

09

Evaluating Visualization Techniques

Notice

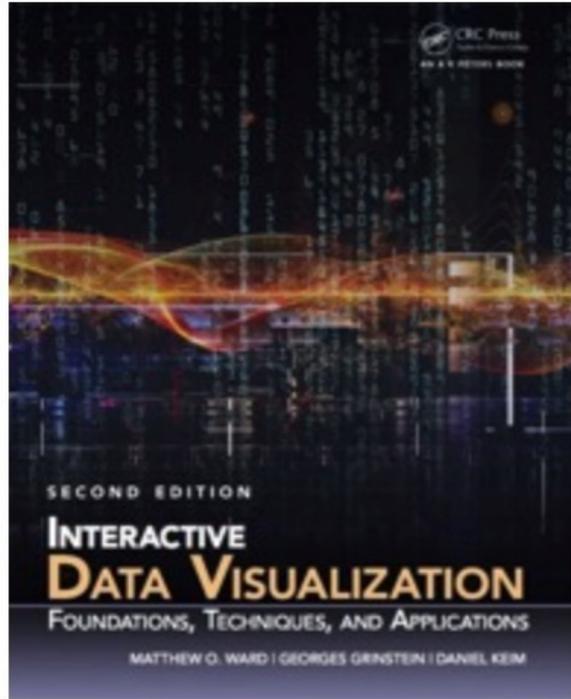
- **Author**

- ◆ **João Moura Pires (jmp@fct.unl.pt)**

- **This material can be freely used for personal or academic purposes without any previous authorization from the author, provided that this notice is kept with.**

- **For commercial purposes the use of any part of this material requires the previous authorisation from the author.**

Bibliography ...



Interactive Data Visualization: Foundations, Techniques, and Applications

Matthew O. Ward, Georges Grinstein, Daniel Keim

2015, 2nd Edition

ISBN: 9781482257373

ISBN (e-Book): 9781482257397

Table of Contents

- **Introduction**
- **User Tasks**
- **User Characteristics**
- **Data Characteristics**
- **Visualization Characteristics**
- **Structures for Evaluating Visualizations**
- **Benchmarking Procedures**
- **An Example of Visualization Benchmarking**

Introduction

Which visualization technique(s) should I use?

Which visualization technique(s) should I use?

- It depends on specific task or **tasks** the user wishes to accomplish, the **characteristics of the data**, and the **user's level of experience** in using visualization to help solve problems.

Which visualization technique(s) should I use?

- It depends on specific task or **tasks** the user wishes to accomplish, the **characteristics of the data**, and the **user's level of experience** in using visualization to help solve problems.
- How to determine under what conditions a particular technique is better than other techniques

Which visualization technique(s) should I use?

- It depends on specific task or **tasks** the user wishes to accomplish, the **characteristics of the data**, and the **user's level of experience** in using visualization to help solve problems.
- How to determine under what conditions a particular technique is better than other techniques
- Whether a modification made to a particular method improves or makes worse some aspect of the technique

Which visualization technique(s) should I use?

- It depends on specific task or **tasks** the user wishes to accomplish, the **characteristics of the data**, and the **user's level of experience** in using visualization to help solve problems.
- How to determine under what conditions a particular technique is better than other techniques
- Whether a modification made to a particular method improves or makes worse some aspect of the technique
- Performing rigorous evaluation is difficult and time-consuming !

User Tasks

User Tasks

- **Peter R. Keller and Mary M. Keller. *Visual Cues: Practical Data Visualization*, 1994**
propose a list of actions:

User Tasks

- Peter R. Keller and Mary M. Keller. **Visual Cues: Practical Data Visualization**, 1994
propose a list of actions:
- **identify**: to recognize an object based on the characteristics presented, such as finding a fracture in an x-ray;

User Tasks

- Peter R. Keller and Mary M. Keller. **Visual Cues: Practical Data Visualization**, 1994
propose a list of actions:
- **identify**: to recognize an object based on the characteristics presented, such as finding a fracture in an x-ray;
- **locate**: to establish the position of an object;

User Tasks

- Peter R. Keller and Mary M. Keller. **Visual Cues: Practical Data Visualization**, 1994
propose a list of actions:
- **identify**: to recognize an object based on the characteristics presented, such as finding a fracture in an x-ray;
- **locate**: to establish the position of an object;
- **distinguish**: to determine that an object is distinct or different from another;

User Tasks

- Peter R. Keller and Mary M. Keller. **Visual Cues: Practical Data Visualization**, 1994
propose a list of actions:
- **identify**: to recognize an object based on the characteristics presented, such as finding a fracture in an x-ray;
- **locate**: to establish the position of an object;
- **distinguish**: to determine that an object is distinct or different from another;
- **categorize**: to classify objects into distinct types, such as different land cover or material types;

User Tasks

- Peter R. Keller and Mary M. Keller. **Visual Cues: Practical Data Visualization**, 1994 propose a list of actions:
- **identify**: to recognize an object based on the characteristics presented, such as finding a fracture in an x-ray;
- **locate**: to establish the position of an object;
- **distinguish**: to determine that an object is distinct or different from another;
- **categorize**: to classify objects into distinct types, such as different land cover or material types;
- **cluster**: to group similar objects based on some relationship. A related action is to segment, which consists of separating dissimilar objects;

User Tasks

- Peter R. Keller and Mary M. Keller. **Visual Cues: Practical Data Visualization**, 1994
propose a list of actions:
- **rank**: to place a group of objects in an order, such as numerical or chronological;

User Tasks

- Peter R. Keller and Mary M. Keller. **Visual Cues: Practical Data Visualization**, 1994
propose a list of actions:
- **rank**: to place a group of objects in an order, such as numerical or chronological;
- **compare**: to examine the similarities and differences between two or more objects where ordering is not possible, such as masking the intersection of two data sets to reveal how they differ;

User Tasks

- Peter R. Keller and Mary M. Keller. **Visual Cues: Practical Data Visualization**, 1994 propose a list of actions:
- **rank**: to place a group of objects in an order, such as numerical or chronological;
- **compare**: to examine the similarities and differences between two or more objects where ordering is not possible, such as masking the intersection of two data sets to reveal how they differ;
- **associate**: to draw a relationship between two or more objects, such as linking temperature and location in weather maps;

User Tasks

- Peter R. Keller and Mary M. Keller. *Visual Cues: Practical Data Visualization*, 1994 propose a list of actions:
- **rank**: to place a group of objects in an order, such as numerical or chronological;
- **compare**: to examine the similarities and differences between two or more objects where ordering is not possible, such as masking the intersection of two data sets to reveal how they differ;
- **associate**: to draw a relationship between two or more objects, such as linking temperature and location in weather maps;
- **correlate**: to find a causal or reciprocal relationship between two or more objects, such as determining the relationship between interest rates and economic growth.

User Tasks - degree of accuracy

User Tasks - degree of accuracy

- In deciding if a particular visualization technique is useful for accomplishing one or more such tasks, one needs to have a clear vision of the **degree of accuracy with which it needs to be accomplished.**

User Tasks - degree of accuracy

- In deciding if a particular visualization technique is useful for accomplishing one or more such tasks, one needs to have a clear vision of the **degree of accuracy with which it needs to be accomplished.**
- Is it essential to correctly identify every occurrence of an object?

User Tasks - degree of accuracy

- In deciding if a particular visualization technique is useful for accomplishing one or more such tasks, one needs to have a clear vision of the **degree of accuracy with which it needs to be accomplished.**
- Is it essential to correctly identify every occurrence of an object?
- What rate of classification or ranking error is acceptable?

User Tasks - degree of accuracy

- In deciding if a particular visualization technique is useful for accomplishing one or more such tasks, one needs to have a clear vision of the **degree of accuracy with which it needs to be accomplished.**
- Is it essential to correctly identify every occurrence of an object?
- What rate of classification or ranking error is acceptable?
- At what resolution is the location of an object needed?

User Tasks - degree of accuracy

- In deciding if a particular visualization technique is useful for accomplishing one or more such tasks, one needs to have a clear vision of the **degree of accuracy with which it needs to be accomplished.**
 - Is it essential to correctly identify every occurrence of an object?
 - What rate of classification or ranking error is acceptable?
 - At what resolution is the location of an object needed?
 - How many falsely labeled correlations are acceptable?

User Tasks - degree of accuracy

- In deciding if a particular visualization technique is useful for accomplishing one or more such tasks, one needs to have a clear vision of the **degree of accuracy with which it needs to be accomplished.**
 - Is it essential to correctly identify every occurrence of an object?
 - What rate of classification or ranking error is acceptable?
 - At what resolution is the location of an object needed?
 - How many falsely labeled correlations are acceptable?
 - For what data characteristics is it important that errors be minimized?

User Characteristics

User Characteristics

- **Users can be classified based on their knowledge and skills:**

User Characteristics

- Users can be classified based on their knowledge and skills:
- **Familiarity with domain.** How much expertise does the user have with the domain of the data being explored?

User Characteristics

- Users can be classified based on their knowledge and skills:
- **Familiarity with domain.** How much expertise does the user have with the domain of the data being explored?
- **Familiarity with task.** How much experience has the user had in performing the desired task? Note that this differs from the previous point, as someone could have significant domain experience, but minimal experience in a particular task.

User Characteristics

- Users can be classified based on their knowledge and skills:
- **Familiarity with domain.** How much expertise does the user have with the domain of the data being explored?
- **Familiarity with task.** How much experience has the user had in performing the desired task? Note that this differs from the previous point, as someone could have significant domain experience, but minimal experience in a particular task.
- **Familiarity with data.** Has the user examined this data previously and formed a reasonable mental model of its contents, or is this her first exposure to it? Is it similar to other data sets she has examined?

User Characteristics

- Users can be classified based on their knowledge and skills:
- **Familiarity with the visualization technique.** Is this the user's first attempt to interpret the data using this particular kind of visualization, or has she spent considerable time using the technique?

User Characteristics

- Users can be classified based on their knowledge and skills:
- **Familiarity with the visualization technique.** Is this the user's first attempt to interpret the data using this particular kind of visualization, or has she spent considerable time using the technique?
- **Familiarity with the visualization environment.** Has the user employed the particular tool (or a similar one) in the past, or is it her first exposure? This differs from the previous factor, as a visualization technique can be implemented in several different packages, and aspects of the packages themselves can influence the effectiveness of the exploration of the data.

User Characteristics

- In an ideal **evaluation using human subjects**, we would want the range of characteristics of the **participants to be as similar as possible** to the intended **audience** for the technique
- The results should be evaluated by grouping people with similar characteristics

Data Characteristics

Data Characteristics

- In an ideal assessment, a **large variety of data sets should be tested**, and they should span the range of characteristics found in the domain for which the visualization is being designed.

Data Characteristics

- In an ideal assessment, a **large variety of data sets should be tested**, and they should span the range of characteristics found in the domain for which the visualization is being designed.
- **Type**. In many cases, the data is all of the same type (e.g., floating-point numbers). Often, however, it is a mixture of types, such as text intermingled with numbers. All combinations of types that can occur should be tested.

Data Characteristics

- In an ideal assessment, a **large variety of data sets should be tested**, and they should span the range of characteristics found in the domain for which the visualization is being designed.
- **Type**. In many cases, the data is all of the same type (e.g., floating-point numbers). Often, however, it is a mixture of types, such as text intermingled with numbers. All combinations of types that can occur should be tested.
- **Size**. Data sets in a particular domain can often take on a wide range of sizes, from a few records to thousands or millions of records. Tests should cover the normal size range, as well as extreme values for the size. 

Data Characteristics

- **Dimensionality.** While in many scientific domains the data generally have fixed dimensionality (one, two, or three), it is generally useful to test a visualization with all possible subsets of dimensions (slices or projections), as well as, with the normal dimensionality.

Data Characteristics

- **Dimensionality.** While in many scientific domains the data generally have fixed dimensionality (one, two, or three), it is generally useful to test a visualization with all possible subsets of dimensions (slices or projections), as well as, with the normal dimensionality.
- **Number of parameters.** The normal number of parameters in a data set varies widely from one domain to another. Visualization techniques should be tested with both univariate and multivariate data sets (unless only univariate data exists); indeed, a common test is to ascertain the maximum number of parameters that can be effectively displayed by a technique.

Data Characteristics

- **Structure.** In a data set, the structure of the data can be simple (e.g., a uniform grid or a table) or complex (e.g., hierarchical). Visualization techniques should be tested using all commonly encountered structures within the domain

Data Characteristics

- **Structure.** In a data set, the structure of the data can be simple (e.g., a uniform grid or a table) or complex (e.g., hierarchical). Visualization techniques should be tested using all commonly encountered structures within the domain
- **Range.** Objects in data sets can take on a wide range of values. Testing should involve exercising the entire range of possible values, including the extremes of the range.



Data Characteristics

- **Structure.** In a data set, the structure of the data can be simple (e.g., a uniform grid or a table) or complex (e.g., hierarchical). Visualization techniques should be tested using all commonly encountered structures within the domain
- **Range.** Objects in data sets can take on a wide range of values. Testing should involve exercising the entire range of possible values, including the extremes of the range. ★
- **Distribution.** Data can be uniformly or non-uniformly distributed, both in values and in attributes (such as spatio-temporal position). Tests should include some extreme cases. ★

Data Characteristics

- **Structure.** In a data set, the structure of the data can be simple (e.g., a uniform grid or a table) or complex (e.g., hierarchical). Visualization techniques should be tested using all commonly encountered structures within the domain
- **Range.** Objects in data sets can take on a wide range of values. Testing should involve exercising the entire range of possible values, including the extremes of the range. 
- **Distribution.** Data can be uniformly or non-uniformly distributed, both in values and in attributes (such as spatio-temporal position). Tests should include some extreme cases. 
- **Real vs. Synthetic.** For some types of evaluation (e.g., size tests) synthetic data is fine, but in general it is more convincing to use real data.

Data Characteristics

- Few domains have made concerted efforts to gather large numbers of data sets for testing:
 - The machine learning archive at the University of California
 - StatLib repository at CMU
- Some Links:
 - <http://lib.stat.cmu.edu/datasets/>
 - <https://github.com/caesar0301/awesome-public-datasets>
 - <https://aws.amazon.com/public-datasets/>
 - <https://archive.org/details/datasets>

Visualization Characteristics

Visualization Characteristics

- **Aspects of the visualization that can be evaluated:**

Visualization Characteristics

- Aspects of the visualization that can be evaluated:
- **Computational performance.** How quickly can the visualization be generated, using data sets of various sizes?

Visualization Characteristics

- Aspects of the visualization that can be evaluated:
- **Computational performance.** How quickly can the visualization be generated, using data sets of various sizes?
- **Memory performance.** How much computer memory is required to generate the visualization?

Visualization Characteristics

- Aspects of the visualization that can be evaluated:
- **Computational performance.** How quickly can the visualization be generated, using data sets of various sizes?
- **Memory performance.** How much computer memory is required to generate the visualization?
- **Data limitations.** What are the upper and lower bounds for the size and complexity of the data that can be visualized with this technique? At what point does the amount of information extractable from the visualization stabilize or decrease with increased data size/complexity? At what point does the error rate for performing the task increase to an unacceptable level?

Visualization Characteristics

- **Degree of occlusion.** What is the likelihood that some subset of the data to be displayed will be occluded by other parts of the visualization? How much is normally occluded?

Visualization Characteristics

- **Degree of occlusion.** What is the likelihood that some subset of the data to be displayed will be occluded by other parts of the visualization? How much is normally occluded?
- **Degree of complexity.** What is the normal learning curve for the technique? How many parameters does the user need to set in order to generate views? How much knowledge is needed to set and adjust these parameters effectively?

Visualization Characteristics

- **Degree of occlusion.** What is the likelihood that some subset of the data to be displayed will be occluded by other parts of the visualization? How much is normally occluded?
- **Degree of complexity.** What is the normal learning curve for the technique? How many parameters does the user need to set in order to generate views? How much knowledge is needed to set and adjust these parameters effectively?
- **Degree of usability.** How easy is it to perform the task? How intuitive is the interpretation of the visualization? How intuitive are the controls for interactions?

Visualization Characteristics

- **Degree of occlusion.** What is the likelihood that some subset of the data to be displayed will be occluded by other parts of the visualization? How much is normally occluded?
- **Degree of complexity.** What is the normal learning curve for the technique? How many parameters does the user need to set in order to generate views? How much knowledge is needed to set and adjust these parameters effectively?
- **Degree of usability.** How easy is it to perform the task? How intuitive is the interpretation of the visualization? How intuitive are the controls for interactions?
- **Degree of accuracy.** How frequently is the user successful or unsuccessful in performing the desired task with this technique? Under what conditions are errors made, and how bad are the errors (i.e., distance from correctness)?

Structures for Evaluating Visualizations

Structures for Evaluating Visualizations

- Forms of evaluation for interactive systems have been developed within the field of **human-computer interaction**, applicable to data and information visualization:

Structures for Evaluating Visualizations

- Forms of evaluation for interactive systems have been developed within the field of **human-computer interaction**, applicable to data and information visualization:
- **Usability tests**. These evaluations concentrate on “the Five E’s”: **effective**, **efficient**, **engaging**, **error tolerant**, and **easy to learn** [398].

Structures for Evaluating Visualizations

- Forms of evaluation for interactive systems have been developed within the field of **human-computer interaction**, applicable to data and information visualization:
- **Usability tests**. These evaluations concentrate on “the Five E’s”: **effective**, **efficient**, **engaging**, **error tolerant**, and **easy to learn** [398].
- Observing users attempting to perform tasks, and noting the types of **difficulties they are having**, the **features they commonly use**, and their level of **comfort/satisfaction** with the tool.

Structures for Evaluating Visualizations

- Forms of evaluation for interactive systems have been developed within the field of **human-computer interaction**, applicable to data and information visualization:
- **Usability tests**. These evaluations concentrate on “the Five E’s”: **effective**, **efficient**, **engaging**, **error tolerant**, and **easy to learn** [398].
- Observing users attempting to perform tasks, and noting the types of **difficulties they are having**, the **features they commonly use**, and their level of **comfort/satisfaction** with the tool.
- Often, the evaluation **starts with some usability goals** or requirements, and the result of the evaluation indicates whether the goal or requirement has been met or not.

Structures for Evaluating Visualizations

- **Field tests.** Unlike usability tests, which are often carried out in a controlled environment over a short period of time to better enable measurement, field tests are performed in the natural environment of the typical user and may last for weeks or months.

Structures for Evaluating Visualizations

- **Field tests.** Unlike usability tests, which are often carried out in a controlled environment over a short period of time to better enable measurement, field tests are performed in the natural environment of the typical user and may last for weeks or months.
- **Field tests** attempt to assess the degree to which the new technique or tool becomes an integral part of a user's activities once the initial curiosity and learning curve have been overcome.

Structures for Evaluating Visualizations

- **Field tests.** Unlike usability tests, which are often carried out in a controlled environment over a short period of time to better enable measurement, field tests are performed in the natural environment of the typical user and may last for weeks or months.
- **Field tests** attempt to assess the degree to which the new technique or tool becomes an integral part of a user's activities once the initial curiosity and learning curve have been overcome.
- The results of field tests are often qualitative, and may change significantly over time. They can be effective, especially if users are encouraged to submit questions and critiques that can lead to clarifications and improvements in functionality.

Structures for Evaluating Visualizations

- **Case studies and use cases**

- **Rather than using experts, users, or student volunteers in evaluation, some visualization researchers attempt to validate the effectiveness of their techniques by showing real (or sometimes contrived) examples of how their method can be used in solving a particular problem or performing a given task.**
- **The key to this sort of evaluation is to ensure that the case studies are sufficiently realistic so that someone with a particular task to perform can be convinced that at least one of the case studies is**

Benchmarking Procedures

Benchmarking Procedures

- Benchmarking is a **formal procedure** for **evaluating the performance** of some object or set of objects:

Benchmarking Procedures

- Benchmarking is a **formal procedure** for **evaluating the performance** of some object or set of objects:
 - **Evaluating** the speed of a car or user-friendliness of an interface.

Benchmarking Procedures

- Benchmarking is a **formal procedure** for **evaluating the performance** of some object or set of objects:
 - **Evaluating** the speed of a car or user-friendliness of an interface.
 - Benchmarks can be **quantitative** (resulting in a number) or **qualitative** (resulting in a relative judgment).

Benchmarking Procedures

- Benchmarking is a **formal procedure** for **evaluating the performance** of some object or set of objects:
 - **Evaluating** the speed of a car or user-friendliness of an interface.
 - Benchmarks can be **quantitative** (resulting in a number) or **qualitative** (resulting in a relative judgment).
 - In general, **quantitative benchmarking experiments are easier to set up and execute** than qualitative experiments.

Benchmarking Procedures

- Benchmarking is a **formal procedure** for **evaluating the performance** of some object or set of objects:
 - **Evaluating** the speed of a car or user-friendliness of an interface.
 - Benchmarks can be **quantitative** (resulting in a number) or **qualitative** (resulting in a relative judgment).
 - In general, **quantitative benchmarking experiments are easier to set up and execute** than qualitative experiments.
- **Steps:**
 - ◆ Formulate a hypothesis
 - ◆ Design the experiments
 - ◆ Execute the experiments
 - ◆ Analyze the results and validate the hypothesis

Benchmarking Procedures

- **Formulate a hypothesis**

- **A specific statement about one or more attributes of the object being assessed, such as “System A allows novice users to more easily identify clusters in data sets containing 5 – 10 dimensions and 1000 to 10000 data points than system B.”**

Benchmarking Procedures

- **Formulate a hypothesis**
 - A specific statement about one or more attributes of the object being assessed, such as “System A allows novice users to more easily identify clusters in data sets containing 5 – 10 dimensions and 1000 to 10000 data points than system B.”
 - Care must be taken in **designing a hypothesis that isn’t too general**, stating as a hypothesis that system A is “better” than system B is very difficult to prove or disprove.

Benchmarking Procedures

- **Formulate a hypothesis**

- A specific statement about one or more attributes of the object being assessed, such as “System A allows novice users to more easily identify clusters in data sets containing 5 – 10 dimensions and 1000 to 10000 data points than system B.”
- Care must be taken in **designing a hypothesis that isn’t too general**, stating as a hypothesis that system A is “better” than system B is very difficult to prove or disprove.
- It is a good idea if the **hypothesis includes information about task, user, and data.**

Benchmarking Procedures

- **Design the experiments**
 - **Create tests that vary only a **single attribute** at a time.**

Benchmarking Procedures

- **Design the experiments**
 - **Create tests that vary only a **single attribute** at a time.**
 - **Measuring the usability of a tool would require the subjects to perform identical tasks that exercise only the usability issue, and not peripheral issues such as the color scheme or the hardware speed.**

Benchmarking Procedures

- **Execute the experiments**

Benchmarking Procedures

- **Execute the experiments**
 - **Factors such as the **training of human subjects are critical****
 - **Each participant should get a similar amount of instruction on performing the experiments, including identifying the procedure they should follow, the expected format of their responses, and the pace at which they should proceed.**

Benchmarking Procedures

- **Execute the experiments**
 - **Factors such as the training of human subjects are critical**
 - **Each participant should get a similar amount of instruction on performing the experiments, including identifying the procedure they should follow, the expected format of their responses, and the pace at which they should proceed.**
 - **The amount of time provided should be reasonably constrained, as users often will respond differently, given differing time constraints**

Benchmarking Procedures

- **Execute the experiments**
 - **Factors such as the training of human subjects are critical**
 - **Each participant should get a similar amount of instruction on performing the experiments, including identifying the procedure they should follow, the expected format of their responses, and the pace at which they should proceed.**
 - **The amount of time provided should be reasonably constrained, as users often will respond differently, given differing time constraints**
 - **Audience should be of sufficient size to make the results statistically significant.** For testing a single attribute of a visualization, at least 15–20 subjects with similar backgrounds.

Benchmarking Procedures

- **Analyze the results and validate the hypothesis.**

Benchmarking Procedures

- **Analyze the results and validate the hypothesis.**
 - **It is then necessary to ascertain whether**
 - **the hypothesis is supported,**
 - **the hypothesis is refuted, or**
 - **there is insufficient evidence to support or refute the hypothesis.**

Benchmarking Procedures

- **Analyze the results and validate the hypothesis.**
 - It is then necessary to ascertain whether
 - the hypothesis is supported,
 - the hypothesis is refuted, or
 - there is insufficient evidence to support or refute the hypothesis.
- Generally, one is looking for **results that are statistically significant**, e.g., they are far enough from random to indicate credibility.

Benchmarking Procedures

- **Analyze the results and validate the hypothesis.**
 - It is then necessary to ascertain whether
 - the hypothesis is supported,
 - the hypothesis is refuted, or
 - there is insufficient evidence to support or refute the hypothesis.
- Generally, one is looking for **results that are statistically significant**, e.g., they are far enough from random to indicate credibility.
- For quantitative variables of the experiments, such as data set size, it is useful to plot results against the variable.

An Example of Benchmarking

- Using social semantic knowledge to improve annotations in personal photo collections, PhD from Nuno Datia, 2015
- The Logical Day Event Segmentation (LDES) algorithm uses the creation time of photos to fulfill a base segmentation. The spatial information is used to fine-tune the limits of each segment.
- The LDES was tested by a set of volunteers.
- **The research statement of the study is that users accept the segmentations suggested by LDES, with minor changes**

The LDES was tested by a set of volunteers. The research statement of the study is that users accept the segmentations suggested by LDES, with minor changes.

The empirical study involved 14 participants who have provided us with some of their rolls (35 in total). There was a balance in terms of gender, with eight males and six females. Their age ranged from 21 to 55 years, covering a wide range of ages (see Figure 6.9). Most of them are computer science students. We provided them with some guidelines to help the roll selection, namely:

1. the photos should have location information;
2. the temporal range of each roll should have more than 1 day;
3. the rolls should reflect real sequences of photos, without a pre-selection.

The participants were responsible for selecting the rolls. For the record, we used all of the provided sets, even though some do not fulfil all three guidelines.

6.1.2.2 Test Design

The experimental unit in this empirical user test is the pair $\langle participant, roll \rangle$. There are some participants that share some rolls. Two rolls are shared between two participants each, and four rolls are shared between two participants. Those rolls represent situations where the participants are simultaneously the photographer and the subject. The segmentations are presented independently to each user. Thus, the 14 participants, interacting with 35 different rolls, makes the 43 experimental units in the test.

Figure 6.12 describes the flow used in the experimental user tests, made using a web application. It includes:

1. a *questionnaire*;
2. a *training phase* and;
3. the *test* itself.

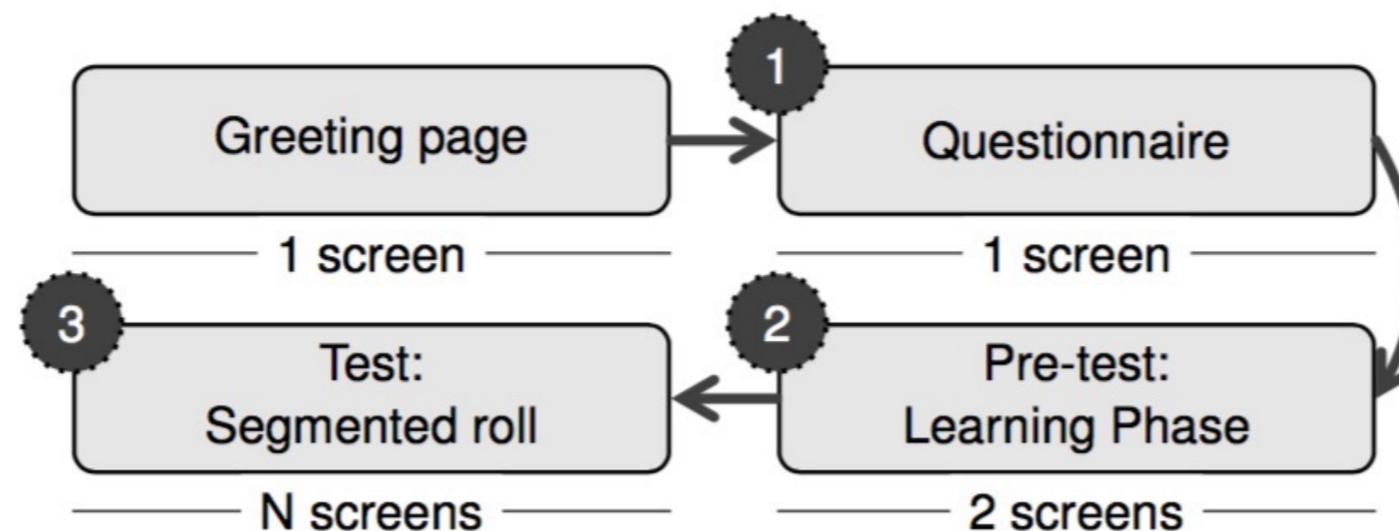


Figure 6.12: Representation of the test flow.

Step 2: The training phase After the questionnaire, the participant passes to a *pre-test learning phase*. The goal is to let the participant learn how photos are presented in the segments and how to change the segmentation, permitting an exploratory interaction. The learning phase has two steps, that share a similar layout:

1. a first one, introduces some simple terms and calls the participant attention to the way the segmentation is presented, showing the basic interaction, and
2. a second one, where the participant can freely interact with the user interface (UI), changing the segmentation at will. This includes the creation of new segments, and changing photos from one segment to another.

A participant can repeat the learning steps, as it is possible to go back and forth thorough them. Figure 6.13 depicts the first step of the learning phase. The left-hand side, marked as **a**, displays a representation of a segmentation. The tooltips indicate the locations of the labels, what a segment is, and identifies the contents of a segments — photos. The photos are labelled with letters, so the order can be checked, after the participant changes

Training phase

ROLL NAME: STEP1

Mon 1 Apr, 2013 12:00 - 12:03

A B C D

Segment
The light gray rectangles represent segments of pictures!

Label
The white label describes a segment, using temporal descriptors!

Tue 2 Apr, 2013 17:00 - 19:00

E F G

Pictures
Inside the segments are pictures, that form a coherent sequence to you.

INSTRUCTIONS

1. Pay attention to the tooltips that are now visible.
2. Close the tooltips.
3. Drag picture B to segment below. Watch what happens to pictures C and D. Notice that a segment is a *continuous time ordered set* of pictures.
4. Now see that a dark gray label appears above the lower segment. Every time you click a picture, that label shows detailed information about time and space (if available).
5. Try selecting another picture by **clicking** on it.

next skip instructions

a b

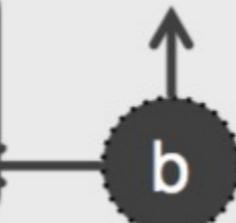
Figure 6.13: Interface for the learning phase.

Thu 25 Dec, 2008 17:49 - 18:09

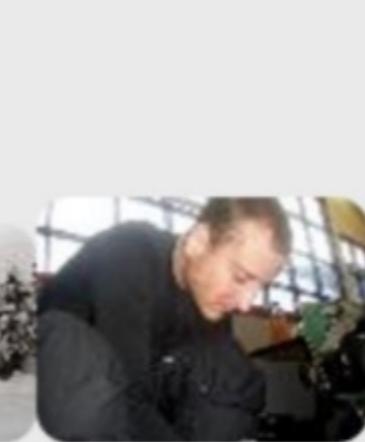


Fri 26 Dec, 2008 09:57 - 10:23

Fri 26 Dec, 2008 @ 10:23:19, probably at Anchorage Municipality



Fri 26 Dec, 2008 14:44 – Sat 27 Dec, 2008 12:33



Sat 27 Dec, 2008 16:11 - 18:36



6.1.5 Experiment analysis

As stated earlier, the research statement that we want to assess in this empirical study is that users will accept the temporal coherent segmentation provided by LDES, with none or minor changes. During the empirical study, we collected several data, namely:

1. the participant's modified segmentation;
2. the stream of actions made on the segmentation, by the participant;
3. the perceived quality of the segmentation.

Experiment Analysis

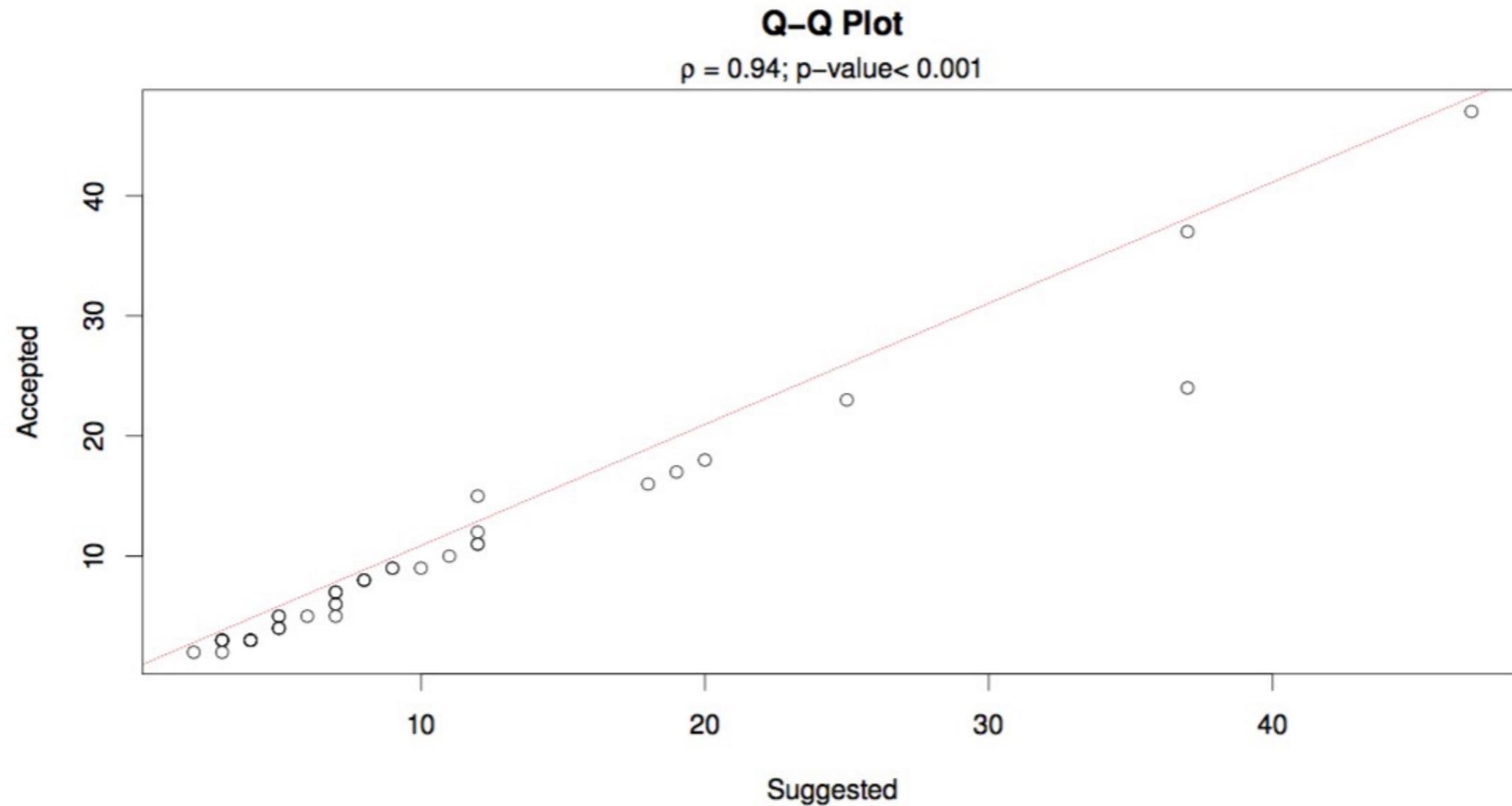


Figure 6.20: Quantile-Quantile plot for the no. of segments for the suggested segmentations and user-modified segmentations.

Another Example of Benchmarking

Another Example of Benchmarking

- A set of experiments for assessing.
 - ◆ The strengths and weaknesses of three multivariate visualization techniques, namely **scatterplot matrices**, **parallel coordinates**, and **star glyphs**,
 - ◆ Two distinct tasks: **cluster analysis** and **outlier detection**.

Another Example of Benchmarking

- **Outlier Detection and Measurement Experiments**
 - ◆ **Stage 1.** develop a quantifiable definition for an outlier and create an algorithm capable of labeling data points appropriately

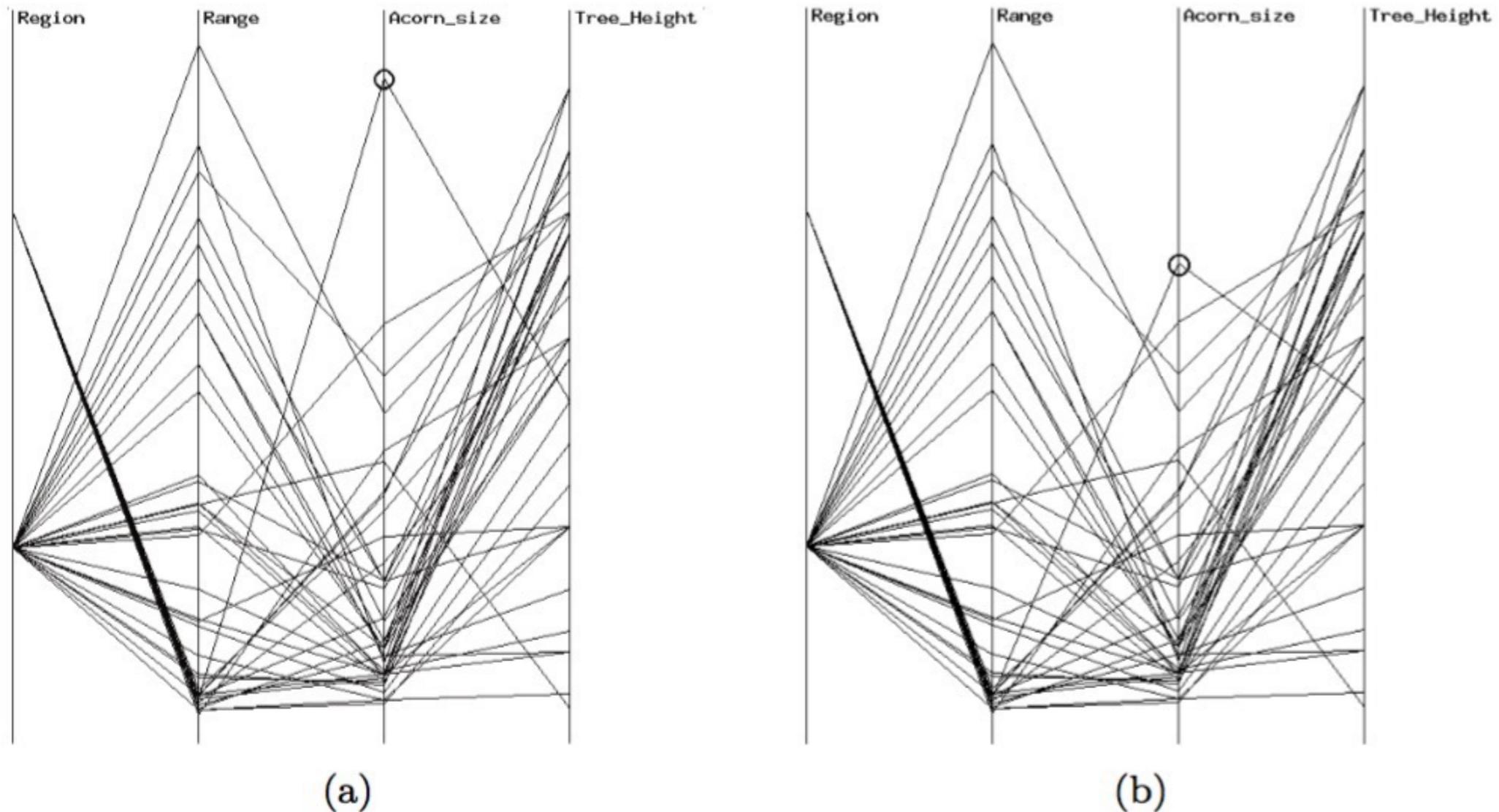
Another Example of Benchmarking

- **Outlier Detection and Measurement Experiments**
 - ◆ **Stage 1.** develop a quantifiable definition for an outlier and create an algorithm capable of labeling data points appropriately
 - ◆ **Stage 2.** Data sets, both real and simulated, were then acquired or generated that contained outliers, according to the definition in Stage 1. For some data sets, the degree of separation between the outliers and the main bodies of data points were varied to test subject sensitivity

Another Example of Benchmarking

- **Outlier Detection and Measurement Experiments**
 - ◆ **Stage 1.** develop a quantifiable definition for an outlier and create an algorithm capable of labeling data points appropriately
 - ◆ **Stage 2.** Data sets, both real and simulated, were then acquired or generated that contained outliers, according to the definition in Stage 1. For some data sets, the degree of separation between the outliers and the main bodies of data points were varied to test subject sensitivity
 - ◆ A total of **six outlier experiments** were designed, and **each was repeated three times** for the **three visualization techniques** being tested.

Another Example of Benchmarking



Parallel coordinates view of data set describing acorn attributes, with a single outlier (circled, in the acorn size dimension) (a) in its original position and (b) with the distance artificially shortened [442]. (Image © 1997 IEEE.)

Another Example of Benchmarking

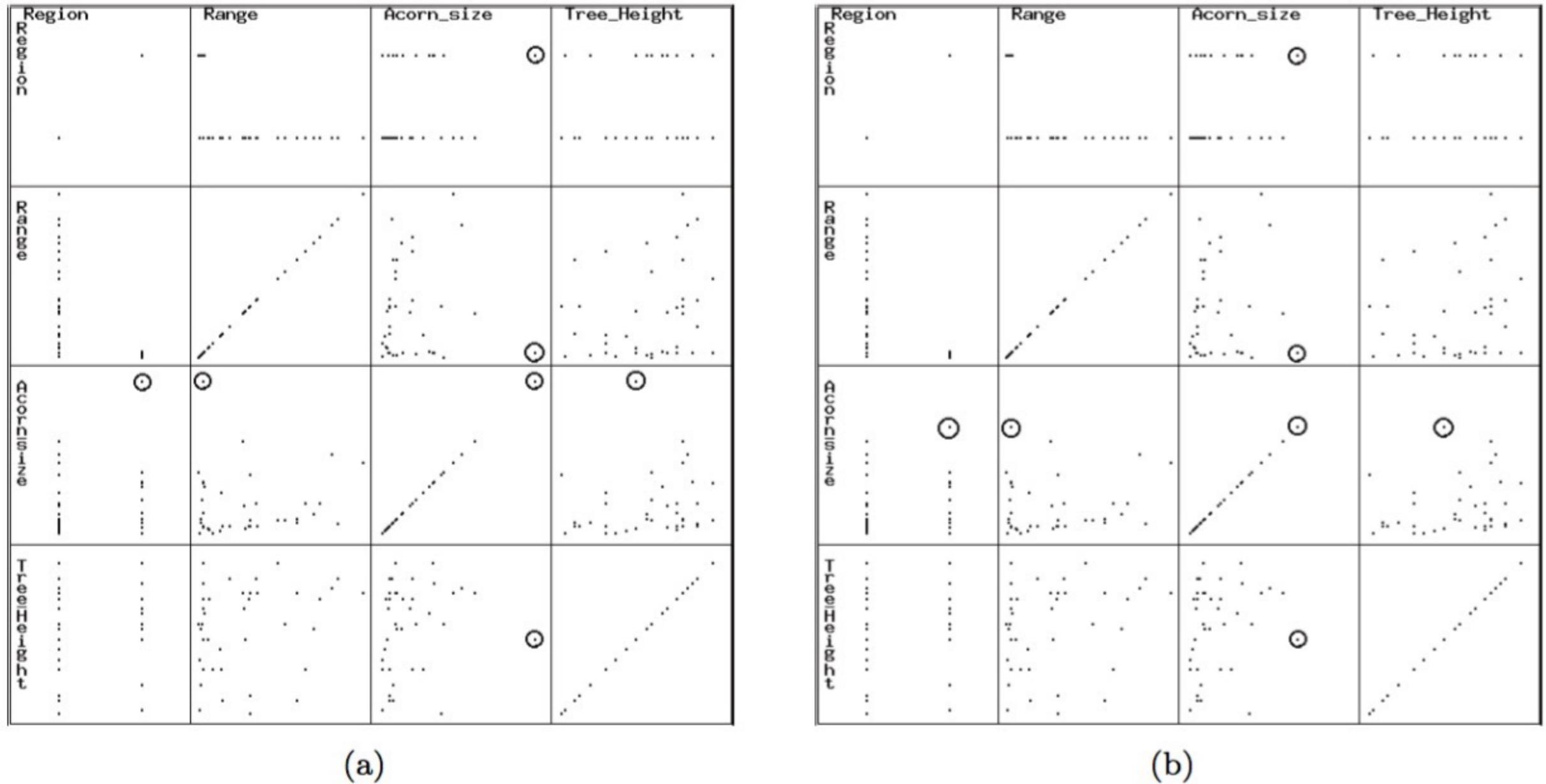


Figure 14.2. Identifying outliers with scatterplot matrices: same data as previous figure, using scatterplot matrices [442]. (Image © 1997 IEEE.)

Another Example of Benchmarking

■ **Stage3.**

- ◆ **The subjects (19 computer science graduate students with minimal exposure to data visualization) were trained to interpret the visualization technique(s) to be assessed and given examples of data sets with identified outliers(approximately one hour of training).**

Another Example of Benchmarking

■ Stage3.

- ◆ The subjects (19 computer science graduate students with minimal exposure to data visualization) were trained to interpret the visualization technique(s) to be assessed and given examples of data sets with identified outliers (approximately one hour of training).
- ◆ They were shown a set of **18 images of data sets containing between 0 and 6 outliers**. The subjects were asked not to spend more than one minute per image. The tasks given to them were:

Another Example of Benchmarking

■ Stage3.

- ◆ The subjects (19 computer science graduate students with minimal exposure to data visualization) were trained to interpret the visualization technique(s) to be assessed and given examples of data sets with identified outliers (approximately one hour of training).
- ◆ They were shown a set of **18 images of data sets containing between 0 and 6 outliers**. The subjects were asked not to spend more than one minute per image. The tasks given to them were:
 - **Determine** if an image contains one or more outliers.

Another Example of Benchmarking

■ Stage3.

- ◆ The subjects (19 computer science graduate students with minimal exposure to data visualization) were trained to interpret the visualization technique(s) to be assessed and given examples of data sets with identified outliers (approximately one hour of training).
- ◆ They were shown a set of **18 images of data sets containing between 0 and 6 outliers**. The subjects were asked not to spend more than one minute per image. The tasks given to them were:
 - **Determine** if an image contains one or more outliers.
 - **Identify** the points believed to be outliers.

Another Example of Benchmarking

■ Stage3.

- ◆ The subjects (19 computer science graduate students with minimal exposure to data visualization) were trained to interpret the visualization technique(s) to be assessed and given examples of data sets with identified outliers (approximately one hour of training).
- ◆ They were shown a set of **18 images of data sets containing between 0 and 6 outliers**. The subjects were asked not to spend more than one minute per image. The tasks given to them were:
 - **Determine** if an image contains one or more outliers.
 - **Identify** the points believed to be outliers.
 - **Estimate** the degree of separation of each outlier on a 5-point scale.

Another Example of Benchmarking

■ Stage 4.

- ◆ Given the subject responses, the **usefulness of each visualization method tested was assessed** in terms of outlier detection and measurement across data sets with different characteristics.
- ◆ The **percentages of correctly and incorrectly detected outliers** were tallied.
- ◆ The **average error in estimating the degree of separation**.

Another Example of Benchmarking

- See the Cluster Detection and Measurement Experiments on page 442
- For both tasks,
 - ◆ scatterplot matrices generally fared best,
 - ◆ followed by star glyphs positioned by principal component analysis,
 - ◆ and lastly by parallel coordinates

One more example for Data VIS

Visualização de dados para auxiliar a defesa da floresta contra incêndios

- Manuel de la Cueva -

LINK

"Visualização de dados para auxiliar a defesa da floresta contra incêndios"

Manuel de La Cueva Couto Henriques

Contexto

Esta tese de mestrado tem como objectivo criar visualizações interactivas de dados que apresentem informação sobre incêndios florestais. Esta informação pode estar relacionada com as previsões meteorológicas, elevação do terreno, tipo de vegetação, o perímetro do incêndio e até a sua gravidade. Estes dados são muito importantes para compreender o que está a ocorrer no terreno e tomar decisões mais conscientes.

Para testar a qualidade destas visualizações foi criado este website. Aqui será submetido a uma avaliação experimental que consiste num conjunto de tarefas que precisará de responder recorrendo a diferentes tipos de visualizações. Serão registados os tempos e respostas dadas por cada tarefa.

Não deve voltar atrás ou actualizar as páginas deste site, isso introduzirá erros na plataforma e terá de começar do início. Por esse motivo, se tiver alguma dúvida ou problema, poderá contactar-me directamente através do meu número de telemóvel: **969476629**.

Sugiro a utilização do **Google Chrome**, pois este site apenas foi testado para esse browser.

Por favor preencha todos os campos abaixo apresentados. Quando estiver preparado(a) para começar a avaliação experimental clique no botão "Começar".

O nome servirá apenas para o contactar no caso de dúvida, para um questionário ou outra eventualidade que possa surgir relacionada com esta avaliação.

Nome

Idade

Sexo

Feminino

Masculino

Tem alguma experiência em combate a incêndios?

Sim

Não

Tem alguma experiência com o MacFire?

Sim

Não

Tem alguma experiência com SIG (Sistemas de Informação Geográfica)?

Sim

Não

Tem alguma experiência com Visualizações Interativas de Dados?

Sim

Não

Começar

Conceitos básicos de combate a incêndios florestais

Como seleccionou que não tinha experiência em combate a incêndios florestais, esta página contém alguns conceitos básicos que deve conhecer para completar as tarefas da avaliação experimental. Por favor, leia com atenção os seguintes tópicos e não hesite em colocar alguma questão.

Frentes de fogo

As frentes de fogo são a frente principal de propagação do fogo. De todo o perímetro do fogo, são as zonas mais intensas e a que se propaga mais rapidamente de todo o incêndio.

Geralmente, a direcção de propagação é igual à direcção do vento.

Flancos do incêndio

Os flancos do incêndio são as laterais do incêndio. São maiores em extensão do que as frentes de fogo mas menos intensas.

Quando o vento muda de direcção os flancos podem passar a ser frentes de fogo, sendo uma situação bastante perigosa. Quando isto acontece, estas novas frentes de fogo são bastante extensas e devastam grandes quantidades de terreno em pouco tempo.



Meteorologia

As condições meteorológicas têm um papel importantíssimo no comportamento do fogo. Geralmente, influenciam o incêndio da seguinte maneira:

A **direcção do vento** define a direcção de propagação do incêndio.

A **velocidade do vento** define a velocidade da propagação do incêndio. A situação fica bastante grave quando a velocidade atinge valores superiores a **30 Km/h**.

A **humidade relativa do solo** define o estado de secura da vegetação. Quando mais seca a vegetação é, mais suscetível é ao aparecimento e à propagação de incêndios. A ocorrência do incêndio é bastante grave quando a humidade atinge valores inferiores a **30%**.

A **temperatura** está directamente relacionada com a humidade relativa. Durante o dia a temperatura sobe e a humidade diminui, acontecendo o oposto à noite. A gravidade do incêndio é bastante acentuada quando a temperatura do vento atinge valores superiores a **30° celsius**, sendo por isso o Verão a estação com maior número de fogos florestais.

Topografia

A elevação do terreno afecta o combate a incêndios florestais de duas maneiras possíveis:

Pode impedir o acesso de veículos terrestres ou de pessoas a determinadas zonas, devido a inclinações muito acentuadas.

Define a velocidade de propagação do incêndio. O fogo propaga-se mais rapidamente a subir uma encosta, pois consegue apanhar com facilidade as folhas das árvores mais à frente. Contrariamente, o fogo propaga-se mais devagar a descer encostas pois tem mais dificuldade em chegar às folhas das árvores seguintes.



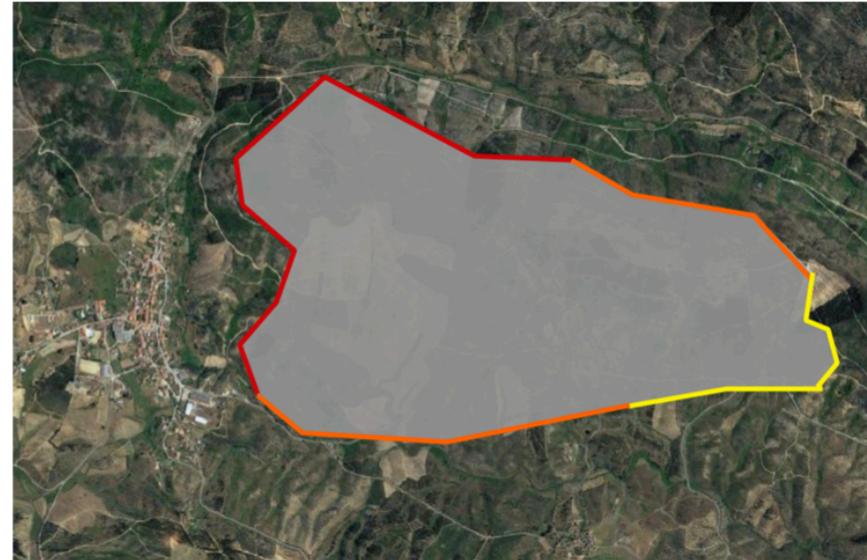
Conceitos básicos de Macfire

Como seleccionou que não tinha experiência com MacFire, esta página contém alguns conceitos básicos que deve conhecer para completar as tarefas da avaliação experimental. Por favor, leia com atenção os seguintes tópicos e não hesite em colocar alguma questão.

Mapas digitais

O Macfire é um sistema informático que permite visualizar informação do terreno directamente num mapa. Estes mapas podem ser imagens de satélite (por exemplo, o Google Maps) ou cartografia mais complexa com informação topográfica ou de utilização do terreno.

Uma das informações mais importantes nestes mapas é a indicação da localização de frentes de fogo, flancos e área ardida.



Estado do incêndio

O estado do incêndio diz respeito à localização de frentes de fogo, flancos e área ardida. Esta informação é adicionada no MacFire à mão por agentes situados no terreno. Esta informação tem uma simbologia específica.

A **área ardida** é representada através de **polígonos** e tem uma cor **cinzenta**.

As **frentes de fogo** e os **flancos** são representadas através de **linhas**. Estas linhas podem ter 3 cores diferentes:



Linhas **vermelhas** são frentes de fogo **muito activas**. São extremamente difíceis de apagar e geralmente não são combatidas directamente com recurso a água.



Linhas **laranjas** são frentes de fogo **activas**. São frentes difíceis de apagar mas possibilitam o combate directo.



Linhas **amarelas** são frentes de fogo **pouco activas**. São fáceis de apagar.

Continuar

Tarefa 0

Esta tarefa serve apenas para compreender o processo de avaliação. Não será registado o tempo ou a resposta dada.

Questão:

Após clicar no botão "Mostrar Visualização", indique qual o número que aparece na janela de visualização.

Próxima Pergunta

Antes de Começar 1/5

Tarefa 24

Encontra-se no dia 22 de Julho de 2019. Certifique-se que procura informação para a resposta com referência a este dia.

A primeira coisa a fazer é olhar para o canto superior esquerdo do ecrã. Aqui terá a contextualização da tarefa para o guiar na resposta e nas visualizações disponibilizadas.

Próximo

botão.

Tarefa 0

Esta tarefa serve apenas para compreender o processo de avaliação. Não será registado o tempo ou a resposta dada.

Questão:

Após clicar no botão "Mostrar Visualização", indique qual o número que aparece na janela de visualização.

Após ler a tarefa e a questão, clique neste botão.

Mostrar Visualização

Próxima Pergunta

Tarefa 1/47

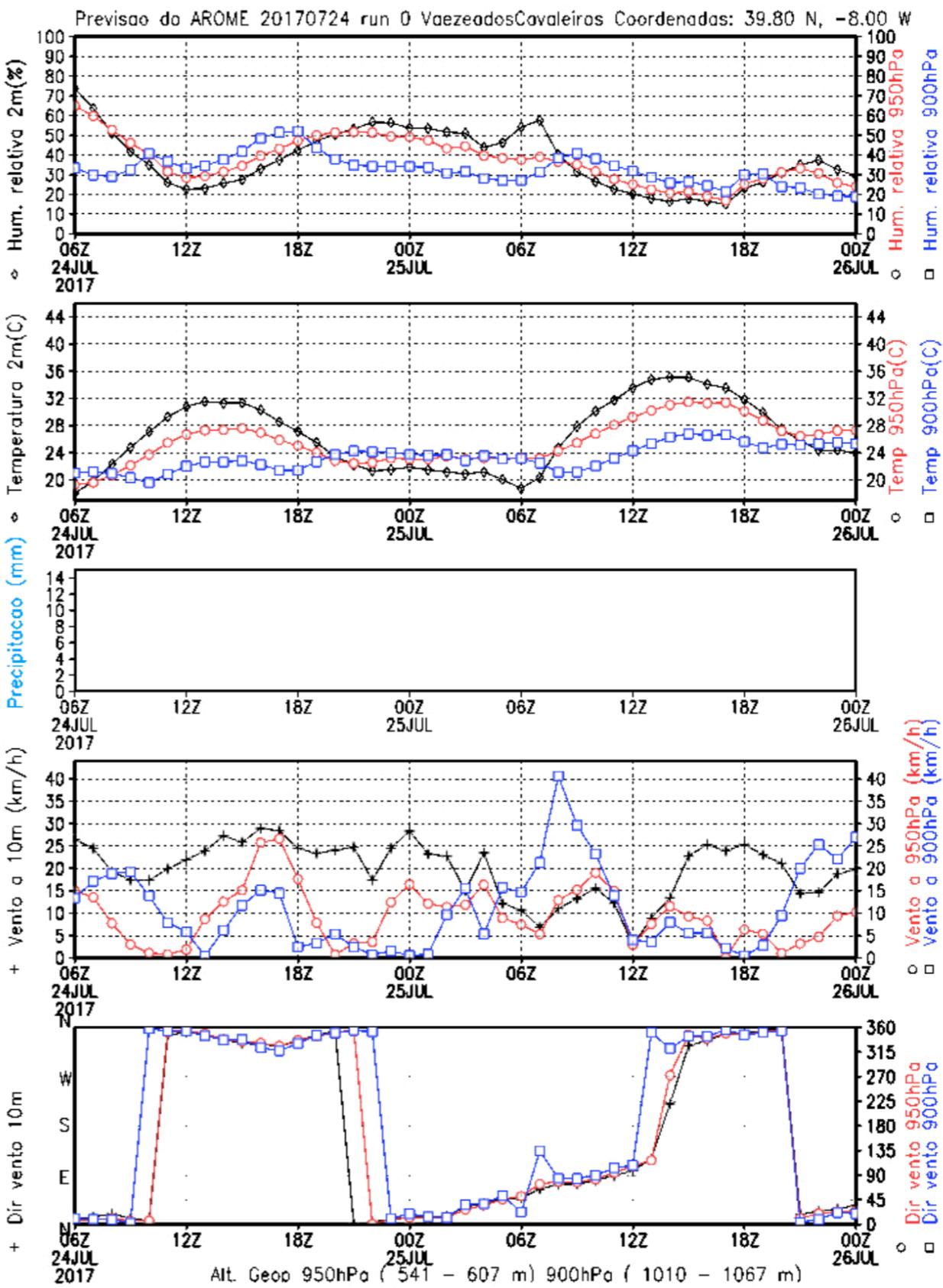
Encontra-se no dia **25 de Julho de 2017**. Certifique-se que procura informação para as respostas com referência a este dia.

A visualização do lado direito utiliza 3 linhas com cores diferentes. Utilize a **linha preta** para responder às perguntas.

Questão:

Qual a velocidade do vento máxima entre as 12 e as 18 horas?

Próxima Pergunta



Tarefa 5/47

Encontra-se no dia 22 de Maio de 2020.

Contudo, procure informação para o dia 24 de Maio de 2020. Pode interagir com a visualização para escolher o dia.

Questão:

Qual a temperatura máxima entre as 12 e as 18 horas?

Próxima Pergunta



[Quem somos](#) | [Onde estamos](#) | [Conselho científico](#) | [Recrutamento](#) | [Projetos](#) | [eBiblioteca](#) | [Enciclopédia ipma](#) | [Contactos](#) | [English](#)

Tempo
Mar
Sismos
Clima
Bivalves
Pescas e aquacultura
Aeronáutica
Agricultura
Navios
Espaço
Fogos Rurais

[IPMA > Tempo > Previsão localidade horária](#)

Selecione Distrito/Ilha:
Selecione a Cidade/Vila:

Lisboa, Lisboa

22 Sex	23 Sáb	24 Dom	25 Seg	26 Ter	27 Qua	28 Qui	29 Sex	30 Sáb
17° 27°	18° 29°	18° 28°	17° 30°	18° 32°	20° 33°	20° 33°	20° 32°	19° 30°
N	N	N	N	NW	E	NW	NW	NW
0%	0%	0%	0%	0%	0%	0%	0%	0%
UV 10	UV 10	UV 9						

22 Sex

8H	9H	10H	11H	12H	13H	14H	15H	16H	17H	18H	19H	20H	21H	22H	23H								
Probabilidade de ocorrência de precipitação, igual ou superior a 0.3 mm, nas últimas 3 horas																							
0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%								
Temperatura / Temperatura sentida																							
19°	21°	23°	24°	25°	26°	27°	26°	25°	24°	23°	22°	20°	19°	20°	19°	19°	19°	19°	19°	18°	19°	19°	
Vento (Intensidade: quilómetros por hora. Direcção: classes de rumo, N, NE, E, SE, S, SW, W, NW)																							
10	11	12	14	15	16	17	20	23	24	22	16	16	18	16	16	13	14	14	16	15	14	14	16
N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Estado do mar (Mar Total: metros, direcção da onda e temp. da água)																							
0.9	0.9	0.9	0.9	1.0	1.0	1.1	1.2	1.2	1.3	1.4	1.4	1.5	1.6	1.7	1.7	1.8	1.8	1.8	1.8	1.8	1.8	1.9	1.9
W	W	W	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW
17°	17°	17°	17°	18°	18°	18°	18°	18°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°

23 Sáb

0H	1H	2H	3H	4H	5H	6H	7H																
Probabilidade de ocorrência de precipitação, igual ou superior a 0.3 mm, nas últimas 3 horas																							
0%	0%	0%	0%	0%	0%	0%	0%																
Temperatura / Temperatura sentida																							
19°	21°	23°	24°	25°	26°	27°	26°	25°	24°	23°	22°	20°	19°	20°	19°	19°	19°	19°	18°	19°	19°		
Vento (Intensidade: quilómetros por hora. Direcção: classes de rumo, N, NE, E, SE, S, SW, W, NW)																							
10	11	12	14	15	16	17	20	23	24	22	16	16	18	16	16	13	14	14	16	15	14	14	16
N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Estado do mar (Mar Total: metros, direcção da onda e temp. da água)																							
0.9	0.9	0.9	0.9	1.0	1.0	1.1	1.2	1.2	1.3	1.4	1.4	1.5	1.6	1.7	1.7	1.8	1.8	1.8	1.8	1.8	1.9	1.9	
W	W	W	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW	NW
17°	17°	17°	17°	18°	18°	18°	18°	18°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°	17°

Avisos Meteorológicos

Para mais informações de avisos meteorológicos consulte [aqui](#).

Informação sobre a previsão:

- As previsões disponibilizadas são obtidas automaticamente através de processamento estatístico das previsões de dois modelos numéricos (ECMWF e AROME), sendo atualizadas duas vezes por dia. Em situações de diferenças significativas entre a previsão e a realidade, as previsões disponibilizadas para as capitais de distrito e ilhas podem ser corrigidas por meteorologistas.
- Devido à natureza dos fenómenos meteorológicos, o estado do tempo, em algumas situações, poderá não ser reproduzido adequadamente pelos modelos numéricos. Deste modo, esta previsão poderá ser distinta da previsão descritiva elaborada pelos meteorologistas, pois estes têm ao seu dispor informação adicional e mais atualizada, pelo que o utilizador deverá consultar também essa previsão e eventuais avisos meteorológicos.
- Para aferir a incerteza da previsão em termos de precipitação, são apresentados dois valores de probabilidade de precipitação:
 - na previsão diária é apresentada a probabilidade de precipitação (igual ou superior a 1mm/24h), a qual é um indicador útil para averiguar se poderá ocorrer chuva e/ou neve durante esse dia;
 - na previsão horária/diária é apresentada a probabilidade de precipitação (igual ou superior a 0.3 mm) nas últimas três horas. Este valor permite identificar se/

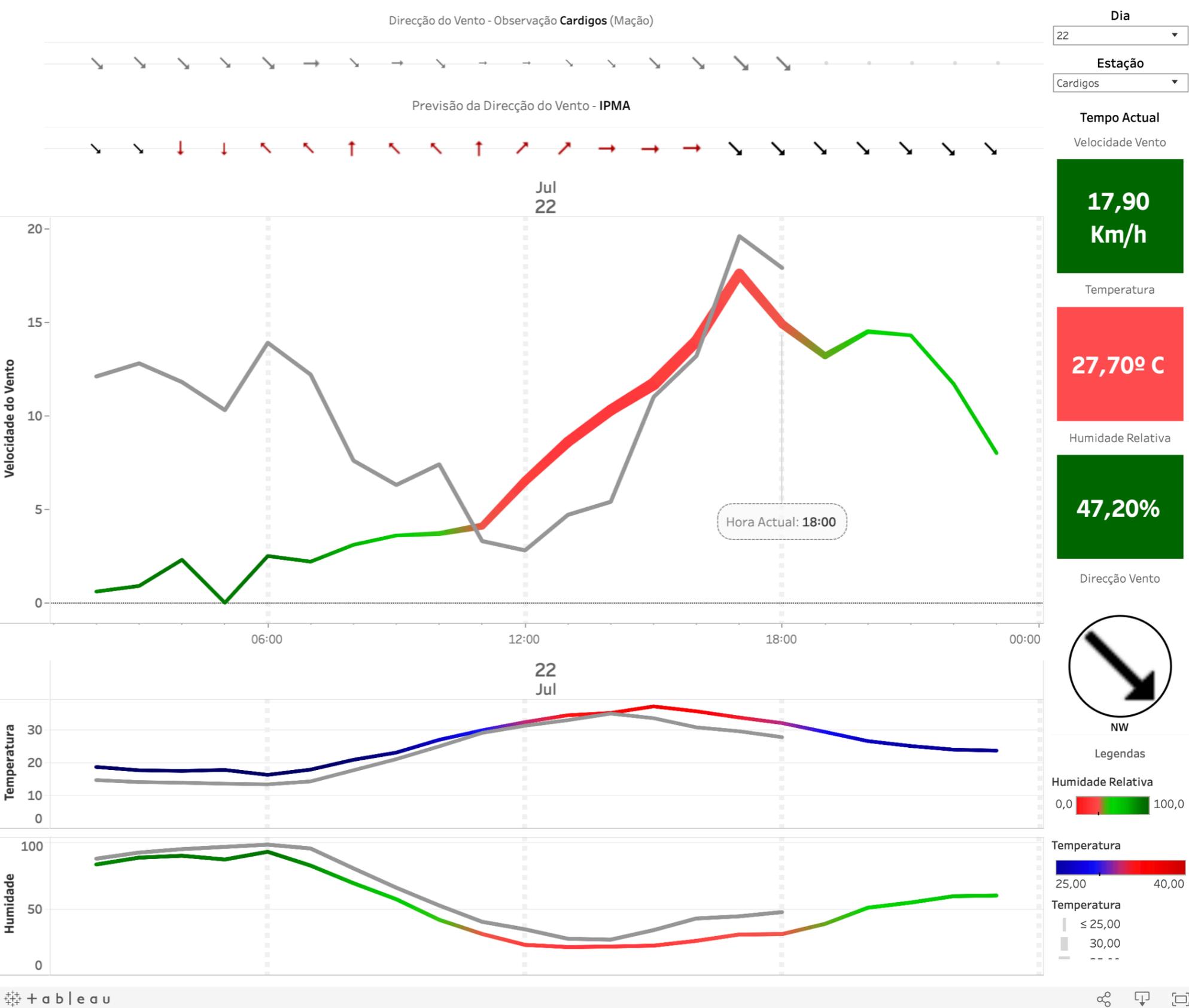
Tarefa 6/47

Encontra-se no dia 22 de Julho de 2019. Certifique-se que procura informação para a resposta com referência a este dia.

Questão:

Qual a velocidade do vento máxima entre as 12 e as 18 horas, segundo as previsões do IPMA?

Próxima Pergunta



Tarefa 46/47

Encontra-se no dia **25 de Julho de 2017 às 18:00 horas**. Existe um incêndio que começou às **16:00 horas** e propagou-se no sentido Sul, como poderá ver na visualização do lado direito.

Questão:

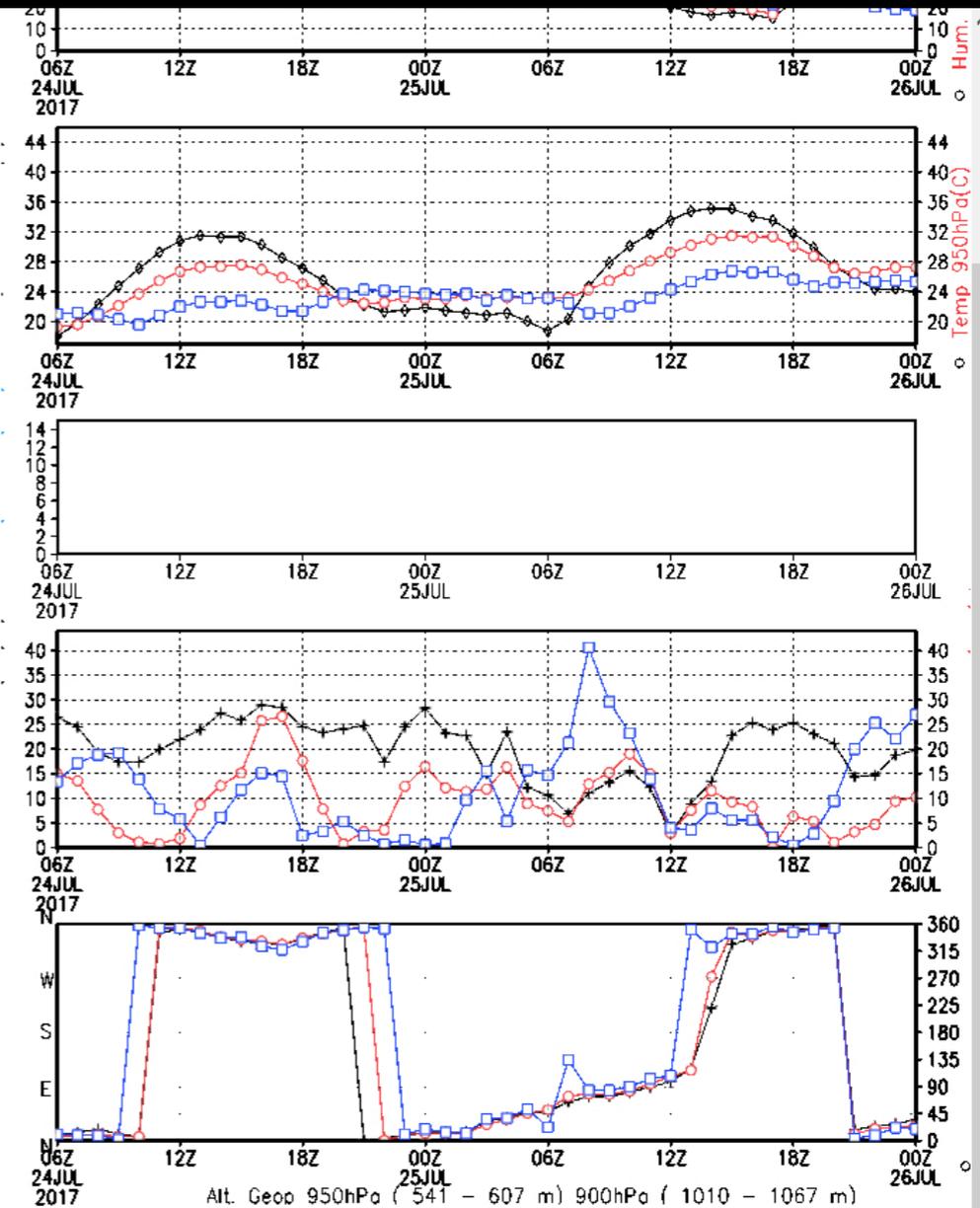
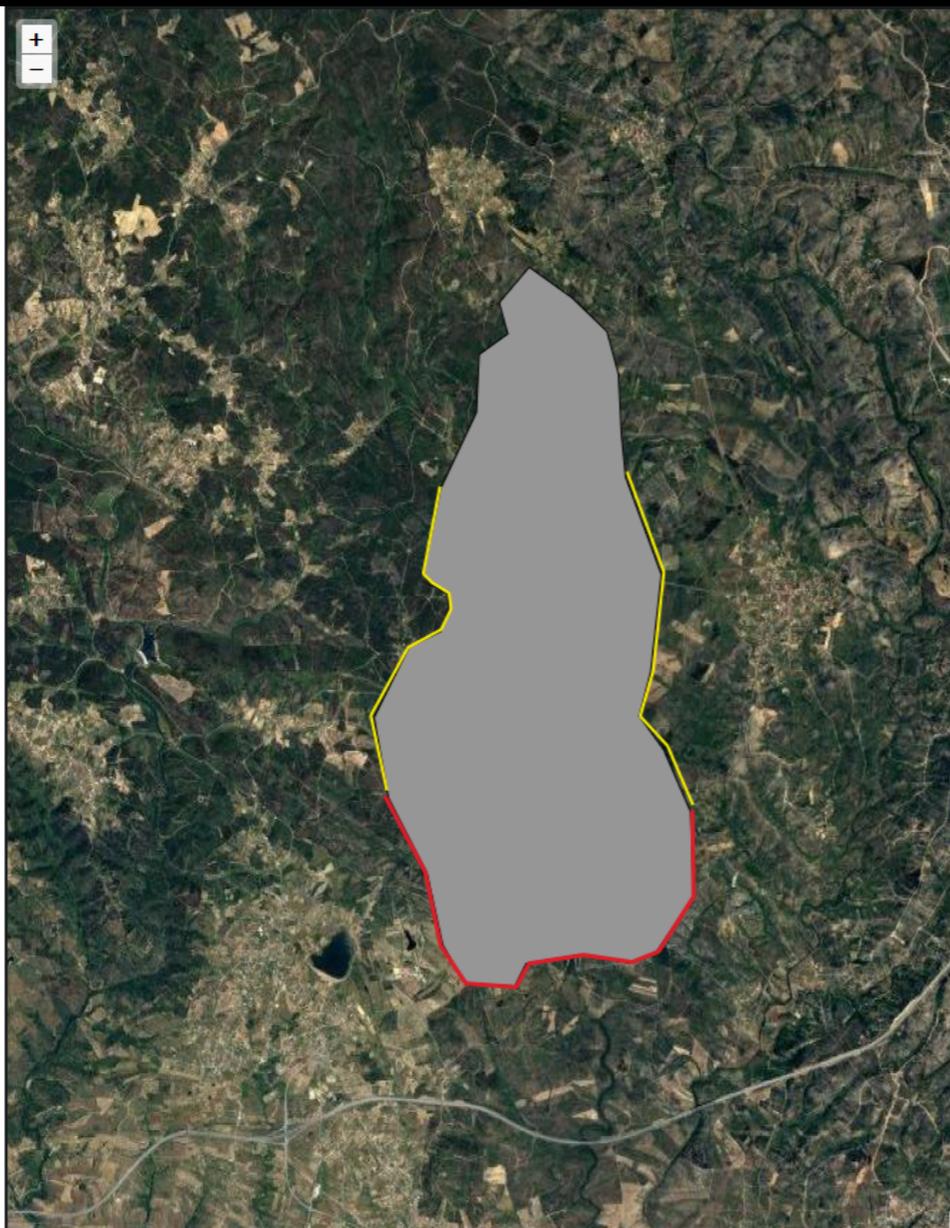
Tendo em conta as previsões da direcção do vento, os flancos do fogo vão passar a ser frentes de fogo até ao final do dia?

- Sim
 Não

Legenda:

	Frente de Fogo Pouco Activa
	Frente de Fogo Muito Activa
	Área Ardida

Próxima Pergunta



Tarefa 45/47

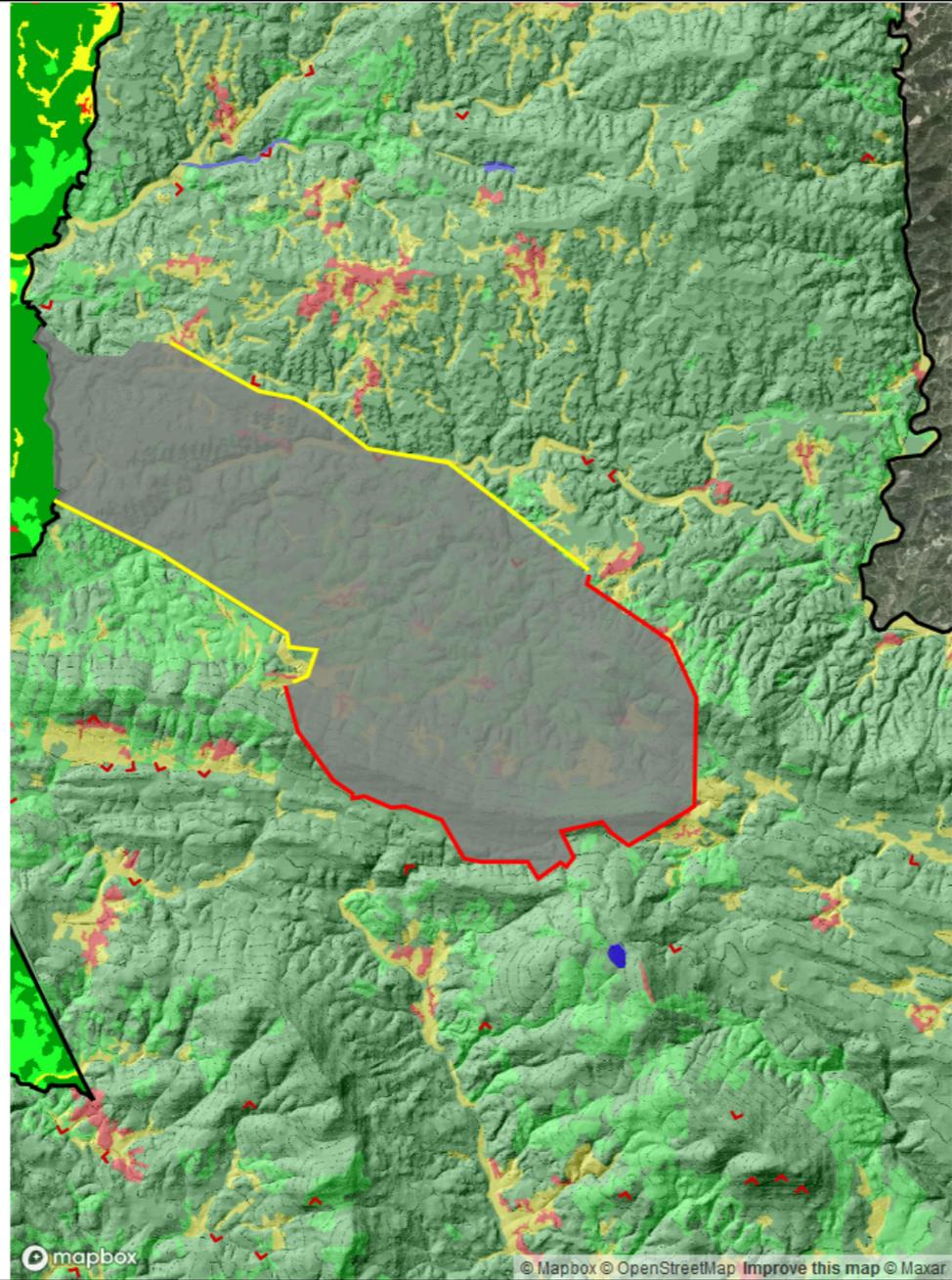
Encontra-se no dia **22 de Julho de 2019 às 18:00 horas**. Existe um incêndio que começou às **12:00 horas** e propagou-se no sentido Sudeste, como poderá ver na visualização do lado direito.

Questão:
Tendo em conta as previsões da direcção do vento, os flancos do fogo vão passar a ser frentes de fogo até ao final do dia?
 Sim
 Não

Legenda:

	Frente de Fogo Pouco Activa
	Frente de Fogo Muito Activa
	Área Ardida
	Território Artificializado
	Terreno Agrícola e Pastagens
	Superfícies Agroflorestais e Mato
	Floresta
	Espaços com Pouca Vegetação
	Água
	Direção da Subida de Encosta Acentuada

Próxima Pergunta



Direcção do Vento - Observação Cardigos (Mação)

Previsão da Direcção do Vento - IPMA

Dia: 22

Estação: Cardigos

Tempo Actual

Velocidade Vento: **17,90 Km/h**

Temperatura: **27,70° C**

Humidade Relativa: **47,20%**

Direcção Vento:

Jul 22

22 Jul

Humidade

Legendas

Humidade Relativa: 0,0 100,0

Temperatura: 25,00 40,00

Temperatura: ≤ 25,00 30,00

Further Reading and Summary

Further Reading

- **Pag 431 - 445 from Interactive Data Visualization: Foundations, Techniques, and Applications, Matthew O. Ward, Georges Grinstein, Daniel Keim, 2015**

- **See also Chapter 16 Research Directions in Visualization, 475 - 487**