interpreted the visualization as something else than Warming Stripes.

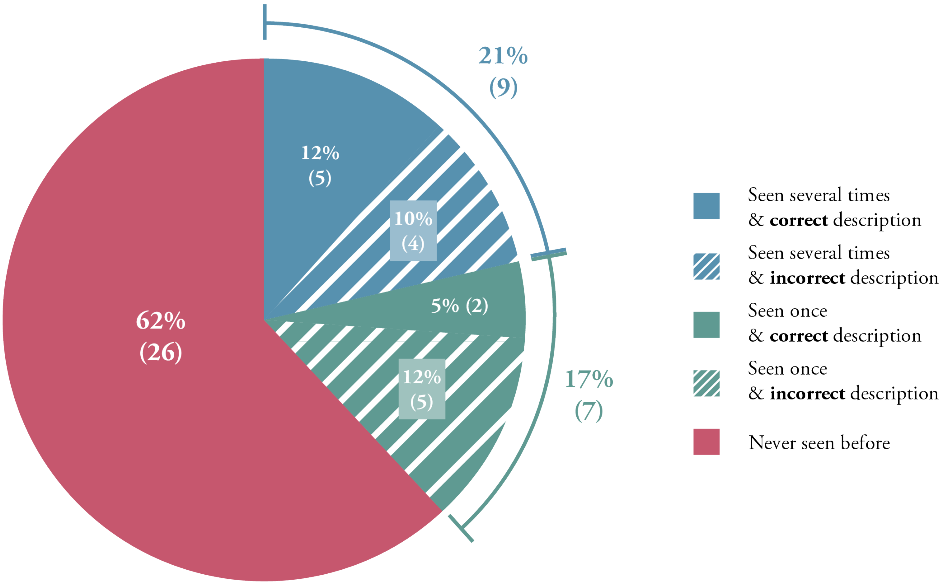


Figure 21: Familiarity with visualization and knowledge of Warming Stripes of survey participants

The open-ended question on the description of the visualization was asked to all participants, regardless of their familiarity with the graph. Nonetheless, for the evaluation of the answers the responses were considered respectively. The results for the participants who are familiar with the graph can be seen in Figure 22. More than half of this group described that the graph contains temperature data, but not all of them defined Warming Stripes as something else related to temperature. The second most common interpretation was classified as a spectral visualization labeled as a “spectrogram” or “splitting of light”. Also, one person interpreted the visualization as a movie barcode in which each frame of a film is represented as an elongated, thin stripe (Raup, 2011). Two participants who claimed to have seen the visualization were unable to give an interpretation. At least 7 out of the total of 16 participants within this group gave an incorrect interpretation of Warming Stripes, which resulted in a lack of awareness regarding the visualization.

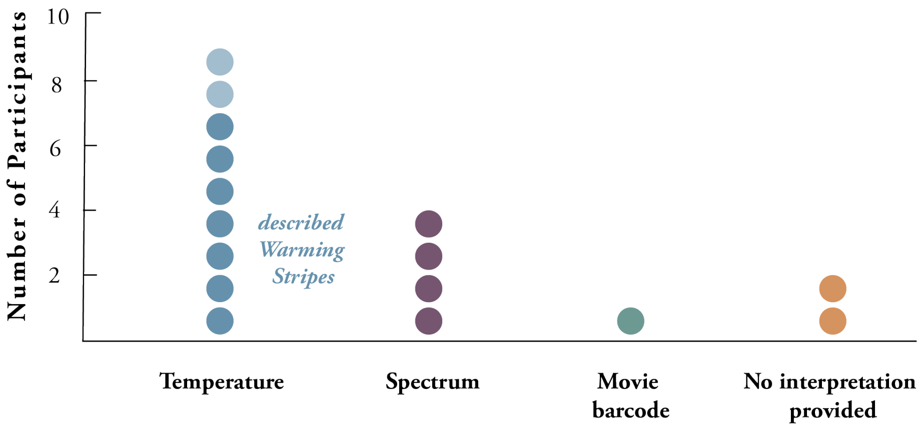


Figure 22: Description of visualization of participants with familiarity with the visualization

Among the participants who had not seen the visualization before (see Figure 23) a more diverse trend was visible within the answers. The two most frequently mentioned answers were still temperature and spectrum, although in this group they were mentioned with equal frequency. Further categories were barcodes and codes which were categorized into one class and as many people thought the graph was related to colors. Two people noted a connection between two variables and interpreted red as one variable and blue as another. Just as many named chemical analyses like "ph-analysis" and "substance analysis", while two other people did not give any interpretation. It is noteworthy that the two most frequently mentioned answers, irrespective of the familiarity with the graph were temperature and spectrum. It is also worth noting that most of the people provided descriptions on the colors rather than on the stripes specifically. Approximately 23% of the participants interpreted the visualization as temperature data but also just as much interpreted it as a spectrum. Therefore, it is safe to say that the visualization is not as intuitive as advertised and requires further explanation. The results might have looked different if a title, legend or labels of the stripes were included.

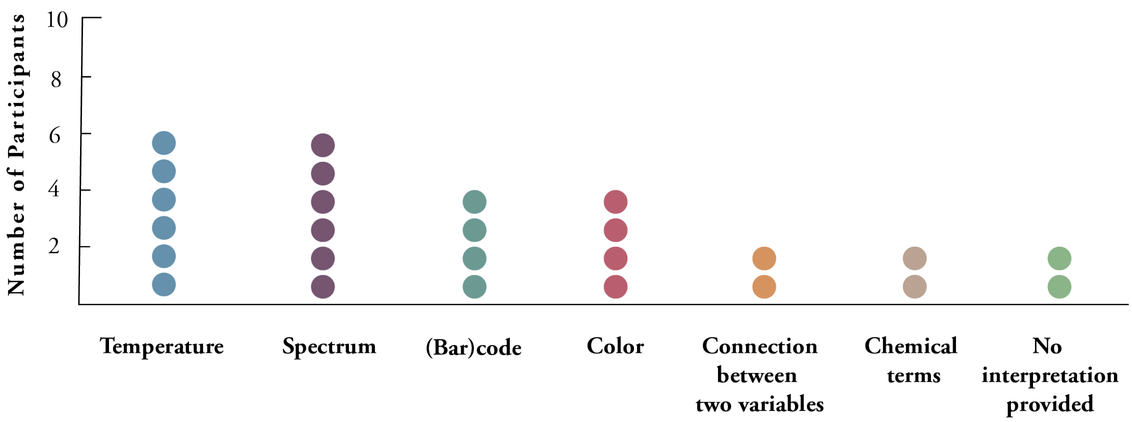


Figure 23: Description of visualization of participants without familiarity with the visualization

Regarding the question on where people have seen the visualization before that was asked to everyone answering they had seen Warming Stripes somewhere before, half of the answers were not included in the results because people gave a description of the visualization itself and not where they had seen it. This suggests a need for improved clarity in the wording of the question, such as providing examples for participants to choose from. Figure 24 shows a word cloud with the mentioned places of observation. A larger font size indicates two mentions while a smaller font size indicates one mentioning. Given the limited sample size, no definite conclusions can be drawn other than that most of the double mentioned locations involve media channels.

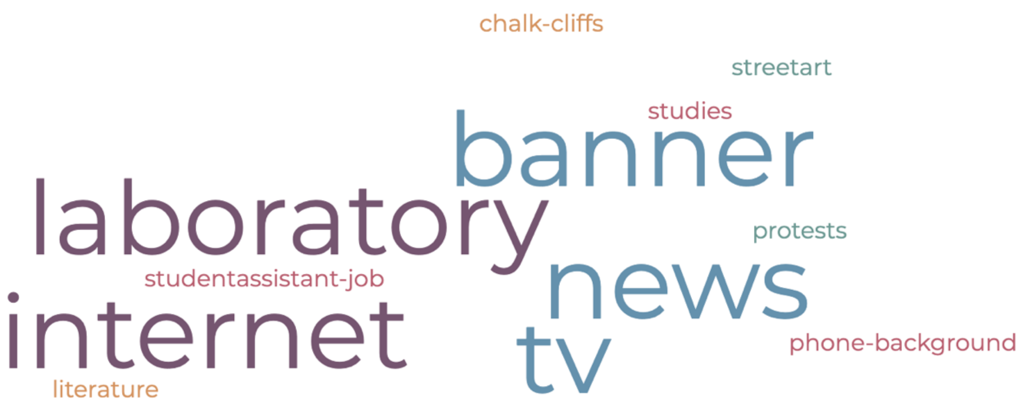


Figure 24: Places of observation of participants with familiarity with the visualization

3.1.1 Main findings

The main aim of the survey was to determine if the public's awareness of Warming Stripes is what its reputation says it is. Based on the results of this survey, this cannot be confirmed since only 17% of the participants knew this visualization and could describe it accurately. Around 22% of the people thought that they had seen the visualization before but did not give a correct description of the graph. Since most of them could provide a description, it is reasonable to presume that people confused the graph with other visualizations.

Around 23% of the people who had not viewed the visualization before believed they had observed temperature data. So, for only for around a quarter of the people the visualization was intuitive in terms of the data represented. An increase or a trend following from one color to the next could only be depicted by the people who described Warming Stripes (17%) and probably of those who stated to see a spectrum (26%). Nonetheless, even if only a quarter saw a connection to temperature data, given to the fact that no other information than the stripes was given, it can be concluded that the colors are a suitable choice for the depiction of temperature because people are familiar with this combination of colors in the context of temperature data depiction. However, due to the small sample size per age group and educational level no trends could be drawn from those classes.

This questionnaire gives answers to Research Question 2 and addresses the corresponding Research Objective 2.

3.2 Preliminary Visualization

Based on the trend for FWI that did not show clear increases in anomalies (see Figure 25), doubts were raised regarding the visualization. Moreover, the vast number of biomes made comparisons between them difficult. Thus, a common classification was chosen to allow comparison between the Blocks. Otherwise, a legend for each biome would have been needed, but the unified classification leads to negligence of individual trends within the biomes. Since the trends in FWI and LFMC differ regionally, not like for TEMP where global increasing trends are notable, a more regional inspection of the trend is suitable. This also leads to the assumption that biomes are regions that are too big for this type of visualizations. Also, the adapted Warming Stripes are not considered as a suitable method for the depiction of LFMC and FWI because the global depiction of these variables would not make sense due to the regional differences. Only with further modifications by taking smaller areas as basis of the stripes, a result can be achieved which also adds up to the research of these variables.

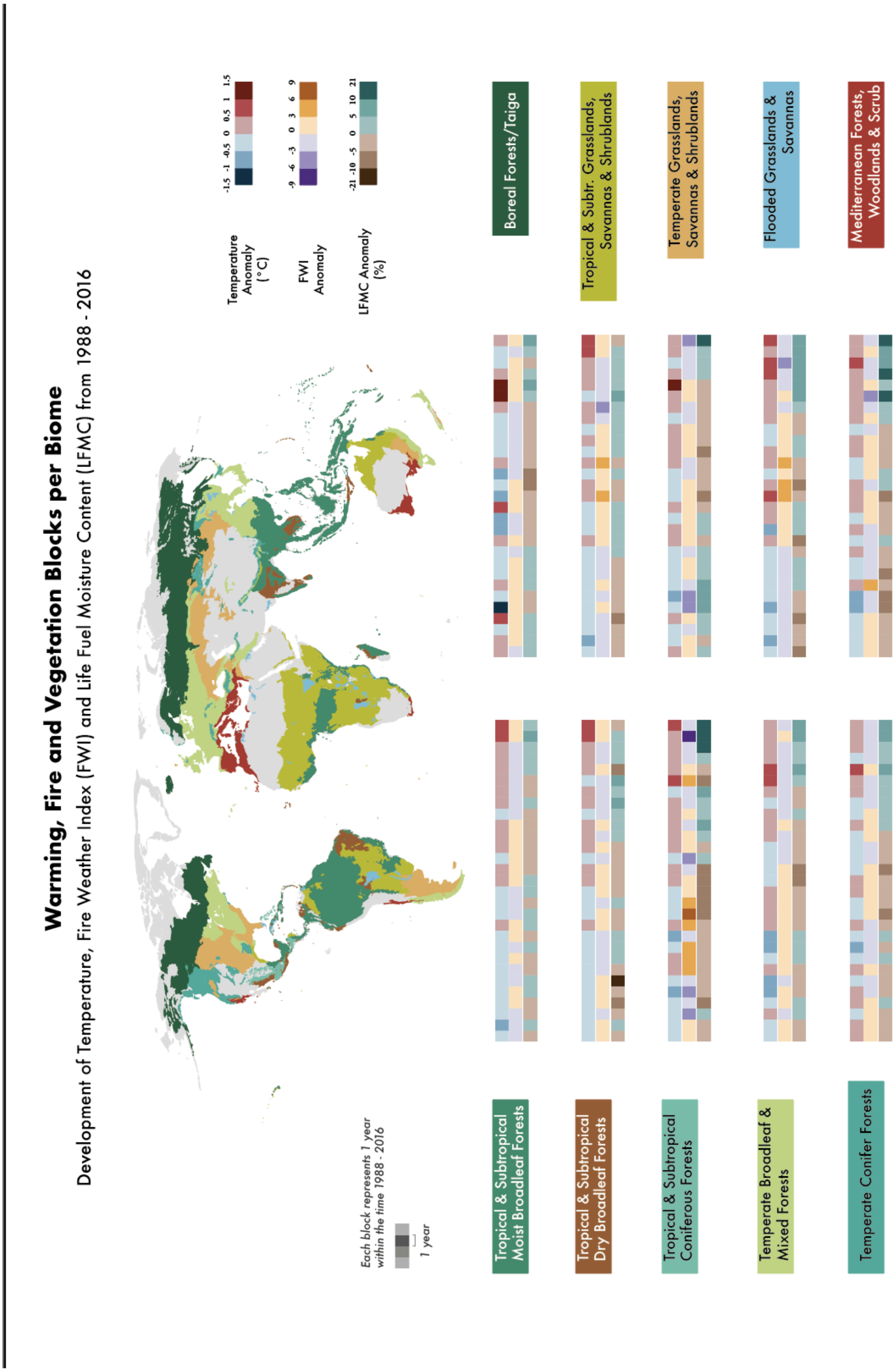


Figure 25: Adaptation of Warming Stripes with Biomes

3.3 Visualization 1: Modification of Warming Stripes

Visualization 1 (see Figure 26) uses small multiples as a visualization type to show the overall development of LFMC, FWI and TEMP over the years 1988 to 2016. Yearly aggregated plots of each variable's anomaly value are presented on a map that solely uses color shading within each cell to indicate the value. A diverging color scheme is utilized for each variable to represent positive and negative anomaly values. Because only colors were used to display the values, it was unnecessary to depict the individual maps in detail, which allowed all 87 maps to fit on one page. The maps are arranged in a linear manner and are accompanied by a stacked bar chart for each point in time, indicating the burned area in km^2 of selected Southern European countries within the respective year. The colors of the stacked bar chart were selected as various shades of gray to prevent confusion with anomaly values and to not overload the audience with excessive colors. This allows for regional differences to be observed, analyzed, and compared with other variables, due to the depiction of anomaly values per pixel. However, it should be noted that the transition from Warming Stripes to Visualization 1, significantly increased the complexity of the visualization. Instead of showing simple stripes, the audience is now exposed to a multitude of maps that also depict regional variations. While arranging the data, an approach was tested in which a time series of equal intervals but with gaps of years was portrayed. This approach was considered as unsuccessful as the variables do not show a linear trend in development. Therefore, skipping years results in the loss of valuable information. While there has been a shift in TEMP from predominantly blue cells to mostly red cells in recent years, FWI does not show a distinct trend towards a majority of high anomaly values. Nonetheless, it is evident that FWI values were significantly lower in the first half of the time series compared to the second half. Meanwhile, LFMC anomaly values show developments to higher values and therefore a greening within this area.

When visually comparing FWI to TEMP of the same region, most of the years show similar spatial patterns with high TEMP anomaly values and corresponding high FWI anomaly values, and vice versa. Whereas for LFMC and FWI, the similarity is not as explicit but also when comparing those two variables, similar patterns can be seen with low FWI anomaly values in the same region as high LFMC anomaly values.

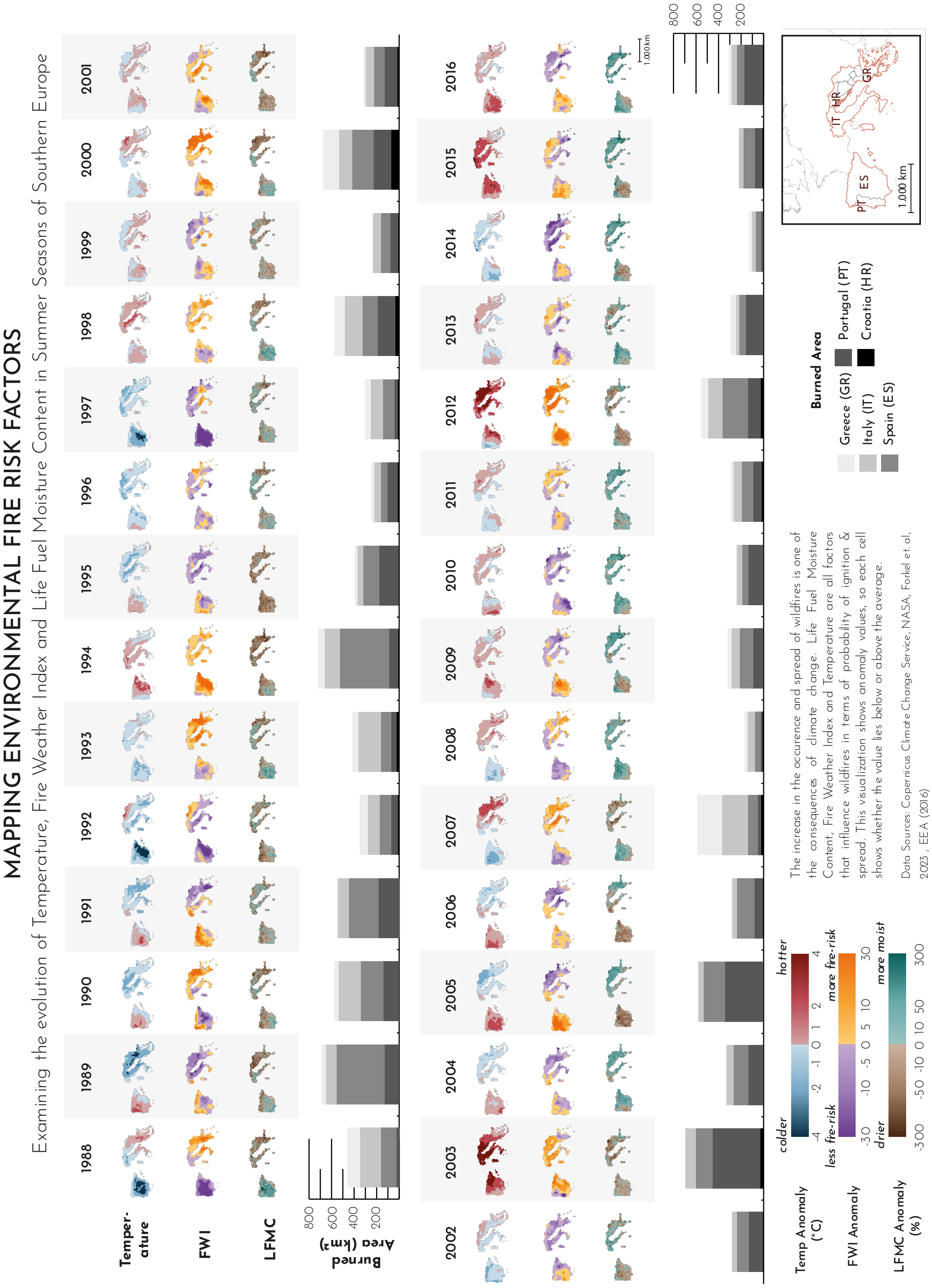


Figure 26: Visualization 1 - Modification of Warming Stripes

3.3.1 Evaluation

The evaluation of Visualization 1 will be based on the fulfilment of the key messages of each visualization which were defined in Chapter 2.2.3. For Visualization 1, the key messages of displaying regional variability, comparison of all three variables, and an overall trend of all three variables were defined. Since regional variability in form of a choropleth map did not result in clear trends, one key message of showing regional variability was updated to regional variability on pixel level. Furthermore, other selected visualization types from the literature review were examined on whether they meet the criteria.

Visualization 1 depicts an overall pixel-based trend of all three variables within summer seasons in Southern Europe from 1988 – 2016. With this depiction, Visualization 1 meets all the criteria that were defined beforehand. We will now further investigate how a spatio-temporal visualization for LFMC, FWI and TEMP could look like and whether these visualization types would fulfil the criteria. It is important to state that the suggested visualizations are not the only possibilities to display one specific visualization type. For each visualization type one visualization possibility that is found most suitable will be evaluated. The evaluation can also be found in Table 10.

For the depiction as a line plot, a pixel-based depiction within a static visualization would not be feasible within this spatial extent and especially not for three variables. Thus, the visualization would need to include a spatial aggregation (e.g., regions within Southern Europe). If the number of regions is not too high, the line plots could also be placed on a choropleth map which shows the regions chosen. Each region would therefore be accompanied by one line plot that shows the overall trend of LFMC, FWI and TEMP in one graph. Showing the variables in a shared-space visualization facilitates the comparison between them. Due to the simplicity of three lines for one graph, a line graph can also be relatively small. Similar would a visualization based on horizon graphs look like. Since a horizon graph for every single pixel would be too cluttered and too much information would be depicted within a static map, also here data must be aggregated. Each region would include three respective trends of LFMC, FWI and TEMP which can be easily compared due to distinctive color schemes. Because of the horizontal mirroring of the negative data values, the visualizations would take up less space. What applies to line plots and horizon graphs also apply to a modification of Warming Stripes. Since a pixel-based depiction with Warming Stripes would not be possible due to cluttering, also in this case a spatial aggregation would be needed to fit the information of the three distinctive variables on a map in form of stripes. It needs to be tested if the visualization with Warming Stripes would clearly give trends. The visualization type of spatially aggregated Climate Spirals would give a clear depiction on the trend of each variable, even though a comparison of the three variables within one map would not be feasible as will be drawn as a conclusion from the creation of Visualization 2. Helix icons on a map would check all of the key messages that are included within this scenario. However, this visualization type was created as an interactive visualization and this visualization was planned to be static therefore, a solution needs to be found to prevent occluded parts of the helices.

	Combination with map	Regional variability on pixel level	Comparison of three variables	Overall trend
Visualization 1	Yes	Yes	Yes	Yes
Line Plot	Yes	No	Yes	Yes
Horizon Graph	Yes	No	Yes	Yes
Warming Stripes	Yes	No	Yes	Yes
Climate Spiral	Yes	No	No	Yes
Helix Icons	Yes	No	Yes	Yes

Table 10: Summary on Evaluation of Selected Visualizations for Visualization 1

3.4 Visualization 2: Modification of the Climate Spiral

Visualization 2 illustrates the monthly mean values of each variable (see Figure 27, Figure 28, Figure 29) respectively across nine biogeographical regions from 1988 to 2016. A comparison of all three variables on one single map covering one A4 page could not be achieved. Therefore, three distinctive visualizations were created which incorporated the development of one variable across all biogeographic regions in Europe.

The spatio-temporal visualizations display spiral graphs based on the Climate Spiral as icons on a choropleth map. One spiral is represented for each of the 10 biogeographical regions of Europe. Monthly values are represented within the graph and each value except for the starting and the end point are connected to the previous and consecutive data point. Therefore, the mean monthly values of each variable are represented within a line that follows the shape of a spiral.

The Climate Spiral employs a blue to yellow color scale. Each variable is assigned a distinct color scheme. TEMP is indicated by a blue to red scheme, LFMC by a brown to green scheme, and FWI by a blue to orange scheme. Although FWI is represented by a different shade of blue than TEMP, the use of distinct blue shades is negligible since the variables are not displayed together. The legend for each variable displays two distinct labels indicating that the color represents the years and not the values. Despite the colors not representing the values, a scheme was chosen that could be mistakenly interpreted as such. For TEMP, the variable where confusion between years and values can occur, an increasing trend of temperature over time is visible and therefore the colors only enhance the depiction. For the other variables no confusion is expected.

The depiction clearly displays the development of each variable over time. Even the lines that indicate the start of the time series are easily distinguishable, and their covering by more recent lines indicates the absence of changes within the anomalies. Some spirals exhibit data gaps resulting from the exclusion of layers with high missing values earlier on. The visibility of these data gaps provides crucial insight into the data availability.

Nonetheless, also disadvantages need to be mentioned. Comparability between the spiral graphs should be facilitated by the decision of choosing data classifications that are able to cover the whole data range. Since some biogeographical regions show changes in anomalies within smaller scales, these regions appear to have no changes at all. Especially for the variable FWI, some regional variation across years was not visible and therefore data was normalized but this operation resulted in similar results with other regions having undiscernible anomaly changes.

A prominent trend which can be seen within LFMC data is that for some regions, the peak of high values changes from one month to another with time. The steppic biogeographical region for example displays a peak during earlier years in August, while this changed to a peak in April in more recent years.

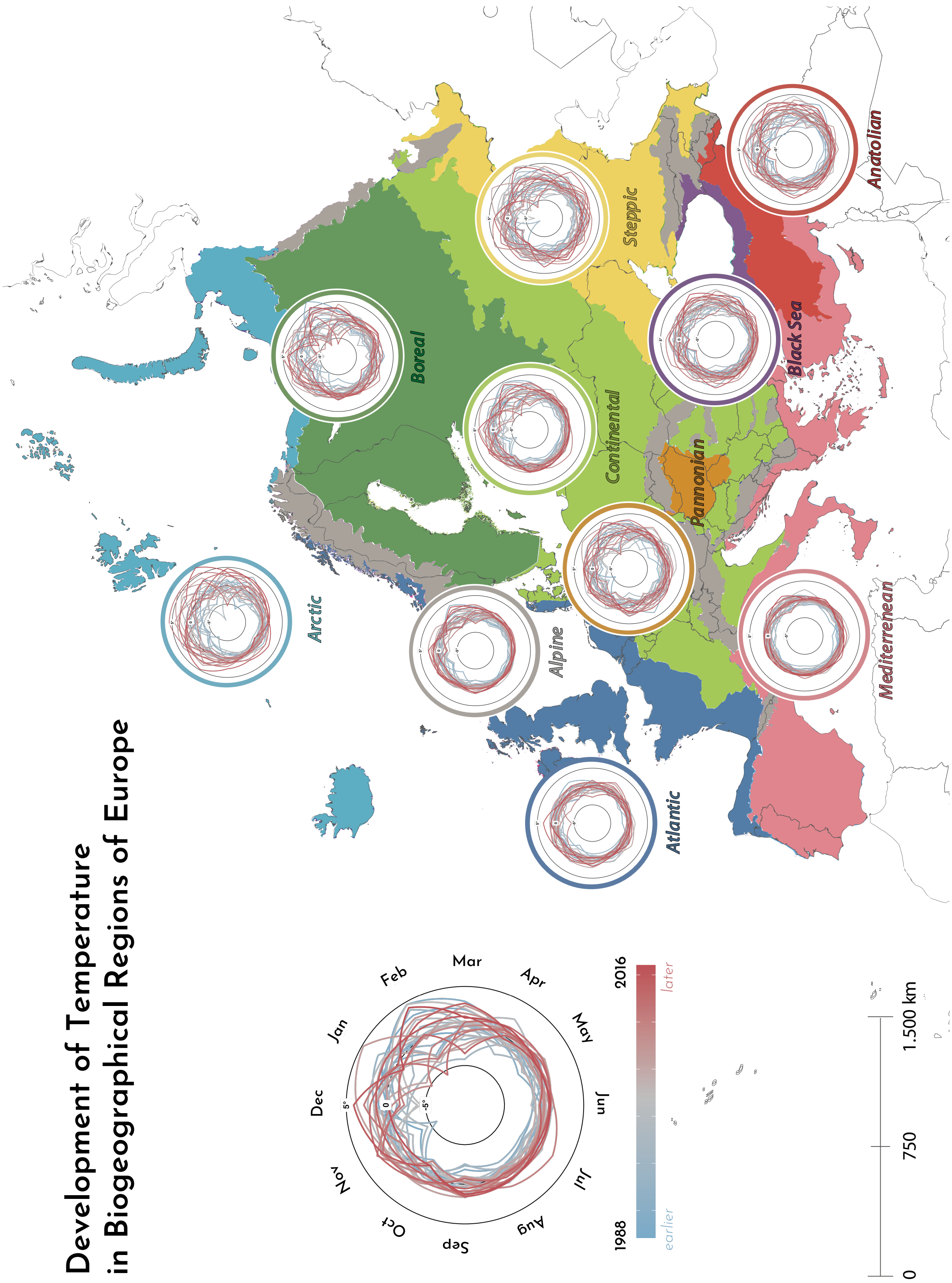


Figure 27: Modification of the Climate Spiral with Temperature

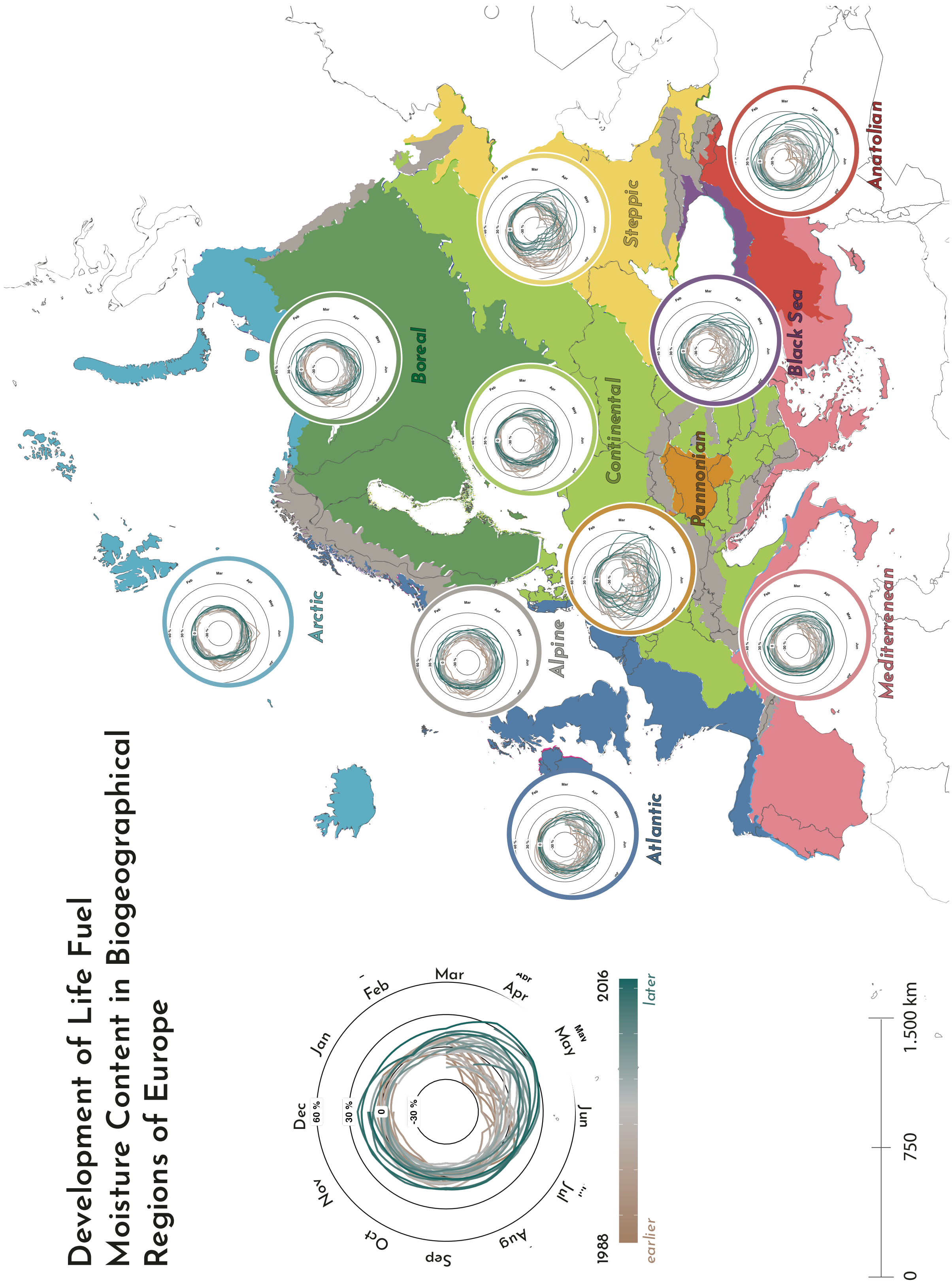


Figure 28: Modification of the Climate Spiral with Life Fuel Moisture Content

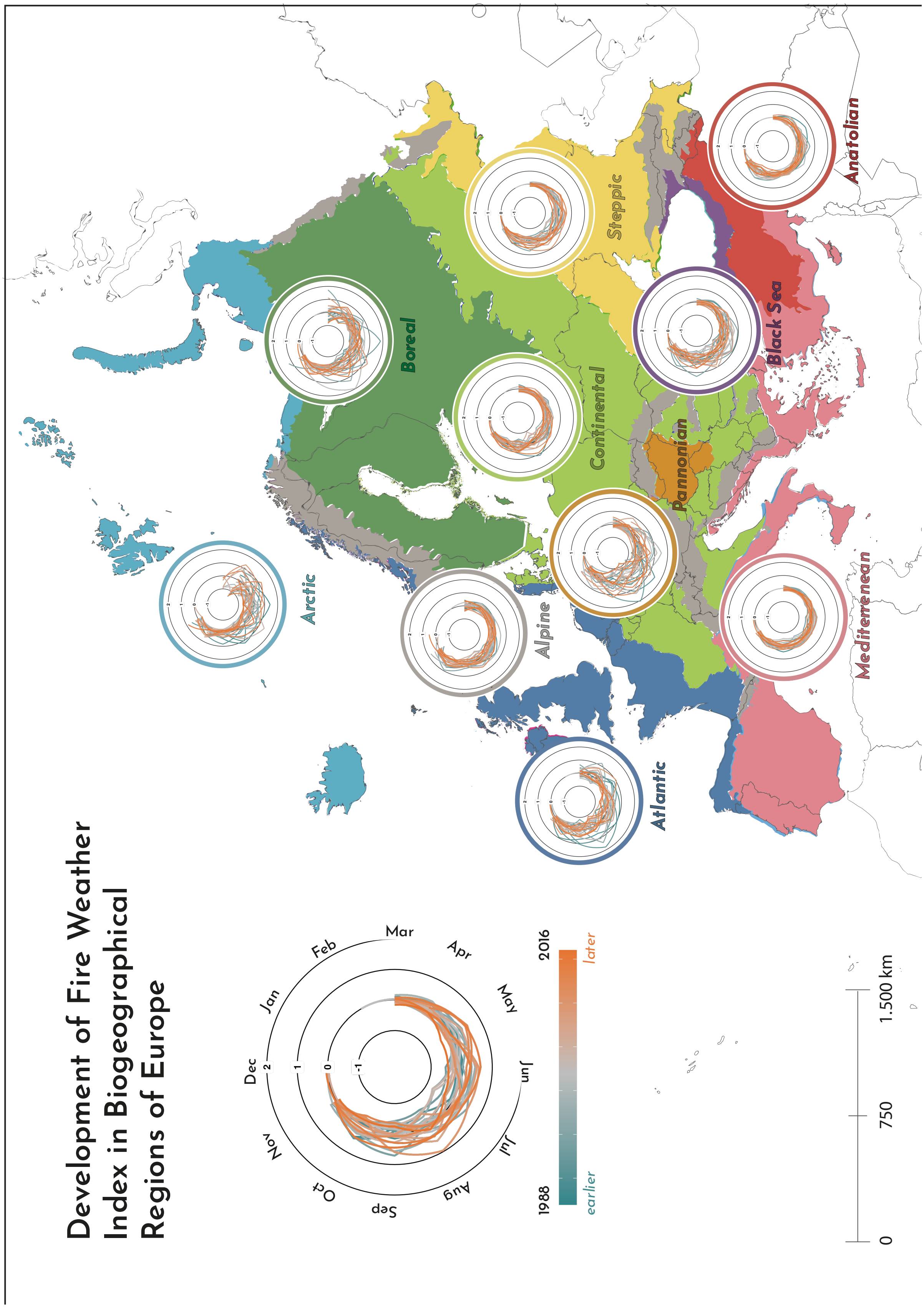


Figure 29: Modification of the Climate Spiral with Fire Weather Index

3.4.1 Evaluation

Also, for the second visualization based on the Climate Spiral, the key messages of the intended visualizations (see Table 3) are mentioned again: periodic patterns of monthly data, comparison of three variables and the visualization of nine biogeographical regions. Like Visualization 1, selected visualizations will be evaluated based on whether they meet the criteria of the key messages.

Unfortunately, Visualization 2 does not meet all of the defined criteria by not showing a comparison for the three variables. Nonetheless, the other criteria are met by the visualization which include the incorporation of a map, periodic patterns for monthly data and the depiction of nine biogeographical regions.

One point that will be considered as more important than the others is the depiction of monthly values and the display of seasonality that goes with it because this is a point that distincts the visualization from others. Therefore, no data aggregation will be excepted for the other visualizations. A line plot is not appropriate for a visualization that involves monthly values over 29 years due to the excessive amount of data points, which would make the plot extremely long and impossible to fit on a map. However, we will extend the idea of considering a line plot without spatial representation in the form of a map. Spatial representation can be achieved by labeling the lines, yet comparing all three variables would require plotting three line plots for each variable, causing illegibility because of the vast amount of data points. Moreover, a cyclic plot allows for greater visibility of the periodic pattern. Without data aggregation, the only feasible approach for a horizon graph would entail creating 30 unique charts, with 10 for each region and 3 distinct variables. Examining these charts for patterns and trends, however, would be impractical. The same limitations apply to Warming Stripes, as generating 30 individual stripes, each with 348 monthly values, would be excessively demanding and would not allow for meaningful comparison. The use of small multiples is not feasible without data aggregation as the resulting series would be too long. The helix icons are the only viable option for this visualization scenario among the selected options. These helix icons can be integrated into a biogeographical choropleth map, where each helix ribbon represents the three variables. The only downside is that the visualization can only show all data values when being interactive. This leads to the conclusion that the key messages that were defined in the beginning are not achievable with the chosen visualization types and within a static visualization that should only cover an A4 page.

	Combination with map	Periodic pattern of monthly data	Comparison of three variables	Nine biogeographical regions
Visualization 2	Yes	Yes	No	Yes
Line Plot	No	No	No	Yes
Horizon Graph	No	No	No	Yes
Warming Stripes	No	No	Yes	Yes
Small Multiples	No	No	No	No
Helix Icons	Yes	Yes	Yes	Yes

Table 11: Summary on Evaluation of Selected Visualizations for Visualization 2

With the creation of the evaluation frameworks for Visualization 1 and Visualization 2 based on defined key message Research Objective 1 was addressed and Research Question 1 can be answered.

3.5 Limitations and Future Work

This thesis covered suitable modifications of novel visualizations Warming Stripes and the Climate Spiral to a spatio-temporal visualization incorporating a map. Therefore, these visualizations were created and adapted for the specific data, and tested for whether a modification was feasible and if not, which visualizations would be suited for the data. Therefore, the thesis did focus on the right visualization type and how to choose it but did not go in depth on whether the visualized data corresponds with existing research in the field of remote sensing time series of LFMC and FWI.

Therefore, an evaluation of the visualizations on their accuracy by comparing the outcome with findings in literature can be assessed in the future. Other visualization types than the ones presented within the literature review can be tested on whether they can depict the key messages of the visualizations. These results can be compared to the results within this study.

Additionally, the visualizations should be evaluated for clarity. The visualizations should be tested for their effectiveness with a user group that fits the criteria of the defined target audience. The testing should include if the key messages could be conveyed to the audience.

Furthermore, the survey was intentionally designed for people without additional information, as more familiarity with the graph was expected from the participants. Further research should focus on providing context to the people by mentioning the title, adding labels for the years, or a legend. This could lead to different findings. Another aspect that should be considered for further investigations is the location of the conduction of the survey. The survey was conducted at the main auditorium of the long night of sciences in which booths from several institutes were located. It would be valuable to assess if questioning of people in the building of the theme "environment & nature" would have led to different outcomes. A comparison of groups of interests could be executed in the future.

4 Conclusion

Throughout this research different visualizations on temporal as well as on spatio-temporal data were inspected. It is noteworthy that the visualizations presented and created within this thesis only show an excerpt of (spatio)-temporal visualizations. These visualizations were chosen based on their suitability for the data but also this does not suggest that these are the only visualizations suitable for the spatio-temporal depiction of LFMC, FWI and TEMP.

Not only traditional approaches which are often mentioned in literature were considered but also so-called novel visualizations like Warming Stripes and the Climate Spiral were inspected, even though no empirical evidence on their effectiveness were given. This study connected to this research gap and tested Warming Stripes for their awareness and effectiveness. The results of the survey did not add up to the reputation of the visualization because awareness and effectiveness could not be proven within the sample group.

Furthermore, an evaluation framework was created which involved the key messages for the visualization of the variables LFMC, FWI, and TEMP. Historic and novel visualizations were inspected and tested for their suitability within these key messages. The visualizations created that are based upon the novel visualizations of Warming Stripes and the Climate Spiral were found to be the most suitable for the defined key messages and the constraint of showing a spatial representation in form of a map within a static visualization.

The created visualizations give insight into a multivariate dataset on environmental fire risk factors which had not been depicted in that way. Nonetheless, the visualizations need to be further investigated by testing their effectiveness as well as their accuracy.

References

Aigner, W., Miksch, S., Schumann, H., & Tominski, C. (2011). *Visualization of Time-Oriented Data*. Springer London. https://doi.org/10.1007/978-0-85729-079-3

Andrienko, N., & Andrienko, G. (2006). *Exploratory analysis of spatial and temporal data: A systematic approach*. Springer.

Augustyn, A. (2023). *Biome*. Britannica. https://www.britannica.com/science/biome

Bertin, J. (1983). *Semiology of Graphics*. University of Wisconsin Press.

Chuvieco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M. P., Vilar, L., Martínez, J., Martín, S., Ibarra, P., de la Riva, J., Baeza, J., Rodríguez, F., Molina, J. R., Herrera, M. A., & Zamora, R. (2010). Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. *Ecological Modelling*, *221*(1), 46–58. https://doi.org/10.1016/j.ecolmodel.2008.11.017

Deacon, J., Fan, A., Fannin, B., Levine, J., Quigley, K., & Yu, P. (2020). *Introduction to data visualization*. Casualty Actuarial Society E-Forum, Summer 2020.

Dempsey, C. (2021, December 5). *Terrestrial Ecoregions GIS Data*. GIS Lounge. https://www.gislounge.com/terrestrial-ecoregions-gis-data/

Dixon, D. (2022). In the breach: Feeling the heat of climate change. *Scottish Geographical Journal*. https://www.tandfonline.com/doi/abs/10.1080/14702541.2022.2157867

Dumble, S. (2021, October 28). *Warming Stripes: A problem solving Task in R*. https://stats4sd.org/blog/67

Ellis, T., Bowman, D., Jain, P., Flannigan, M., & Williamson, G. (2022). Global increase in wildfire risk due to climate-driven declines in fuel moisture. *Global Change Biology*, *28*, 1544–1559. https://doi.org/10.1111/gcb.16006

European Centre for Medium-Range Weather Forecasts. (2020, March 16). *Fact Sheet: Earth System Data Assimilation*. https://www.ecmwf.int/en/about/media-centre/focus/2020/fact-sheet-earth-system-data-assimilation

European Centre for Medium-Range Weather Forecasts. (2023, August 23). *Fact Sheet: Reanalysis*. https://www.ecmwf.int/en/about/media-centre/focus/2023/fact-sheet-reanalysis#:~:text=Reanalysis%20data%20provide%20the%20most,as%20'maps%20without%20gaps'.

European Environment Agency. (2021). *Forest Fires in Europe* [dataset]. https://www.eea.europa.eu/en/analysis/indicators/forest-fires-in-europe?activeAccordion=309c5ef9-de09-4759-bc02-802370dfa366

European Environment Agency. (2023). *Biogeographical Regions* [dataset]. https://www.eea.europa.eu/en/datahub/datahubitem-view/11db8d14-f167-4cd5-9205-95638dfd9618?activeAccordion=334349

Field, R. D., Spessa, A. C., Aziz, N. A., Camia, A., Cantin, A., Carr, R., de Groot, W. J., Dowdy,

A. J., Flannigan, M. D., Manomaiphiboon, K., Pappenberger, F., Tanpipat, V., & Wang, X. (2015). Development of a Global Fire Weather Database. *Nat. Hazards Earth Syst. Sci.*

Forkel, M., Schmidt, L., Zotta, R., Dorigo, W., & Yebra, M. (2023). Estimating leaf moisture content at global scale from passive microwave satellite observations of vegetation optical depth. *Hydrology and Earth System Sciences*, *27*, 39–68. https://doi.org/10.5194/hess-27-39-2023

Fridays For Future. (2023). *Fridays for Future – How Greta started a global movement*. Fridays For Future. https://fridaysforfuture.org/what-we-do/who-we-are/

Grainger, S., Mao, F., & Buytaert, W. (2016). Environmental data visualisation for non-scientific contexts: Literature review and design framework. *Environmental Modelling & Software*, *85*, 299–318. https://doi.org/10.1016/j.envsoft.2016.09.004

Hawkins, E. (n.d.). *Show Your Stripes*. Retrieved November 6, 2023, from https://showyourstripes.info/

Hawkins E. [@ed_hawkins]. (2016, May 9). *Spiralling global temperatures from 1850-2016 (full animation)* [Tweet]. Twitter. https://twitter.com/ed_hawkins/status/729753441459945474

Hawkins, E., Fæhn, T., & Fuglestvedt, J. (2019). The Climate Spiral Demonstrates the Power of Sharing Creative Ideas. *Bulletin of the American Meteorological Society*, *100*(5), 753–756. https://doi.org/10.1175/BAMS-D-18-0228.1

Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Biavati, G., Dee, D., Thépaut, J., & Rozum, I. (2023). *ERA5 hourly data on single levels from 1940 to present* [dataset]. https://doi.org/10.24381/cds.adbb2d47

Jabbari, A., Blanch, R., & Dupuy-Chessa, S. (2018). Composite Visual Mapping for Time Series Visualization. *2018 IEEE Pacific Visualization Symposium (PacificVis)*, 116–124. https://doi.org/10.1109/PacificVis.2018.00023

Jones, M., Abatzoglou, J., Veraverbeke, S., Andela, N., Lasslop, G., Forkel, M., Smith, A., Burton, C., Betts, R., Werf, G., Sitch, S., Canadell, J., Santín, C., Kolden, C., Doerr, S., & Le Quéré, C. (2022). Global and Regional Trends and Drivers of Fire Under Climate Change. *Reviews of Geophysics*, *60*. https://doi.org/10.1029/2020RG000726

Knaflic, C. N. (2015). *Storytelling with Data: A Data Visualization Guide for Business Professionals* (1st ed.). Wiley. https://doi.org/10.1002/9781119055259

Kraak, M.-J. (2014). *Mapping Time: Illustrated by Minard's Map of Napoleon's Russian Campaign of 1812*. Esri Press.

Landeshauptstadt Dresden. (2023). *About Us: Dresden Science Night*. Dredner Lange Nacht Der Wissenschaft. https://www.wissenschaftsnacht-dresden.de/en/translate-to-englisch-lndwdd

MacEachren, A. M. (1995). *How maps work: Representation, visualization, and design*. The Guilford Press.

Mikecz, J. M. (2019, January 1). *Visualizing Data (within and beyond Python)*. https://jeremymikecz.com/498/5_wbdata_viz.html

Monmonier, M. (1989). Geographic Brushing: Enhancing Exploratory Analysis of the Scatterplot Matrix. *Geographical Analysis*, *21*(1), 81–84. https://doi.org/10.1111/j.1538-4632.1989.tb00879.x

Monmonier, M. (1990). Strategies For The Visualization Of Geographic Time-Series Data. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *27*(1), 30–45. https://doi.org/10.3138/U558-H737-6577-8U31

Munzner, T. (2015). *Visualization Analysis & Design*. CRC Press.

NASA's Jet Propulsion Laboratory. (2023). *The Effects of Climate Change*. https://climate.nasa.gov/effects

National Centers for Environmental Information. (n.d.). *Dead Fuel Moisture*. Retrieved November 5, 2023, from https://www.ncei.noaa.gov/access/monitoring/dyk/deadfuelmoisture

National Centre for Atmospheric Science. (2018, September 18). *Ed Hawkins' warming stripes add colour to climate communication*. National Centre for Atmospheric Science. https://ncas.ac.uk/ed-hawkins-warming-stripes-add-colour-to-climate-communication/

Natural Resources Canada. (n.d.). *Canadian Forest Fire Weather Index (FWI) System*. Retrieved November 5, 2023, from https://cwfis.cfs.nrcan.gc.ca/background/summary/fwi

Nocke, T., Schumann, H., & Böhm, U. (2004). Methods for the visualization of clustered climate data. *Computational Statistics*, *19*(1), 75–94. https://doi.org/10.1007/BF02915277

O'Connor, Z. (2023). Environmental color interventions on a macro scale: Tactical urbanism and issues of global concern. *Color Research & Application*, col.22845. https://doi.org/10.1002/col.22845

Peuquet, D. J. (1994). It's About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. *Annals of the Association of American Geographers*, *84*(3), 441–461. https://doi.org/10.1111/j.1467-8306.1994.tb01869.x

Poole, S. (2022, October 28). Greta Thunberg's Climate Book is a great scientific one-stop shop – despite its anti-capitalist bent. *The Telegraph*. https://www.telegraph.co.uk/books/what-to-read/greta-thunbergs-climate-book-great-scientific-one-stop-shop/

Raup, J. (2011, April 6). *Movie Barcode: An Entire Feature Film In One Image*. https://thefilmstage.com/movie-barcode-an-entire-feature-film-in-one-image/

Reijner, H. (2008). The Development of the Horizon Graph. *Proc. Vis08 Workshop From Theory to Practice: Design, Vision and Visualization*, *3*.

Riffomonas Project (Director). (2022a, May 30). *Using ggplot2 to recreate a line plot of annual temperature anomalies (CC217)*. https://www.youtube.com/watch?v=gOyya2DLVfQ

Riffomonas Project (Director). (2022b, June 2). *How to recreate climate temperature spirals in R with ggplot2 (CC218)*. https://www.youtube.com/watch?v=NYF9ySYSvwQ

Sengupta, S. (2022, June 17). The surprising story of ‘warming stripes.’ *The New York Times*. https://www.nytimes.com/2022/06/17/climate/climate-warming-stripes.html

Theus, M. (2005). Statistical Data Exploration and Geographical Information Visualization. In *Exploring Geovisualization* (pp. 127–142). Elsevier. https://doi.org/10.1016/B978-008044531-1/50424-3

Tominski, C., Schulze-Wollgast, P., & Schumann, H. (2005). 3D Visualization for. *IEEE*.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley.

Yebra, M., Dennison, P. E., Chuvieco, E., Riaño, D., Zylstra, P., Hunt, E. R., Danson, F. M., Qi, Y., & Jurdao, S. (2013). A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products. *Remote Sensing of Environment*, *136*, 455–468. https://doi.org/10.1016/j.rse.2013.05.029

Yebra, M., Scortechini, G., Badi, A., Beget, M. E., Boer, M. M., Bradstock, R., Chuvieco, E., Danson, F. M., Dennison, P., Resco De Dios, V., Di Bella, C. M., Forsyth, G., Frost, P., Garcia, M., Hamdi, A., He, B., Jolly, M., Kraaij, T., Martín, M. P., … Ustin, S. (2019). Globe-LFMC, a global plant water status database for vegetation ecophysiology and wildfire applications. *Scientific Data*, *6*(1), 155. https://doi.org/10.1038/s41597-019-0164-9

Appendix: Survey

TECHNISCHE UNIVERSITÄT DRESDEN

Umfrage zur Bekanntheit und Interpretation einer Visualisierung / Survey on Familiarity & Interpretation of a visualization

Im Zuge meiner Masterarbeit des Studiums *Msc Cartography* beschäftige ich mich mit Visualisierungsformen in Kombination mit Karten. Die folgende Umfrage bezieht sich auf eine bestimmte Visualisierungsform und ihrer Bekanntheit und Interpretation.

Ihre Antworten sind anonym und werden ausschließlich für akademische Zwecke verwendet. Alle erhobenen Daten werden streng vertraulich behandelt und in zusammengefasster Form ausgewertet, somit werden keine persönlich identifizierbaren Informationen weitergegeben oder veröffentlicht. Indem Sie mit der Umfrage fortfahren, erklären Sie Ihr freiwilliges Einverständnis zur Teilnahme.

Das Ausfüllen des Fragebogens wird etwa 5 Minuten dauern.

Ich danke Ihnen im Voraus für Ihre Teilnahme!

Bei etwaigen Fragen kontaktieren Sie bitte:
monika_lalita.krautschneider@mailbox.tu-dresden.de

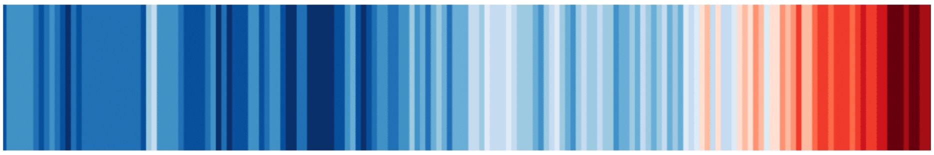
1. Wie alt sind Sie?
How old are you?

- unter 15 *under 15*
- 16 – 30 Jahre *16 – 30 years*
- 31 – 45 Jahre *31 – 45 years*
- 46 – 60 Jahre *46 – 60 years*
- Über 60 Jahre *over 60 years*

2. Was ist Ihr höchster Schul- oder Hochschulabschluss?
What is the highest degree or level of education you have completed?

- Kein abgeschlossener Schulabschluss *No completed formal education*
- Grund-/Hauptschulabschluss *Primary school*
- Realschule/Mittlere Reife
- Abitur oder gleichwertiger Abschluss *Secondary school*
- Abgeschlossene Ausbildung *Completed technical training/apprenticeship*
- Bachelor-Abschluss *Bachelors degree*
- Master-Abschluss *Masters degree*
- Doktorgrad *Doctorate degree*
- Andere *Others*

3. Haben Sie die folgende Visualisierung schon einmal gesehen?
Have you ever seen the visualization below?



- Ja, bereits mehrmals *Yes, several times*
- Ja, ein Mal *Yes, once*
- Nein *No*

Falls **ja**, führen Sie bitte Beispiele an wo Sie die Visualisierung bereits gesehen haben:
*If **yes**, please give examples of where you have already seen this visualization:*

Falls **ja**, geben Sie bitte eine kurze Beschreibung darüber, was diese Visualisierung zeigt.
*If **yes**, please provide a brief description of what you remember this visualization represents.*

Falls **nein**, geben Sie bitte eine kurze Beschreibung darüber, was diese Visualisierung Ihrer Meinung nach darstellen könnte (z.B. Was ist die Bedeutung der Streifen und der Farben? Welche Daten werden hier präsentiert?)
*If **no**, please provide a brief description of what you think this visualization might represent (e.g., What is the significance of the stripes and colors? Which data is presented here?)*

Haben Sie diese Visualisierung bereits mit anderen Daten als Temperatur gesehen?
Have you seen data, other than temperature, depicted with this visualization?

- Ja *Yes*
- Nein *No*

Falls ja, wissen Sie noch welche Daten (anders als Temperatur) dargestellt wurden?
If yes, do you remember which topic (other than temperature) was depicted?

____