Interactive Correlation Panels for the Geological Mapping of the Martian Surface

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

im Rahmen des Studiums

Visual Computing

eingereicht von

Rebecca Nowak, BSc
Matrikelnummer 00626227

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ.Prof. Dipl.-Ing. Dr.techn. Eduard Gröller
Mitwirkung: Dr.techn. Thomas Ortner, MSc

Wien, 9. August 2021

Rebecca Nowak
Eduard Gröller

Informatics

Interactive Correlation Panels for the Geological Mapping of the Martian Surface

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Visual Computing

by

Rebecca Nowak, BSc
Registration Number 00626227

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Prof. Dipl.-Ing. Dr.techn. Eduard Gröller
Assistance: Dr.techn. Thomas Ortner, MSc

Vienna, 9th August, 2021

Rebecca Nowak
Eduard Gröller

Erklärung zur Verfassung der Arbeit

Rebecca Nowak, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Rebecca Nowak

Wien, 9. August 2021

Rebecca Nowak

Acknowledgements

This work was enabled by VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH. VRVis is funded by BMK, BMDW, Styria, SFG, Tyrol, and Vienna Business Agency in the scope of COMET - Competence Centers for Excellent Technologies (879730) which is managed by FFG. A special thanks goes to the Imperial College London, and specifically to Professor Sanjeev Gupta, who invited me for a research stay at the Department of Earth Science & Engineering.

I would like to thank VRVis for the opportunity to work in an area as exciting as Mars exploration. I want to thank my colleagues at VRVis for their support, and I am also tremendously grateful for the moral support provided by my family and friends. Last but not least I want to thank my supervisors Thomas Ortner and Eduard Gröller, who helped me achieve my goal with remarkable patience and active support.

Kurzfassung

In den letzten Jahren hat sich die Verwendung von *digitalen Aufschlussmodellen* etabliert, um geologische Untersuchungen am Computer durchzuführen. Diese hochauflösenden, 3-dimensionalen Modelle von Aufschlüssen werden auch für die Erforschung des Mars erstellt. Mit spezialisierter Software können GeologInnen auf digitalen Aufschlussmodellen geologische Attribute annotieren, wie zum Beispiel die Grenzen zwischen verschiedenen Gesteinsschichten. Nach der Annotation erstellen GeologInnen *Schichtprofile*, eine graphische Beschreibung der Gesteinsschichten. Um ein geologisches Modell einer größeren Region aufzustellen, werden die Gesteinsschichten in mehreren Schichtprofilen korreliert. In einem *Korrelationsdiagramm* werden die korrelierten Schichten der Schichtprofile graphisch verbunden. Das Erstellen dieser Korrelationsdiagramme ist sehr aufwändig, in der Regel werden sie per Hand mit Zeichenprogrammen erstellt. Durch diese Einschränkung werden die Diagramme am Schluss des Interpretationsvorgangs erstellt, um ein aufwändiges Editieren im Nachhinein zu vermeiden. Auch geht durch den Wechsel in ein Zeichenprogramm die Verbindung zwischen Ursprungsdaten und den kodierten Daten im Diagramm verloren. Diese Arbeit ist Teil einer Design-Studie mit dem Ziel die Erstellung von Korrelationsdiagrammen zu automatisieren und aus einer statischen Illustration eine interaktive Applikation zu machen, die in den Interpretationsprozess integriert werden kann. Nach einer Einführung in die relevanten Themengebiete analysieren wir in dieser Arbeit publizierte Korrelationsdiagramme um den Entwurfsraum aufzuspannen. Mit den Ergebnissen der Analyse in Kombination mit ExpertInnenmeinungen, die wir in Workshops und über einen Forschungsaufenthalt am *Imperial College London* einholen konnten, beschreiben wir mögliche Designentscheidungen, und schließlich die Minimalanforderungen an einen Prototyp. Der Prototyp, der im Zuge dieser Arbeit entstanden ist, wurde erweitert und in einem Artikel, das die gesamte Design-Studie umfasst, präsentiert.

Abstract

In recent years, *digital outcrop models* have become a popular tool to carry out geological investigations on the computer. These high-resolution, 3-dimensional models of outcrops are also created for the exploration of Mars. With specialized software, geologists can annotate geological attributes on digital outcrop models, such as the boundaries between different rock layers. After annotating, geologists create *logs*, a graphic description of the rock layers. To establish a geological model of a larger region, corresponding layers are correlated in multiple logs. The correlated layers of the logs are graphically linked in a *correlation panel*. Creating correlation panels is very time-consuming, and they are usually created by hand with drawing programs. Due to this restriction, the diagrams are created at the end of the interpretation process to avoid time-consuming editing afterwards. When switching to a drawing program, the connection between the original data and the encoded data in the diagram is also lost. This work is part of a design study with the aim of automating the creation of correlation panels, and turning a static illustration into an interactive application that can be integrated into the interpretation process. In this work, after a short introduction to the exploration of Mars with the help of geology, we analyse published correlation panels to explore the design space of these illustrations. In addition to that analysis we conducted workshops and a research stay at *Imperial College London* with our domain collaborators. Using the information gained from the analysis and our collaborators, we describe possible design choices, and extract the minimum requirements for a prototype. The prototype created in the course of this work was later extended and presented in a paper that encompasses the whole design study.

Contents

CHAPTER 1

Introduction

The question whether there is, or was, life on other planets has occupied humankind for a long time. Mars has turned out to be a prime candidate for this search, and has been the destination of numerous spacecraft since the first successful fly-by in 1965. With the advent of high-resolution cameras mounted on rovers, scientists can nowadays study Mars' surface in great detail.

Geology is the science integral to finding signs of life on Mars, so called *biosignatures*. Geologists analyse rock layers to gain information about the history of the planetary

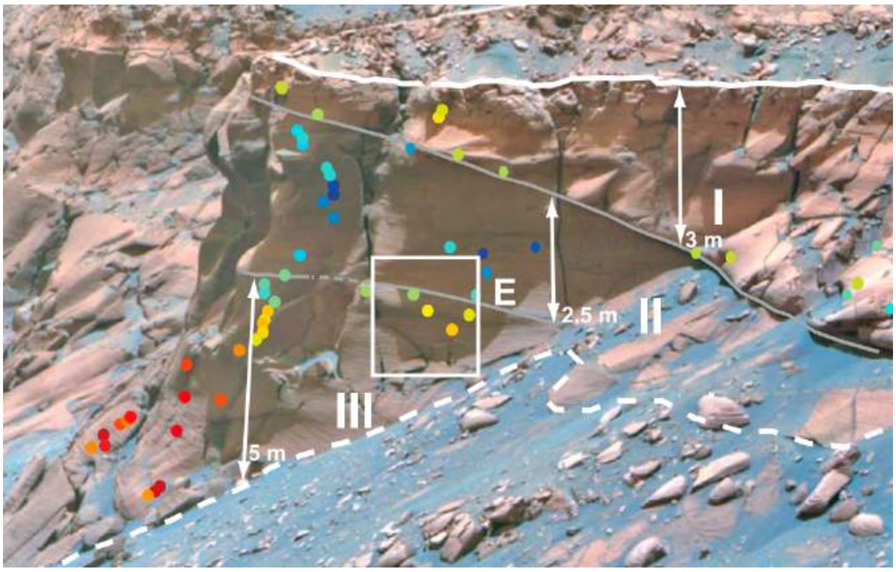


Figure 1.1: Geological annotations on a false-colour image of Cape Desire, Mars. [HGE$^+$11]

surface (Figure 1.1). On Earth, geologists can examine outcrops, where rock layers are visible, by walking (or climbing, as the case may be) up to them. They can look closely at rocks, touch them, or even put them in their mouths to examine their texture. On Mars, at least for now, this is not possible. Geologists have to rely on image material and other sensor data captured by satellites, rovers, and stationary landers. Rover imagery is of particular interest, as rovers can specifically target outcrops for image collection by driving past them, and capture images in a much higher resolution than possible from orbit. In recent years, *digital outcrop models* have become a popular tool for remote geology. Digital outcrop models are 3D-reconstructions of outcrops experts can study instead of, or additionally to examinations in person. Planetary geology, where being *remote* is a given, and also remote locations on Earth profit from these models. They are also used in the oil and mining industry in addition to data gathered by drilling wells [MHd+20]. Using specialised software, geologists can examine, annotate, and measure features in a digital 3D version of the outcrop. The annotations and analysis of the outcrop are then used to create a so-called *correlation panel*. A correlation panel is a geological diagram used to describe and correlate geological layers between different locations, thereby creating a depositional model of a wider area.

This work is part of a three-stage design process which was conducted at VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH. The results of the whole design process have been published in a paper: *InCorr: Interactive Data-Driven Correlation Panels for Digital Outcrop Analysis* [OWN+20]. We shall refer to the whole three-stage process as the *InCorr design study* throughout this work. This work details the second stage of the InCorr design study: gathering domain knowledge and expert opinions, identifying key features, and creating an extensible prototype implementation.

1.1 Motivation

Correlation panels are generally created manually, using drawing software such as Adobe Illustrator. This process is cumbersome and error prone, and changing a correlation panel once completed is very time consuming. If additional data becomes available, or interpretation continues in a way that necessitates a change in the correlation panel, this becomes a problem. For this reason, experts create correlation panels at the very end of the interpretation process. This makes it difficult to use correlation panels as a tool for collaboration or analysis during the interpretation process.

Traditionally, correlation panels are a static visualisation. However, with annotated digital outcrops, the data they are based on is already available digitally. Correlation panels visually encode annotations on digital outcrop models, but the link between corresponding data items is lost when exported to a separate drawing tool.

Correlation panels can look quite different depending on who created them and for what purpose. Experts have strong personal preferences concerning visualisation choices; the same geological feature might have different visual encodings. Which geological attributes are displayed in a correlation panel varies with use-case and author preference.

It makes communication and collaboration more difficult if authors choose different visual encodings.

1.2 Aim of this Work

The overall goals of the InCorr design study are the following:

- make geological interpretation faster and less error prone
- create correlation panels in a data-driven way from expert annotations
- allow experts to create and update correlation panels quickly in a data-driven way during the interpretation process
- enhance correlation panels by making them interactive, and linking original data to their encodings in the correlation panel

The goal of the second stage of the InCorr project and therefore the goal of this work is to create the basis for an implementation that leaves experts with enough freedom to choose their own styles, but also promotes collaboration by using predefined encodings where possible. The implementation needs to be similar to existing correlation panels and retain the advantages of the static representation. Experts need to be allowed some freedom or they will reject the tool. However, the experts we talked to in the course of this work were also in favour of more standardisation to improve collaboration.

For a data-driven generation of correlation panels the information from which a certain feature of the correlation panel is derived needs to be present in the data. The *data* in this context are digital outcrop models and the annotations experts add to them. A prerequisite for generating a correlation panel is an *annotation system* which combines visual annotations with semantic context, so the information can be encoded into the correlation panel. Therefore, designing and implementing the prototype of such a semantic annotation system is also part of this work.

1.3 Methodology

To place this work into the context of the InCorr design study, we consider the nine-stage design study methodology framework devised by Sedlmair et al [SMM12] (Figure 1.2). The first stage is (*learn*): studying visualisation literature. The authors stress that a firm grasp of the visualisation literature is a necessity when it comes to conducting a design study. This and the next two stages, namely *winnow* and *cast*, fall into the category of *precondition*. *Winnow* and *cast* deal with finding suitable collaborations with domain experts.

Finding collaborators was part of the first phase of the InCorr design study. A collaboration with planetary scientists was already established by the time phase two of InCorr,

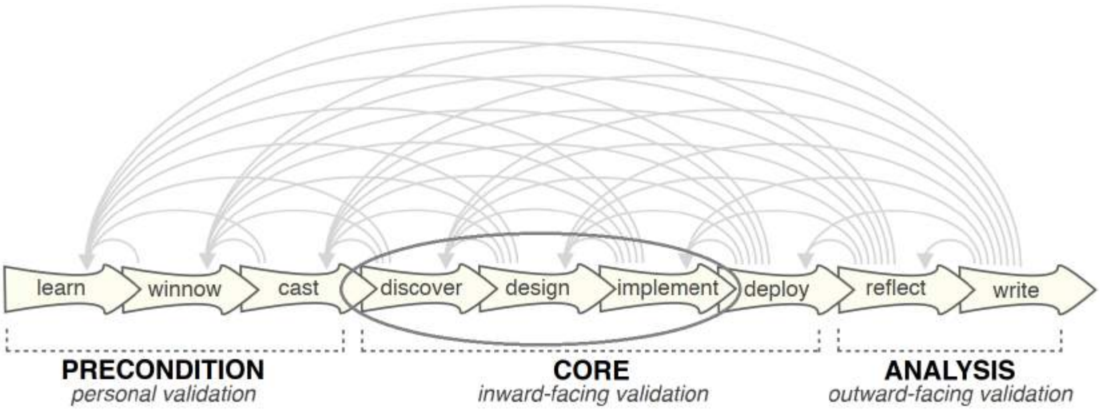


Figure 1.2: The nine-stage design study methodology framework devised by Sedlmair et al. [SMM12] The stages of the InCorr design study this work encompasses are *discover*, *design*, and *implement*.

which this work is part of, started. This work contributes to the *discover*, *design*, and *implement* stages in the *core* category of Sedlmair et al.'s methodology.

In the course of a week-long research stay at the Imperial College London as well as occasional informal talks and workshops, we collected expert opinions and feedback. In addition to this we analyse correlation panels published by different experts to acquire a holistic understanding of the visualisation device that is at the core of this work. The questions relevant to this analysis are the following:

- which geological attributes do correlation panels display?
- how are they encoded?
- what are the key features of a correlation panel?
- which features can be derived from data and can therefore be part of the automatic generation of a correlation panel?

Using the knowledge gathered from experts and the analysis of published correlation panels we identify

- the range of possible visual encodings for the relevant geological features
- which visual encodings should be preferred
- which features need to allow authors to choose their own encodings
- which features should use predetermined encodings to facilitate collaboration

- which interactions are necessary to realise each feature

As part of the *implementation* stage of the InCorr design study, we create an easily extensible prototype that contains at least the features identified as necessary.

1.4 Structure of this Work

To understand the domain, it is necessary to acquire knowledge in the fields of planetary exploration and geology. What is a correlation panel? How do geologists currently create correlation panels? What challenges does planetary geology face as opposed to terrestrial geology in this context? These are among the questions we answer in Chapter 2, where we explain the geological concepts behind correlation panels and discuss relevant aspects of planetary exploration. In the last section of Chapter 2 we describe the problems faced by experts creating correlation panels that we address in this work.

In Chapter 3 we present an overview of related papers in the visualisation community and existing software tools for (planetary) geologists, which relate in some way to this work. In Chapter 4 we first devise a task abstraction, describe our data model, and evaluate the design space by analysing correlation panels taken from different publications. We use this information and input from the domain experts we collaborated with to characterize the design of a semantic annotation system and of data-driven, interactive correlation panels.

Chapter 5 presents our visualisation and interaction design for a basic but extendable implementation of the features identified as most important in Chapter 4. We present our results in Chapter 6 and further and future work in Chapter 7.

CHAPTER 2

The Exploration of Mars with the Help of Geology

Geology is a science that can help us understand what certain environments were like a long time ago. It has been practised on Earth for a long time, but geology is also one of the key fields to understanding the history of other solid bodies in the solar system.

The field of *planetary geology* evolved from using geological principles in the context of planetary exploration. Applying geological principles that hold true on Earth to Mars, by looking for *analogies*, is an integral aspect of planetary geology. On Earth, often both geological features and the geological processes that produced them are known. When trying to form a hypothesis about the conditions under which geological features have formed on Mars, similar geological features on Earth might lead to a better understanding of Mars' past [RvG18].

In Section 2.1 we present a short history of the exploration of Mars. Section 2.2 of this chapter introduces basic geological principles that are used to explore not only Earth but also, for example, Mars or Venus. We need these concepts to understand what correlation panels are and what they represent. Section 2.3 ties those two topics together, making the connection from outcrops on Mars to the experts on Earth who investigate them, and discuss where correlation panels fit into this process.

2.1 The Exploration of Mars

Missions to explore Mars have been carried out as early as the 1960ies, with the first successful flyby undertaken by the Mariner 4 spacecraft in 1965. In 1971, Mariner 9 orbited Mars for a full year, sending back a wealth of data including information on the planet's atmosphere, surface composition, and topography. It mapped the entirety of the surface, discovering Mars' large volcanos and canyon systems, as well as ancient river beds. More spacecraft were sent to orbit and land on Mars in the following years. In 1976, the landers of Viking 1 and Viking 2 collected and analysed soil samples, whereas the corresponding orbiters found evidence of the composition of Mars' pole caps. The north cap is made of water ice which does not melt over the course of the Martian year, whereas the south cap consists of carbon dioxide which evaporates during the summer, adding to the planet's atmospheric pressure [Age].

In the following decades a number of missions were conducted, sending orbiters and landers to Mars. The first rover to collect useful data was Sojourner, a part of the Mars Pathfinder mission. Separating from the lander, which delivered it to the surface, it explored the surroundings of the landing site. Rounded pebbles and cobbles found on the surface are presumed to have been formed in a warmer era by flowing water [AAj].

The Odyssey Orbiter (launched in 2001) discovered evidence of large deposits of underground water ice, and continues its work (as of July 2021) to this day, making it NASA's spacecraft which operates for the longest period of time exploring Mars. Among many other discoveries, the Odyssey Orbiter measured radiation in low orbit, finding it to be twice as high as the one found on Earth - valuable information for potential human habitation [AAi].

The next rover to land on Mars was MER (Mars Exploration Rover) Spirit (2004), followed closely by her twin, MER Opportunity (also 2004). In addition to cameras for navigation and hazard avoidance, they were equipped with a microscopic imager and a panoramic camera (Pancam, Figure 2.1). The rovers found evidence that there was a wet period of time in the history of Mars when microbial life could have existed [JPLAb][JPLAa].

In 2006 the Mars Reconnaissance Orbiter arrived at Mars to learn "about Mars' changing climate, geologic history and potential ability to harbour life" [AAg]. It has sent an unprecedented amount of data back to Earth in the just over 15 years it has been in orbit. The orbiter is still operational, working as a data relay for other missions. [AAd]

In 2008 a stationary lander with a robotic digging arm called the *Phoenix Mars Lander* analysed a soil sample and confirmed what the Odyssey Orbiter found in 2002: The soil sample contained water ice [AAk].

In August 2012 the Mars Science Laboratory (MSL) [AAb], also called *Curiosity*, started to explore Mars. With its width of 2.8 metres and length of 3 metres the MSL was the biggest rover to explore the red planet up to that point. Curiosity found evidence that ancient Mars was an environment in which microbes could have lived by drilling for and analysing a powder sample [AAc].

In February 2021 the NASA rover *Perseverance* landed successfully. Perseverance is not only equipped with superior cameras (Figure 2.2) but is also capable of collecting rock samples for a sample-return mission.

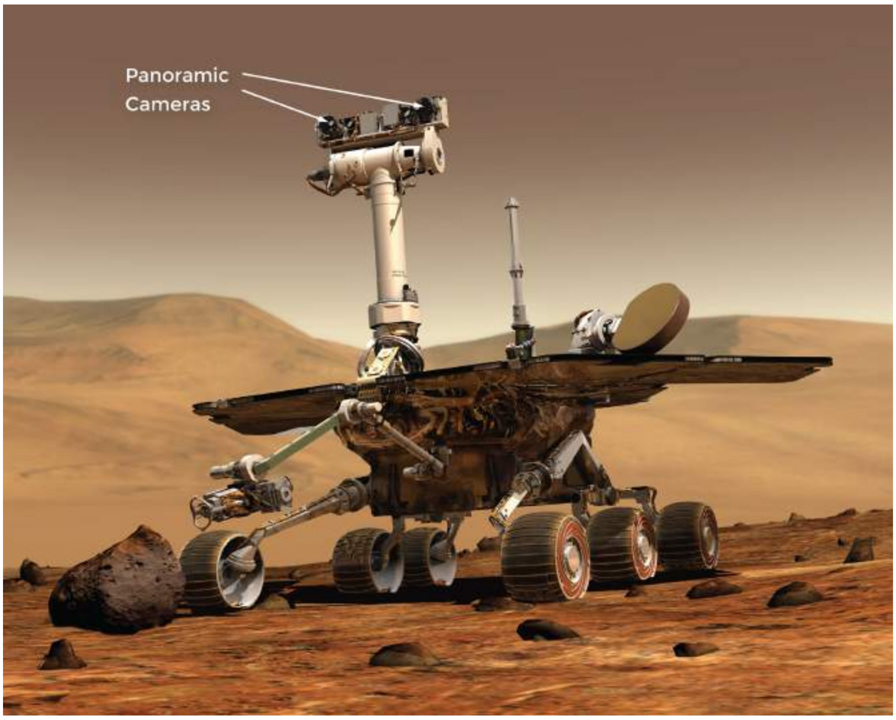


Figure 2.1: Pancam (Panoramic Cameras) on th MER Mars rovers [AAh].

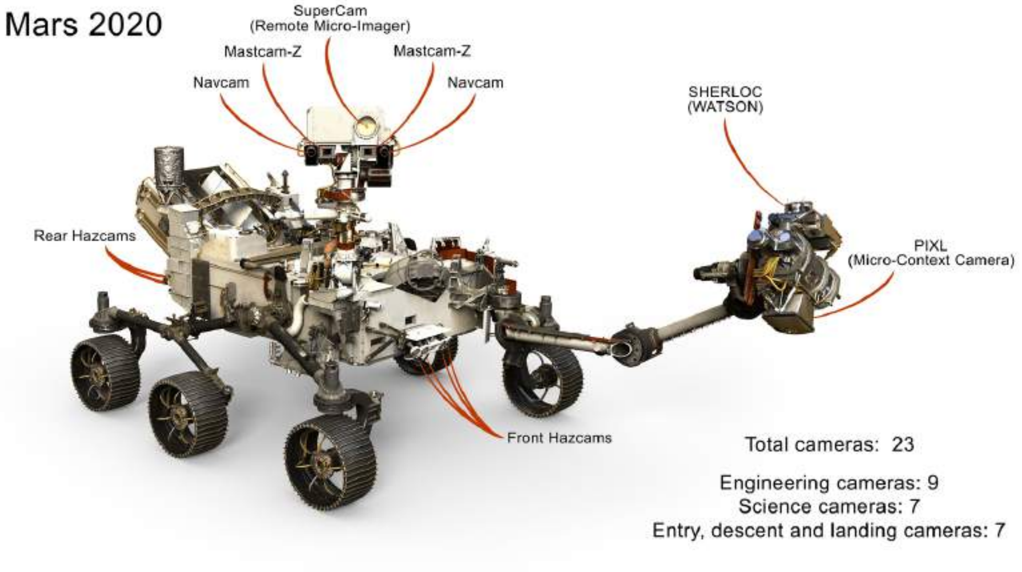


Figure 2.2: Cameras on the Mars rover Perseverance.[AAa]

2.2 Geological Concepts

Correlation panels are a highly specific and complex geological illustration. Acquiring the knowledge necessary to understanding the concepts behind such an illustration is therefore crucial. In Sedlmair et al.'s design study methodology this step is part of the stage *discover*, the first stage of the *core* category. The experts we collaborated with recommended literature for this purpose, pointing us in the right direction in our investigation. We use the recommended material and additional literature to present a short introduction to the underlying geological principles and terminology used in the course of this work.

Geologically speaking, a *rock* is a combination of *minerals*, which are solid homogenous crystalline substances. Rocks can be assigned a type according to their origin: *Igneous rocks* are formed by other rocks melting and solidifying again. *Metamorphic rocks* require high temperatures and high pressure to be formed. Lastly, *sedimentary rocks* are created from loose particles and deposited in layers (*sediments*). Sediments can be either formed by mechanical processes (*siliciclastic sediments*), or by chemical and biological processes (*chemical sediments, biological sediments*). Figure 2.3 depicts the process from the creation of sediments to the eventual deposition in layers (*bedding*) and conversion to solid rock [GJ14].

The *lithology* of a sediment refers to its composition or mineralogy [Tuc03]. Examples of this include the grain size or colour of a rock.

Diagenesis or *diagenetic* processes are what happens to sediments after deposition. These processes depend on factors like the depth of buried sediments, and the resulting change in temperature and pressure [CMT06].

A layer thicker than one centimetre is called a *bed*, thinner layers are called laminae. *Strata* might be used as a synonym for beds, but might also refer to a cluster containing many beds. A homogenous layer of sediments means that the circumstances responsible for the deposition, the *depositional conditions*, of that bed did not change significantly during the period it was formed. An abrupt change in physical and/or chemical circumstances will lead to a relatively clearly defined boundary between neighbouring beds. These sharp boundaries are called *bedding planes* or *bounding planes* [CMT06]. A more general term for the boundary between two rock types is *contact* [All13].

A secondary layer (i.e. a layer within another layer) might be angled differently to its primary (or *main*) layers' bedding surfaces. These secondary layers are called *cross strata. Cross beddings* and *cross laminae* can both be called *cross strata.* A *bedset* is a group of similar beds [CMT06].

A *stratum* is a lithological term for a rock layer. Unlike *bed*, it does not tell us how thick that layer is [All13]. Stemming from the same Latin word *stratum*, *stratigraphy* is a general term, meaning the characterisation of rock bodies and how they are categorised into distinct units depending on their attributes [RvG18].

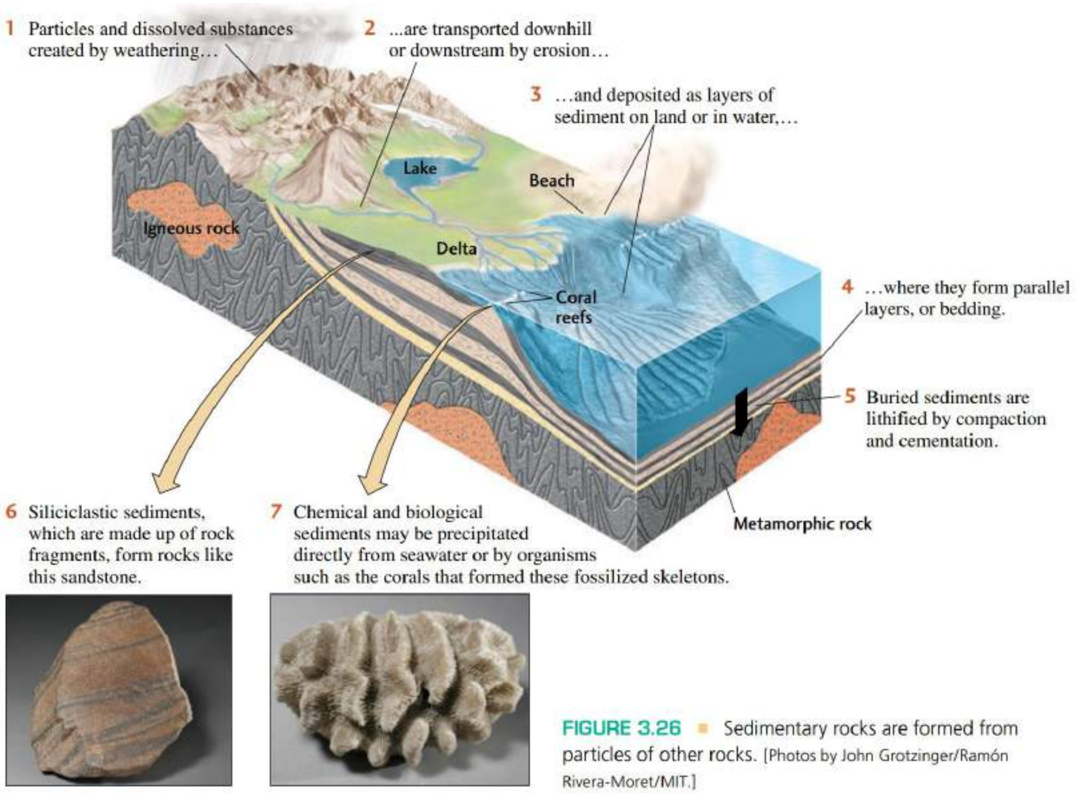


Figure 2.3: The creation of sedimentary rocks ([GJ14] page 77).

One such unit is called a *stratigraphic unit*. If the unit is defined by its lithology then it is called *lithostratigraphic unit* [All13].

Another concept we mention in this work is *dip and strike* (Figure 2.4). Geologists perform dip and strike measurements in their analysis of rock layers. A dip and strike measurement tells us the angle of a layer in relation to the horizontal plane that would be formed by the surface of a body of water. The *strike* or *strike line* is the horizontal line that lies on that layer. The magnitude of the dip is measured between the horizontal plane and the direction of the layer (the dip). *Dip azimuth* refers to the direction of the dip as projected onto the horizontal plane and compared to the cardinal direction north.

Sketches are an essential element of geological field work. According to Coe [Coe10], sketches are even more valuable as a resource than photographs or verbal descriptions: A sketch is more succinct, the geologist selects relevant features and omits irrelevant ones, meaning that it already contains a certain amount of interpretation. The pictorial representation of features can be grasped (and produced) more quickly than a corresponding verbal description. A *visual log* is such a sketch, that can be found in field notebooks as well as in publications (Figure 2.5). Coe asserts that *geological logs* are the best method to record any stratigraphic information [Coe10].

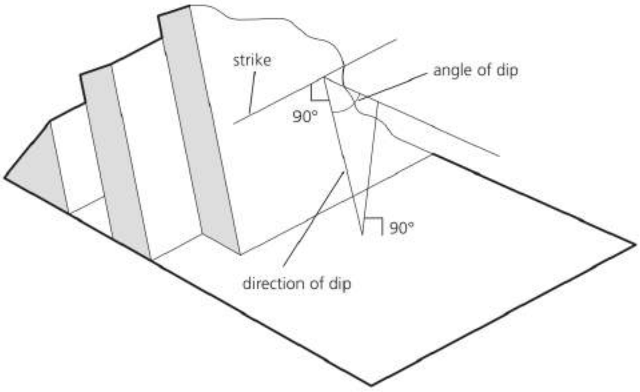
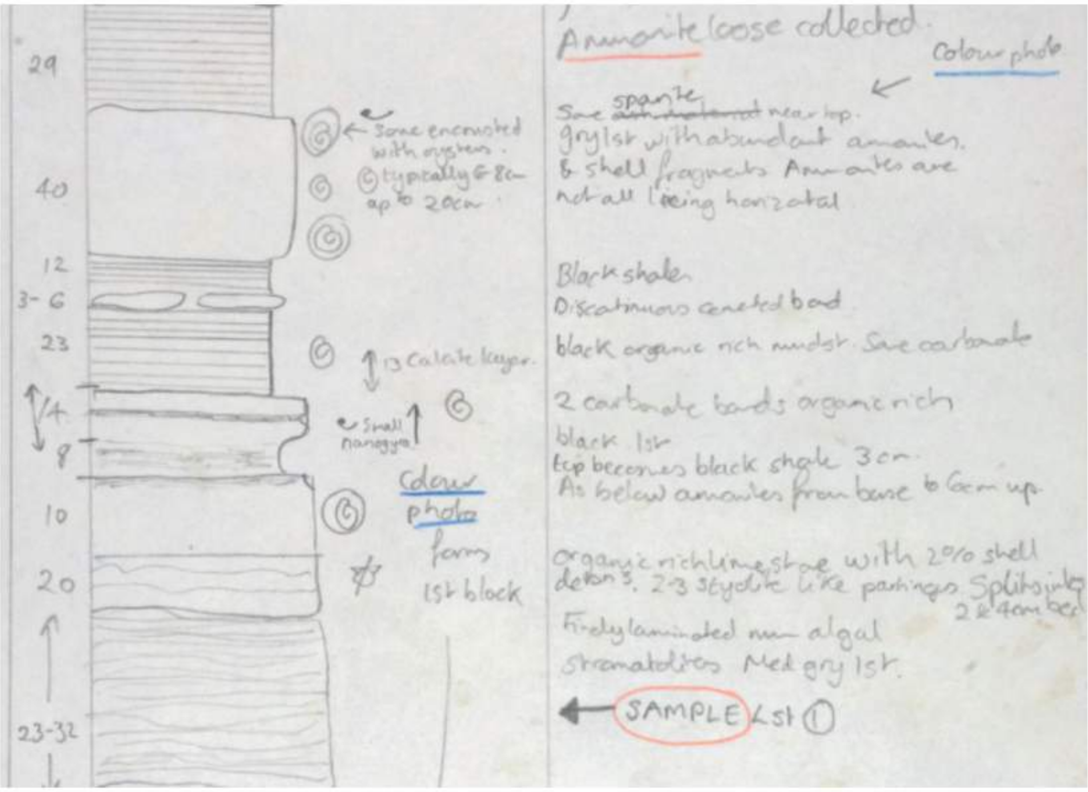


Figure 2.4: Dip and strike ([All13] page 170).



Ammonite loose collected.

Some ~~[illegible]~~ sparite near top.
gry lst with abundant ammonites.
& shell fragments Ammonites are not all lying horizontal

Black shale.
Discontinuous cemented band
black organic rich mudst. Some carbonate

2 carbonate bands organic rich
black lst
top becomes black shale 3cm.
As below ammonites from base to 6cm up.

organic rich limestone with 2% shell debris. 2-3 stylolite like partings. Splits into 2 & 4cm beds
Finely laminated mm algal stromatolites Med gry lst.

SAMPLE Lst ①

Figure 2.5: A section of a graphical geological log from a field notebook (Notebook of Angela L. Coe, The Open University, UK., from [Coe10] page 58).

Before we discuss the topic of geological logs, we give a few clarifying notes. The term *geological log* might sometimes cause confusion as there are different phrases that refer to the same concept. The above mentioned *visual log* is one of them. Instead of *visual* log, *graphic* log is used by some authors (for example Nichols [Nic09]). In the context of geology, the adjective *geological* is usually omitted. However, adjectives like *stratigraphic*, *lithological*, or *sedimentary* might be used to qualify a log. Logs are also sometimes called *stratigraphic columns* (as in [HGE$^+$11]) or *stratigraphic sections*. The word *log* itself might also lead to confusion, given its specific meaning in computer science and mathematics. Therefore we want to make it clear that when we use the word *log* in the context of this work, the geological visualisation of rock layers is meant.

An example of a log drawn during fieldwork can be found in Figure 2.5. The drawing on the left hand side represents the geological strata of an outcrop. Each roughly rectangular shape encodes a layer in that outcrop. One layer is distinguished from another by its geological attributes. The width of the shape corresponds, in general, to the grain size of the corresponding layer. The height of each layer corresponds the actual thickness of that layer at the location represented in the log. The shapes in a log are often filled with patterns which represent different lithological attributes. In addition to these patterns, which fill a shape, symbols may be used to represent sedimentary structures and fossils. Figures A.1, 4.4, and A.2 show the patterns and symbols used by different authors. A log with descriptions of the different elements can be found in Figure 4.4. The border or transition between two strata, or to use the geological term, the *contact*, can be shaped in many different ways. It might be a sharp border, or a gradual transition, or one of many other contact types. Tucker shows visual representations for these types of contacts in [Tuc03] (Figure 4.10).

Knowing the grain size of strata is an integral part of creating a geological log. To determine the grain size of a layer in the field, geologists use so-called *grain-size charts* (Figure 2.6). Grain size categories include *clay*, *silt*, *sand*, *granules*, and *gravel*. *Sand* is usually further divided into categories, for example *very fine sand*, *fine sand*, *medium sand*, *coarse sand*, *very coarse sand*. Which categories are used depends on the data represented in a log. If no stratum in a log is made of rock with larger grain sizes, that side of the scale is often left out. How finely the categories are divided also depends on the data. An example of this practice are the details of three logs in Figure 2.7. Each has different categories for the grain size based on the data represented.

Sometimes pre-printed tables are used to aid in the creation of logs in the field. One such (filled in) table can be seen in Figure 2.8. In the table, *texture* is used as an alternative term for grain size.

As Coe remarks in her book *Geological Field Techniques* [Coe10], logs come in a variety of styles, according to the personal preference of the author as well as the subject matter being illustrated.

Logs are not only used to record findings in the field, but also for interpretation and to illustrate those findings in a publication. In general, a log in a publication will be quite

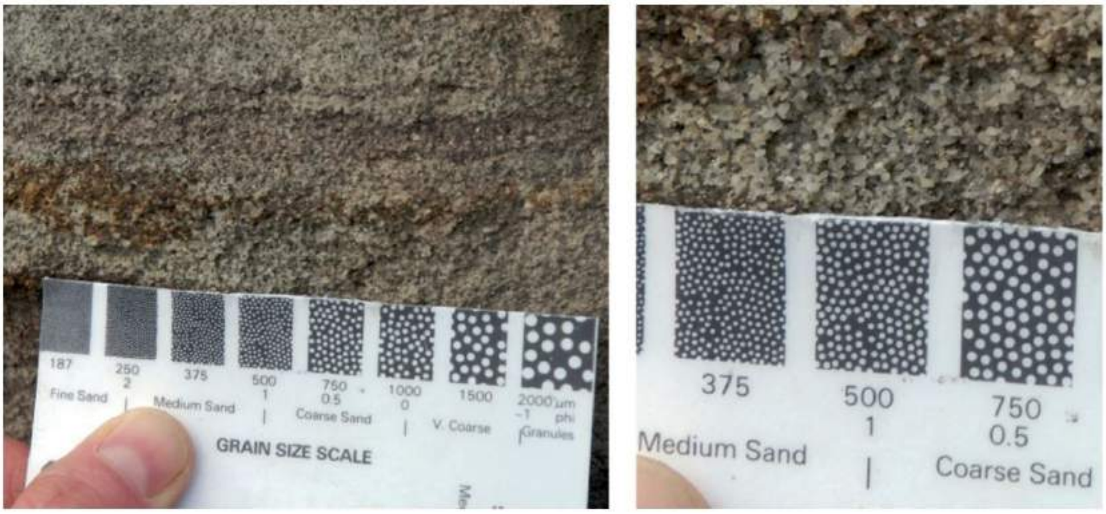


Figure 2.6: A so-called *grain-size chart* being used to determine the grain size of a geological layer. In this case, as the author points out, the average is 500 µm, ranging from 375 to 740 µm [Coe10].

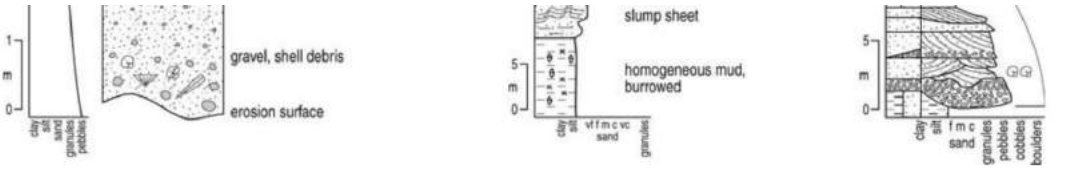


Figure 2.7: Detail of three logs by the same author with different grain size scales [CMT06].

similar to one found in a field notebook. Neither type of log is standardised. The reasons for this lack of standardisation cited by Tucker in his book *Techniques in Sedimentology* [Tuc88] lie in the variety of the data that might be contained in a log. One factor is scale. A log might cover a few decimetres or even hundreds of metres. One stratum in the log might encode any type of layer, be it a bed, a bedset, or even a facies [Tuc88]. A facies comprises multiple properties of sedimentary rock, and was created in a specific depositional environment [Tuc03].

Logs represent attributes of strata at specific locations, where they can be observed. On Earth we do not have to rely on outcrops alone, where strata are visible. In the gas and oil industry, wells are drilled in locations of interest, and logs are then created from the resulting data. Whether logs stem from natural outcrops or not, they only represent the geological strata at certain locations. What form the strata take in-between these locations, what a coherent model of a whole area might look like, has to be inferred by experts. *Correlation panels* are a tool to visualise and support the inference of this model.

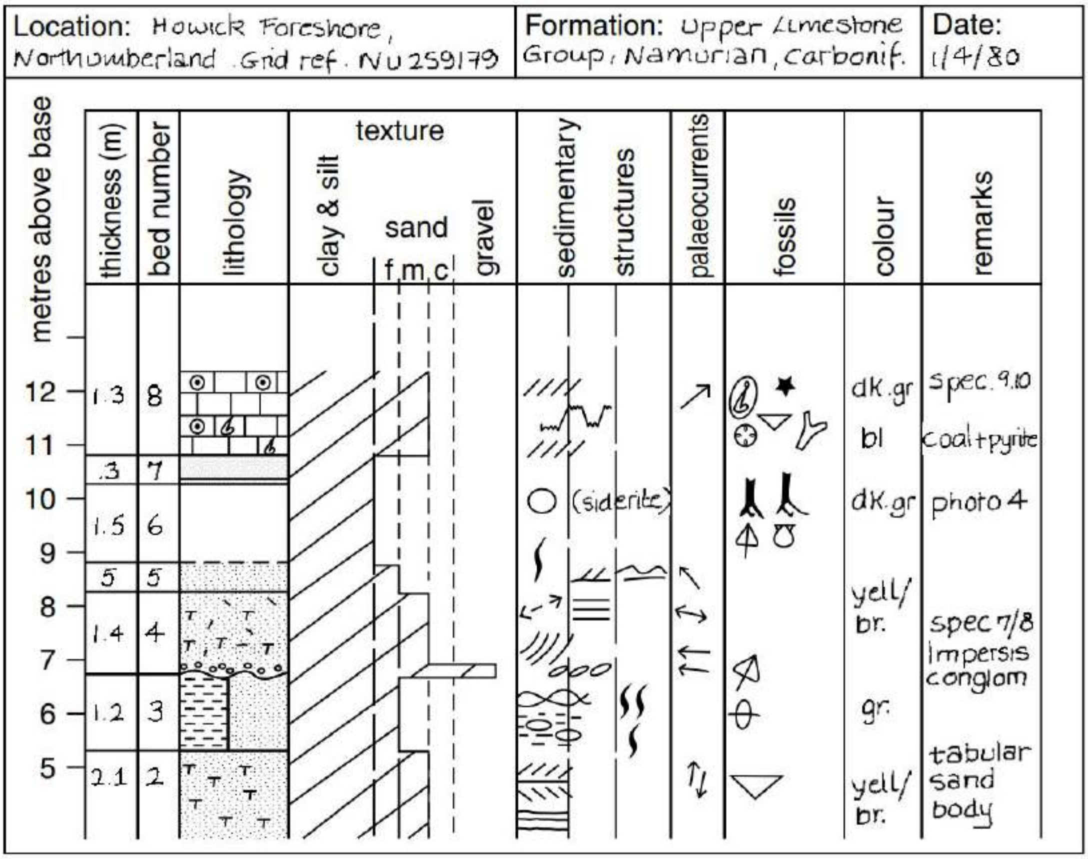

Location: Howick Foreshore, Northumberland Grid ref. NU 259179
Formation: Upper Limestone Group, Namurian, Carbonif.
Date: 1/4/80

Figure 2.8: A graphic log form, which can be used in the field to systematically record information about outcrops. The symbols are explained in Figure A.1 (*Sedimentary Rocks in the Field 3rd Edition* by Maurice Tucker[Tuc03], page 11).

As the Oxford University Press *Dictionary of Geology and Earth Science* puts it, a correlation in the context of stratigraphy is "the establishment of a correspondence between stratigraphic units" [All13]. Following this definition a correlation panel illustrates "probable stratigraphic equivalence from place to place" [All13]. Visually, a correlation panel is an illustration where multiple logs are placed next to each other and contacts (and therefore strata) are *correlated* by connecting them visually with lines. As the definition above implies, correlated contacts represent the *same* contact present at two different locations (two different logs). The locations might be a few metres apart on the same outcrop, hundreds of metres apart on the edge of a crater, or might cover even larger distances (as the correlation panel in Figure 4.14). In Section 4.3.2 we analyse multiple correlation panels used in publications by experts, where it will become clear how varied these illustrations can be. Although the underlying concept of placing logs next to each other and correlating contacts (or strata) is observed in all correlation panels, which additional information is added and how it is encoded, depends very much on the use-case and an author's preferences.

2.3 From Rocks on Mars to Correlation Panels

As we can see from the short overview of the exploration of Mars in Section 2.1, there have been many expeditions to that planet, and all successful missions add to the data we on Earth have available for research. Of course, this leads us to the obvious difference between conducting geological research on Earth and planetary science. Rather than travelling to, and examining rock formations in person, experts have to rely only on data. This data is also much harder (and more expensive) to acquire than, for example, data of remote locations on Earth.

Rovers like Curiosity and Perseverance are important to the study of outcrops. Firstly, they are capable of capturing outcrops in resolutions high enough that experts can use the resulting images for geological analyses. Secondly, their course across the surface of Mars can be adapted depending on previous data. They can and do get close to features of interest to obtain more detailed images of those features. However, when selecting a rover's path there are also constraints not connected with the scientific interest of locations. The safety of the rover is paramount and plays a big part in selecting a route or landing spot, which means that a rover may have to keep a certain distance from an interesting geological feature, even if driving closer to it would yield better data.

Once the data is collected it has to be sent back to Earth. A ground station on Earth sends data to coordinate the operations of orbiters, rovers, landers, and other platforms like probes. Orbiters might be used as relay stations to convey data to surface-based platforms. Rovers and landers receive this data, and send back data they collect. For this they might again use orbiters as relay stations. Figure 2.9 illustrates how data is collected and transferred to Earth in planetary exploration.

Another reason the data collected by rovers are vital in the context of this work is that the Mars rovers Curiosity and Perseverance are equipped with high resolution stereo cameras (Mastcam [AAe] and Mastcam-Z [AAf]). The images taken by these cameras, but also by other cameras of the rovers or even orbiter images can be used to create photomosaic images and 3D point clouds. For this, various 3D vision algorithms are used, including algorithms to first register images from different sources. After 3D reconstruction, the digital outcrop model can be examined by geologists. They use them to examine the stratigraphy and different sedimentary features, and gain information about impact craters, and the environment in the distant past [PBG+18].

After familiarising themselves with the data set, experts generally *annotate* a digital outcrop model, that is, if the software tool they are using allows this. Chapter 3 introduces some tools used for this purpose. Annotations are part of the interpretation process. Experts detect features on the outcrop and annotate them with visual markers. Annotations are, of course, not limited to digital outcrop models. Experts also annotate two-dimensional images with drawing software. Figure 2.10 shows such an example from

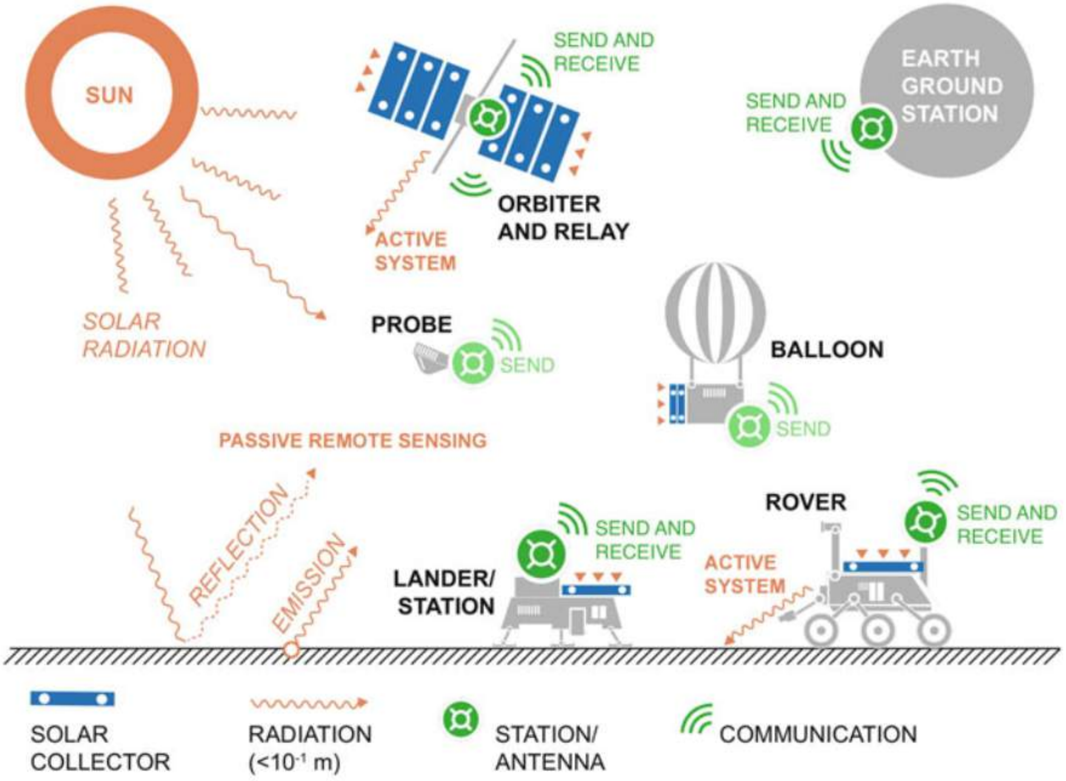


Figure 2.9: The different platforms used for planetary exploration. A ground station on earth sends data to coordinate the operations of orbiters, rovers, landers, and other platforms like probes. Orbiters might be used as relay stations to convey data to surface-based platforms. Rovers and landers receive this data, and send back data they collect. For this they might again use orbiters as relay stations [RvG18].

a paper on the first mudstone found on Mars. An annotated image is shown next to a log with corresponding contacts connected.

After annotating outcrops, experts create correlation panels manually in software such as Adobe Illustrator. The correlation panels are then used in publications to present experts' findings in a succinct fashion.

2.4 Domain Specific Terms: Reference

In this section we list and define the most important geological terms and concepts we use in this work in table 2.1.

Figure 2.10: The use of geological annotations in planetary geology with an annotated image (right) and corresponding log (left). This is part of a figure taken from Schieber et at. [SBC+17], where geologists present their findings on the first mudstone found on Mars.

Term	**Definition**
sedimentology	"The scientific study, interpretation, and classification of sediments, sedimentary processes, and sedimentary rocks" [All13]
lithology	The composition of a sediment. [Tuc03]
layer	A rock layer that can be distinguished from its neighbours in some way.
bed	A rock layer thicker than one centimetre.
stratum (pl. strata)	The lithological term for a rock layer. It is a more general term than bed, as it does not indicate a certain extent. According to Allaby [All13], *bed* and *stratum* are sometimes used interchangeably, but they are not synonyms.
stratigraphic unit	"A body of rock forming a discrete and definable unit" [All13]. Stratigraphic units are defined with respect to certain attributes of the rock layers. This means, that the same outcrop might be divided into different stratigraphic units depending on which attributes are used to define them.
contact	The boundary between two rock layers.
facies	The attributes of a sedimentary rock, being the product of a particular depositional environment or process [Tuc03].

Table 2.1: Definitions of domain specific terms used throughout this work.

2.5 Problem Statement

Drawing correlation panels is a time-consuming process. For this reason, experts relegate this task to the very end of the interpretation process. This ensures that they will not have to change the correlation panel after creating it. Correcting or editing a correlation panel with drawing software to reflect new insights is, again, very time consuming. This approach ignores one important use case of the correlation panel: that it is not only a tool for presentation, but also a tool for *interpretation.* A correlation panel might inform the interpretation process, for example indicating that annotations should be added in a certain area, or that certain features should be annotated in greater detail. These changes might then again impact the correlation panel itself, necessitating changes in the visualisation. When drawing correlation panels manually, this flexible approach becomes prohibitively expensive in terms of effort.

Let us assume that experts have access to digital outcrop models and software that allows them to annotate these models, like our collaborators. In this case, hand-drawn correlation panels also fail to exploit the fact that at least part of the data encoded in them is already present in a digitised form, i.e. the digital outcrop model and the annotations added to it. Taking advantage of this data can not only automate some of the manual labour involved in creating a correlation panel, it also becomes possible to preserve the connection between encoded data and their source. This connection of data items opens up the possibility of *interaction*, leading us to the vision of *data-driven, interactive correlation panels.*

CHAPTER 3

Related Work and State of the Art

In this chapter we present other research that deals with geological visualisation and visualisation tools currently used by geologists and specifically planetary geologists.

3.1 Geological Visualisation in the Oil and Mining Industry

Correlation panels are routinely used in the oil and gas industry for well correlation. The rich data gathered when drilling wells can also be used to automatically generate logs and suggestions for correlations between them. Different techniques from dynamic programming algorithms to neural networks [ASA19] have been proposed to automatically correlate wells. Well correlation algorithms can be divided into pairwise-well correlation algorithms and multi-well correlation algorithms [LSW$^+$19].

As early as 2011, Höllt et al. [HBG$^+$11] devised a framework that allows the interpretation of well data with the help of multiple linked 2D and 3D views (Figure 3.1). The problems described by Höllt et al. were somewhat similar to the ones faced by geologists when constructing correlation panels. The interpretation relies on tracing horizons in seismic reflection volumes which is a time-consuming task. Höllt et al. introduce a semi-automatic method to generate these horizons by combining seismic data and well data. They use seed points and a cost function in combination with user-defined constraints to guide the global minimization algorithm that traces horizons. The resulting horizons are visualised in 3D and 2D views. Horizons can be edited by clicking and dragging. 2D and 3D views are linked: the 3D view allows picking, and active items are highlighted in all views.

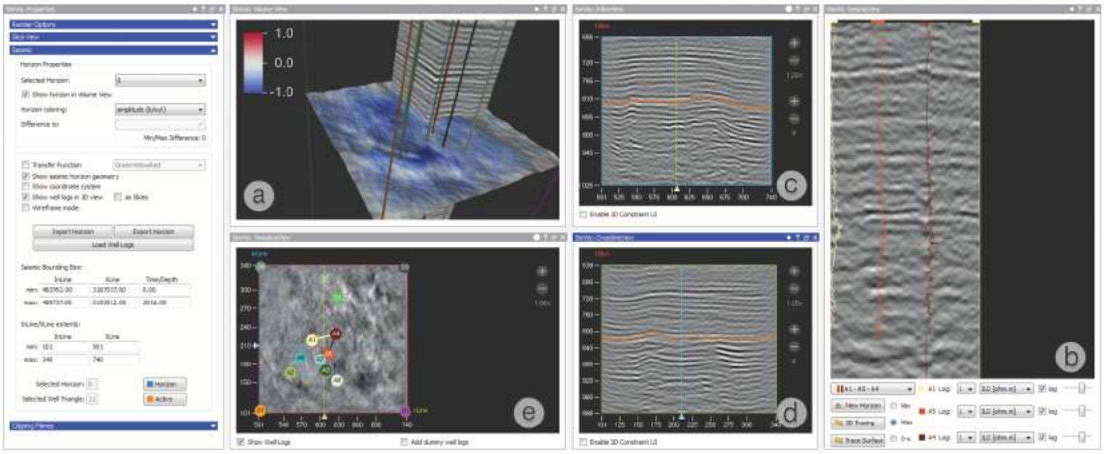


Figure 3.1: The multiple views of Höllt et al.'s framework for interactive seismic interpretation [HBG+11].

Höllt et al. [HBG+11] use an interaction diagram to visualise the linked views of their framework (Figure 3.2). Each view is listed with its actions and a list of artefacts it shows. Arrows indicate the flow of data between the views.

Wu et al. [WSFL18] present a multi-well correlation algorithm. They use the *geological distance* between logs to order and correlate well logs. Starting with longer logs, they order them on a *geologically reasonable path* along continuous structures, then correlate logs sequentially. Correlating longer logs (containing more data) in areas with fewer discontinuities first, and using them as references for subsequent correlations achieves better results than sequentially correlating logs in a random order. Correlations are further weighted with *correlation confidences*, which estimate the certainty of a correlation. Wu et al. focus on the calculation of well correlations rather than their visualisation. Figure 3.3 shows their visualisation of well logs correlated by their algorithm.

One major problem when correlating wells fully automatically is the uncertainty of results. Liu et al. [LSW+19] propose an interactive visual analytics system that addresses this issue. They point out that even the most advanced automatic correlation algorithms produce incorrect results, which necessitates time-consuming manual checks and corrections.

The first part of their contribution is a new pairwise well-log correlation method (Figure 3.4). One well log consists of multiple data channels for each depth value, resulting in around 200.000 values per log. The values are normalised, smoothed, and organised in a matrix (Step 1). Principle component analysis is applied, then a correlation coefficient matrix is constructed, and finally the multiple data channels per log are combined into one value with a weighted summation (Step 2). An *activity curve* is constructed from that data. A value in the activity curve describes the variance of neighbouring values.

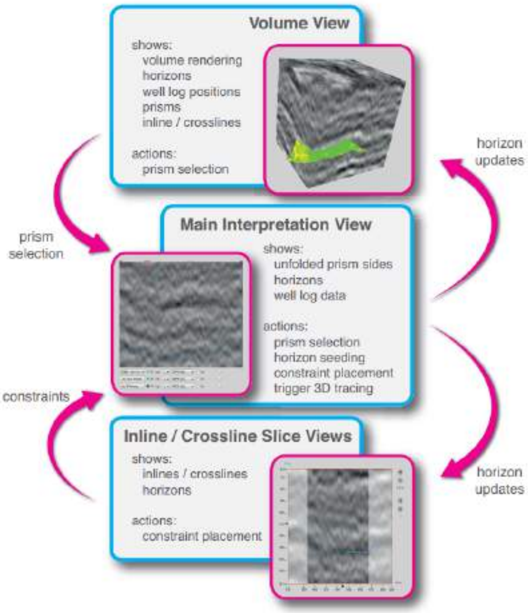


Figure 3.2: An interaction diagram for the linked views of Höllt et al.'s framework [HBG+11].

Values above a threshold on the activity curve are interpreted as boundaries between layers (Step 3). For each layer various statistical features are calculated (mean, variance, thickness, etc.). The similarity between two layers is calculated using the weighted sum of differences between these features. The weights of features can be adjusted interactively by the users. A *match matrix* consists of the similarities of all possible combinations of layers in two wells. Layers in wells are matched by finding an optimal path through the match matrix (Step 4).

The second part of Liu et al.'s contribution is the interactive visualisation platform they present (Figure 3.5). It consists of four coordinated views: The *map view*, where well log locations are displayed in their geographic context; the *correlation view* with detailed views of correlation results of selected well logs; the *matrix view*, which aims to communicate the reasoning behind a certain correlation, and the *attribute view*, which allows users to inspect the original well log data. In the map view, each point represents one well log. A triangulation net is used to connect neighbouring well logs. The connections formed by one layer between logs can be displayed using red lines (*connection net*), and the depth of that layer is shown via a coloured depth contour. Liu et al. designed the correlation view by looking at existing geological visualisations, and choosing a similar encoding. They argue that a familiar visualisation will be more intuitive and reduce users' cognitive burden. Well logs are encoded with rectangles, each correlated layer uses a different colour. Correlated layers in different logs are connected in that same colour. Layers

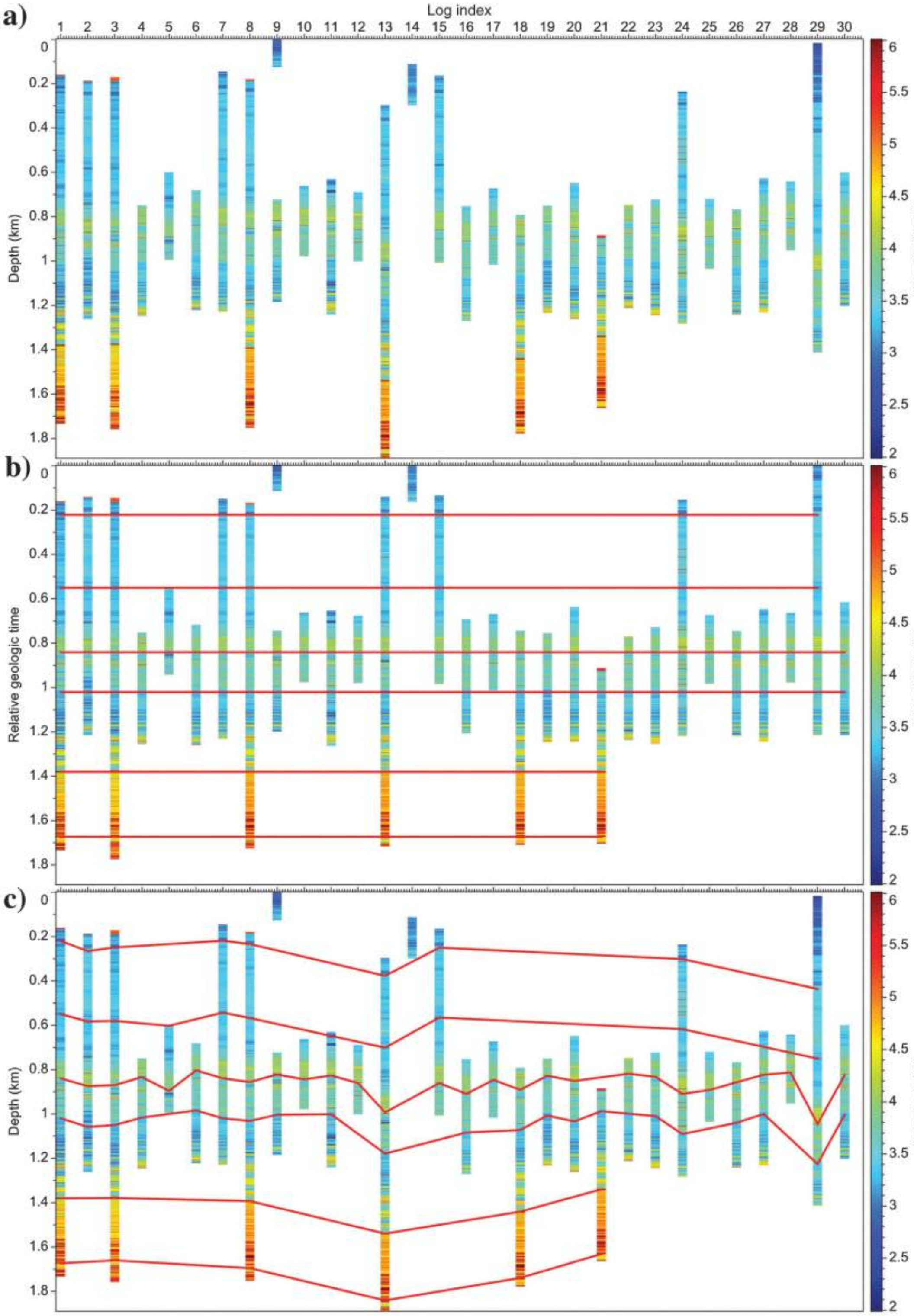


Figure 3.3: Velocity well correlations by Wu et al. [WSFL18]. Uncorrelated logs are (a) ordered using their geological distance, and (b) correlated sequentially. The red lines are the correlations, shown in (c) on the original logs.

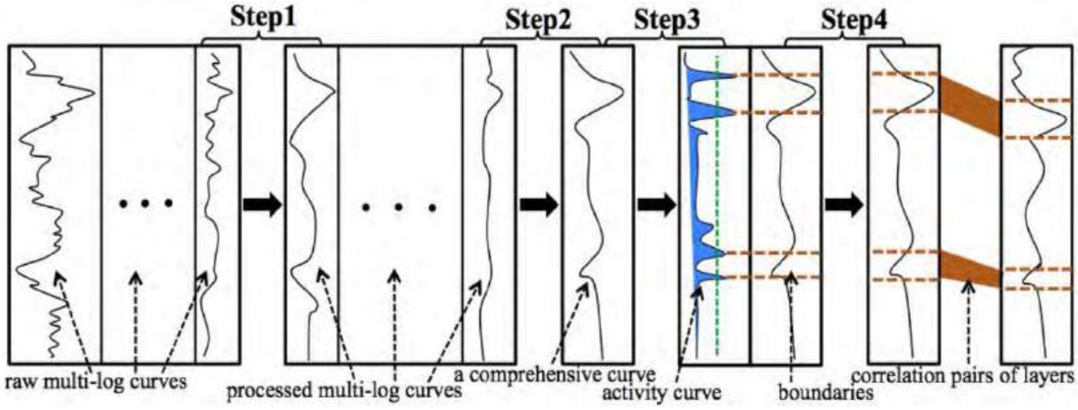


Figure 3.4: The four steps of Liu et al.'s well correlation method [LSW+19].

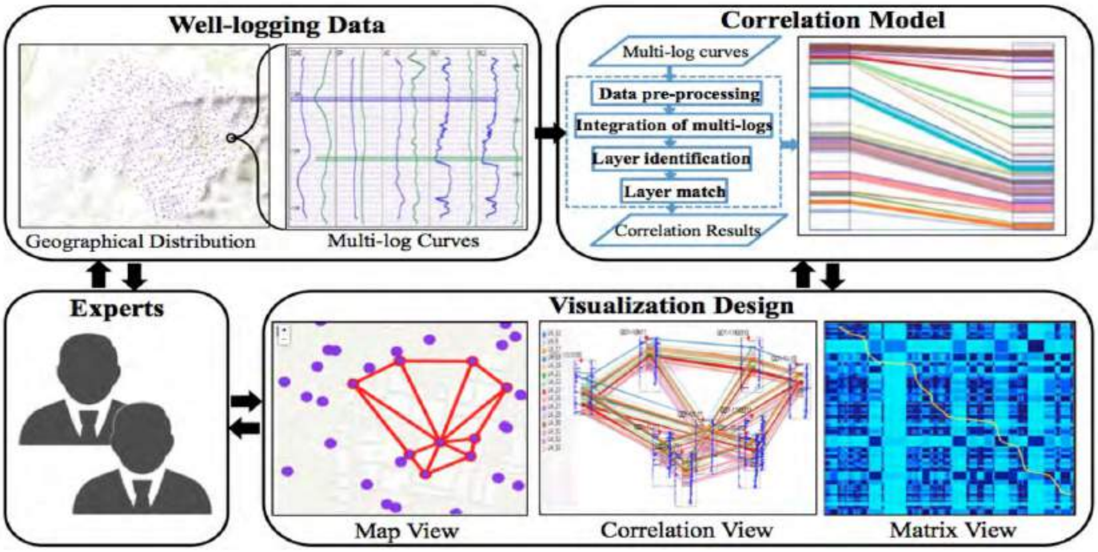


Figure 3.5: The pipeline of the pair-wise well log correlation method by Liu et al. [LSW+19].

that have been incorrectly calculated by the algorithm can be split up or merged by the user in the attribute view and the matrix view respectively. Correlations can be deleted or added by the user as well. Liu et al. presented their visualisation platform to three experts who analysed the wells on an oilfield with the software. After the analysis they were interviewed. One advantage experts reported was that to accomplish the analysis tasks they did with the software they would usually have to use multiple different tools. They also suggested that users should have more control over the correlation process by exposing more parameters of the underlying model [LSW$^+$19].

Liu at al.'s work takes an approach similar to the InCorr design study, but there are some major differences that stem from the application to well logs in the context of oil and mining as opposed to logs of outcrops in the context of planetary geology. Well logs typically cover hundreds or thousands of metres, with one set of measurements taken approximately every 10 centimetres. One oil field might contain as many as 2.000 logs, leading to vast amounts of data. In remote geology, there is, for now at least, no log data derived from drilling bore holes hundreds of metres deep. This means that many methods used for analysing well logs simply cannot be applied to planetary use-cases. Although some aspects of the visualisation are similar (the general principle of the correlation panel is the same), the challenges are quite different. When working with well logs, the main challenge is to use the dense data to automate finding layers and correlations, and to communicate the reliability of these calculations. In planetary geology the data is too sparse to reliably apply the same methods for these tasks, and experts generally trace layer boundaries by hand, and correlate layers manually.

One tool often mentioned in the context of geological analysis and visualisation in the oil and gas industry is Schlumberger's software Petrel [Lim18]. It is a large software suite providing tools specifically for the petroleum industry. Petrel provides a so-called *well-correlation-module*, which can display well logs and connected data (for example seismic data). Logs can be generated automatically using a *well log calculator*, or manually using a *log editor*. New logs can even be estimated by trained neural networks [Lim21].

While the software provides tools for logs and correlation panels, it is tailored to analyse wells rather than outcrops. It also assumes an abundance of available data, which is very far from the reality of planetary science. The visualisation is therefore quite different to the visualisations used in the context of outcrop analysis (compare Figure 3.6 with the correlation panels and logs in Sections 4.3.1 and 4.3.2).

3.2 Visualisation Tools for Digital Outcrop Models and Planetary Data

In this section we will present tools that are used for visualisation and interpretation of digital outcrop models.

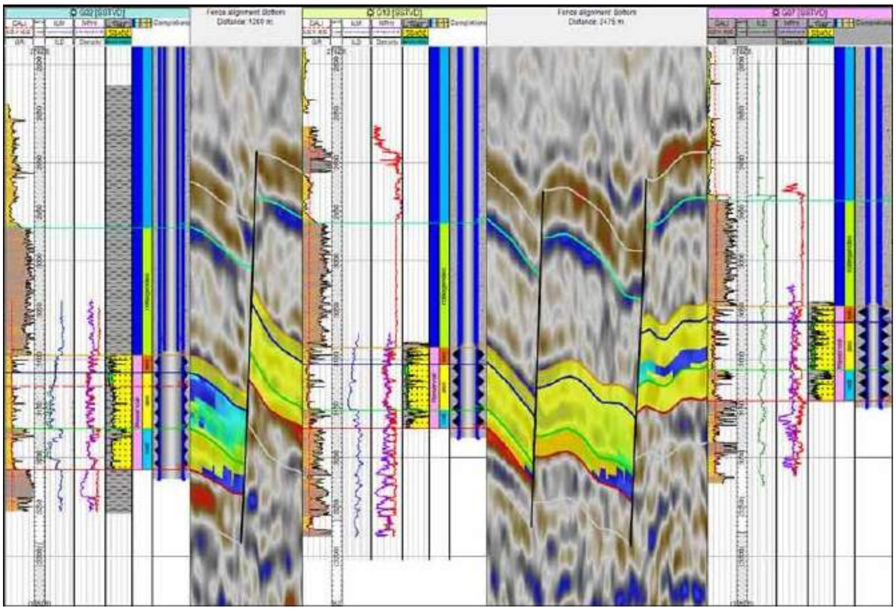
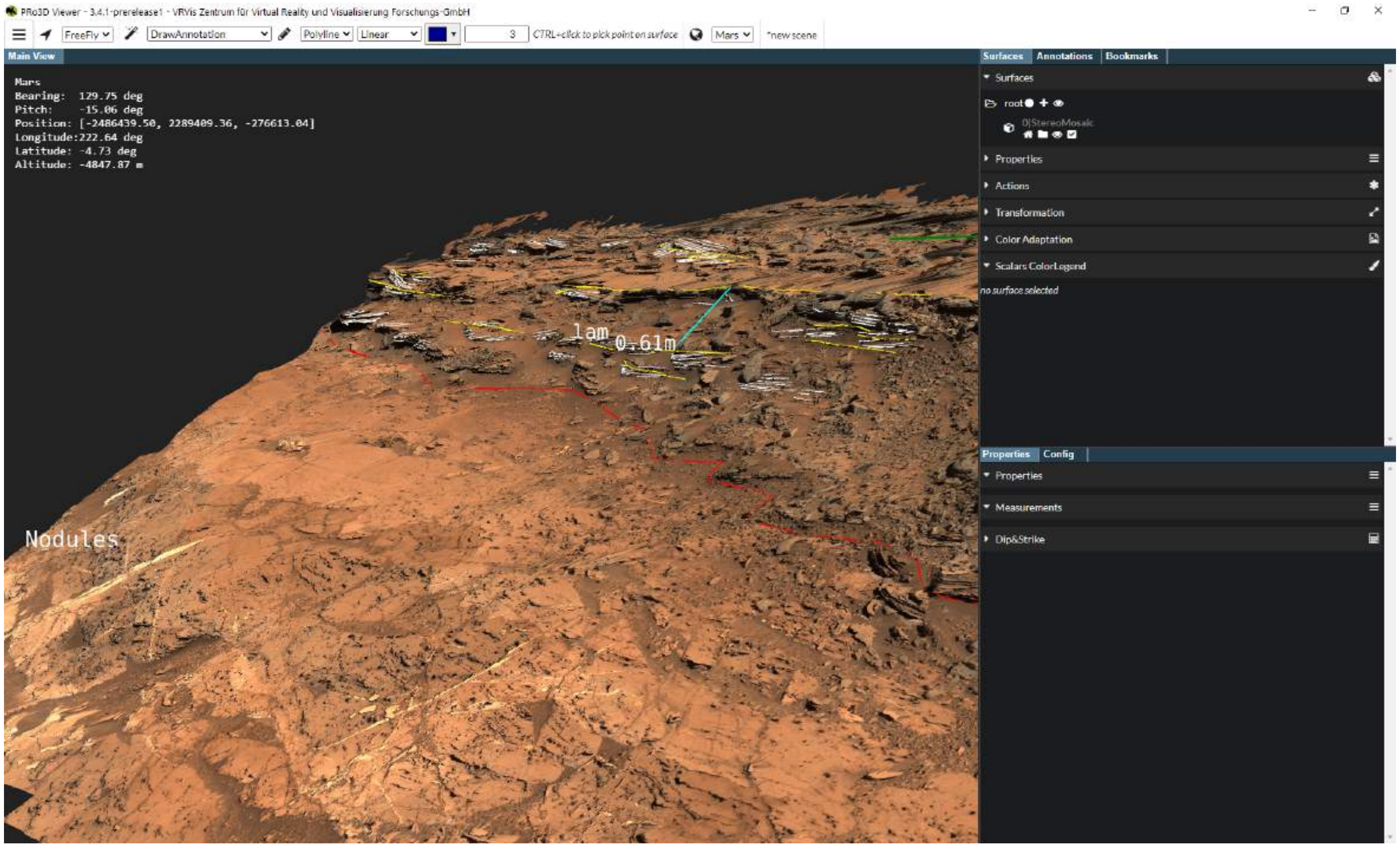


Figure 3.6: The well correlation module in Petrel[Lim21].

3.2.1 PRo3D

PRo3D (Planetary Robotics 3D-Viewer) is a software built specifically to explore and interpret digital outcrop models generated from imagery captured with the stereo-cameras of Martian rovers. Multiple spatially referenced data sets (representing different locations of rovers) can be loaded at once, and can also be combined with orbiter imagery. Digital outcrops can be explored, measured, and annotated in real-time [BGT^{+}18].



```
Mars
Bearing:  129.75 deg
Pitch:    -15.06 deg
Position: [-2486439.50, 2289409.36, -276613.04]
Longitude:222.64 deg
Latitude: -4.73 deg
Altitude: -4847.87 m
```

Figure 3.7: An outcrop loaded in PRo3D with annotations. The data is available online [VRVb] courtesy of NASA/JPL/CalTech/ASU, 3D data processing by Joanneum Research [mbH].

PRo3D is part of the PRoViDE (Planetary Robotics Vision Data Exploitation) framework [PMT+13], [PMT+15]. PRoViDE was devised to give experts access to 3D vision products for data collected on planetary missions, such as digital outcrop models or photomosaic images.

PRo3D can handle large amounts of data and still allows real-time interaction. Although OBJ files can be used in PRo3D, it was designed to work with files in the OPC (Ordered Point Cloud) format, which are created with the PRoViP (Planetary Robotic VIsion Processing) pipeline [PBG+18]. PRoVip is used to generate 3D models from image data collected by stereoscopic cameras as mounted on various rovers, like PanCam, Mastcam, and Mastcam-Z (Figures 2.1 and 2.2).

PRo3D has been used by planetary scientists to analyse Martian terrain through annotating stratigraphic boundaries and sedimentary structures, and making dip and strike measurements [BGG+15],[BGT+18]. Among other use-cases, it has also been extended to generate simulation data for the detection of geological artefacts via deep learning systems [BKP+20],[PTN+20],[TFN+20], and to simulate fly-by sequences [OHB+20].

The prototype created in the course of this work was later integrated into *PRo3D* [VRVb].

3.2.2 Lime

Lime [BRN+19] is a software tool that allows the visualisation and annotation of digital outcrop models in 3D. Annotations can be used to trace line features, create closed polygons, and create three-point planes for measuring and representing orientations. Lime provides extensive tools for adding additional data (especially 2D image data) to the 3D models. Examples for additional data are georeferenced maps or satellite images, subsurface data, or multi-sensor data. 2D images are integrated using projection planes, so-called *Panels*, to project 2D-images onto the outcrop. Lime uses texture layers whose transparency can be set by the user to visualise multiple layers of data at the same location [BRN+19].

Of special interest for this work is that a sedimentary log can be projected onto the outcrop. As opposed to the approach used in this work, which addresses a unified approach of annotation and log creation in the same tool, a log for use in Lime has to be created with a different software. The log is imported into Lime as a 2D image, and projected onto the scene. The placement is selected manually by the user. Figure 3.8 is taken from Buckley et al. [BRN+19], and shows a log positioned in the 3D scene and projected onto the outcrop.

3.2.3 VRGS

Virtual Reality Geological Studio (VRGS) [HGWR07], [FPHR10],[RVLH+14],[Hod17] is another tool for geoscientists who work with digital outcrop models. VRGS works with point clouds rather than meshes. It allows scientists to inspect and analyse digital

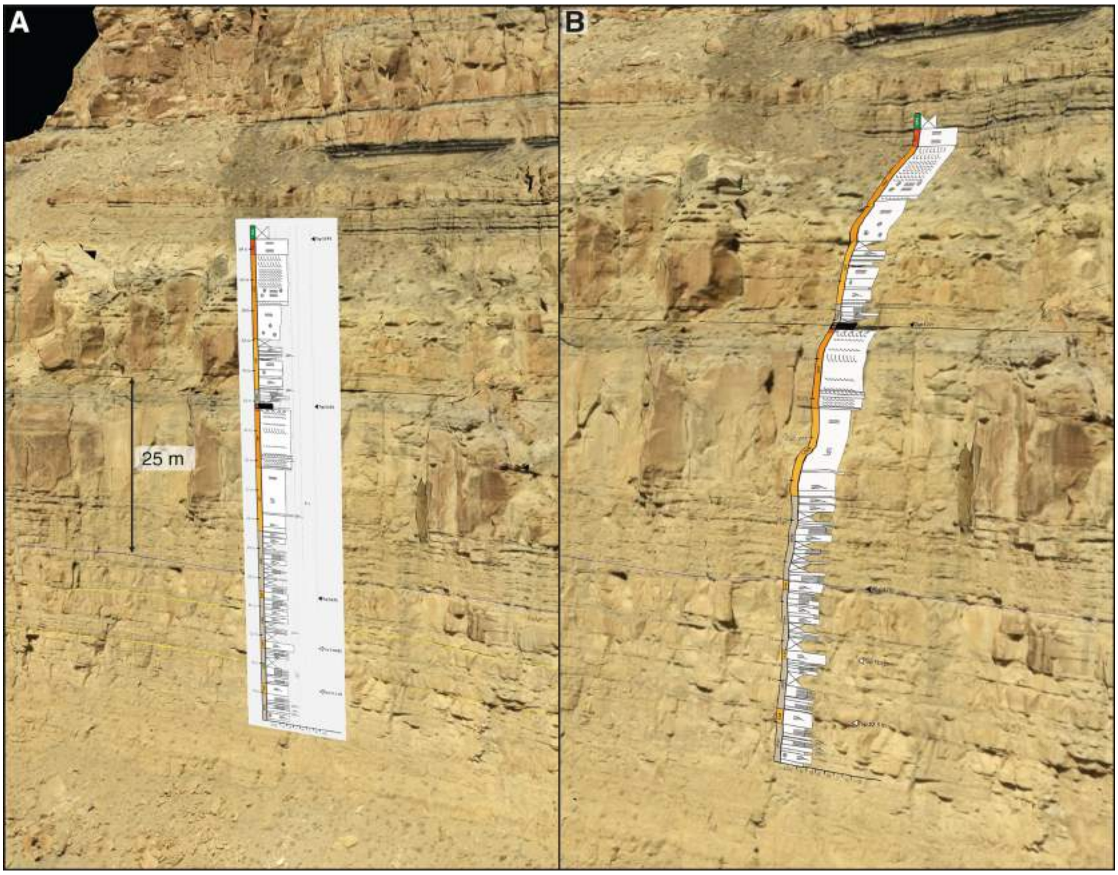


Figure 3.8: Integration of a sedimentary log in the Lime software. An image can be positioned in the 3D scene (A) and projected onto the surface of an outcrop (B) [BRN+19].

outcrop models. The software includes the visualisation of logs on a 3D view of an outcrop (Figure 3.9).

Rarity et. al. [RVLH+14] produce 3D-views including sedimentary logs similar to the one in Figure 3.9. As in Lime [BRN+19], the sedimentary log is imported into VRGS, rather than created in the software. The log has to be georeferenced correctly beforehand by the user in order to achieve the correct placement in the 3D scene. The placement of individual units then needs to be corrected by linking the units of the log to the units in the digital outcrop model. This means that, opposed to the approach of Lime, where logs carry no semantic information, the logs in VRGS are semantically linked to the digital outcrop model.

3.2.4 MOSIS

The selling point of the Multi Outcrop Sharing and Interpretation System (MOSIS) is that it uses virtual reality (VR) to provide an immersive experience for geoscientists

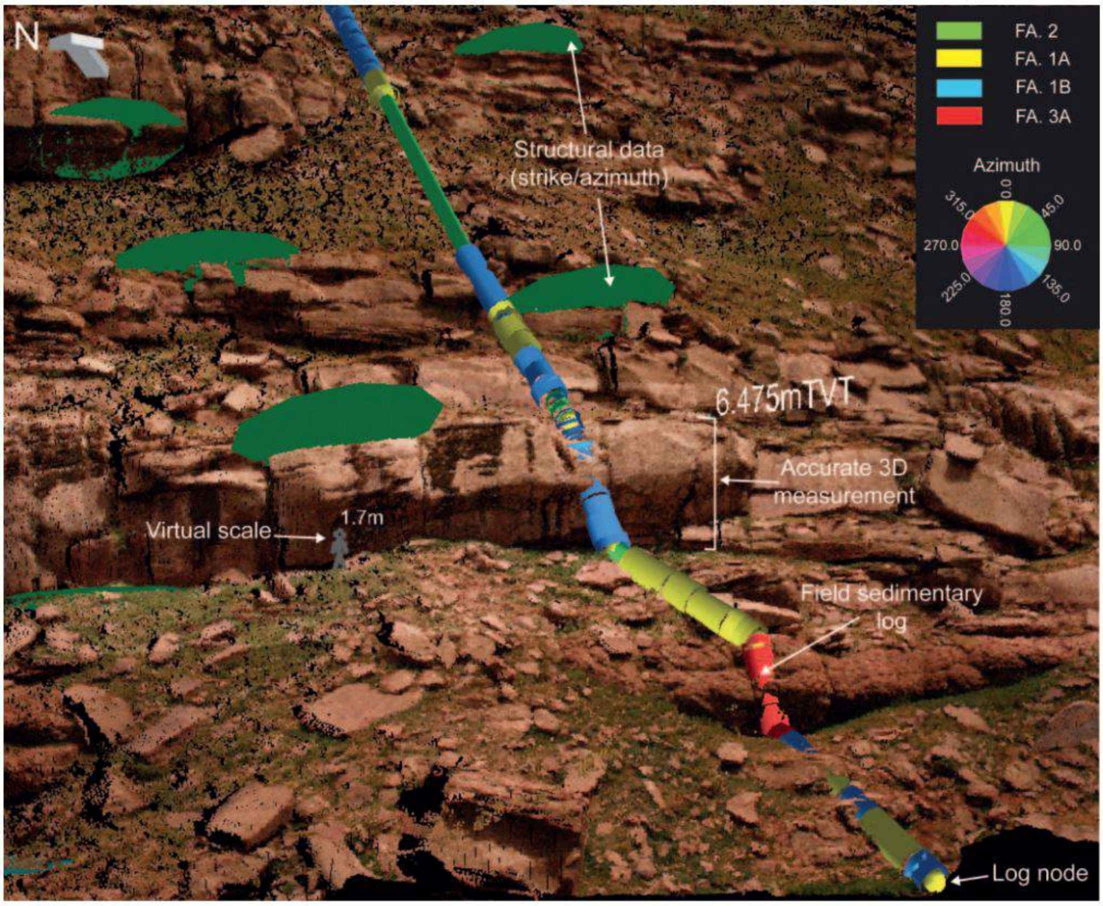


Figure 3.9: A sedimentary log in the VRGS software [FPHR10]. The authors use cylinders to represent the log in 3D, the width of the cylinder indicating grain size. The colour encodes the facies type.

(Figure 3.10). What sets MOSIS apart from the tools discussed above, in addition to its focus on virtual reality using head mounted displays, is that it includes not only 3D reconstruction from LiDAR and multiple 2D images, but also tools for sharing digital outcrop models via the internet. Tools for interpretation include drawing lines on the surface of the digital outcrop model, measuring distances, visualising planes, calculating dip and dip direction, and adding markers [GVA+17].

3.2.5 Flattening and Unfolding of 3D Data

In a wider sense, a log could be seen as a flattened visualisation of an outcrop. Flattening or unfolding 3D data into a 2D visualisation has been used in many areas, for example molecular visualisation [BLMG+15], vascular studies [KFW+02], and gastroenterology [HGQ+06]. Byška et al. [BLMG+15] present a 2D representation of protein tunnels

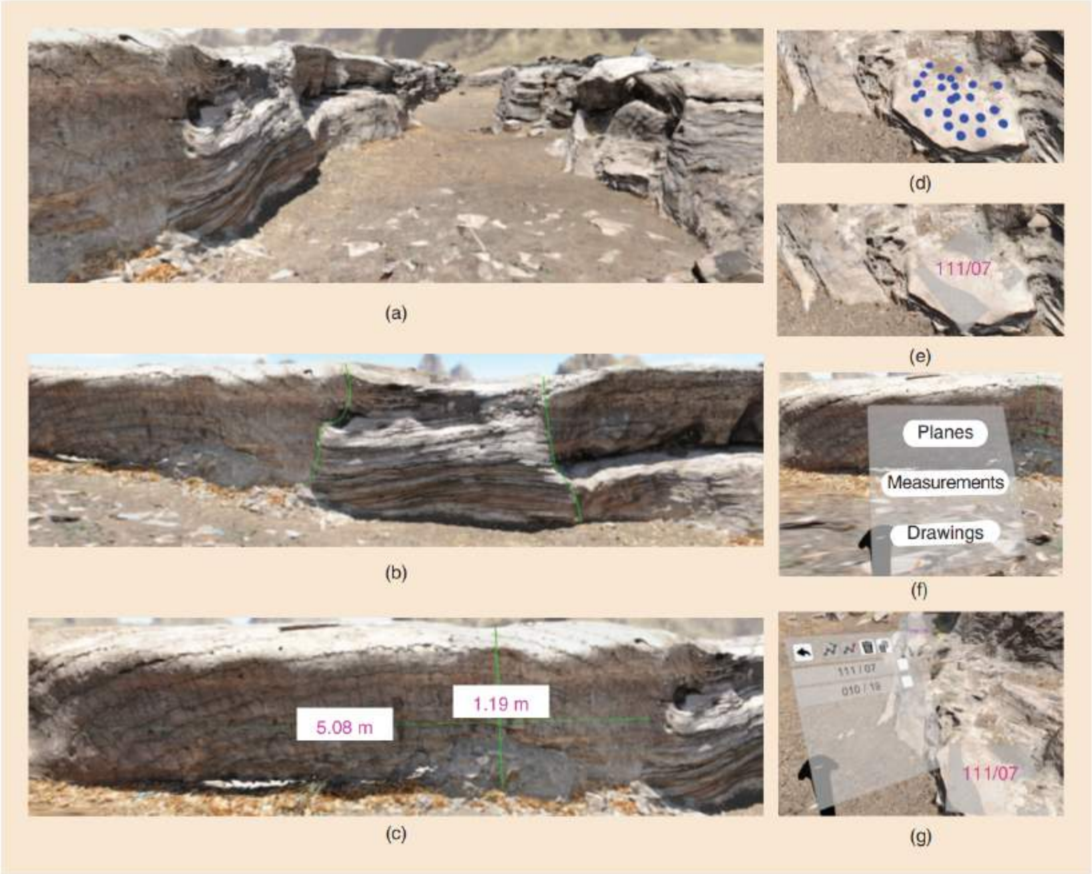


Figure 3.10: Multi Outcrop Sharing and Interpretation System (MOSIS): (a) An outcrop can be (b) annotated and (c) measured. By (d) selecting multiple points on a surface a plane (e) can be interpolated. Figures (f) and (g) show the user interface of the VR application [GdSJVK+18].

over time. Kanitsar et al. [KFW+02] present different methods for *Curved Planar Reformation*, where cross-sections of tubular structures are displayed in a curved surface. Medicine and chemistry are areas in which the technique of flattening is especially of interest [KMM+18],[ZS09].

The degree to which original data is preserved or transformed varies strongly. As opposed to the examples above, the 2D encoding of an outcrop is not derived directly and only from the digital outcrop model, but relies heavily on annotations added to the digital outcrop model by experts. The actual surface shape of an outcrop is of little interest in the context of a geological log, and the appearance is encoded using patterns or colours rather than original image data [HGQ+06].

Kreiser et al. [KMM+18] use a coding system to classify flattening techniques in medicine (Figure 3.11). The most important aspect of that system is the concept of *preservation* *characteristics*. For their use-cases they distinguish techniques by whether they preserve *areas* or *angles*. The preservation characteristics determine what type of measurement can be performed on the encoded data.

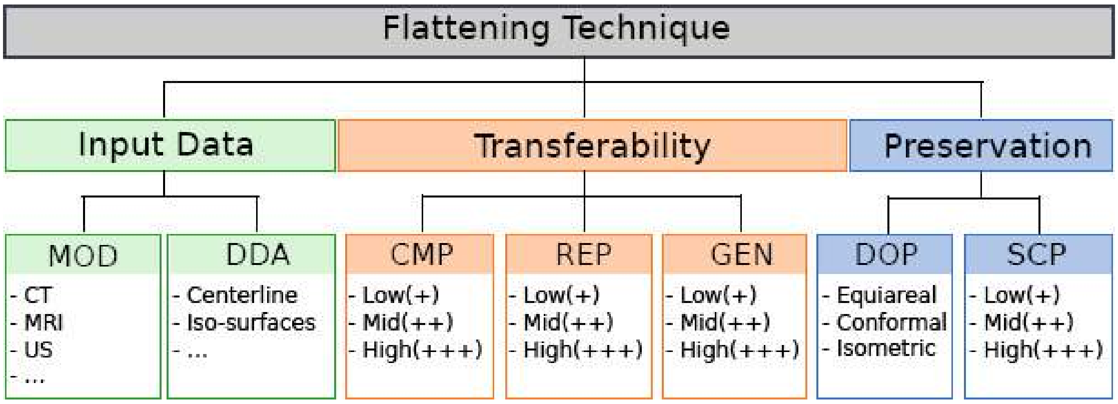


Figure 3.11: The taxonomy of the coding system for flattening techniques devised by Kreiser et al. [KMM+18]. The abbreviations from left to right: Data Acquisition Modality, Derived Data Attributes, Comparability, Reproducibility, Generalizability, Degree of Preservation, and Spatial Context Preservation.

The taxonomy by Kreiser et al. [KMM+18] was created specifically for the field of medicine. Generalizability, for example, refers to whether a technique can be used for different organs. It might be valuable to generalise it for 3D-to-2D visualisation techniques. Considering the projection from 3D data to 2D representations in terms of which attributes are *preserved*, for example, could be an approach to achieve some classification and characterisation of flattening techniques.

CHAPTER 4

Towards Interactive Data-Driven Correlation Panels

In this chapter, we present a task abstraction, chart the design space, and determine the essential elements of a viable prototype. We will then present our visualisation and interaction design in Chapter 5.

Taking the design study methodology of Sedlmair et al. [SMM12], this chapter deals with the *discover* and *design* stages. During the *discover* stage we aim to understand and characterise the problem as domain scientists see it, and start working towards a solution in the design stage. As Sedlmair et al. put it, "*design* at this stage is the generation and validation of data abstractions, visual encodings, and interaction mechanisms [SMM12]".

4.1 Task Abstraction

Geologists create correlation panels to build a geological model of a wider area. To do this, they first interpret outcrops by annotating them. On digital outcrop models, this means, for example, to trace contacts with an annotation tool. Once some annotations are present on a digital outcrop, they can use them to derive logs. Once at least two logs are present, the geologists can correlate strata, thereby creating a correlation panel. The tasks we address in this work are the following:

- **T1** Create new semantic annotation types.
- **T2** Annotate outcrop.
- **T3** Create a log and assign grain sizes.
    - **T3a** Select log positions and create log.
    - **T3b** Assign grain sizes to strata.

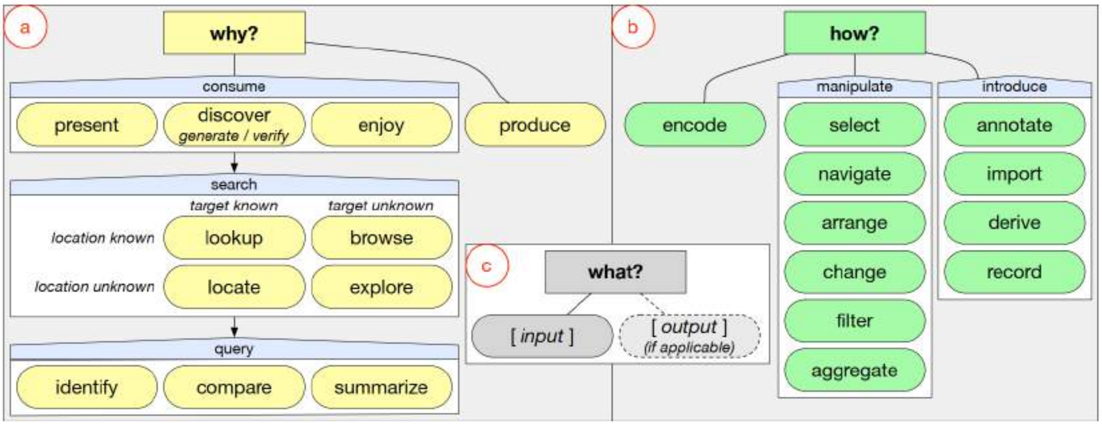


Figure 4.1: Brehmer et al.'s multi-level typology of abstract visualisation tasks [BM13]. The typology contains three categories: *why a task is performed* (yellow), *how a task is executed* (green), and the task *inputs and outputs* (grey).

- **T4** Order logs and correlate strata.
    - **T4a** Order logs.
    - **T4b** Correlate strata.

Brehmer et al. defined a multi-level typology of abstract visualisation tasks (Figure 4.1). We follow this typology to abstract the tasks T1 to T4 (Figure 4.2).

In T1 geologists create new semantic annotation types, which they then use to annotate the outcrop (T2). We split T3 into two subtasks. First, geologists identify and select suitable points on annotations to create a log (T3a), then they assign grain sizes to strata (T3b). Next geologists arrange the logs in the correct order (T4a), and lastly correlate strata between logs (T4b).

4.2 Data Model

We use the principles of functional domain driven design [Wla18] to create our data model. Therefore we name our data types according to the domain. We now enrich the task abstraction with our data model. The geologist creates multiple instances of the domain type `Semantic` (T1). Next the geologist creates multiple instances of the domain type `Annotation` (T2). Each annotation references a `Semantic`. We denote this by `Annotation * Semantic`. In T3a the geologist selects points on some of the previously created `Annotations`: `Point * Annotation * Semantic`. We call these points `Contacts`. From them we can derive a `Stratum` in the log with its upper and lower contact: `Contact * Contact`. All strata derived from the geologist's selection in T3a form a `Log`. They are stored in a tree data structure reflecting the hierarchical nature of strata. In T3b the geologist assigns grain sizes to strata: `Stratum * GrainSize`. In

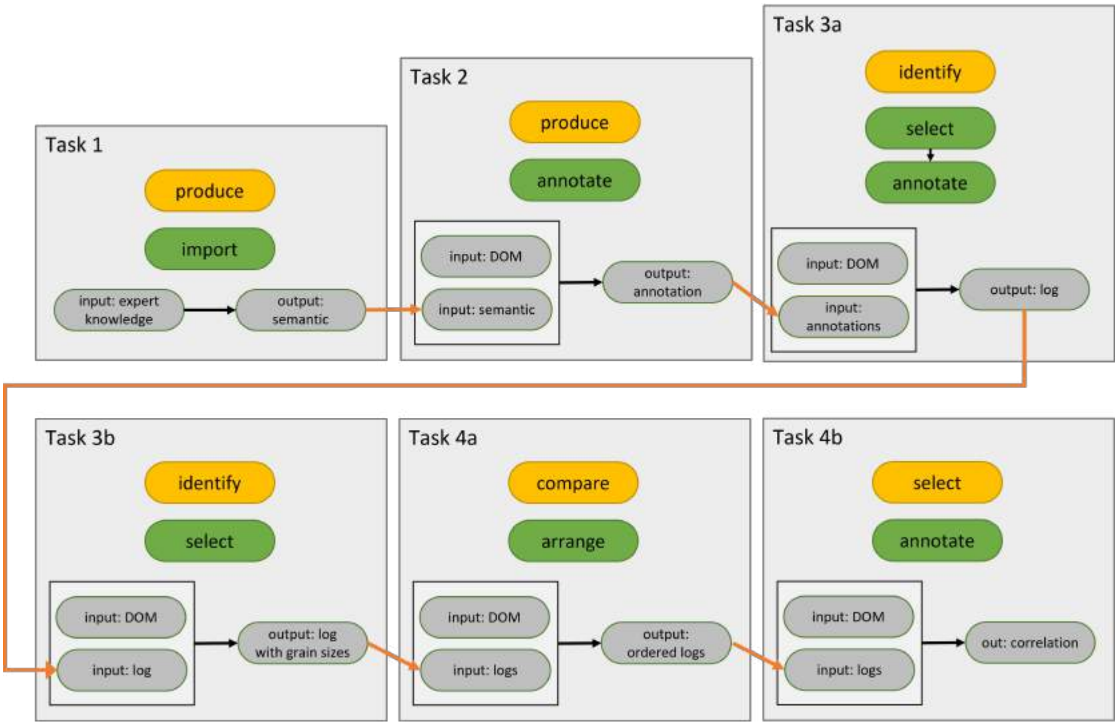


Figure 4.2: Task abstraction following Brehmer et al.'s multi-level typology of abstract visualisation tasks [BM13]. In T1 geologists create new semantic annotation types, which they then use to annotate the outcrop (T2). We split T3 into two subtasks. First, geologists identify and select suitable points on annotations to create a log (T3a), then they assign grain sizes to strata (T3b). Next geologists arrange the logs in the correct order (T4a), and lastly correlate strata between logs (T4b).

T4a they order the created logs: `Log * Index`. Correlating `Contacts` between logs (T4b) leads us to a list of the type `Contact * Contact`.

4.3 Design Space Exploration

One consideration for the design of interactive correlation panels is that the visual encodings are, to a degree, predefined. The correlation panel is an established visualisation, and deviating from established abstractions and encodings is not in the interest of domain experts. Therefore, the focus of this work must be to find the best solutions following these constraints, rather than inventing a new visualisation. To be able to do so, we need a broad understanding of the visualisation as it is used today.

Geological logs and correlation panels come in a wide variety of styles. The granularity and types of information vary as well as the visualisation style, or the encoding, thereof. This makes it necessary to investigate which features of a log or correlation panel are

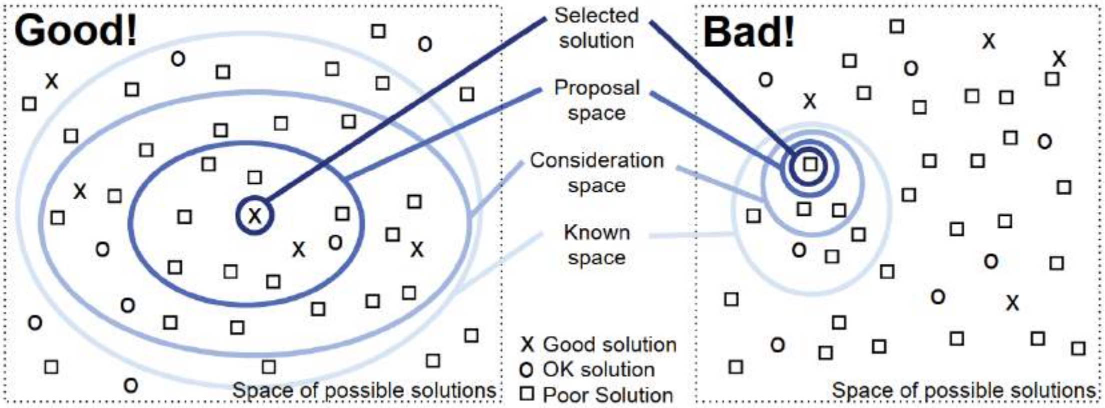


Figure 4.3: The perils of starting out on the design process before having extended the *known space* sufficiently [Mun14].

indispensable, which are common, and which can be attributed to an author's idiosyncratic approach. Some features can also be expected to have little use for data derived from rover images. Fossils for example are unlikely to play any role in the analysis of image data captured on Mars.

To put it into the context of visualisation research, this section aims to extend the *known design space*. Munzner explains the need to start with a known space as large as possible succinctly with a figure in her book *Visualization Analysis and Design* (Figure 4.3) [Mun14]. The smaller the known space compared to the space of possible solutions is, the higher is the risk of not even considering options that might be superior to the ones present in the known space.

In Sections 4.3.1 and 4.3.2 we present and analyse examples of logs from different sources. The first source are books aimed at geology students, which give samples of logs and explain how they are created. The second source are published papers, where the authors use logs to communicate their findings.

In Sections 4.3.3 and 4.3.4, building on the analysis and information gathered by consulting experts, we aim to answer the following questions concerning geological features:

- How commonly is a feature used? Is it indispensable or nice-to-have?
- Which visualisation should be chosen if a feature is visualised in different ways by different authors?
- Should a feature use a pre-defined style or allow users to select from multiple styles, or even let users generate their own styles?
- Where is it necessary or advisable to restrict users to a common style to facilitate team-work and interpretation?

To answer these questions, we evaluate the features of the logs and correlation panels from section 4.3. We also discuss high-level implementation options where applicable. In section 4.3.4 we provide a structured and coherent list of features which can then be used as a requirements catalogue for the implementation of a minimal viable prototype. This catalogue contains not only necessary features, but also features that can be used to extend the prototype.

4.3.1 Logs

The log form in Figure 2.8 uses the elevation above base (of the outcrop) on the vertical axis. It is labelled on the left hand side of the form. The thickness of each stratum (in this case beds) and the bed number are the first two columns. Next is the lithology, which is illustrated with patterns. A legend for the patterns and symbols used in the log can be found in Figure A.1 in the appendix. The grain size is plotted in the column headed *texture*, with grain sizes ranging from small (clay & silt) to large (gravel). Patterns are used to further illustrate the nature of sediments. These patterns are explained in Figure A.1 under the heading *siliciclastic sediments.* In the column *sedimentary structures*, various symbols are used to illustrate features, again explained in Figure A.1. Palaeocurrents are encoded by simple arrows in a separate column. To indicate the presence of certain fossils, the author uses symbols which are also explained in Figure A.1. The colour of beds is described with abbreviations. The right-most column is labelled *remarks.* Noteworthy for our purposes is the reference to photographs in this column.

The log in Figure 4.4 contains multiple columns. The first two columns contain vertical labels covering one or multiple strata. In this log, strata correspond to beds. The third column contains a horizontal label with a number for each bed. Next is the graphic description of strata. The vertical axis is labelled on the left and encodes the scale. An annotation describes the vertical scale of the log. The horizontal axis encodes the grain size of strata, with labelled vertical lines below the graphic illustration specifying individual grain sizes. The strata themselves contain patterns for grain size as well as two other geological features: Ripples, and stratification. The demarcations of strata are irregular, and might be relatively straight and horizontal, oblique, or even curved, signifying, as the annotations tell us, sharp planar boundaries or sharp erosive boundaries. A gradual change in grain size (again, according to the annotations) is visualised by rounding the right-hand side of a stratum.

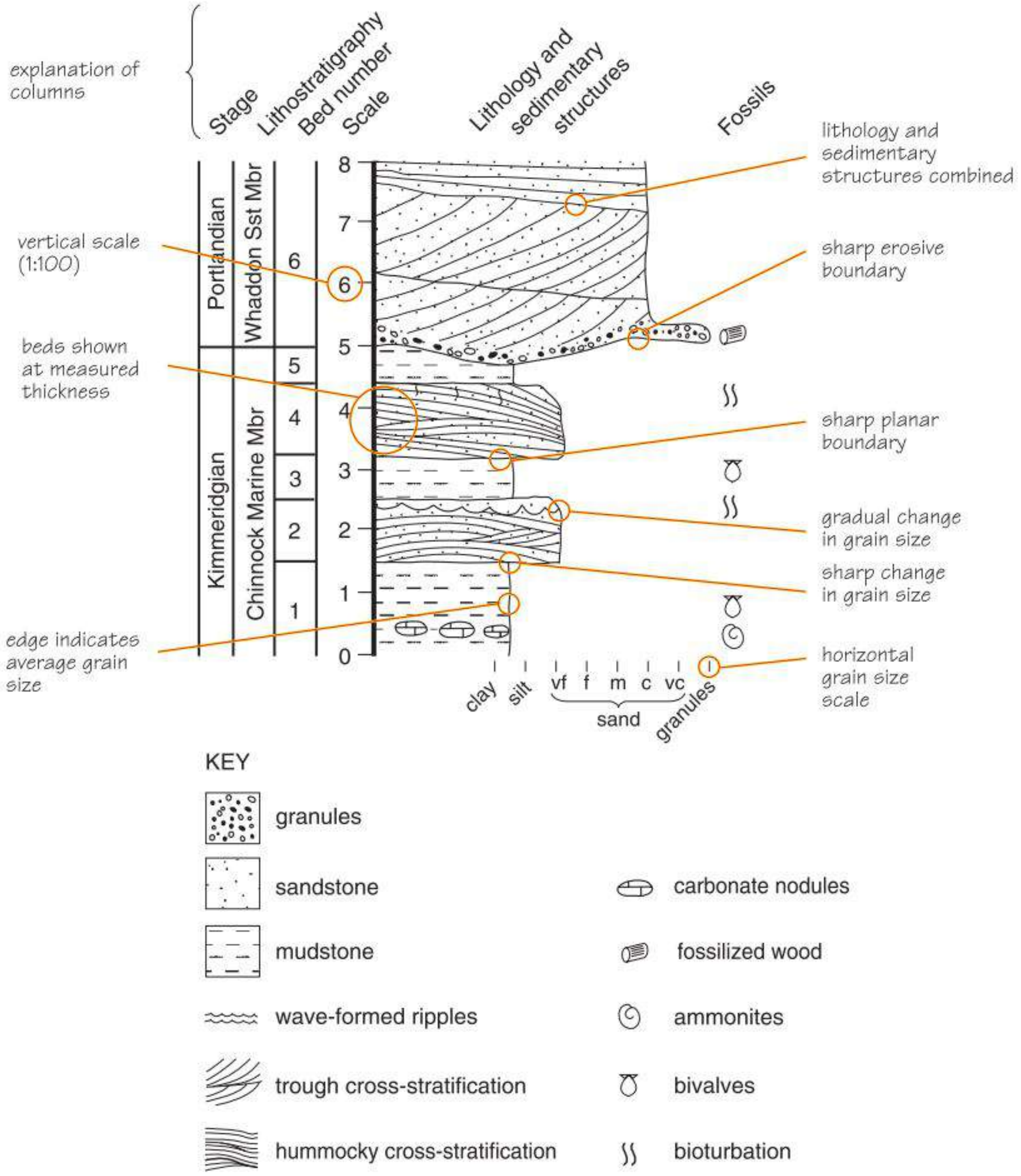


Figure 4.4: A graphic log with explanatory notes. The author points out that this is a *neat* version of a log rather than one created in the field. The field log might contain more columns, notably with photographs and links to other information. (Geological Field Techniques by Angela Coe [Coe10], page 118)

Collinson et al. [CMT06] provide examples of graphic logs which contain less detailed information than the ones presented in Figure 2.8 and Figure 4.4. The three logs in Figure 4.5 only contain two (a, c) and one (b) column. The log in Figure 4.5a is divided into a column for the grain size, and a column for lithological structures which are presented graphically with additional textual descriptions. The log in Figure 4.5b combines these two columns into one, using the grain size values to adapt the width of the column containing lithological structures. For the log in Figure 4.5c a column to illustrate the proportion of sand and mud was added on the left-hand side. Arrows pointing upwards to the right of strata illustrate that grain sizes get smaller with higher elevation. Unlike logs (a) and (b), log (c) does not contain textual comments. The scale of each log in Figure 4.5 is indicated by a small labelled ruler at the bottom. The labelled horizontal axis (the grain size) is adapted for each log according to the grain sizes relevant for that log. For example, whereas the first log (Figure 4.5a) uses the label *sand*, and the second and third logs distinguish *very fine, fine, medium, coarse*, and *very coarse sand* (Figure 4.5b), and *fine, medium*, and *coarse sand* (Figure 4.5c) respectively.

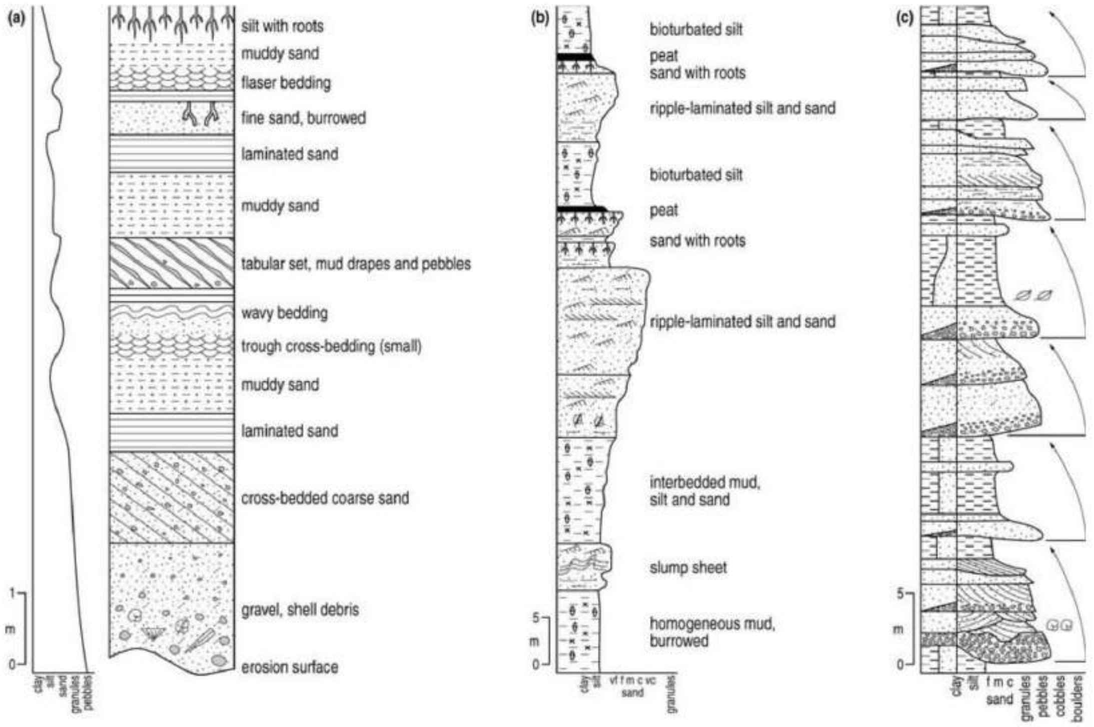


Figure 4.5: Three examples of visual logs. Log (a) separates the visualisation of grain size and other lithology, devoting one column to each. Log (b) combines these two properties in one column. Log (c) contains an additional column on the left-hand side to indicate the proportion of grain types in strata, and replaces the textual comments in logs (a) and (b) with symbols and patterns. A legend for these symbols and ornaments can be found in Figure A.2 ([CMT06], page 245).

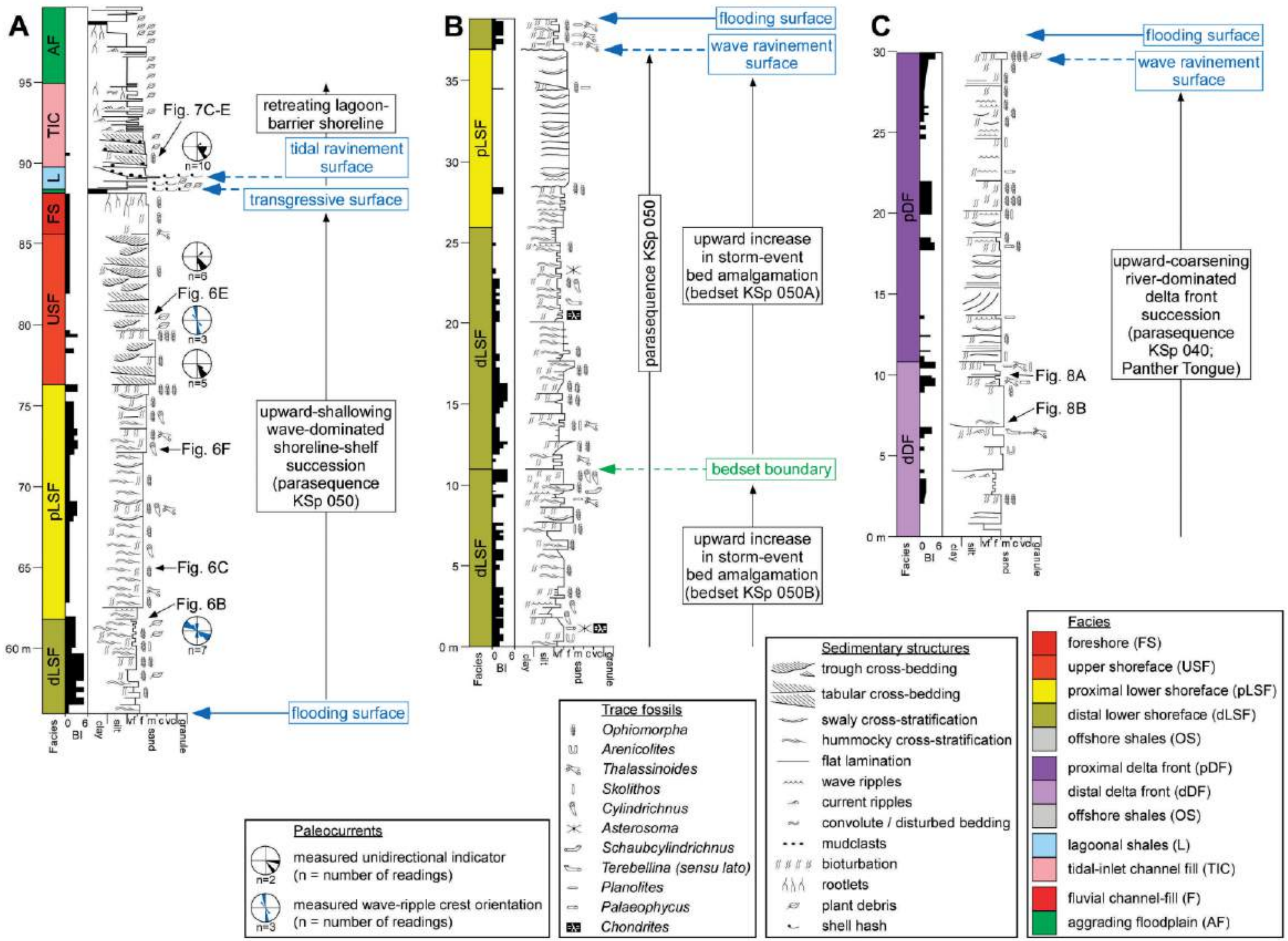


Figure 4.6: An example of three visual logs published in a paper by Hampson et al. [HGS$^+$11]. Although three logs are presented next to each other, this is not a correlation panel. Notably, these logs contain an additional column with facies information. Glyphs encoding the palaeocurrent flow are placed to the right of the main column of the log.

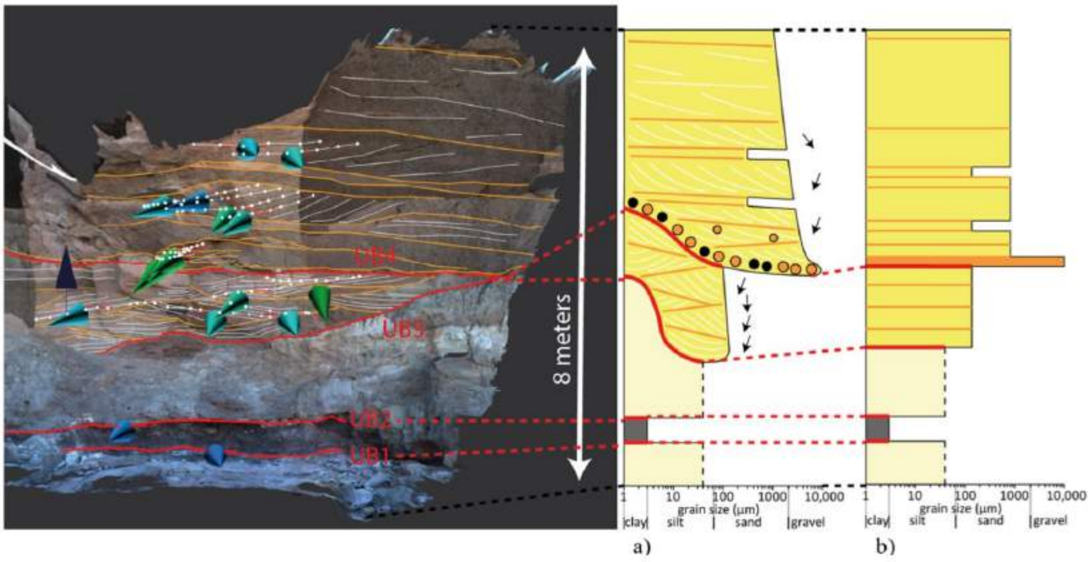


Figure 4.7: A log drawn by one of our collaborators for the purpose and at the beginning of this project. On the left-hand side we can see an annotated digital outcrop model. (a) To the right of the digital outcrop model is a log as the expert would draw it for a publication, (b) on the right-hand side a simplified version thereof. The simplified version is what the expert imagines a computer generated log might look like.

The logs presented in Figure 4.6 again present grain size and lithology in one column. Palaeocurrents are illustrated with rose diagrams, which are positioned to the left of the corresponding sections of the combined grain size and lithology column. References to figures are realised as floating labels with arrows pointing to the exact spot in the corresponding column. Two additional columns are used in these logs. The first column is coloured and contains abbreviations referencing facies. The second column plots the *bioturbation intensity* (labelled *Bl*) vertically. The area covered by the biodiversity function is coloured in black. Legends concerning abbreviations, symbols, and colours can be found at the bottom of the figure. To the right of each log, framed notes add information to points or areas in the diagrams. The vertical axis of the logs encodes the elevation, labelled every five metres to communicate the scale of the logs.

Figure 4.7 was created by a domain expert for the project of this thesis. On the left-hand side we can see a digital outcrop model. This digital outcrop model has been annotated. The red and orange lines encode contacts. The orange lines encode subordinate contacts, i.e., the contacts of strata within larger strata. These larger strata are delineated by red lines.

A log drawn by hand for publication (a) is placed next to the digital outcrop model. The grain size is encoded using colours and the width of strata. Some of the contacts between strata in the log are drawn using curved lines. The black and orange circles

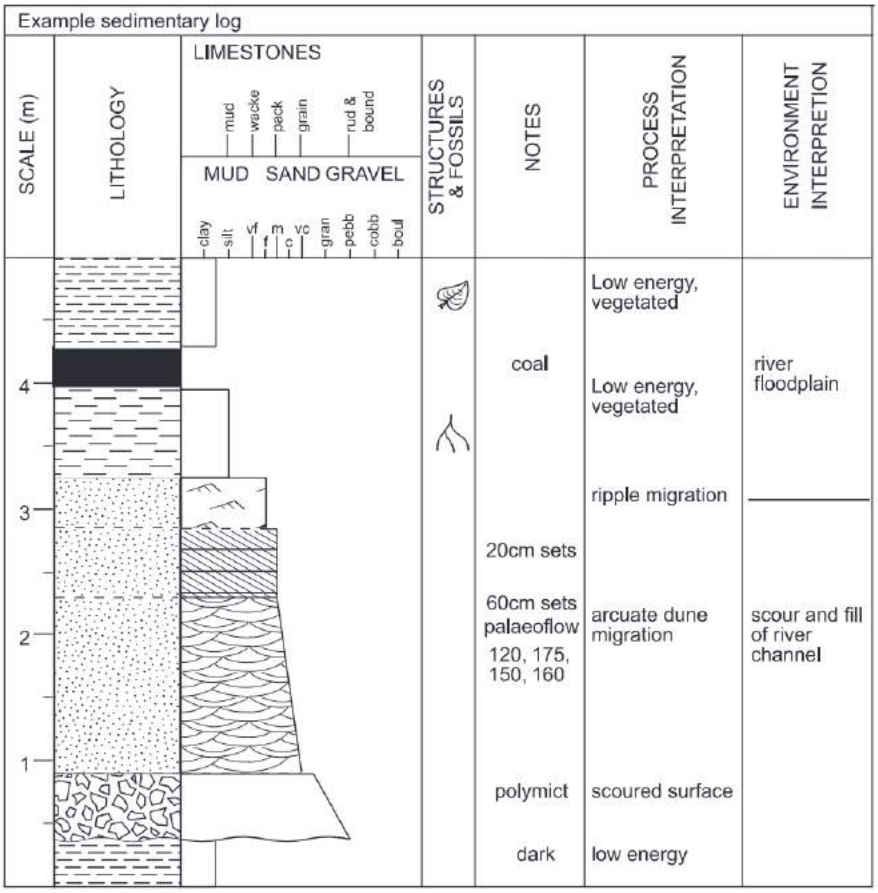


Figure 4.8: An example of a log from Sedimentology and Stratigraphy, 2nd Edition [Nic09], page 71. Lithology and grain size are presented in separate columns. Four columns with additional information are added on the right-hand side.

on the log indicate a river channel; they encode pebbles or pieces of gravel. The log on the right-hand side (b) was drawn (again, by hand) by our collaborator as a simplified version, that he imagines could be computer generated. We can see that he omitted the symbols used to encode additional geologic attributes. He also replaced the curved borders of strata (contacts as well as the gradual changes in grain size) with straight borders. The grain sizes are labelled below the logs ranging from clay to gravel.

In *Sedimentology and Stratigraphy* by Nichols [Nic09] we find another log with lithology and grain size in separate columns (Figure 4.8). The scale is marked on the left-hand side of the log. The lithology column uses patterns to symbolise rock types. Unlike the log in Figure 4.5a, patterns and symbols are also used in the grain size column of this log. Another feature that sets this log apart from the ones before it, is the use of two different scales for grain size. Above the standard scale there is a separate scale for limestones. Symbols for structures and fossils are presented in a separate column rather than floating next to the grain size column as in Figure 4.6 or superimposed on the lithology column (Figure 4.5a) or the combined grain size and lithology column (4.5b). The legend for the patterns and symbols used by Nichols can be found in Figure A.3. Next to a column for general notes, the log also provides two columns for interpretation: *process interpretation* and *environment interpretation*. The interpretation takes the form of textual notes.

Comparing the legends presented in Figures A.2, A.1, and A.3, it is interesting to note that while there are certainly commonalities, this is another area that lacks strict standardisation. For one, categories or category names differ. Many ornaments are similar or even identical, but sometimes a completely different visualisation is chosen. Tucker for example uses no pattern for his category *clay, mudstone*, while Collinson et al. [CMT06] use a line pattern for their category *mudstone/clay stone*. Nichols [Nic09] (Figure A.3) uses separate categories for mudstone and clay stone. Figure 4.9 shows the four different patterns next to each other.


(a) (b) (c) (d)

Figure 4.9: Different patterns used for the same category of clay and mudstone by Collinson et al. [CMT06] (a), and Tucker [Tuc03] (b). Nichols [Nic09] uses separate categories for clay (c) and mudstone (d).

Other examples where Collinson et al.'s and Tucker's styles diverge are the pattern for *matrix-supported conglomerate* and the symbols for *asymmetrical* and *symmetrical ripples*.

Many of the logs presented up to this point contain irregular (non-horizontal) stratum boundaries. In reality, the contacts between strata are not necessarily sharp, straight boundaries. This may or may not be reflected in visual logs. Compare the logs drawn by our collaborators in Figure 4.7: in the more detailed log (a) some contacts are curved. The simplified version (b) has horizontal borders between strata. The detail in which properties of contacts are depicted, and whether this is done via text labels or graphical means varies as well (compare for example Figures 4.5a, b, and c). The simplest method of encoding contacts are horizontal lines 4.5a. However, many logs use curved and wavy lines as well where applicable. Tucker[Tuc03] presents a collection of visual encodings for different types of contacts 4.10.

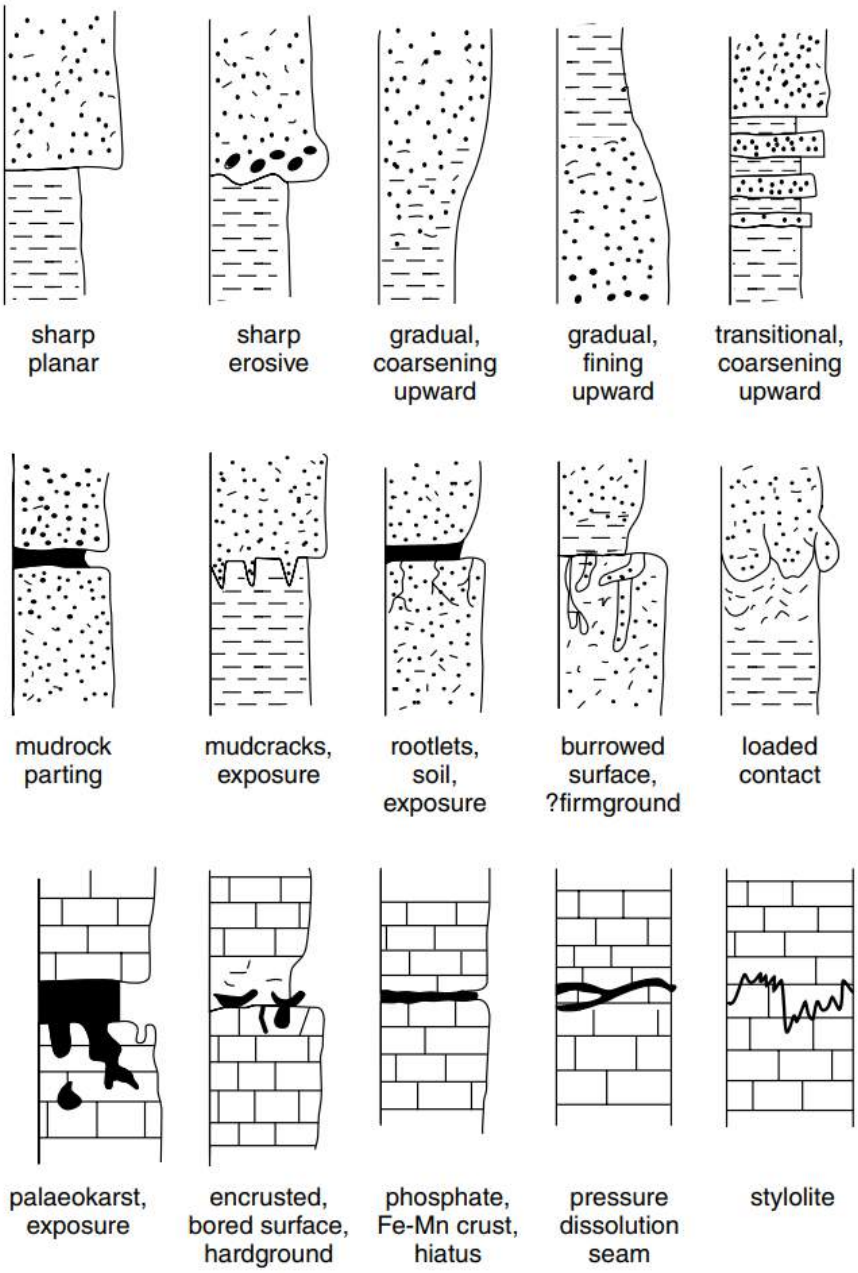


Figure 4.10: Different encodings for types of contacts in visual logs. From *Sedimentary Rocks in the Field, 3rd Edition* [Tuc03], page 90.

4.3.2 Correlation Panels

When investigating geological features that extend over large areas, geologists create multiple geological logs. Logs are placed next to each other with a visual encoding of the local variations in geological strata that occur in one area. Geologists analyse the attributes of the strata within the logs to identify strata that occur in multiple logs. The contacts of these strata are connected with lines to indicate that they are one and the same stratum observed at different locations. The resulting Figure is called a *correlation panel*. As with geological logs, correlation panels can vary greatly in appearance, depending on the stylistic preferences of the author, but also in the amount or detail of data included in the illustration. However, in general we can assume that logs in correlation panels will contain slightly less detailed information than when presented on their own. This is by no means a rule, but the trend is undeniable.

Correlation panels are regularly combined with other visualisations, as for example maps, showing the location of each log in the correlation panel. Figure 4.11 [COBR$^+$12] is an example of this approach. The correlation panel in Figure 4.11 is comparatively sparse. On the left-hand side of the correlation panel, geological epochs are aligned to the strata of the first log. Each log is labelled with a number. The logs are ordered by their location on the map from south-west (SW) to north-east (NE). The vertical alignment of the logs is based on the fact that they are of bore-holes rather than outcrops: All logs are aligned to the zero-depth at the top, the depth values being labelled on the vertical axis. Additionally each log is labelled with a maximum depth value below its bottom-most stratum. Contacts are connected with dashed or continuous lines. Some correlation lines are labelled with U for unconformity and an index to identify and reference that specific correlation. On the horizontal axis, the correlation panel uses the distances of logs between each other. A labelled ruler is included at the bottom of the correlation panel. The logs themselves contain only one column of uniform width. Each stratum is filled with a pattern that corresponds to a certain sedimentary type. These patterns are also used to fill the areas between correlation lines. Question marks are used where no data is available, for example in log (7), where correlation lines are drawn below the extent of the log. Data might not be available for a variety of reasons, but often it is simply that an outcrop only exposes a certain subset of the strata as they occur, and different outcrops show different subsets.

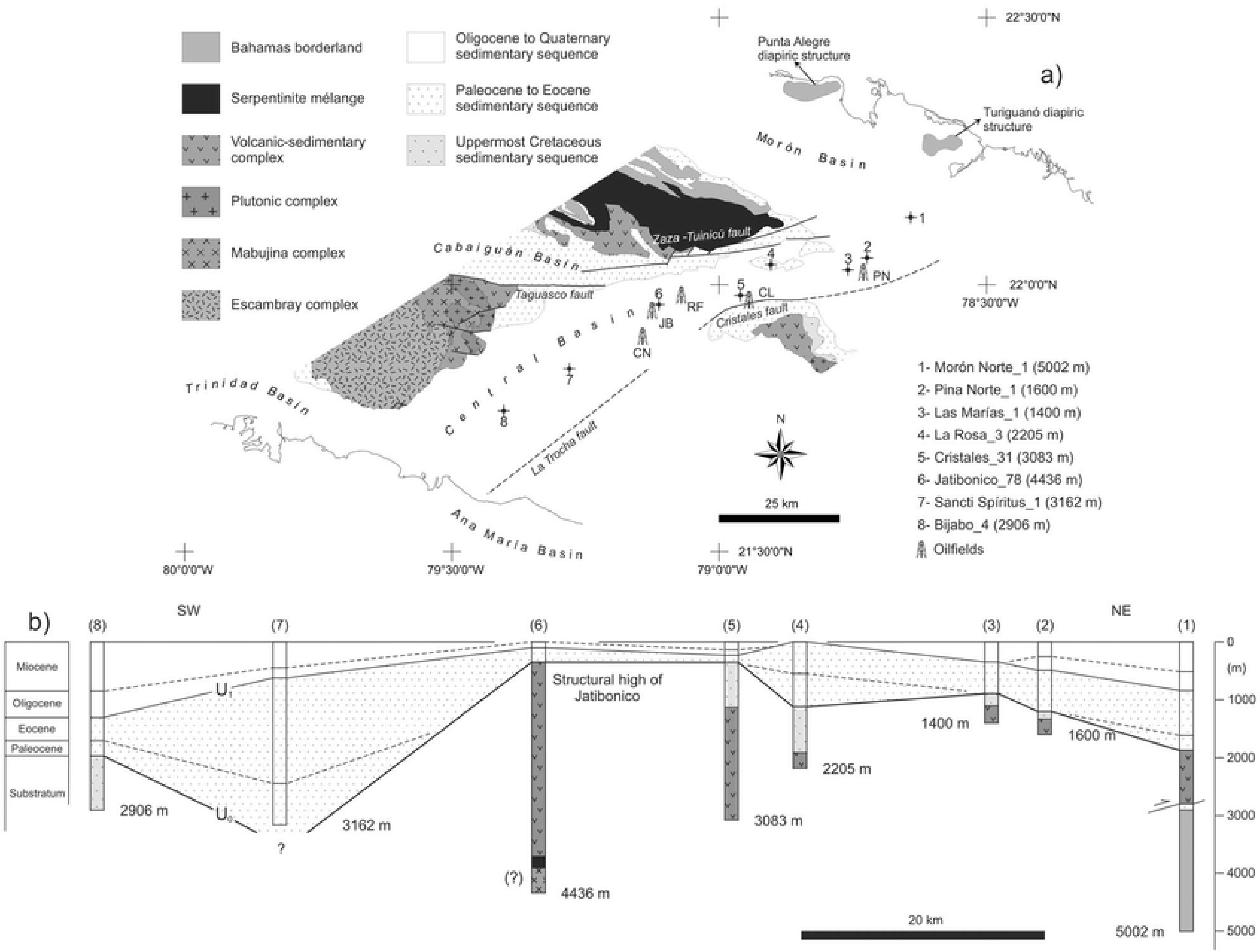


Figure 4.11: An example of a correlation panel (b) adjacent to a map (a) [COBR+12].

Figure 4.12 is an example from Hayes et al. [HGE+11], which is a planetary geology paper dealing with outcrops in the Victoria crater on Mars. A small map of the crater with the outcrop positions marked and a ruler to indicate the scale is included below the correlation panel on the right-hand side. Below that we can see so-called *layer tracings* of the outcrops. The logs use one column with ornaments, which are explained in a legend below them. Letters above the logs are used to identify them, and the name of the area they are taken from is written below each log. Each log ends in a wavy line which is explained with a label and an arrow below logs (B) and (C) as the "Furthermost extent of exposed stratigraphy" [HGE+11]. Another label with an arrow is used in log (A) to add information to a stratum. Dashed and solid lines encode different types of contacts and the correlations between them. Dashed lines encode *diagenetic contacts* and correlations between *diagenetic contacts*. Solid lines are used to encode *bounding surfaces* and correlations between *bounding surfaces*.

In this correlation panel each log is labelled with elevation values, and the logs are vertically aligned at zero metres. The logs are of equal width and spaced evenly and without indication that the horizontal axis has any semantic significance. An additional visualisation approach contained in this correlation panel are glyphs indicating the direction of the bedding dip azimuth next to strata. In this case the dip azimuth is visualised as an arrow inside a circle with the cardinal direction north marked as N at the top of the circle. One of these glyphs uses dashed lines for the arrow as well as the circle. The grain size is combined with the description of the lithology in the legend below the logs. The pattern used for strata is explained here. The categories are more specific than the ones presented in the previous example because they were chosen to fit the data.

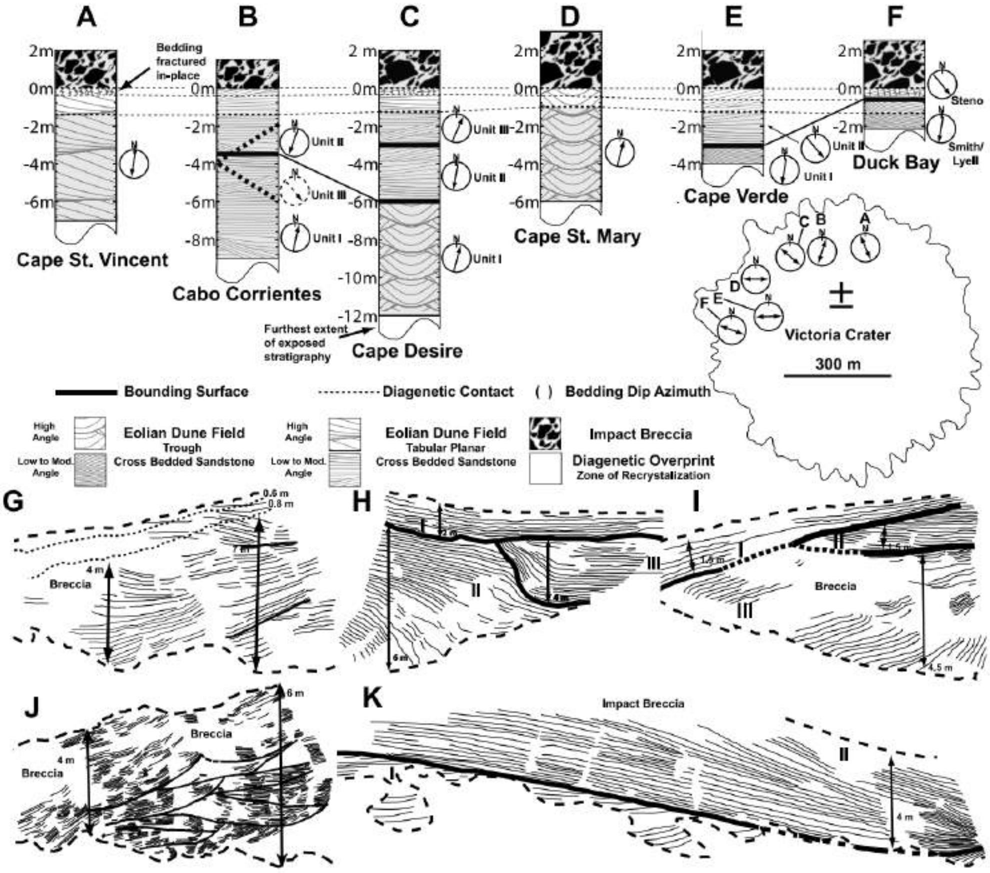


Figure 4.12: An example of a correlation panel with illustrations of the outcrops from [HGE+11].

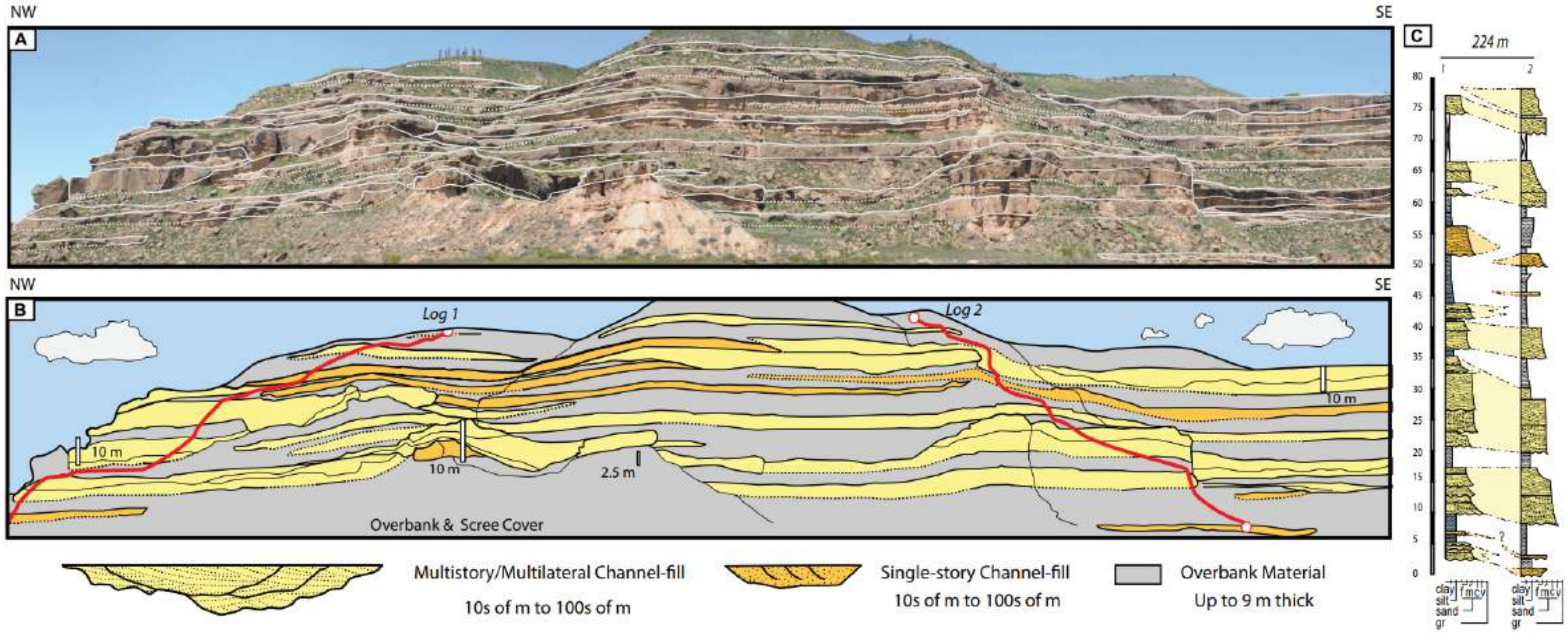


Figure 4.13: An example of a correlation panel [BH19].

In Figure 4.13 a correlation panel of two logs taken from the same outcrop is combined with an annotated photograph and a sketch of the outcrop. In the sketch and the logs, strata are coloured consistently according to the legend at the bottom of the Figure. The logs are drawn in the sketch as red lines with circles at the start and end points. The logs use one column dedicated to lithology and grain size. The grain size is used on the horizontal axis and therefore shapes the right borders of strata. Below each log a ruler is labelled with the specific grain sizes. The used abbreviations are explained in the original caption of the Figure. A labelled ruler on the left side of the correlation panel indicates that the values on the vertical axis are the elevation. The stratum height is given in metres, which can be deduced from the context, but is not made explicit in the correlation panel itself. The distance between the two logs is labelled at the top of the correlation panel. Correlations between strata and the continuation of strata between the two logs are drawn with dashed lines. A question mark indicates uncertainty between two strata in Log 1 and Log 2.

The correlation panel in Figure 4.14, published by Cain in his thesis [Cai09], demonstrates that correlation panels can contain numerous logs. In this case 22 sedimentary logs are shown together in one Figure. Cain points out that only 22 logs were selected for the sake of visual clarity from over 80 logs which could have been used in the correlation panel. The logs are scaled and aligned vertically based on their elevation. A ruler on the left-hand side of the correlation panel is labelled with elevation values in metres. Next to the ruler, labels divide the vertical space into two geological eras. The boundary between those eras is based on the boundary between different facies. This correlation is traced throughout the correlation panel with a green line, with different background colours above and below the correlation line. The background colours of the correlation panel encode the type of facies. A legend with facies associations can be found on the right-hand side at the bottom of the correlation panel. Cardinal directions are given in the top left and top right corners of the correlation panel. Question marks are used to indicate uncertainty in areas between logs. A detailed view of the logs is presented in the appendix of Cain's thesis. Notably, the logs used for the correlation panel do not use symbols and patterns but the logs presented separately in the appendix do. It stands to reason that these features were left out of the correlation panel as they would be of little benefit at the small size to which the logs have to be scaled down. In the correlation panel, the grain size is used on the horizontal axis of the logs to determine the width of strata. However, only the logs in the appendix provide the reader with a labelled ruler that maps intervals to grain size categories.

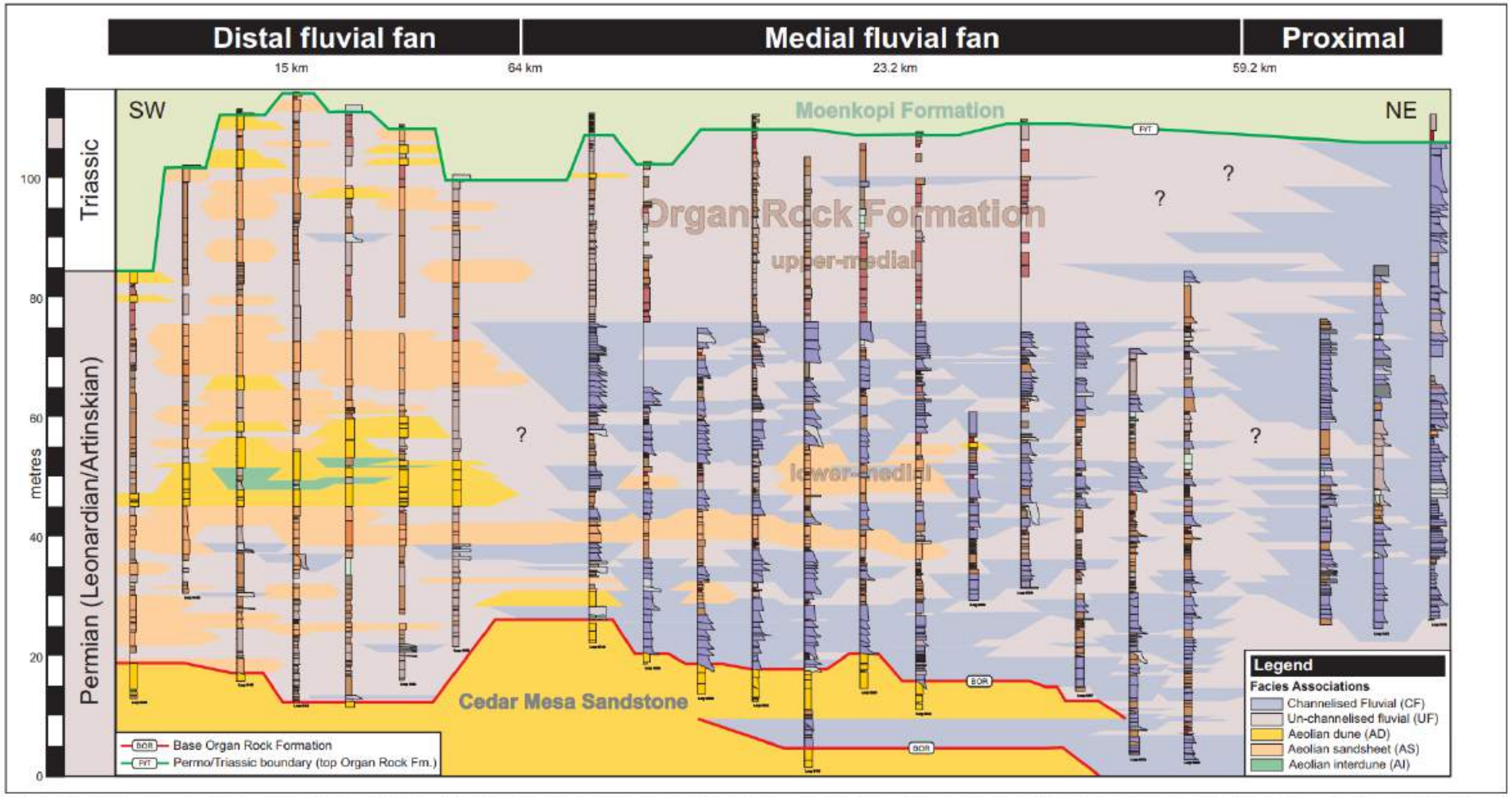


Figure 4.24: Regional correlation panel illustrating regional variations in gross-scale sedimentary architecture throughout the proximal, medial and distal parts of the Organ Rock Formation. See figure 4.1(a) for panel transect. This panel is tied to over 80 sedimentary logs, though for purposes of clarity only 22 are presented here. See section 3.6 for explanation of facies associations.

Figure 4.14: An example of a correlation panel from Cain 's work [Cai09], page 183.

In figure 4.15 we can see how Nichols [Nic09] uses logs with a higher degree of detail for his correlation panel than Cain (Figure 4.14). Nichols' correlation panel only includes four logs to Cain's 22. Unlike all other correlation panels presented up to this point, the logs in the correlation panel by Nichols contain more than one column. The logs are similar to the one in Figure 4.8, which is a representative example of a log from Nichols' book: Lithology and grain size are presented in separate columns, and both of those columns use patterns to symbolise different geological attributes. Each log has a third column containing information about the depositional facies. There is no scale for the vertical axis in the figure(s). The horizontal axis of the grain size column in each log can be found above the column and is labelled with grain size categories ranging from clay to boulders.

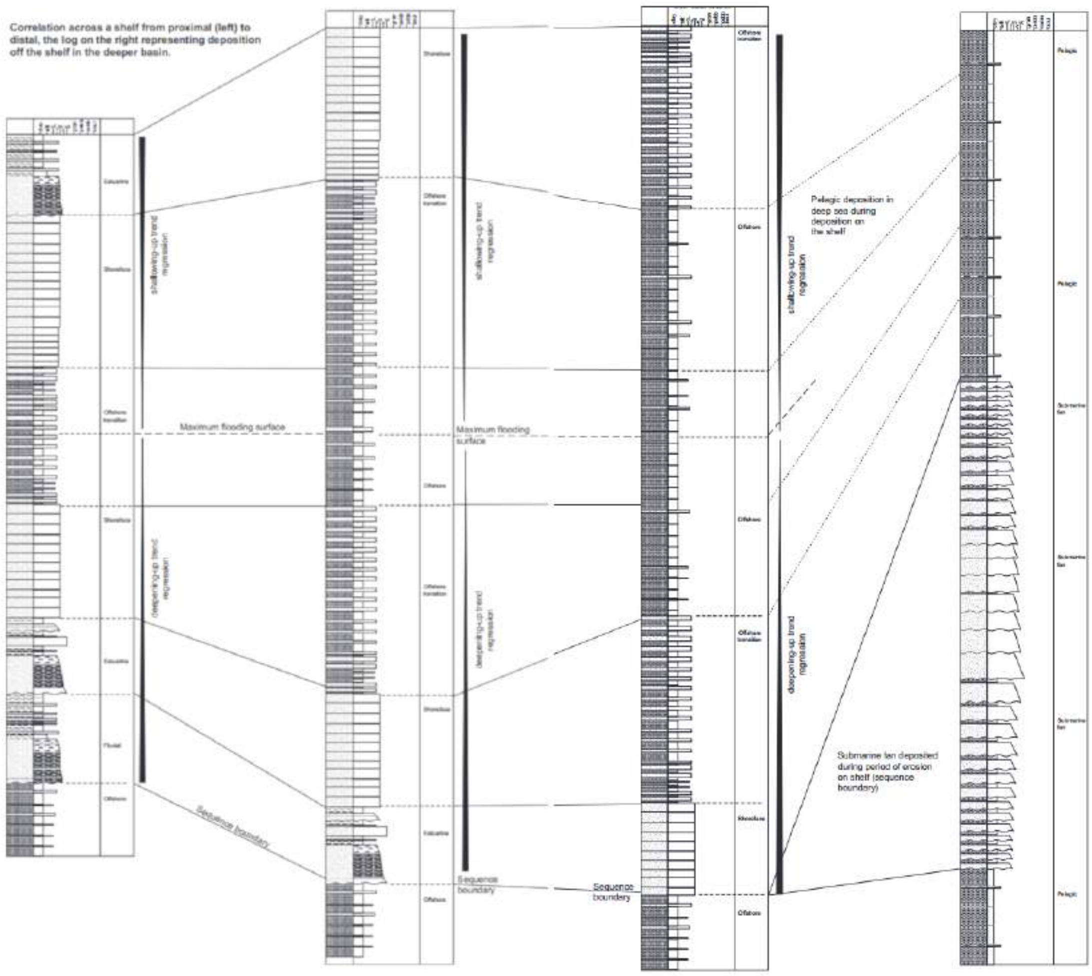


Figure 4.15: An example of a correlation panel from *Sedimentology and Stratigraphy* [Nic09], pages 374 and 375. Originally presented as two figures on subsequent pages, here the two halves of the correlation panel are placed next to each other.

The correlation panel in Figure 4.16.A contains ten logs. The correlation panel is presented alongside three other sub-figures and a legend. Figure 4.16.B visualises the relative locations of all logs. The cardinal directions are depicted using a labelled arrow pointing north. The palaeocurrent flow directions are labelled using oriented triangles, and a labelled line is used to indicate the scale. Figures 4.16C and 4.16D are photographs of outcrops, the scale is indicated with a labelled line. The logs and areas corresponding to these photographs are labelled in the correlation panel using thick red lines which are annotated with a reference to the figures.

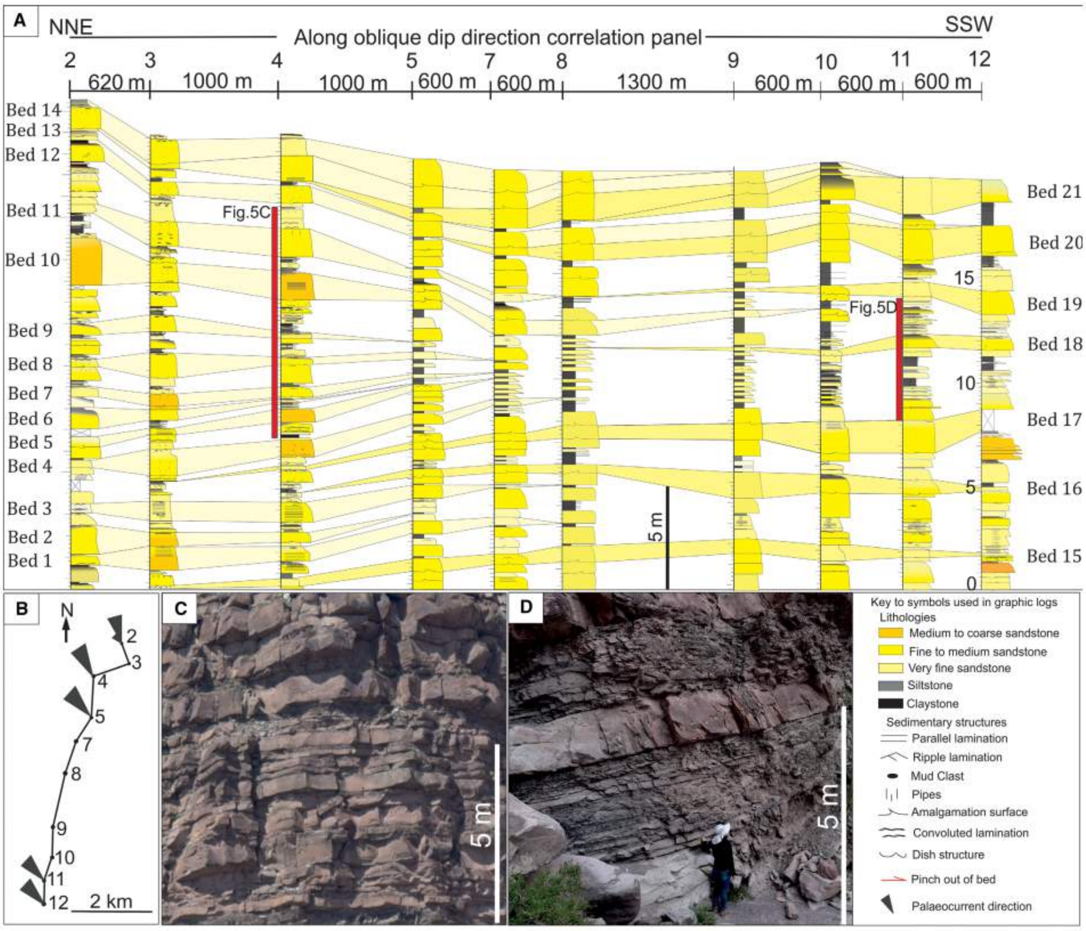


Figure 4.16: In the correlation panel by Liu et al. [LKF+18] (A) we can see how strata are correlated using coloured planes. Alongside the correlation panel the relative locations of the single logs, with palaeocurrent directions marked with triangles (B), images of outcrops (C, D), and a legend are given.

The logs in the correlation panel in Figure 4.16A encode the grain size and sedimentary structures in the same column as the log in Figure 4.5b. The stratigraphic strata used are beds. The grain size is encoded using colours and the width of logs. Sedimentary structures are encoded using symbols. The height of strata is relative to the height of the corresponding strata in the outcrops. The scale is indicated in the same way as in the other sub-figures, with a labelled line. The horizontal distance between the logs in the correlation panel is relative to the distance of the locations where the logs were taken. It is labelled on a line above the logs. Above this line, the logs are annotated with numbers. Above the numbers, the reasoning behind the order of the logs is given. On the left and right edges of the correlation panel, beds with correlations are denoted by "Bed x" where x is the number of each bed.

4.3.3 Semantic Annotations on Digital Outcrop Models

Annotations on digital outcrop models are the basis for creating correlation panels in a data-driven way. They are additional data artefacts that experts produce while interpreting an outcrop. Examples for annotations are contacts between strata or grain sizes. Which aspects of a correlation panel can be automatically generated, depends on which types of annotations are available. A *type of annotation* encodes a specific geological aspect. Contacts between strata, for example, might be annotated by drawing lines along the contacts. From these annotations we can derive strata for a log, as two contacts delineate a stratum. If the lines have no semantic meaning apart from the simple concept of representing a contact however, we have no way to infer the hierarchy of these contacts. This means that the generation of correlation panels depends heavily on - and is inherently linked to - the *annotation system* available to experts for annotating digital outcrop models.

With the annotation system proposed here, we want to encourage the use of annotation data not only to generate correlation panels, but also other visualisations that can be created using annotation data. The consideration that led to this decision is the following. The correlation panel is only one of multiple visualisations that geologists create based on the same data. The data, in this case, are outcrops, and the annotations and measurements associated with these outcrops. We believe it would be the wrong way to focus solely on annotations necessary for the creation of correlation panels, if a few generalisations allow the system to be used in a broader context. In our design of the annotation system, we keep this in mind, providing the basis for the data-driven generation of other visualisations as well.

Burnham et al. [BH19], for example, use several different visualisations for the sandstone body widths, which can be annotated (and measured) using an annotation system. In their paper, the tool VRGS (see Chapter 3) was used to annotate the sandstone bodies and measure body widths. The correlation panel in figure 4.13 is from the same paper. Figures 4.17 and 4.18 are two examples of additional visualisations.

There are two basic approaches to annotations on digital outcrop models. Considering an annotation as simple geometry (a line, a plane, etc.), or enriching annotations with semantic meaning. Although contacts and cross laminations are both annotated as polylines on the outcrop, they are different entities and, for example, their encoding in a

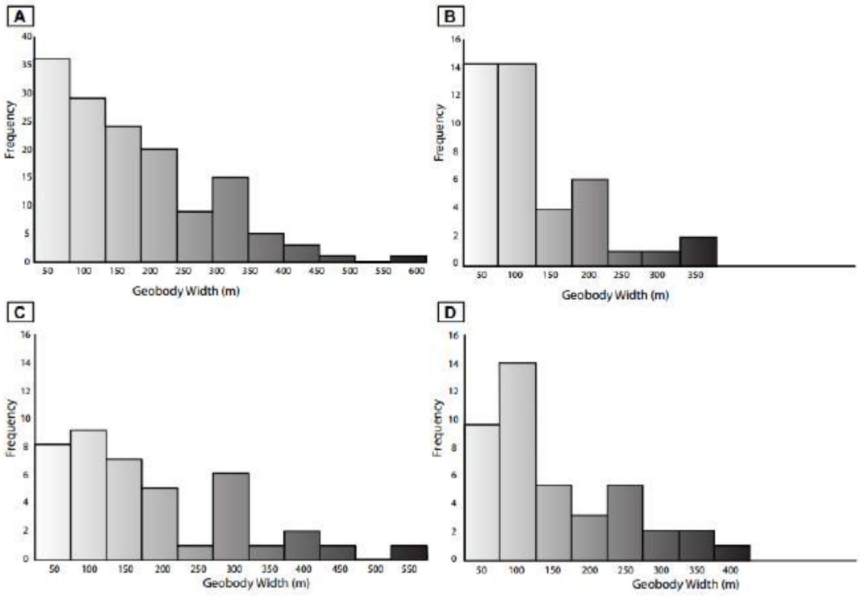


Figure 4.17: In addition to correlation panels, experts use other visualisations based on data derived from outcrops. Here bar charts encode sandstone body widths. [BH19].

correlation panel differs. As we have discussed in Chapter 2, contacts are hierarchical by nature. In general this is encoded using the width of lines when annotating. The thinner the lines the lower a contact is in the hierarchy. In a system without semantic meaning it would be left to the expert using it to ensure that the line widths are consistent. The expert has to remember, document, or check which line widths were used for which level of the hierarchy. Keeping consistency not only within one scene but over multiple scenes makes this even more challenging.

The challenge when providing experts with pre-defined, semantically meaningful annotation types is that this would lead to a large number of types (Chapter 2, Section 4.3), or restrict experts to a limited selection. In discussions with our collaborators it became clear that the second option would lead experts to abandon the tool, and that restricting the types of annotations would not be accepted. On the other hand, experts *wanted* the possibility to add semantic meaning to annotations, going further than simple geometry would allow. Semantic meaning is also crucial to generate correlation panels: As mentioned with the example of contacts and cross laminations, which encoding to use for an annotation cannot always be inferred from the geometry type of the annotation. That being said, in the case of contacts and cross laminations we could assume that cross laminations are annotated using thinner lines, but in the end this is just a guess and we cannot be sure what the geologist intended to encode.

To address this issue, we present a scheme to allow the use of semantic annotations without limiting the number of annotation types. Discussing this idea with our collaborators,

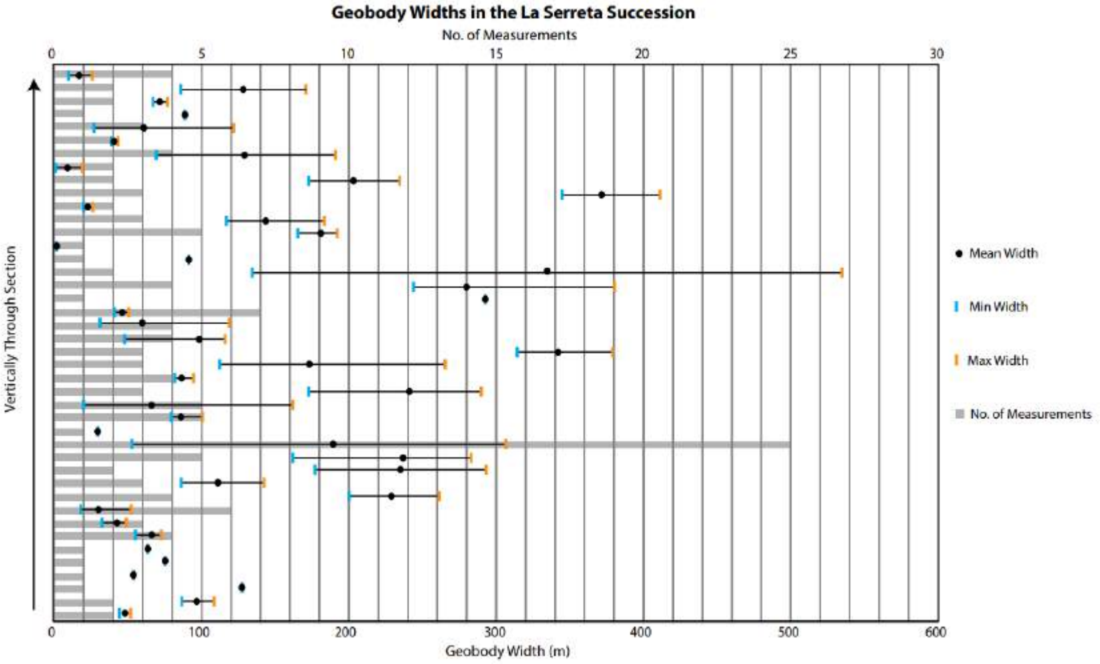


Figure 4.18: Another visualisation of the sandstone body width from the same paper as the correlation panel in Figure 4.13, and the bar charts in Figure 4.17 [BH19].

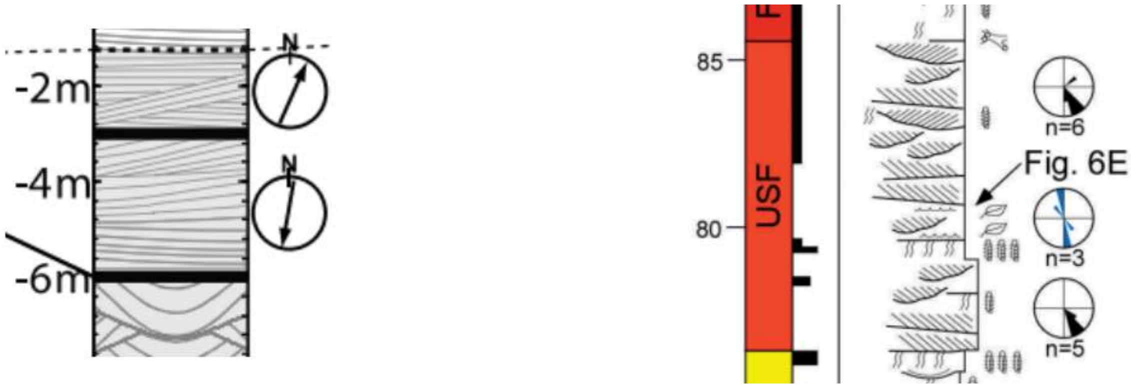
they advised us to keep the system as generic as possible.

When annotating with simple geometry, the colour and width can usually be adjusted. The available types of geometry depend on the tool, but lines are a basic feature. Outcrop annotation software also often provides the user with a tool to measure and annotate dip-and-strike (e.g. PRo3D [VRVb], Lime [BRN$^+$19], and VRGS [Hod17]). For a dip-and-strike annotation the expert selects (at least) three points on the surface of the digital outcrop model, which are then used to calculate and visualise the best fitting plane. The plane is used to calculate the angle of the dip.

We identified three attributes an annotation type needs to provide. A geometry type, and the visual attributes width and colour. The geometry type determines which geometric primitive is used to encode a geologic feature on the digital outcrop model. The weight and colour are needed to distinguish different features with the same geometry type. Another reason for including the width is, that the hierarchy of contacts is often encoded using the width of contacts. As this is an implicit encoding, we add an attribute that is solely responsible for defining their place in the hierarchy. We call this the *level* of an annotation type.

The next consideration is how a certain annotation is encoded in a correlation panel (or other visualisation). We chose three categories: *Hierarchical*, *angular*, and *metric*. Hierarchical annotations are encoded as the contacts of strata in the log, using the *level* of that annotation type to determine its place in the hierarchy. Angular and metric annotations are always leaf nodes in that hierarchy. They can be accumulated, and the result can be encoded as an additional diagram or glyph. How annotations are accumulated depends on the *semantic type*. An accumulation function could, for example, be the maximum of the length of all annotations in a stratum, as is the case for the grain size. We choose the semantic type *metric* for a grain size measurement, as it represents a metric attribute of a stratum. Another example is the palaeocurrent, for which we would choose the semantic type *angular*. Angular annotations can be accumulated and encoded in angular diagrams like rose diagrams, which would not make sense for metric annotations like the grain size. The grain size of Figures 4.17 and 4.18 show encodings for metric annotations.

For the encoding of angular measurements, consider, for example, the encoding of the dip azimuth in Figure 4.12, or the encoding of palaeocurrents in Figure 4.6. Details of those two figures are shown in Figure 4.19.



(a) (b)

Figure 4.19: Examples how angular measurements may be encoded in correlation panels. In (a) the dip azimuth is encoded, in (b) the palaeocurrent flow. (The original figures were taken from Hayes et al. [HGE+11] (a), and Hampson et al. [HGS+11] (b) respectively.

In these examples, angular measurements are encoded using small diagrams next to the corresponding strata in the logs. Geologists often use rose diagrams to show orientations within a stratum in an aggregated fashion (Figure 5.15).

Table 4.1 lists the attributes of a semantic annotation type in our system, and example values for each attribute for the type *grain size*. The analysis in Section 4.3 shows that the width of a stratum in a log is often derived from the grain size within that stratum in the outcrop. To measure the grain size, experts draw lines on representative grains. The lengths of these lines can be aggregated within each stratum (by taking the maximum), and the grain size of the stratum is calculated from its corresponding grain size values. This can only be done if grains are visible on the digital outcrop model. If the grain size cannot be measured, experts estimate the grain size, so they need to be able to set this

Attribute	**Edit**	**Example**
label	Yes	grain size
width	Yes	1.0
colour	Yes	yellow
level	Yes	0
semantic type	No	metric
geometry	No	line

Table 4.1: The semantic annotation system, and which properties can be edited for existing annotation types without leading to inconsistencies. An existing annotation type implies that there might be annotations present which use that type.

value for each stratum manually.

Another matter we need to consider is which properties a user can change on existing annotation types, and which should not be changed after an annotation type has been created. We assume that an existing annotation type has annotations associated with it, i.e. there are annotations of that type. When a user changes the label, weight, and colour of an annotation type, all associated annotations should be changed accordingly. The same is true for the level. This property has implications for the correlation panel (more precisely, the logs therein), changing the level of hierarchical annotations does not lead to inconsistencies, as long as the correlation panel is updated to reflect the changes. The geometry of an annotation type, however, cannot be changed without invalidating existing annotations: If the geometry of an annotation type was changed from point to line, for example, all existing annotations of that type would be inconsistent, at least when assuming a line has at least two points. This leads us to the conclusion that the geometry of an annotation type should not be changed after the type has been created. Whether the semantic type of an annotation type can be changed after creation is less clear. It can be argued that existing annotations do not lose consistency by changing this attribute, because the semantic type only affects their encoding in a different visualisation. However, the semantic type is, in essence, an attribute that describes what an annotation type can be used for and therefore what it represents. Whether it represents a hierarchical boundary between strata, or, for example, a dip and strike measurement. In other words, if this attribute changes, it can be assumed that the expert has changed what that annotation type represents. And in that case, creating a new annotation type should be the method of choice, rather than repurposing an old one. For this reason, we suggest that the semantic type of an annotation type should not change after creation.

4.3.4 Encodings for Logs and Correlation Panels

We use the word *artefact* for visual entities or attributes like the appearance of a stratum, but also the alignment or order of entities. When analysing which aspects of a correlation panel to include in an implementation that creates the panels in a data-driven way, it is necessary to also analyse which artefacts can be annotated on an outcrop. Another aspect we want to explore is which artefacts experts add to the correlation panel independently of the annotation process. An example for artefacts that are added to the correlation panel without being annotated on the outcrop are the correlations themselves. Nevertheless, correlating contacts is an interaction that a correlation panel must provide. To draw a clear line between our implementation and drawing programs like Adobe Illustrator, we focus on artefacts that can be derived from available data.

We use a top-down approach to order artefacts in this section. We start with the most high-level view on a correlation panel, and then move on to lower-level elements. We also list the most basic artefacts first and then describe more detailed or complex artefacts.

The basic elements of a correlation panel are two or more logs connected by correlations. There may or may not be various labels, labelled axes, legends, or other elements surrounding the logs, but no correlation panel can be without logs and correlations.

Logs can be positioned in a correlation panel in different ways. The order usually reflects the geographical locations of logs. We have to distinguish between vertical and horizontal alignment. Apart from their order, the horizontal alignment of logs does not necessarily encode information, as in Figure 4.12 where they are evenly spaced. In some cases, the horizontal spacing encodes the actual distances between the logs although not necessarily to scale (as in Figure 4.11). The distance between logs might also be added to the correlation panel as a text label (Figure 4.13). The vertical alignment of logs is done by first finding one specific contact that is present in all logs. Now the logs are aligned vertically so that the correlation line connecting this contact is horizontal. The selected contact may not be present in all logs. In this case the expert estimates a vertical alignment for those logs. Experts also suggested the use of multiple contacts with assigned priorities for vertical alignment. If the contact with the highest priority is not present in two neighbouring logs, the contact with the next lower priority is used to align the logs, and so on.

The order of logs is usually related to their relative geographical locations. The correlation panel in Figure 4.11 is presented alongside a map (a) where the logs (b) are ordered starting in the north-east and progressing in a south-westerly direction. Another example is the correlation panel in Figure 4.12 where the logs are ordered anti-clockwise along the wall of the Victoria Crater on Mars. These two examples might lead to the assumption that the log order could be inferred in a data-driven way without input from the user, but the experts we talked to emphasised that having manual control over the order of logs is essential. The narrative of the publication, the geological environment, or other factors might impact the order of logs in a way that cannot be derived from their relative locations alone. Letting experts order logs manually also allows them to re-order them dynamically during the interpretation process. Table 4.2 lists the artefacts we identified, different options for encodings, and the source of the data relevant for each artefact.

Correlation Panels

Artefact	Visual Encoding	Source
logs	Table 4.3	mixed
horizontal alignment	(1.a) logs are evenly spaced	-
	(1.b) the distance between logs in the correlation panel encodes the geographic distance between logs	data
	(2) distance between logs is labelled	data
vertical alignment	(a) tops of logs are aligned	-
	(b) user defined	user
	(c) use contact	mixed
	(d) use multiple contacts with priorities	mixed
order	(a) user defined	user
	(b) automatic sorting based on geographic location with manual control	mixed
correlations	(1.a) straight lines connect contacts	user
	(1.b) curved lines connect contacts	user
	(2) different line styles encode uncertainty	user
	(3.a) areas between correlation lines are coloured in	data
	(3.b) estimated strata are drawn between logs	user
vertical axis labelling	(a) none	-
	(b) elevation	data
	(c) labelled scale bar	data

Table 4.2: Visualisation options for correlation panels based on Section 4.3 and informal discussions with experts. We list each artefact with possible visual encodings and the source of the relevant data. Alternative encodings are listed with letters, additional encodings that might be used at the same time are listed with numbers. The *source* indicates whether the visualisation can be derived from the digital outcrop model or annotations (*data*), the user needs to create an artefact (*user*), or the artefact is created by combining available data and decisions of the user (*mixed*).

As discussed in Chapter 2, a *correlation* between two contacts means that these contacts are in fact the delineation of the same sedimentary strata observed in two different locations. Correlations are usually visualised using lines. In addition to these lines, the area between correlation lines might be coloured to visualise the estimated shape of strata in the region between logs. The correlation panel by Cain [Cai09] (Figure 4.14) does without lines altogether, using colours to represent different strata not only in the logs but throughout the correlation panel. As the correlations between logs are inferred by experts, a certain degree of uncertainty might be involved with a specific correlation. Our collaborators emphasised that is was important to express this uncertainty. The most common way to encode uncertainty in this context is to use different line styles. A dashed line implies more uncertainty than a continuous line (see Figure 4.15 for an example). Another way is to place question marks in places with a high amount of uncertainty (as in Figure 4.14).

In Section 4.3.1 we discussed that logs come in a variety of shapes and forms. In the same way as above with correlation panels, we now investigate what the necessary ingredients for a log are.

Having pointed out the lack of standardisation for visual logs, there are some artefacts that vary rarely. In most instances, the horizontal and vertical axis encode the grain size and vertical position of strata, respectively [Tuc88].

Further, logs in correlation panels are usually labelled, so they can be referred to individually. In the correlation panel in Figure 4.13 the logs are numbered from left to right. The example in Figure 4.12 uses letters above the logs, and the locations as text labels below the logs. The logs in Figure 4.15 are not labelled individually. Table 4.3 lists encodings for correlation panels. We list each artefact with possible visual encodings and the source of the relevant data. Alternative encodings are listed with letters.

Logs

Artefact	**Visual Encoding**	**Source**
strata	see Table 4.4	mixed
labelling	(a) one text label	user
	(b) one id label and one text label	mixed
horizontal axis labelling	(a) none	-
	(b) labelled horizontal axis (grain size)	data
vertical axis labelling	(a) none	-
	(b) labelled vertical axis (elevation)	data

Table 4.3: Visualisation options for logs based on Section 4.3 and informal discussions with experts. We list each artefact with possible visual encodings and the source of the relevant data. Alternative encodings are listed with letters. The *source* indicates whether the visualisation can be derived from the digital outcrop model or annotations (*data*), the user needs to create an artefact (*user*), or if the artefact is created by combining available data and decisions of the user (*mixed*).

Strata are the building blocks of logs. There are many different ways to visualise a stratum. There are also many different types of information that can be added to one stratum as we have seen in Section 4.3. Based on the logs and correlation panels in Section 4.3, the simplest encoding of a stratum is a rectangle with colour, texture, or both, representing geological (lithologic) attributes (as an example see Figure 4.12).

Strata

Artefact	Visual Encoding	Source
stratum shape	(a) strata are represented by simple rectangles	data
	(b) strata can be complex polygons to encode different contact types	mixed
stratum width	(a) strata have uniform width	data
	(b) stratum width derived from grain size	mixed
stratum appearance	(a) stratum colour derived from grain size	data
	(b) textures used instead of simple colours to encode geological features	mixed
	(c) glyphs/symbols placed inside strata to encode geological features	user
columns	(1.a) grain size	data
	(1.b) grain size and lithology	mixed
	(2) hierarchical columns	data
labels	(a) none	-
	(b) labelled strata	user

Table 4.4: Visualisation options for strata based on Section 4.3 and informal discussions with experts. We list each artefact with possible visual encodings and the source of the relevant data. Alternative encodings are listed with letters, additional encodings that might be used at the same time are listed with numbers. The *source* indicates whether the visualisation can be derived from the digital outcrop model or annotations (*data*), the user needs to create an artefact (*user*), or the artefact is created by combining available data and decisions of the user (*mixed*).

Most logs encode the grain size of a stratum in the outcrop as stratum width in the correlation panel. There are logs where a separate column is dedicated to the grain size, the other lithology is encoded using rectangles of equal width (compare the logs in Figure 4.5). The more common variety is the combined visualisation. This is especially true for correlation panels, where the horizontal space each log may take up can be limited.

The domain experts collaborating with us on this project suggested a multi-column design. Strata are by nature hierarchical. A log might show strata at a lower level in the hierarchy, or an accumulating view of strata high in the hierarchy. The logs in Figure 4.6 take a similar approach. The logs consist of three columns (and floating glyphs on the right). The grain size and lithology are combined in one column, and on the right hand side a column with the facies is added. A detail is also shown in Figure 4.19b. The facies can be seen as a higher level view of the lithology column.

Important strata might be labelled in correlation panels. The correlation panels in

Figures 4.12, 4.14, and 4.16 show different examples of this. Our collaborators expressed an interest in a related artefact, i.e., labels for contacts. Once a stratum or contact is labelled, different encodings can use these labels, without requiring additional user input. A contact label could, for example, be displayed in the 3D View of the digital outcrop model as well as the correlation panel.

Sedimentary structures and lithology are frequently encoded using various symbols and patterns, respectively. We suggested pre-defined patterns for lithological attributes to our collaborators, but they were not entirely satisfied with that approach. Regarding symbols for sedimentary structures, using a pre-determined set would work for them, but for patterns representing lithology we should take a different approach. As we have seen in Chapter 2, there are variations as to how a certain lithological attribute is encoded by different authors (Figure 4.9). It appears that at least some experts draw their own patterns according to personal preferences. In any case, our collaborators expressed a definite preference for using their own patterns, rather than being restricted to pre-defined ones. Considering this information, and the fact that we encountered differences between authors' encodings in our analysis, we surmise that a solution should provide import facilities for patterns rather than pre-defined patterns.

Tucker, in his *Techniques in Sedimentology* [Tuc88], asserts that standardisation may lead to undesirable restrictions or be may too complex to be of much use. We have seen the same opinion with our collaborators. However, at the same time we observed a desire for consistent encodings to facilitate teamwork. An idea from a workshop with our collaborators was a protected access to certain features for select expert-users. In this scenario, only the team lead has an account with access rights to determine the encodings to be used by all team members. Team members would not be able to access the features used to change encodings. In this way, a team can be sure to use the same set of encodings, thereby removing obstacles to communication, and generating consistent visualisations by default.

CHAPTER 5

Visualisation and Interaction Design

We use the task abstraction devised in Section 4.1 to structure this chapter. The tasks we identified are the following:

- **T1** Create new semantic annotation types.
- **T2** Annotate outcrop.
- **T3** Create a log and assign grain sizes.
  T3a Select log positions and create log.
  T3b Assign grain sizes to strata.
- **T4** Order logs and correlate strata.
  T4a Order logs.
  T4b Correlate strata.

Before we describe the visualisation and interaction design for each task in Sections 5.3 to 5.6, we formulate some design goals in Section 5.1, and give an overview of the system in Section 5.2.

5.1 Design Goals

As described in Chapter 2, there are two typical applications of geological logs. The first application is as a recording and interpretation mechanism during field work. The second application is as a means of communication in publications and presentations. In the

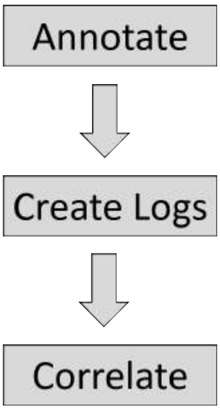


Figure 5.1: Workflow when creating correlation panels by hand. The creation of a correlation panel happens at the end of this workflow, after the interpretation process is finished and logs are finalised.

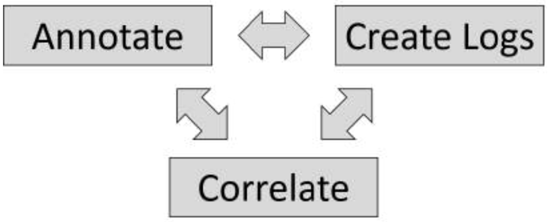


Figure 5.2: Workflow made possible when creating correlation panels semi-automatically. By generating correlation panels easily and quickly, they can be created and re-created or updated at any time during the interpretation process. Using preliminary or unfinished logs is possible.

context of remote geology using digital outcrop models, the field work is replaced by inspecting and annotating the digital outcrop model. Being able to quickly generate and update digital correlation panels could allow experts to move from creating a correlation panel *after* the interpretation (the workflow depicted in Figure 5.1) to a more integrated workflow (Figure 5.2). A log can be generated at any time, and annotations added to an outcrop after creating a log can be integrated easily.

We formulate the following design goals:

- **G1** The user's effort to generate a log should be minimal.
- **G2** Editing (adding, deleting) correlations should be a quick and simple process.

- **G3** Updating or re-generating logs after adding new annotations should be a quick and simple process.
- **G4** Preserve the properties of static correlation panels.

The need to link strata as encoded in logs to the corresponding area on a digital outcrop model is illustrated by the correlation panel in Figure 4.16. In this correlation panel, labelled red lines next to logs link to photographs of the original outcrop that are presented alongside the correlation panel. One conclusion is that a correlation panel and the digital outcrop model should be juxtaposed, or the user should be able to choose viewing both in juxtaposition. Letting the user choose has an advantage. With a large number of logs in a correlation panel, screen space might simply be insufficient to display the digital outcrop model and the whole correlation panel at once.

Another conclusion we draw from the correlation panel in Figure 4.16 is, that visually linking original and encoded strata and contacts is valuable to experts. As Munzner [Mun14] points out, a very common means to link data in *multiple views* is linked highlighting. By selecting a data item in one view the corresponding data is highlighted in linked views as well. In the case of a correlation panel, we want to link strata in the correlation panel to the annotations encoding strata in the digital outcrop model. The annotations tracing contacts between strata are inherently linked to the contacts in the logs, as the logs are generated using exactly those annotations.

5.2 Overview of the Design

The prototype is organised into *views*:

- **3D View** This view displays the digital outcrop model, and is used to explore and annotate.
- **Annotation Type Views**

  Expert View Semantic annotation types are displayed with all attributes and can be edited and selected.

  Simple View Only the name of semantic annotation types is displayed, they can be selected but not edited.
- **Correlation Panel View** Displays the correlation panel and allows the user to interact with it in various ways.
- **Grain Size View** This view displays a list of available grain sizes, which can be selected and edited.
- **Log List View** A list of logs that allows the user to edit the label of a log and delete logs.

The user can resize and move each view according to their preference. Different tasks might suggest different layouts of views. For annotating an outcrop (T2), the *3D view* will be the most important view. An *Annotation Type View* is necessary for T2, but does not require much space. So a layout as in Figure 5.3 could be used. Figure 5.7 shows a layout that lends itself to assigning grain sizes for a geological log (T3b). The correlation panel shows a 2D depiction of the log. Here the strata can be selected. The grain size is assigned using the *Grain Size* View. Lastly, the 3D view shows the log on the outcrop where a user might want to inspect each stratum to make an estimate on the grain size.

We provide a reduced grain size view to achieve design goal G1. By removing user interface elements that are irrelevant for the current task we strive to improve experts' efficiency by maximising screen space for relevant elements and removing distractions.

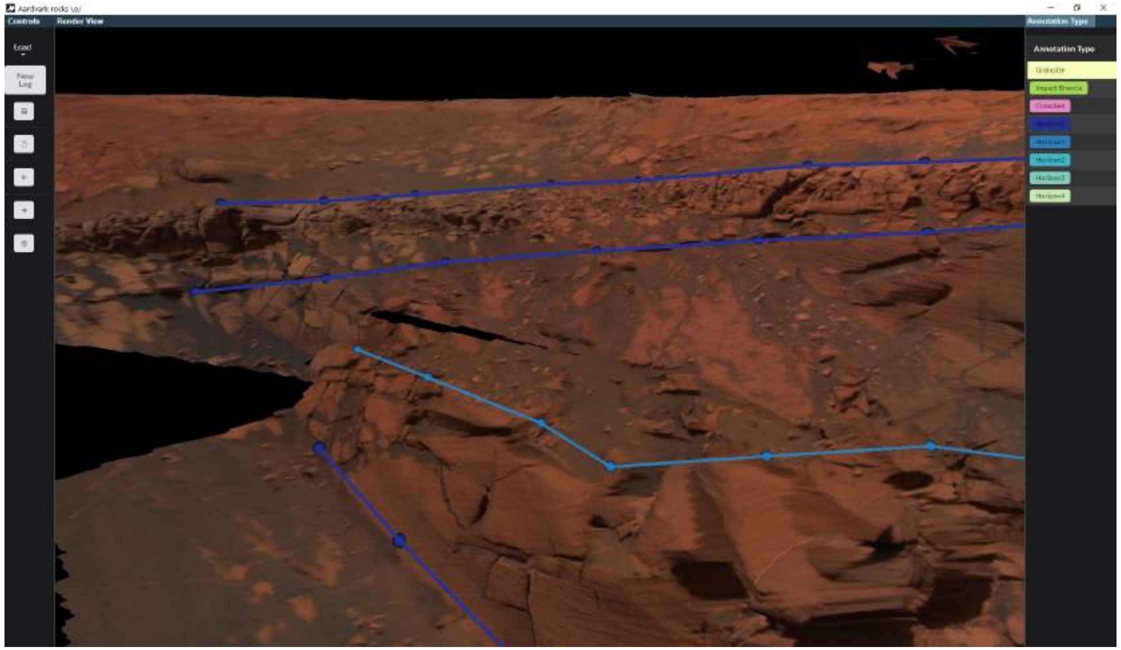
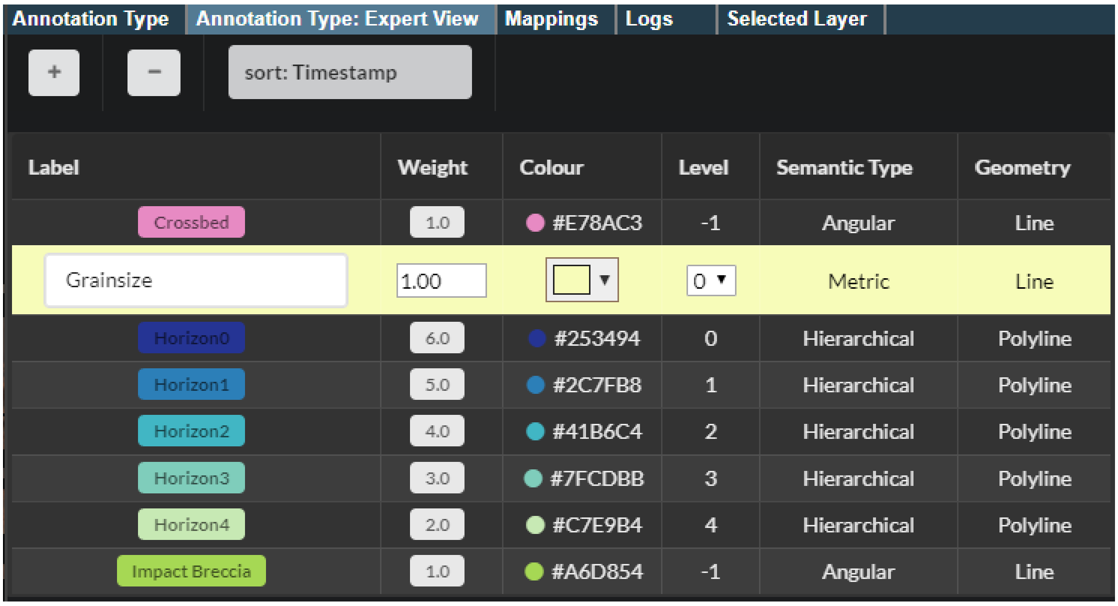


Figure 5.3: Views organised for annotating. The 3D View is given the most screen space, only the menu (left) and the simple view on annotation types (right) are also visible.

The menu on the left side in Figure 5.3 is used to load outcrops, create a log, and save and load annotations.

5.3 T1: Creating Semantic Annotation Types

There are different types of annotations, as for instance cross beds or grain size, which can be selected from the view *Annotation Type*. Each type has various properties and can be edited. New custom types can be created, and unwanted types can be deleted. The basis for the annotation type system is described in detail in Section 4.3.3.



Label	Weight	Colour	Level	Semantic Type	Geometry
Crossbed	1.0	● #E78AC3	-1	Angular	Line
Grainsize	1.00	▼	0 ▼	Metric	Line
Horizon0	6.0	● #253494	0	Hierarchical	Polyline
Horizon1	5.0	● #2C7FB8	1	Hierarchical	Polyline
Horizon2	4.0	● #41B6C4	2	Hierarchical	Polyline
Horizon3	3.0	● #7FCDBB	3	Hierarchical	Polyline
Horizon4	2.0	● #C7E9B4	4	Hierarchical	Polyline
Impact Breccia	1.0	● #A6D854	-1	Angular	Line

Figure 5.4: Annotation Type Expert View. In this view annotation types can be added, deleted, and edited.

According to different planetary geologists we talked to, in certain areas there is no final consensus on many terminological and stylistic issues (see Section 4.3 for examples). An example of this is the way in which geologists annotate geological features on pictures or 3D models of outcrops. Our own investigations and discussions with experts have led us to the same conclusions as stated by Coe [Coe10] and mentioned in Section 2.2. Stylistic differences in visual geological logs (and correlation panels) are vast and very much depend on the preferences of the author. What emerged from informal discussions is that whereas being able to use one's individual style is valued highly, a certain degree of consistency would be beneficial when working in a team. This thought is an integral part of the design of the two *Annotation Type Views*. There are two different views on the same data: The *Simple View* (Figure 5.3) and the *Expert View* (Figure 5.4). One reason for this are the two tasks for annotation types. A user is either annotating (T2), and needs to select annotation types, or they are manipulating (adding, editing, or removing) annotation types (T1). In the first case, the user is focused on the 3D View, and only needs to select types. For this task, maximising the space available to the 3D View means minimising the space the Simple View takes up. The label of each annotation type is the only information necessary for the user to select a type, meaning that all other fields can be omitted for T2. For T1, where new annotations are added, edited, or deleted, all fields are displayed. Decluttering the interface during the annotation process is not the only reason for the existence of the two views. Our collaborators were in favour of the distinction between normal users and expert users, who could determine encodings. Although the prototype presented here does not provide user accounts with different levels of access to functionalities, it has been designed with that concept in mind. In such a system, which distinguishes between normal and expert users, the Expert View could only be accessible to expert users. In this way, many aspects of annotations that often depend on author preference (like the colour of annotations for different purposes) could be set by the team leader and then used consistently by all team members.

The interactions possible in the Expert View are listed in Table 5.1. Additionally, users can sort annotation types by their attributes.

Interactions	**Description**
Add new	Add a new annotation type. The annotation type is added and can be edited in the *New* mode.
Remove	Remove an annotation type.
Change sorting	Cycle through sorting options.

Table 5.1: Annotation Type Expert View: Global Interactions

When creating and editing annotation types we have to consider which properties of an existing annotation can be changed without leading to inconsistencies. We explained our reasoning on which attributes can be edited after a type has been created in Section 4.3.3 (results in Table 4.1). To implement the restrictions in editing semantic type and geometries, we devised a system that allows three different views (associated with three states) on the same data type.

There are three states an annotation type can assume, *New*, *Edit*, and *Display*. Figure 5.5 describes these states, and illustrates their connection. The attributes *Semantic Type* and *Geometry* can only be edited in the state *New*. The state *New* is only accessible once, on creating a new annotation type. As soon as *save* is selected, the annotation type enters the state *Edit*. In the state *Edit* only label, grain size, colour, and level can be edited. Unselected annotation types enter into the state *Display*, in which no changes to its attributes are possible.

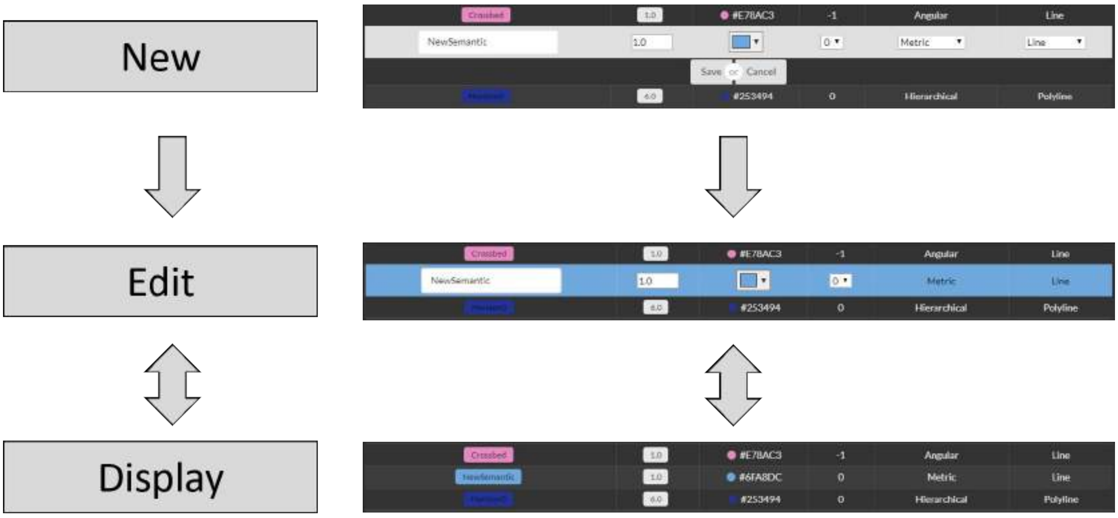


Figure 5.5: The three states of an annotation type. Semantic type and geometry can only be edited in the state *New* and therefore only when creating an annotation type.

Tables 5.2 and 5.3 list the interactions possible with annotation types in the states *Edit* and *New*. In the state *Display*, the only action possible is selection, which puts the annotation type into the state *Edit* (Table 5.4).

Interactions	**Description**
Rename	Change the width of the annotation type.
Change weight	Changing the width of an annotation in this view leads to an update in the 3D View, where the changed weight is applied to all existing annotations of this type.
Change colour	Changing the colour of an annotation in this view leads to an update in the 3D View, where the changed colour is applied to all existing annotations of the corresponding type.
Change level	Changing the level of an annotation has implications for a geological log where this annotation is used, as annotations are grouped according to their place in the hierarchy.

Table 5.2: Annotation Type Expert View: Interactions in *Edit* Mode

Interactions	**Description**
Change semantic type	The semantic type determines the ways in which annotations of this type can be processed.
Change geometry	The geometry type determines which geometric shape the annotation has.

Table 5.3: Annotation Type Expert View: Additional interactions in *New* Mode. All interactions listed in Table 5.2 are also available in *New* Mode.

Interactions	**Description**
Selection	Select an annotation and thereby change its state to *Edit* mode.

Table 5.4: Annotation Type Expert View: The only interaction possible in *Display* Mode is selection.

5.4 T2: Annotating an Outcrop

To select an annotation type, the user clicks on it in the *Annotation Type* view. The row of the selected annotation type is highlighted. To select a point on the surface of the outcrop in the 3D View, the user holds the *control* button and clicks on the place where the first point of the annotation should be placed. Depending on the geometry type of the annotation type, the user keeps selecting points on the surface until the annotation is complete. Pressing *enter* completes an annotation. To create a geological log, at least two hierarchical annotations have to be placed on the outcrop.

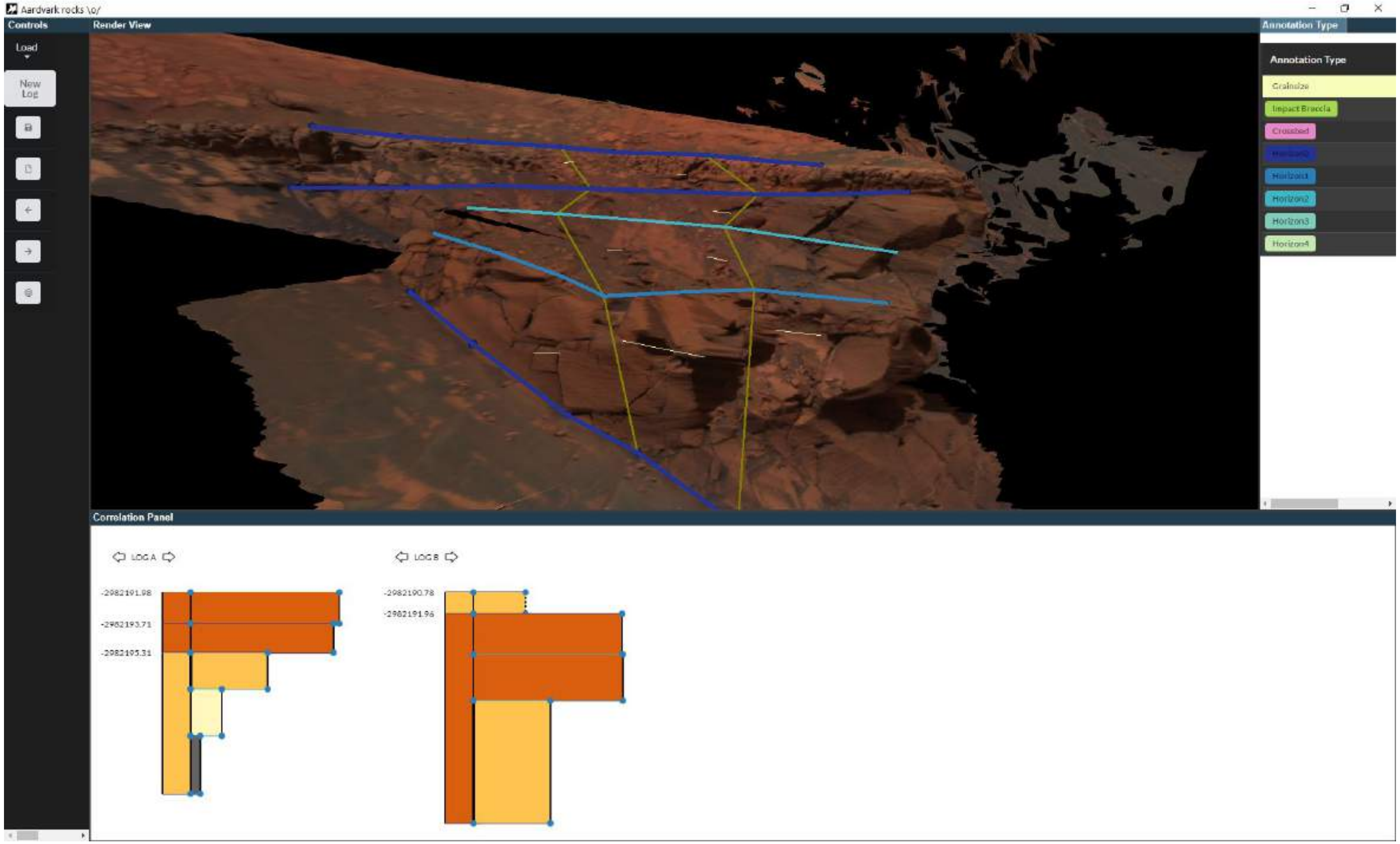


Figure 5.6: 3D View and correlation panel

In the 3D View the user can inspect digital outcrop models in 3D. This view can be used to explore and annotate the data. It is also used to select the position of a log (T3a).

Interaction	**Description**
Navigation	Move the camera
Pick point on surface	The user picks points on surfaces to create annotations.
Hover over annotation	An annotation is highlighted when the user moves the mouse over it.
Pick point on annotation	The user needs to pick points on annotations to create a log.
Complete an annotation	The user determines when an annotation is finished.

Table 5.5: Interactions in the 3D View

How many points can be added to an annotation depends on its geometry type. For example, an annotation with the geometry type *line* is automatically complete after selecting two points. When creating annotations with an arbitrary number of points, the user has to determine when an annotation is finished.

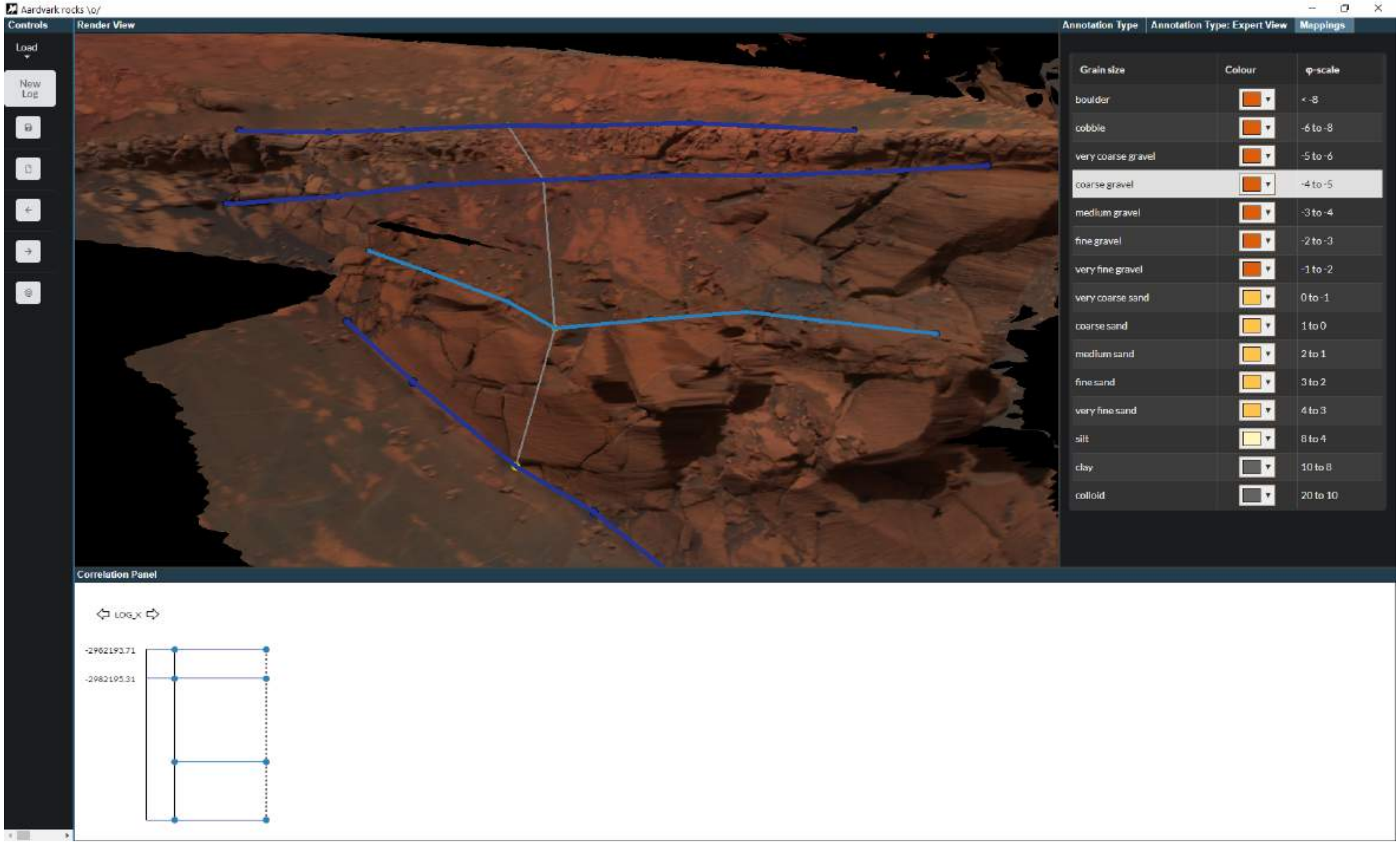
5.5 T3: Creating a Log and Assigning Grain Sizes

To create a log, a location on the outcrop needs to be specified. To this purpose, the user selects points on hierarchical annotations at the desired location. An annotation is highlighted in yellow if the mouse hovers over it. For our prototype only the points at each end and between segments of an annotation can be selected, which are the points the user added when creating the annotation. These points are visualised as spheres. Once a point is selected the corresponding sphere is highlighted in yellow. For one log, only one point can be selected on an annotation, as a log cannot contain one contact twice. If another point on the same annotation is selected, the previously selected point is removed from the selection.

By clicking on *New Log*, a log is generated. It will appear in the *Correlation Panel View*. The grain size and the corresponding width of a stratum can be calculated, if grain size annotations are available for that stratum. If no grain size annotations are available, a default width is selected and the stratum is depicted in white. The grain size can be specified after the log has been created by selecting a stratum in the geological log, and the clicking on the grain size in the *Grain Size* View. If the grain size is not determined or estimated by selecting an item from the provided list, the border on the right hand side of a stratum is dashed to indicate that there is some uncertainty in regards to the grain size.

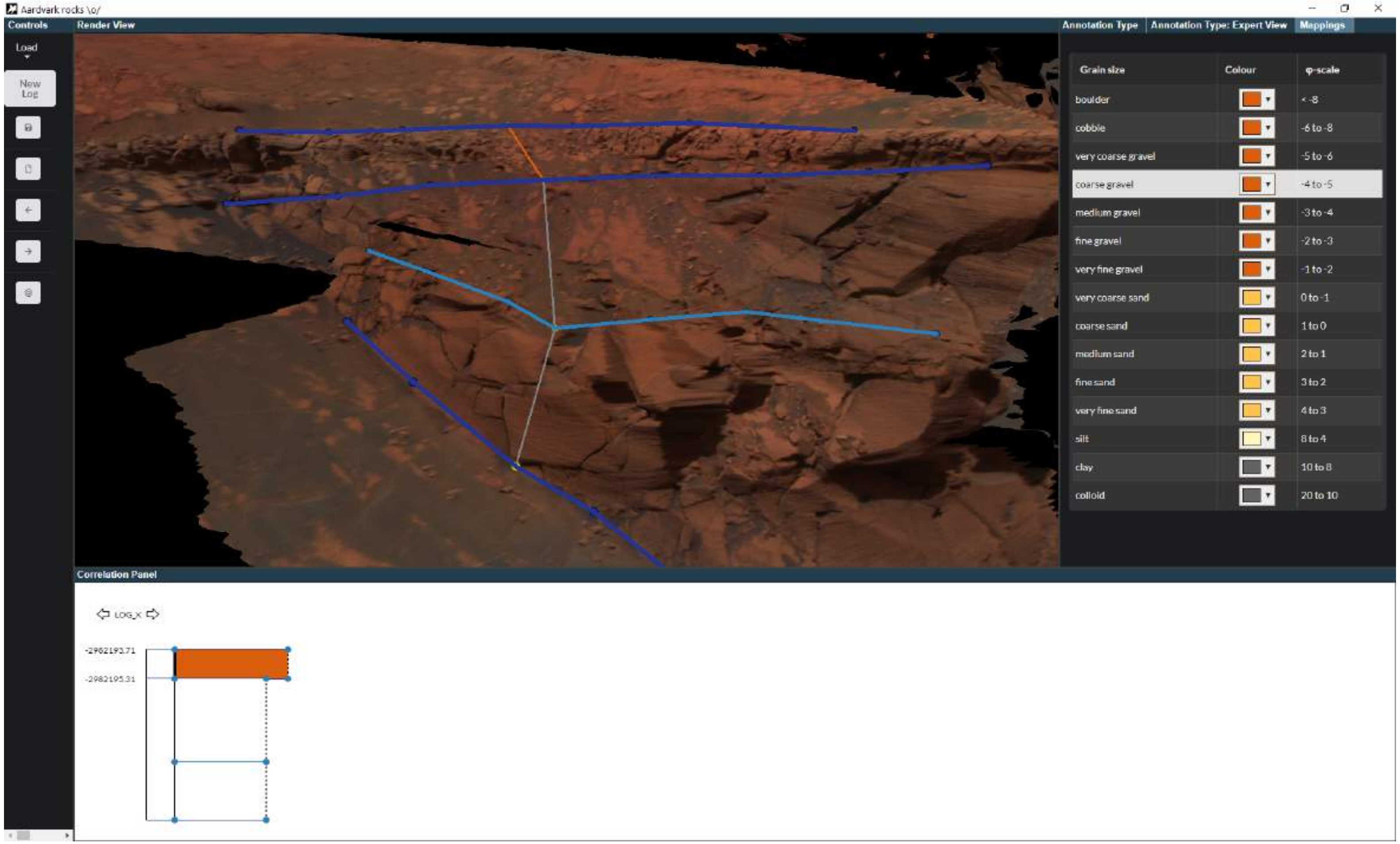
The grain size of a stratum is integral to generating correlation panels, as it is used to determine the width and colour of the encoded representation of strata in logs. Geologists annotate the diameter of a grain with a straight line. The length of the line and therefore the diameter is calculated. The *grain size* of a stratum can be inferred from the individual measurements it contains, by taking the length of the longest annotation of the type grain size in the stratum. What category a grain size falls into is fairly well standardised, depending on the diameter of the grains. Therefore, a value derived from grain size annotations can be used to assign a grain size category. In the Grain Size View, these categories are listed with the φ-scale, a scale geologists use to categorise grain sizes.

The resolution of the data available for Martian (and sometimes also terrestrial) digital outcrops is often not good enough to distinguish individual grains. Also, some types of grain sizes, such as mudstone, are generally impossible to measure on photographic data. In this case, the geologist estimates the grain size. This means that in addition to the automatic calculation of the grain size, a means to input the grain size manually is necessary. This can be done in the Correlation Panel View, by clicking on a stratum to select it, and then clicking on the estimated grain size in the Grain Size View. The possible interactions in the Grain Size View are listed in Table 5.6.



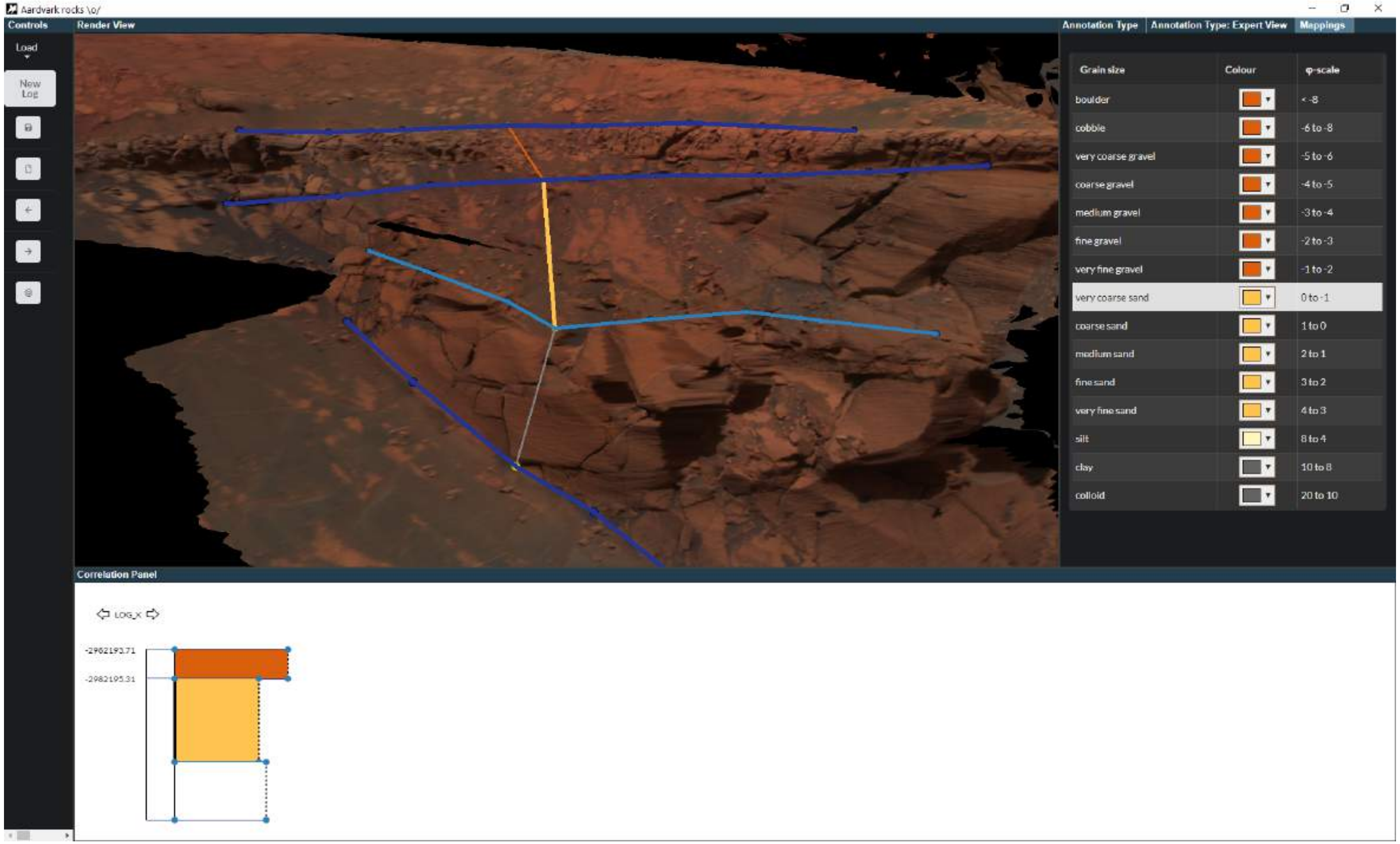
Grain size	Colour	φ-scale
boulder	🟧	< -8
cobble	🟧	-6 to -8
very coarse gravel	🟧	-5 to -6
coarse gravel	🟧	-4 to -5
medium gravel	🟧	-3 to -4
fine gravel	🟧	-2 to -3
very fine gravel	🟧	-1 to -2
very coarse sand	🟨	0 to -1
coarse sand	🟨	1 to 0
medium sand	🟨	2 to 1
fine sand	🟨	3 to 2
very fine sand	🟨	4 to 3
silt	⬜	8 to 4
clay	⬛	10 to 8
colloid	⬛	20 to 10

Figure 5.7: Selecting grain sizes from the list (Step 1/3). If there are no grain size annotations in a stratum, it is not assigned a grain size and the corresponding stratum in the log is assigned a default width.



Grain size	Colour	φ-scale
boulder	🟧	< -8
cobble	🟧	-6 to -8
very coarse gravel	🟧	-5 to -6
coarse gravel	🟧	-4 to -5
medium gravel	🟧	-3 to -4
fine gravel	🟧	-2 to -3
very fine gravel	🟧	-1 to -2
very coarse sand	🟨	0 to -1
coarse sand	🟨	1 to 0
medium sand	🟨	2 to 1
fine sand	🟨	3 to 2
very fine sand	🟨	4 to 3
silt	⬜	8 to 4
clay	⬛	10 to 8
colloid	⬛	20 to 10

Figure 5.8: Selecting grain sizes from the list (Step 2/3). The colour (corresponding to a certain grain size) that is selected for a stratum is also used to visualise the corresponding segment of the log in the 3D View.



Grain size	Colour	φ-scale
boulder	🟧	< -8
cobble	🟧	-6 to -8
very coarse gravel	🟧	-5 to -6
coarse gravel	🟧	-4 to -5
medium gravel	🟧	-3 to -4
fine gravel	🟧	-2 to -3
very fine gravel	🟧	-1 to -2
very coarse sand	🟨	0 to -1
coarse sand	🟨	1 to 0
medium sand	🟨	2 to 1
fine sand	🟨	3 to 2
very fine sand	🟨	4 to 3
silt	⬜	8 to 4
clay	⬛	10 to 8
colloid	⬛	20 to 10

Figure 5.9: Selecting grain sizes from the list (Step 3/3). The correlation panel and the 3D View are linked. If a stratum is selected in the correlation panel (the left border of the selected stratum is depicted with a thicker line style) then the corresponding segment of the log is displayed with a thicker line as well.

Interactions	**Description**
Change colour	Change the colour associated with the corresponding grain size.
Select	Assign a grain size to the selected stratum.

Table 5.6: Interactions in the Grain Size View.

The two-dimensional log in the correlation panel is finished after assigning grain sizes to strata. There is also a three-dimensional representation of the log. It is visualised in the 3D View as a series of line segments connecting contacts, in the same way as in the correlation panel in Figure 4.13. Each segment (corresponding to a stratum) is linked with its encoding in the correlation panel. If a stratum in the correlation panel is selected, the corresponding segment of the log is visualised with a thicker line (Figure 5.9). The colour of the each segment of the log in the 3D View is linked to the colour associated with the grain size of the corresponding stratum.

5.6 T4: Ordering Logs and Correlating Strata

The Correlation Panel View (Figure 5.10) shows the generated logs and allows the user to order them, and correlate strata. Newly generated logs are added to the right of existing logs. The horizontal order of logs can be changed using the arrow buttons above each log. Zooming and panning is possible in the Correlation Panel View. The label sizes are calculated based on the zoom level. A stratum can be selected by clicking on it, which allows the selection of a grain size category, and highlights the selected stratum of the log in the 3D View. A correlation is drawn by clicking on the corner of the rectangle representing a stratum that should be correlated. The style of the line representing the correlation can be changed by clicking on it with the left mouse button. Dashed and solid line styles are cycled through by clicking repeatedly. Clicking on a correlation line with the right mouse button deletes the correlation line.

The vertical scale, i.e. by what factor the actual thickness of a stratum is scaled when creating a log, was selected based on the scale of the strata in the sample outcrop shown in this work. As mentioned in previous chapters, logs might cover hundreds of metres or just a few centimetres in height. To allow for this, the vertical scale of logs can be adjusted by the user using keyboard shortcuts.

In the Log List View (Figure 5.11), all logs present in the correlation panel are listed. The label of each log can be changed via the text field in this view. A log can be deleted

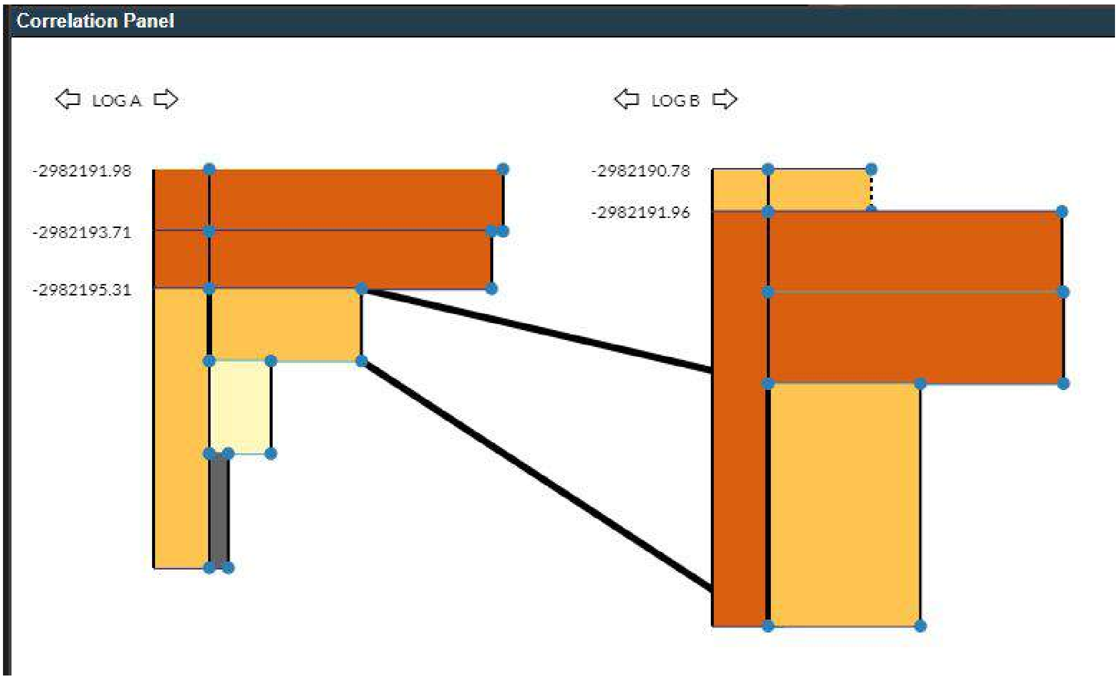
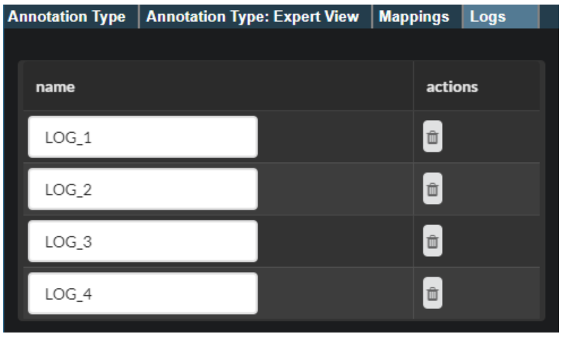


Figure 5.10: The Correlation Panel View.

by clicking on the button with the trash can icon in the relevant row. The interactions are listed in Table 5.8.



name	actions
LOG_1	🗑
LOG_2	🗑
LOG_3	🗑
LOG_4	🗑

Figure 5.11: The Log List View.

Interactions	**Description**
Zoom	The zoom level of the correlation panel can be changed by dragging the right mouse button. Label sizes are adapted according to the zoom level.
Stretch/compress y coordinates	The vertical scale of the logs can be adjusted using y and ctrl+y.
Pan	The correlation panel can be panned by dragging the mouse.
Change order	The order of logs can be changed by using the left and right arrows above each log.
Select stratum	A stratum is selected by clicking on it. The left border of a selected stratum is drawn in a thicker line than the other borders.
Draw correlation	A correlation is drawn by clicking on the corners of strata to be connected.
Change correlation line style	The style of a correlation line can be changed by clicking on it with the left mouse button.
Delete correlation	A correlation is deleted by clicking on it with the right mouse button.

Table 5.7: Interactions in the Correlation Panel View.

Interactions	**Description**
change label	Changes the label of the log. The label displayed on the correlation panel is automatically updated if the label is changed in this view.
delete	Clicking on the button with the dust bin icon deletes the log.

Table 5.8: Interactions in the Log List View.

5.7 Implementation Details

We created the prototype using the Aardvark (An Advanced Rapid Development Visualization and Rendering Kernel) framework [VRVa], which uses the functional programming language F#.

The Aardvark Framework provides multiple libraries for rendering and user interface creation. For 3D rendering, it uses an incremental rendering layer as a virtual machine on top of OpenGL which allows more efficient rendering than an approach that does not use an incremental system [HSMT15]. With this approach the frame rate depends on how much changes from frame to frame rather than the size and complexity of the scene.

The incremental update system is also used for the web-based graphical user interfaces which can be created using the Aardvark.Media library. Aardvark Media uses the Elm Architecture [imp]. The Elm Architecture is based on three key concepts:

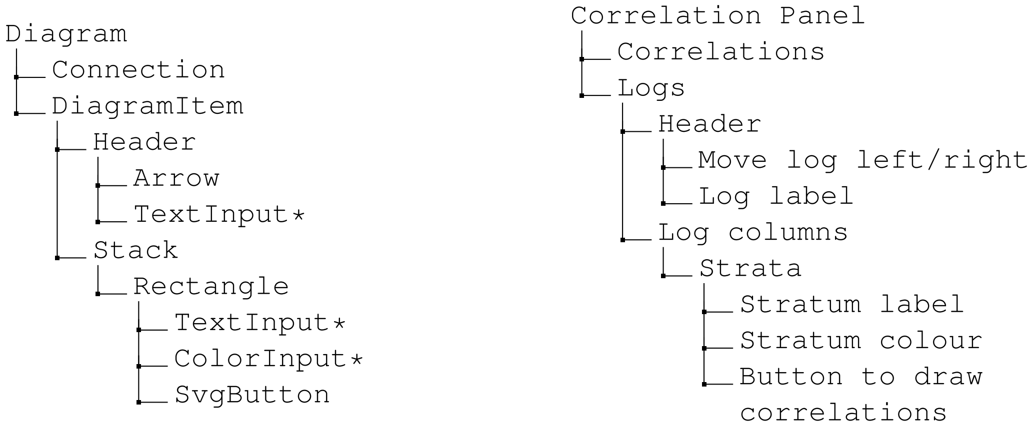
- **Model** The application state.
- **View** A function converting the state into code that is used to display that state to the user (for example HTML or OpenGL), and triggers messages.
- **Update** A function which changes the state via messages.

If a user interacts with the graphical user interface, they trigger messages which are passed to the *Update* function, and change the *Model*. The system is modular and modules (called *apps* in Aardvark.Media) build on each other, beginning with small modules which can be combined to create more complex constructs. When writing applications using Aardvark.Media, we use the same logic.

For instance, the app responsible for displaying the correlation panel consists of the sub-apps in Figure 5.12.

Each of the apps consists of the three parts of the Elm Architecture: The Model is a type in F#. The Update function receives messages triggered by user interactions or elsewhere in the code, and updates the current model directly or by distributing messages to sub-apps to update sub-models. The View function defines the user interface for each app, by calling the View function of a sub-app whose user interface is integrated into the super-app. The code for each app is divided into two files, one for the Model with the type itself and all necessary simple types, and another for the update and view function. Simple types are types that do not require their own update and view logic.

An app can contain multiple instances of other apps, as in the case of the Diagram and many of its sub-apps. Using this modular approach has an advantage which was essential for this work. Changing or extending specific parts of the user interface means locating and replacing a very specific and self-contained piece of code. Even exchanging apps is relatively simple, given their modular nature.



(a) (b)

Figure 5.12: (a) An example of the hierarchical nature of apps in the architecture of the prototype. Hierarchy of apps of the DiagramApp. Apps with an asterisk are part of the Aardvark.Media library. (b) The correlation panel app uses the apps in (a), partly as sub-apps of more domain-specific apps.

The correlation panel plot is realised using SVG. To make the code easy to re-use, extend and replace, all of the underlying drawing code was implemented in a modular fashion. We created F# functions that provide wrappers for simple SVG functionality. On these wrappers we base ever more complex functionality up to a 2D camera which can be used to manipulate the view on the SVG canvas, also scaling labels appropriately. Another example are responsive buttons with mouse-over functionality, and the diagram application that is used to display correlation panels. Figure 5.13 shows the way in which Aardvark libraries and prototype modules are based on each other. The prototype module UI.Plus extends Aardvark.Media's graphical user interface library with multiple apps for user interface elements, as well as an app for keyboard input.

SvgPlus works in the same way as Aardvark.Media, using different layers of abstraction to realise more and more complex constructs (Figure 5.14). At the lowest level, F# wrappers to create simple SVG elements like lines or rectangles. These wrappers are used to create more complex objects with attached event handling, for example a clickable dotted line where colour, stroke width, dash length, dash distance, and beginning and end point are customizable. Wrappers and primitives have incremental and non-incremental versions to allow dynamic and static user interface elements.

An example for an app built on the implemented Svg.Plus system is the RoseDiagramApp. Figure 5.15 shows a rose diagram created with it. The app uses primitives that draw

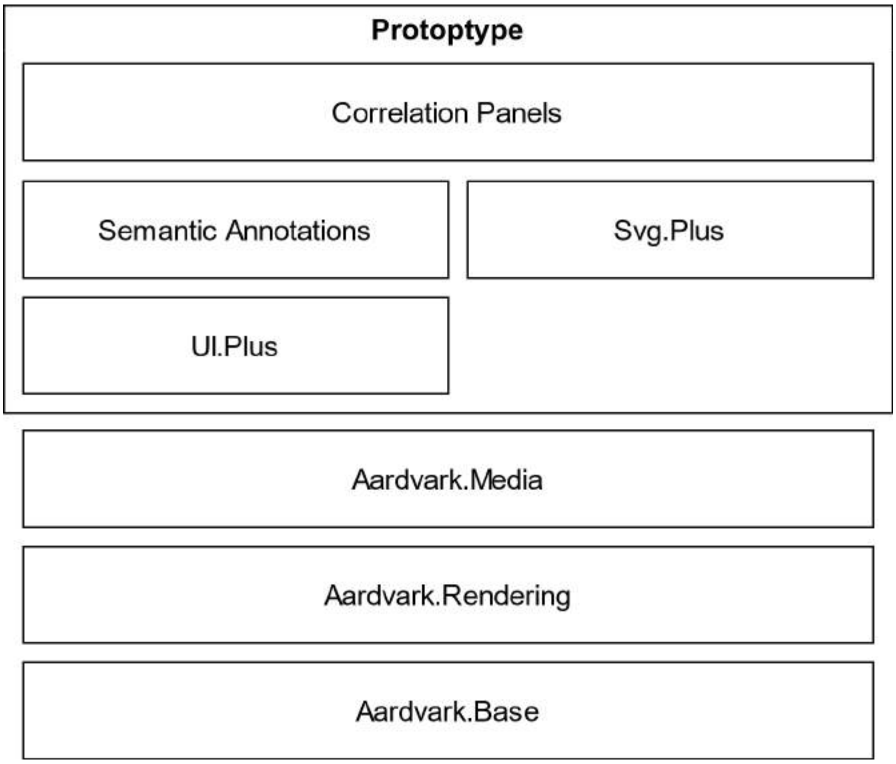


Figure 5.13: The system architecture of the prototype. Based on the Aardvark libraries Base, Rendering, and Media, we implemented extensions for user interface elements (UI.Plus) and SVG (Svg.Plus).

an arbitrary number of concentric circles, a set number of equally spaced radial lines between two circles, and filled circle segments (the single bins in a rose diagram). For creating the filled circle segments, wrapper functions for SVG paths were implemented. Event listeners could be attached to each of these elements (making, for example, each bin clickable).

The data structure selected to store logs is a tree, reflecting the hierarchical nature of strata. Each internal node in the tree represents a stratum in the log. Empty strata (without associated annotations) can be leaf nodes. Non-hierarchical annotations are always leaf nodes. To create a log from existing annotations, the annotations on an outcrop are recursively divided into strata.

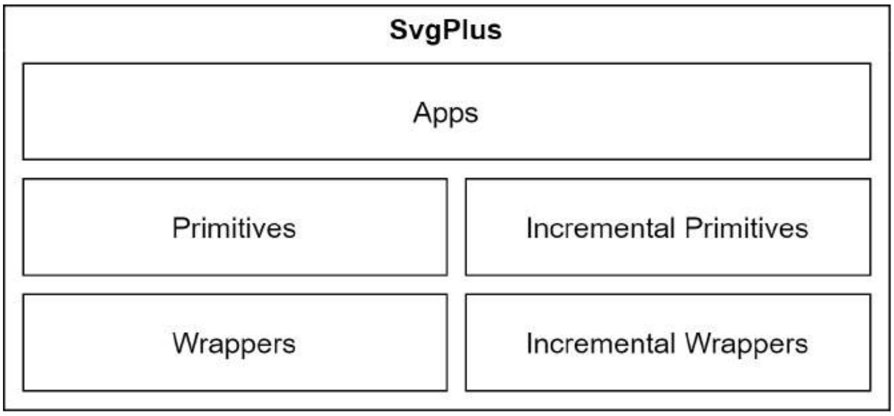


Figure 5.14: The architecture of Svg.Plus.

Figure 5.15: A sample rose diagram implemented using Svg.Plus.

CHAPTER 6

Results

Sedlmair et al. [SMM12] list three possible design study research contributions:

1. A **problem characterization and abstraction**.
2. A **validated visualisation design**.
3. A **reflection** on the design study itself.

The design study InCorr [OWN$^+$20], which this work contributes to, focuses on the second type of contribution, presenting a validated visualisation solution. This work's focus lies on the problem characterisation and abstraction, and based on that, the implementation of the prototype which was extended to create the InCorr visualisation solution. We also present lessons learned from the design process.

6.1 Semantic Annotations

We implemented a prototype for a semantic annotation system (Chapter 5) according to the requirements determined in Section 4.3.3 and presented it to our collaborators.

There was one major issue we had overlooked in our design. Some annotation types require not one, but two (or possibly more) geometric representations. For example, a plane for measuring an angle, and a line for measuring a related distance, or two angles, might be necessary to describe one geologic attribute. Our system should therefore be extended by allowing the user to select multiple geometry types for an annotation type. When drawing an annotation of that type, the tool could prompt the user to draw each geometry in order.

The simple view on annotation types for selecting a type when annotating was well received by our collaborators. When annotating, they prefer to use the majority of the available screen space for the 3D View. Therefore, they appreciated the minimised view showing only the labels of annotation types.

We requested feedback from our collaborators specifically regarding our selection of default semantic types, which the prototype provides. They said which types were provided made little difference. They assume that experts will create their own types anyway.

One minor issue was the use of the phrase *semantic type*. They argued that a better term should be used to make it immediately obvious what it means. We might have been too immersed in the world of visualisation design when choosing that name. We should have remembered to consider the language of the domain experts who would later use the tool.

6.2 Correlation Panels

In Section 4.1 we presented a task abstraction according to the multi-level typology of abstract visualisation tasks devised by Brehmer et al. [BM13]. We presented a design space analysis of geological logs and correlation panels, and various options for encodings, based on that analysis and input from experts in Chapter 4.

Tables 6.1 to 6.3 sum up the visualisation choices made for our prototype based on the work described in Chapter 4. The tables in that Chapter (Tables 4.2, 4.3 and 4.4) contain the results of the analysis in Section 4.3 and the information gathered from our collaboration with domain experts. In our prototype we implemented the design choices printed in bold as listed in Tables 6.1, 6.2, and 6.3. We used the prototype to automatically generate correlation panels, the results can be seen in Figure 5.10.

The reasoning behind most design choices for the prototype was to choose the simplest, but still viable, option. We took precautions to implement encodings in a way which facilitated an update to more complex encodings. Another factor that contributed to our choices were the requirements of our collaborators. For example, we implemented hierarchical columns rather than one simple column specifically because our collaborators favoured that choice.

In Chapter 5.1 we formulated four design goals. We did not verify whether we achieved the design goals in a formal user study, but we evaluated each design goal informally.

G1 The user's effort to generate a log should be minimal. In our prototype, a log can be created with a single click. If there are no grain size annotations for a stratum, the grain size has to be selected manually. This adds an effort of two clicks per stratum without grain size annotations.

G2 Editing (adding, deleting) correlations should be a quick and simple process. Correlations are added with two clicks, one on each contact that should be connected. They can be deleted with one click of the right mouse-button on the correlation.

G3 Updating or re-generating logs after adding new annotations should be a quick and simple process. Logs can be deleted by navigating to the Log List View,

Feature	Visual Encoding	Source
logs	Table 6.2	mixed
horizontal alignment	**(1.a) logs are evenly spaced**	-
	(1.b) the distance between logs in the correlation panel encodes the geographic distance between logs	data
	(2) distance between logs is labelled	data
vertical alignment	**(a) tops of logs are aligned**	-
	(b) user defined	user
	(c) use contact	mixed
	(d) use multiple contacts with priorities	mixed
order	**(a) user defined**	user
	(b) automatic sorting based on geographic location with manual control	mixed
correlations	**(1.a) straight lines connect contacts**	user
	(1.b) curved lines connect contacts	user
	(2) different line styles encode uncertainty	user
	(3.a) the area between correlations is coloured in	data
	(3.b) strata are drawn between logs	user
vertical axis labelling	(a) none	-
	(b) elevation	data
	(c) labelled scale bar	data

Table 6.1: Visualisation options for a correlation panels based on Section 4.3 and informal discussions with experts. Visualisation options selected for the prototype implementation are bold.

and clicking on delete. The prototype has no automatic update of logs, so it is necessary to delete and re-generate logs. Although deleting and re-generating a log is a quick process, automatic updates could reduce the effort for the user further.

G4 Preserve the properties of static correlation panels. We kept the design of the correlation panel created by the prototype close to existing examples of correlation panels. Due to the limited nature of the prototype there are some properties many static correlation panels have and our prototype does not. Published static logs are usually more polished than the correlation panels generated by our prototype. One example are the straight correlation lines in our prototype which might overlap labels. InCorr [OWN$^+$20] addresses this and other issues, like grain size labelling for each log.

In our estimation, we achieved design goals G1 and G2. There is some room for improvement concerning design goals G3 and G4.

Logs

Artefact	Visual Encoding	Source
strata	see Table 4.4	mixed
labelling	(a) **one text label**	user
	(b) one id label and one text label	mixed
horizontal axis labelling	(a) none	-
	(b) labelled horizontal axis (grain size)	data
vertical axis labelling	(a) none	-
	(b) **labelled vertical axis (elevation)**	data

Table 6.2: Visualisation options for logs based on Section 4.3 and informal discussions with experts. Visualisation options selected for the prototype implementation are in bold.

Feature	Visual Encoding	Source
stratum shape	**(a) strata are represented by simple rectangles**	data
	(b) strata can be complex polygons to encode different contact types	mixed
stratum width	(a) strata have uniform width	data
	(b) stratum width derived from grain size	mixed
stratum appearance	**(a) stratum colour derived from grain size**	data
	(b) textures used instead of simple colours to encode geological features	mixed
	(c) glyphs/symbols placed inside stratum to encode geological features	user
columns	**(1.a) grain size**	data
	(1.b) grain size and lithology	mixed
	(2) hierarchical columns	data
labels	**(a) none**	-
	(b) labelled strata	user

Table 6.3: Visualisation options for strata based on Section 4.3 and informal discussions with experts. Visualisation options selected for the prototype implementation are in bold.

6.3 Lessons Learned

In the section on the *discover* stage of their design studies methodology [SMM12], Sedlmair et al. argue that there is a *sweet spot* between an expert-level understanding of the target domain and no knowledge of that domain whatsoever. Some knowledge is necessary but acquiring much knowledge is prohibitively time-consuming. The only advice Sedlmair et al. have on finding this elusive balance is to accumulate enough experience doing design studies.

In the case of this work, in addition to the vast expert knowledge inherent in the domain, we faced a large domain-specific vocabulary that provided additional challenges. Geology is an old science that is still evolving, and its vocabulary evolves as well. Terms might be used differently in the literature a decade or two ago than they are now. This, and the lack of standardisation made it essential to establish the meaning of terms and concepts thoroughly and accurately. Despite our efforts, our domain collaborators occasionally had to correct assumptions we had made about a term or concept. This highlights the importance of collaborating with domain experts, and also the importance of acquiring domain knowledge diligently. It also exemplifies that we need to be conscious of the time it takes to acquire enough domain knowledge, and plan a design study accordingly.

CHAPTER 7

Further and Future Work

7.1 Further Work

This work was part of a long-term design study project, whose results were published in *InCorr: Interactive Data-Driven Correlation Panels for Digital Outcrop Analysis* [OWN^{+}20]. The prototype created in the course of this work was extended and the resulting visualisation tool validated in two ways. First, a correlation panel created with the tool was validated against a manually created correlation panel based on the same data (Figure 7.1), a second small-scale user study was conducted.

7.2 Future Work

In Chapter 4, the possibilities for future work on the data-driven generation of interactive correlation panels are varied and vast. It should also be noted that much of the work lies in the field of engineering rather than scientific investigation.

Improving Teamwork by Standardisation

We have mentioned the lack of standardisation for correlation panels and geological logs in this work. The experts who collaborated with us on this project disagreed with strict standards, but they did find that there was a need for consistency, especially when working in a team. This lead to the idea of restricting access to features that allow a change of encoding to certain users of the tool. In this way, expert users or team-leaders could choose encodings and ensure that the team produces consistent visualisations. Editing or creating annotation types is a feature that could conceivably be restricted when standardisation takes precedence over personal preference. For example, if multiple users annotate the same data, a fixed set of annotation types could be used for the sake of coherence.

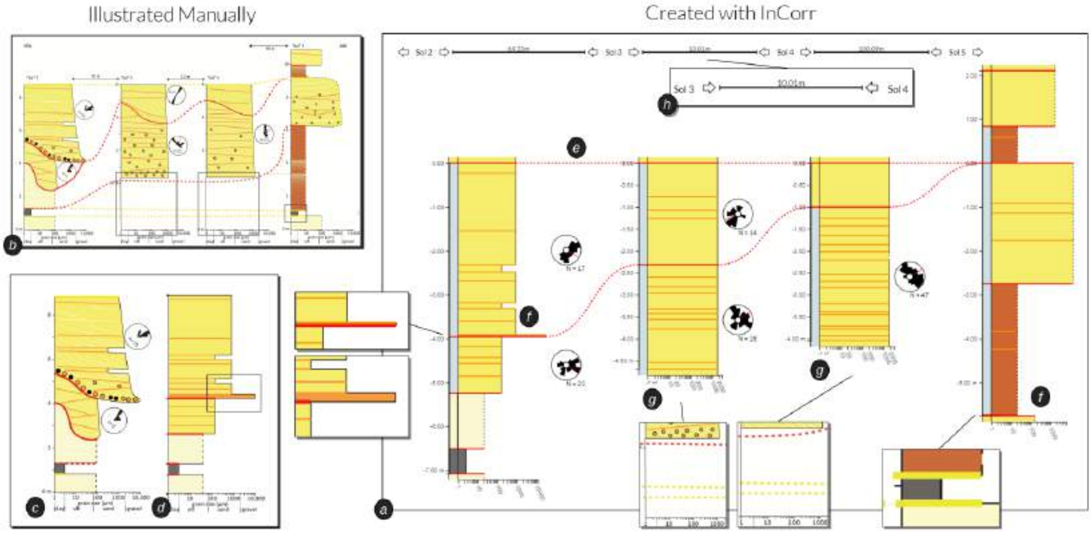


Figure 7.1: Comparison of correlation panels created with InCorr and by hand [OWN+20].

Semantic Zooming for Correlation Panels

One issue some correlation panels fight with is the limited space given to each log, especially if a large number of logs needs to be presented in one correlation panel as in Figure 4.14. The correlation panel presented in Figure 4.14 contains 22 logs, and the logs are accordingly sparse. To alleviate this problem the author adds detailed versions of the logs in the appendix [Cai09]. The implication for an implementation is that the amount of detail of the logs in a correlation panel should be adaptable.

The concept of semantic zooming could be applied to this problem. As explained in [Mun14], with semantic zooming, items are not only made bigger by zooming in, but they might change appearance significantly depending on the number of pixels available for their display.

In the context of correlation panels, differently detailed versions of a log exist, and according to the amount of screen space available to a log, an appropriate level of detail is selected. However, especially for use of a static image in a publication, the user should retain some control over the selected level of detail.

Patterns and Symbols

Our collaborators would like to be able to assign patterns and symbols to strata in the log. Using symbols and patterns to encode geological properties is common practice when drawing single logs. Correlation panels are usually less detailed, but there are also correlation panels that use patterns and symbols (as shown in Section 4.3). The reason correlation panels use patterns and symbols less often is that logs in a correlation panel usually need to be more compact to fit the whole correlation panel into the available space. By using semantic zooming, we can display highly detailed logs in a correlation panel, which includes the use of symbols and patterns. Our collaborators already draw their own patterns with drawing software like Corel Draw for use in their manually drawn correlation panels and logs. Creating a system that allows users to import their own SVG patterns. This feature could be restricted in the same way as annotation types to enhance teamwork and consistency. Assigning patterns and symbols could be done in the same way as assigning grain sizes.

Interactive Maps

One visualisation often combined with correlation panels in publications is an overview map of the area where the logs originate. The logs are marked and labelled on the map. Cruz-Orosa et al. [COBR+12] use this combination discussed in Section 2.2 (Figure 4.11). A dynamic map could be implemented as an additional view, with logs being automatically added when generated. It could also serve as an intuitive tool for navigating between individual logs, by selecting them in the map view. For easy use in publications, an export function could be provided.

Properties of Contacts

Collinson et al. state in their book *Sedimentary Structures* [CMT06], that, for environmental reconstruction, it is just as vital to document contacts and their properties as it is to document strata.

We discussed the variety of forms contacts can take in Section 4.3. Different encodings for various properties of contacts are presented in Figure 4.10. The first step is assigning properties like sharp, erosive, gradual, etc. to contacts. Then the encodings have to be defined, and assigned to properties. There are two options: creating pre-defined encodings, for example similar to Tucker's [Tuc03], or letting experts define their own encodings. The properties of a contact might imply a change in the shape of strata in the log. For example, this is the case for a gradual contact.

Searching

When geologists annotate one, or even multiple digital outcrop models, they might create thousands of annotations. Our collaborators requested a feature that would allow them to quickly search for specific items. For example, geologists name distinctive contacts that can be identified in larger areas. They want to be able to find, and zoom in on these items. The coordinated view system of 3D View and correlation panel could be extended with a system that allows experts to zoom in on items in both views after searching for them.

Extended 3D Log

A log is currently encoded as a polyline in the 3D View. The current system draws logs in the 3D View as lines connecting the selected points on hierarchical annotations. This visualisation could be extended by encoding, for example, the grain size into the visualisation as it is in the log in the correlation panel.

As we discussed in Chapter 3, Buckley et al. [BRN$^+$19] project the picture of a log onto the digital outcrop model. Similarly, more detailed versions of logs (see also semantic zooming above) could be drawn in the 3D View. It would be possible to implement interactions that allow experts to edit properties of logs in the 3D View.

Correlations in the 3D View

Currently, correlations are only visualised in the correlation panel. It would be possible to integrate them into the 3D View as well. The encoding would have to be chosen carefully. Connecting correlated strata with polylines might lead to visual clutter, and would be confusing if outcrops are not right next to each other. Encoding correlations with glyphs that allow the user to navigate between connected strata might be a solution.

Assigning Annotations to Logs and Strata

To create a correlation panel, we first need a method to create a single log. The input we need to create a log is the following:

- at least two hierarchical annotations (usually there will be much more)
- the location of the log
- the grain size of each stratum

A log is not one point in space, but rather a line following the surface of the outcrop as illustrated in Figure 4.13. The annotations divide the outcrop into multiple strata. Other annotations that lie between two annotations are part of the stratum. However, which annotations should be considered when generating a log is not quite clear. Annotations that intersect the line defining the location of the log can be considered part of the log without much controversy, but in general experts want to consider annotations in a wider area. In our prototype we do not implement methods to define a wider area, but take all annotations on an outcrop into account. This approach could be improved by letting the expert set a distance from the log beyond which annotations are not incorporated into that log.

The simplest automatic method to sort annotations into strata is to take the elevation at each point selected on hierarchical annotations to define the location of the log, and assign (non-hierarchical) annotations by comparing their average elevation with the elevations of these points. This approach was used for the prototype. It could be improved by using the planes created by a hierarchical annotation with associated dip and strike measurements (the *contact plane* between two strata) to determine which stratum an annotation should be assigned to. As our collaborators pointed out, there might still be annotations that could not be assigned easily. For example, if an annotation stretches across the boundary of two strata. Our collaborators suggested that the user should be allowed to assign and re-assign annotations to strata by hand, which was implemented in InCorr [OWN$^+$20]. Following feedback from our collaborators, in InCorr we did not automate the process of assigning annotations to strata, but left it to the experts to assign them manually.

CHAPTER 8

Conclusion

To create a geological model of an environment, geologists use a diagram called *correlation panel*. This diagram consists of multiple *geological logs*, which contain abstracted graphic descriptions of rock layers at a certain location. Layers that are present in different logs are connected visually in the correlation panel. Experts usually create these Figures using drawing software, which is a very time-consuming process, and also makes changing a correlation panel once completed difficult. This work is part of a design study with the goal to create interactive correlation panels in a data-driven way. The goal of the design study way to make it possible for experts to use correlation panels during the interpretation stage rather than at its end, and preserve the context of the encoded data by linking them to their origin. In this work, after a short introduction to relevant topics, we analysed published correlation panels to explore the design space. With the results of the analysis and the input of experts we could collect in the course of workshops and a research stay at the *Imperial College London*, we discussed possible design choices, and described which were used as requirements for our prototype. We presented the prototype, and how it can be used to derive from interpretation data generate simple correlation panels. We also presented reflections and lessons learned of our design and implementation process.

APPENDIX A

Legends

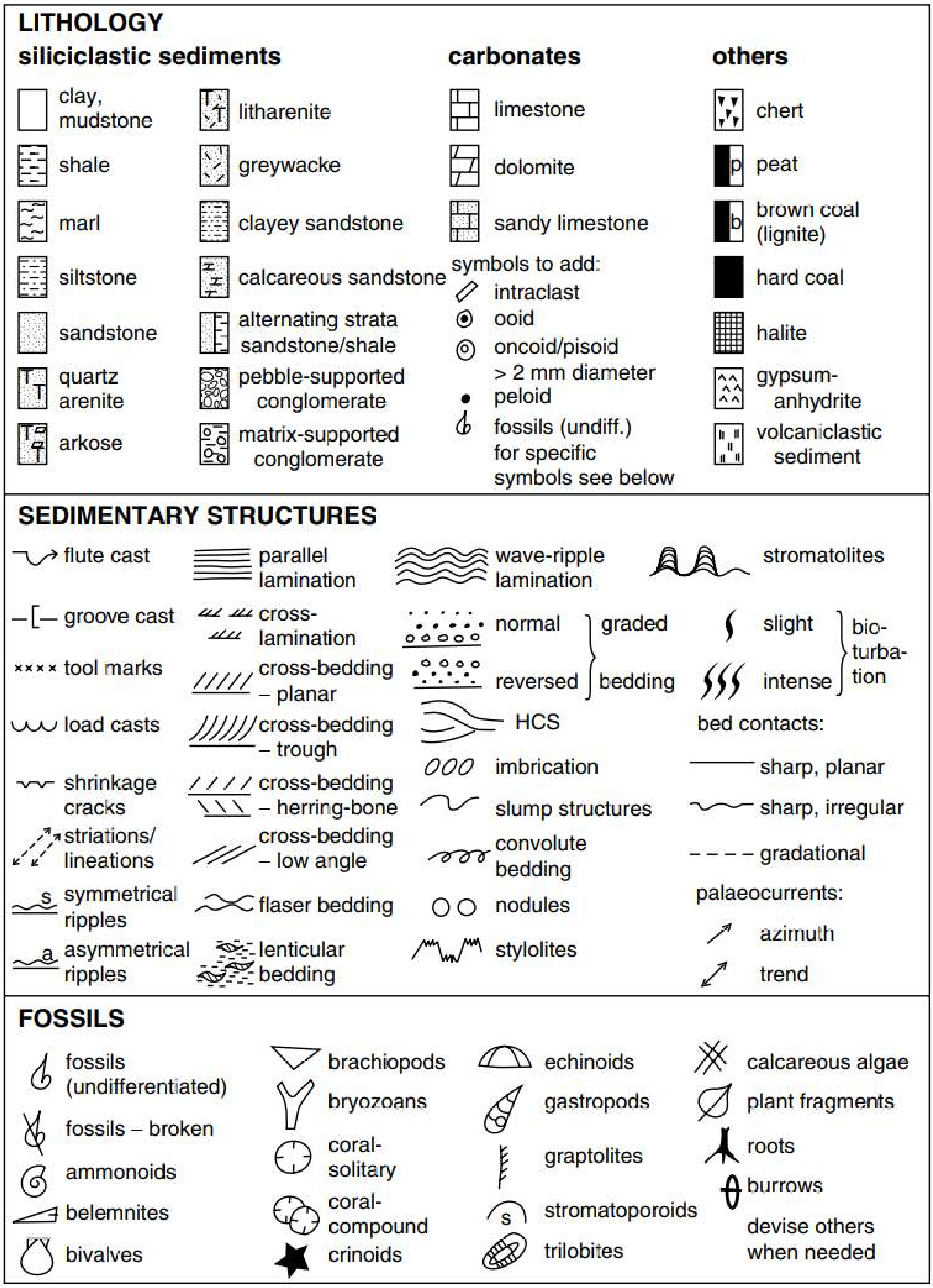


Figure A.1: Symbols for lithology, sedimentary structures, and fossils for use in a graphic log (as used in Figure 2.8) [Tuc03]. When we compare the patterns and symbols used by various authors, we find that although many are similar, sometimes completely different encodings are used for the same property (see Section 4.3).

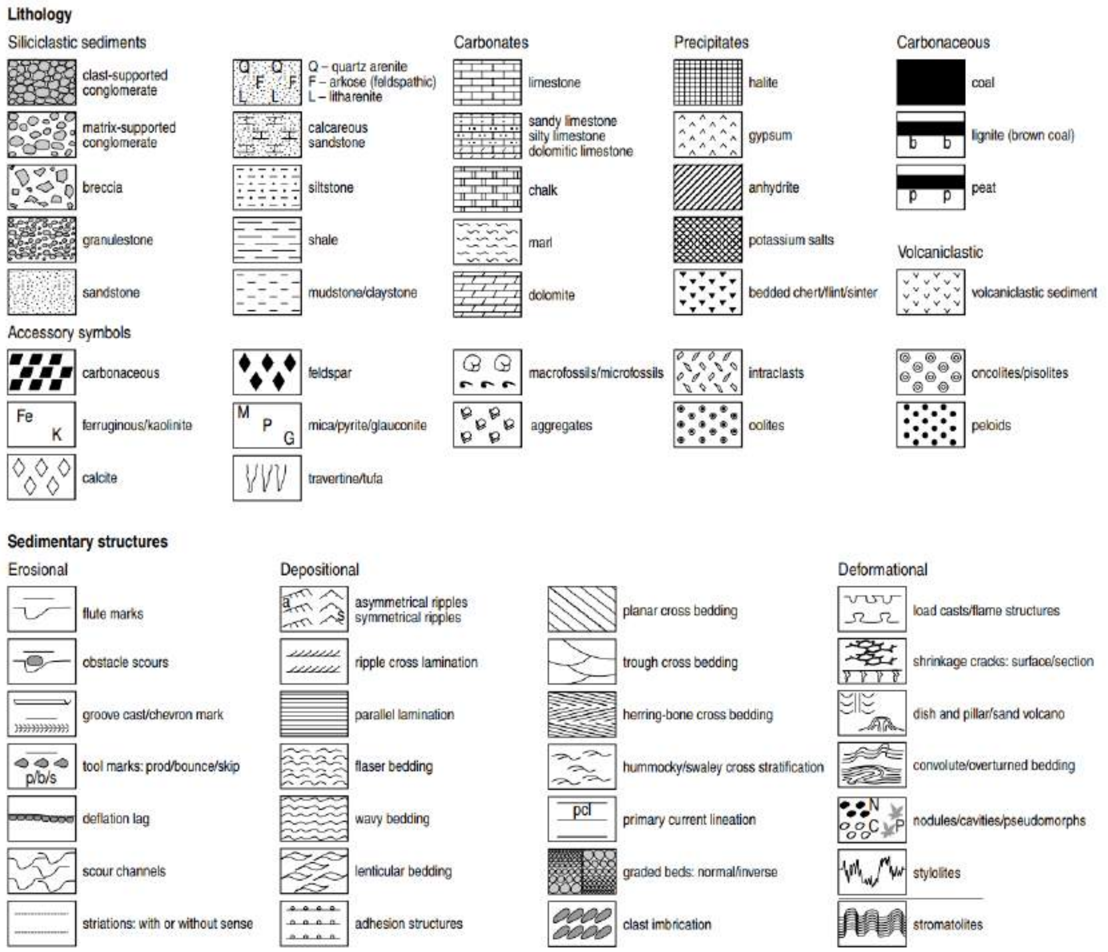


Figure A.2: Symbols for lithology and sedimentary structures used in *Sedimentary Structures, 3rd Edition by Collinson et al.* (used in Figure 4.5) [CMT06].

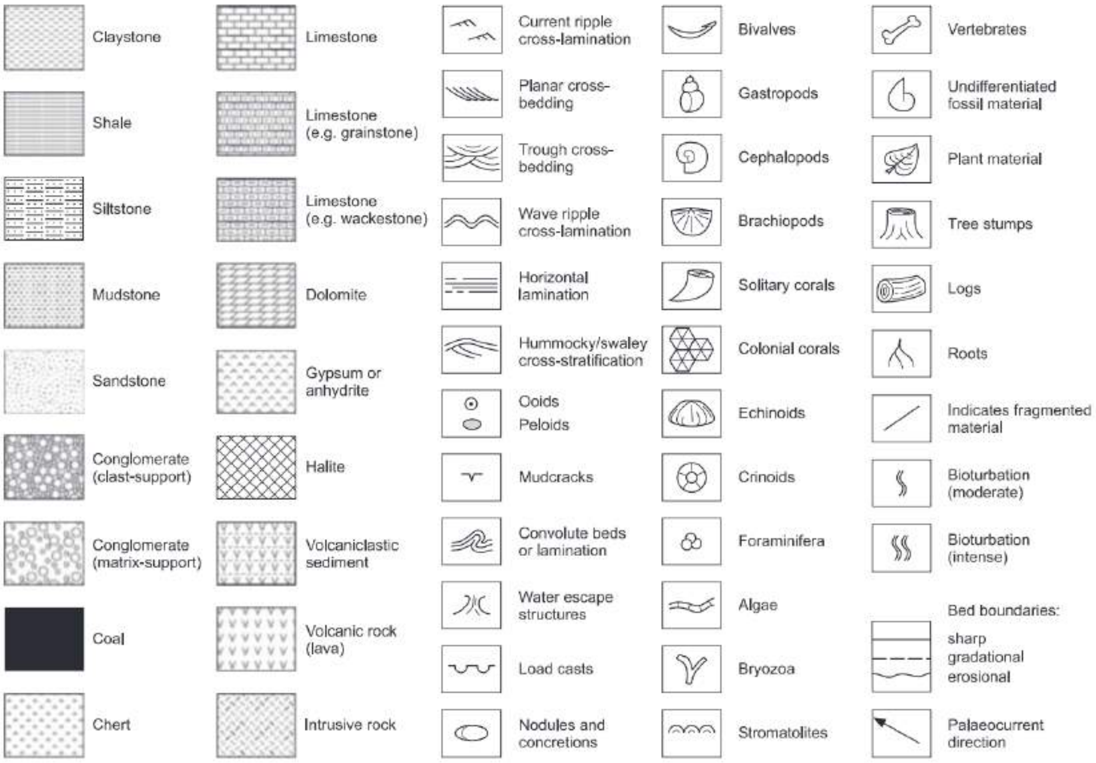


Figure A.3: Patterns and symbols for logs as presented in *Sedimentology and Stratigraphy, 2nd Edition* [Nic09], page 72.

List of Figures

List of Tables

Bibliography

[AAa] National Aeronautics and Space Administration. The cameras on the mars 2020 perseverance rover. https://mars.nasa.gov/mars2020/spacecraft/rover/cameras/. Accessed: 2021-04-05.

[AAb] National Aeronautics and Space Administration. Mars curiosity rover. https://mars.nasa.gov/msl/home/. Accessed: 2021-04-06.

[AAc] National Aeronautics and Space Administration. Mars curiosity rover mission results. https://mars.nasa.gov/msl/mission/science/results/. Accessed: 2021-04-06.

[AAd] National Aeronautics and Space Administration. Mars reconnaissance orbiter. https://mars.nasa.gov/mro/. Accessed: 2020-05-06.

[AAe] National Aeronautics and Space Administration. Mastcam. https://mars.nasa.gov/msl/spacecraft/instruments/mastcam/. Accessed: 2021-03-25.

[AAf] National Aeronautics and Space Administration. Mastcam-z. https://mars.nasa.gov/mars2020/spacecraft/instruments/mastcam-z/. Accessed: 2021-03-25.

[AAg] National Aeronautics and Space Administration. Mer. https://mars.nasa.gov/mer/mission/science/results/. Accessed: 2020-05-04.

[AAh] National Aeronautics and Space Administration. Mer pancam. https://mars.nasa.gov/mer/mission/instruments/pancam/. Accessed: 2020-04-06.

[AAi] National Aeronautics and Space Administration. Odyssey. https://mars.nasa.gov/mars-exploration/missions/odyssey/. Accessed: 2020-05-03.

[AAj] National Aeronautics and Space Administration. Pathfinder. https://mars.nasa.gov/mars-exploration/missions/pathfinder/. Accessed: 2020-05-03.

[AAk] National Aeronautics and Space Administration. Phoenix mars lander. https://www.nasa.gov/mission_pages/phoenix/main/index.html. Accessed: 2021-04-06.

[Age] European Space Agency. Missions to mars. https://exploration.esa.int/web/mars/-/56504-missions-to-mars. Accessed: 2020-05-03.

[All13] Michael Allaby. *A dictionary of geology and earth sciences*. Oxford University Press, 2013.

[ASA19] Yazeed Alaudah, Moamen Soliman, and Ghassan AlRegib. Facies classification with weak and strong supervision: A comparative study. In *SEG Technical Program Expanded Abstracts 2019*, pages 1868–1872. Society of Exploration Geophysicists, 2019.

[BGG$^+$15] Robert Barnes, Sanjeev Gupta, Michele Giordano, Jeremy Morley, Jochen Müller, Yanyan Tao, James Sprinks, Chris Traxler, Gerd Hesina, Thomas Ortner, et al. Geological interpretation and analysis of surface based, spatially referenced planetary imagery data using progis 2.0 and pro3d. *EPSC Abstracts, Vol. 10, 2015*, 10, 2015.

[BGT$^+$18] Robert Barnes, Sanjeev Gupta, Christoph Traxler, Thomas Ortner, Arnold Bauer, Gerd Hesina, Gerhard Paar, Ben Huber, Kathrin Juhart, Laura Fritz, et al. Geological analysis of martian rover-derived digital outcrop models using the 3-d visualization tool, planetary robotics 3-d viewer—pro3d. *Earth and Space Science*, 5(7):285–307, 2018.

[BH19] Brian S. Burnham and David Hodgetts. Quantifying spatial and architectural relationships from fluvial outcrops. *Geosphere*, 15(1):236–253, 2019.

[BKP$^+$20] Andreas Bechtold, Christian Köberl, Gerhard Paar, Christoph Traxler, Rebecca Nowak, and Filippo Garolla. Towards automatic detection of shatter cones on planetary rover images. In *Geological Society of America (GSA) 2020 Annual Meeting*, 2020.

[BLMG$^+$15] Jan Byška, Mathieu Le Muzic, Eduard Gröller, Ivan Viola, and Barbora Kozlikova. Animoaminominer: Exploration of protein tunnels and their properties in molecular dynamics. *IEEE transactions on visualization and computer graphics*, 22(1):747–756, 2015.

[BM13] Matthew Brehmer and Tamara Munzner. A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, 2013.

[BRN$^+$19] Simon J. Buckley, Kari Ringdal, Nicole Naumann, Benjamin Dolva, Tobias H. Kurz, John A. Howell, and Thomas J.B. Dewez. LIME: Software for 3-D visualization, interpretation, and communication of virtual geoscience models. *Geosphere*, 15(1):222–235, 01 2019.

[Cai09] Stephen Cain. *Sedimentology and stratigraphy of a terminal fluvial fan system: the Permian Organ Rock Formation, South East Utah.* PhD thesis, Keele University, 2009.

[CMT06] John Collinson, Nigel Mountney, and David Thompson. *Sedimentary Structures, 3rd Edition.* Terra Publishing, PO Box 315, Harpenden, Hertfordshire AL5 2ZD, England, 2006.

[COBR$^+$12] Israel Cruz-Orosa, Francesc Bat, Emilio Ramos, Lluís Rivero, and Yaniel Taset. Structural evolution of the la trocha fault zone: Oblique collision and strike-slip basins in the cuban orogen. *Tectonics*, 31, 10 2012.

[Coe10] Angela Coe, editor. *Geological Field Techniques.* Blackwell Publishing Ltd in association with The Open University, Walton Hall, Milton Keynes MK7 6AA, United Kingdom, 2010.

[FPHR10] Ivan Fabuel-Perez, David Hodgetts, and Jonathan Redfern. Integration of digital outcrop models (doms) and high resolution sedimentology – workflow and implications for geological modelling: Oukaimeden sandstone formation, high atlas (morocco). *Petroleum Geoscience*, 16(2):133–154, 2010.

[GdSJVK$^+$18] Luiz Gonzaga da Silveira Jr, Mauricio Veronez, Gabriel Kannenberg, Demetrius Alves, Leonardo Santana, Jean de Fraga, Leonardo Inocencio, Laís Souza, Fernando Marson, Fabiane Bordin, Francisco Tognoli, Kim Senger, and Caroline Cazarin. A multioutcrop sharing and interpretation system: Exploring 3-d surface and subsurface data. *IEEE Geoscience and Remote Sensing Magazine*, 6:8–16, 06 2018.

[GJ14] John Grotzinger and Thomas H. Jordan. *Understanding Earth (Fifth edition).* W.H. Freeman and Company, New York, NY 10010, 41 Madison Avenue, seventh edition, 2014.

[GVA$^+$17] Luiz Gonzaga, Mauricio Roberto Veronez, Demetrius Nunes Alves, Fabiane Bordin, Gabriel Lanzer Kannenberg, Fernando P. Marson, Francisco M.W. Tognoli, and Leonardo C. Inocencio. Mosis—multi-outcrop sharing & interpretation system. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 5209–5212. IEEE, 2017.

[HBG+11] Thomas Höllt, Johanna Beyer, Fritz Gschwantner, Philipp Muigg, Helmut Doleisch, Gabor Heinemann, and Markus Hadwiger. Interactive seismic interpretation with piecewise global energy minimization. In *2011 IEEE Pacific Visualization Symposium*, pages 59–66. IEEE, 2011.

[HGE+11] Alexander G. Hayes, John P. Grotzinger, Lauren A. Edgar, Steven W. Squyres, Wesley A. Watters, and Jasha Sohl-Dickstein. Reconstruction of eolian bed forms and paleocurrents from cross-bedded strata at victoria crater, meridiani planum, mars. *Journal of Geophysical Research: Planets*, 116(E7), 2011.

[HGQ+06] Wei Hong, Xianfeng Gu, Feng Qiu, Miao Jin, and Arie Kaufman. Conformal virtual colon flattening. In *Proceedings of the 2006 ACM symposium on Solid and physical modeling*, pages 85–93, 2006.

[HGS+11] Gary J. Hampson, M. Royhan Gani, Kathryn E. Sharman, Nawazish Irfan, and Bryan Bracken. Along-strike and down-dip variations in shallow-marine sequence stratigraphic architecture: Upper cretaceous star point sandstone, wasatch plateau, central utah, usa. *Journal of Sedimentary Research*, 81(3):159–184, 2011.

[HGWR07] David Hodgetts, Robert L. Gawthorpe, Paul Wilson, and Frank Rarity. Integrating digital and traditional field techniques using virtual reality geological studio (vrgs). In *69th EAGE Conference and Exhibition incorporating SPE EUROPEC 2007*, pages cp–27. European Association of Geoscientists & Engineers, 2007.

[Hod17] David Hodgetts. Virtual reality geological studio (vrgs). http://www.vrgeoscience.com/, 2017. Accessed: 2018-01-26.

[HSMT15] Georg Haaser, Harald Steinlechner, Stefan Maierhofer, and Robert F. Tobler. An incremental rendering vm. In *Proceedings of the 7th Conference on High-Performance Graphics*, pages 51–60, 2015.

[imp] The elm architecture. https://guide.elm-lang.org/architecture/. Accessed: 2021-05-11.

[JPLAa] National Aeronautics Jet Propulsion Laboratory and Space Administration. Mer - spirit. https://www.jpl.nasa.gov/missions/mars-exploration-rover-spirit-mer-spirit/. Accessed: 2020-05-04.

[JPLAb] National Aeronautics Jet Propulsion Laboratory and Space Administration. Opportunity. https://www.jpl.nasa.gov/missions/details.php?id=5909. Accessed: 2020-05-04.

[KFW+02] Armin Kanitsar, Dominik Fleischmann, Rainer Wegenkittl, Petr Felkel, and Eduard Gröller. *CPR-curved planar reformation.* IEEE, 2002.

[KMM+18] Julian Kreiser, Monique Meuschke, Gabriel Mistelbauer, Bernhard Preim, and Timo Ropinski. A survey of flattening-based medical visualization techniques. *Computer Graphics Forum*, 37(3):597–624, 2018.

[Lim18] Schlumberger Limited. Petrel e & p software platform. https://www.software.slb.com/products/petrel, 2018. Accessed: 2018-01-26.

[Lim21] Schlumberger Limited. Petrel e & p software platform. petrel well correlation. https://petromehras.com/petroleum-software-directory/geology-software/petrel-well-correlation, 2021. Accessed: 2021-03-22.

[LKF+18] Qun Liu, Ben Kneller, Claus Fallgatter, Victoria Valdez Buso, and Juan Pablo Milana. Tabularity of individual turbidite beds controlled by flow efficiency and degree of confinement. *Sedimentology*, 65(7):2368–2387, 2018.

[LSW+19] Yuhua Liu, Chen Shi, Qifan Wu, Rumin Zhang, and Zhiguang Zhou. Visual analytics of stratigraphic correlation for multi-attribute well-logging data exploration. *IEEE Access*, 7:98122–98135, 2019.

[mbH] Joanneum Research Forschungsgesellschaft mbH. Joanneum research. https://www.joanneum.at/. Accessed: 2021-04-18.

[MHd+20] Ademir Marques, Rafael Kenji Horota, Eniuce Menezes de Souza, Lucas Kupssinskü, Pedro Rossa, Alysson Soares Aires, Leonardo Bachi, Mauricio Roberto Veronez, Luiz Gonzaga, and Caroline Lessio Cazarin. Virtual and digital outcrops in the petroleum industry: A systematic review. *Earth-Science Reviews*, 208:103260, 2020.

[Mun14] Tamara Munzner. *Visualization analysis and design.* CRC press, 2014.

[Nic09] Gary Nichols. *Sedimentology and Stratigraphy, 2nd Edition.* John Wiley & Sons, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK, 2009.

[OHB+20] Laurence O'Rourke, Philip Heinisch, Jürgen Blum, Sonia Fornasier, Gianrico Filacchione, Hong Van Hoang, Mauro Ciarniello, Andrea Raponi, Bastian Gundlach, Rafael Andrés Blasco, et al. The philae lander reveals low-strength primitive ice inside cometary boulders. *Nature*, 586(7831):697–701, 2020.

[OWN+20] Thomas Ortner, Andreas Walch, Rebecca Nowak, Robert Barnes, Thomas Höllt, and Eduard Gröller. Incorr: interactive data-driven correlation panels for digital outcrop analysis. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):755–764, 2020.

[PBG+18] Gerhard Paar, Robert Barnes, Sanjeev Gupta, Christoph Traxler, Matt Gunn, and Christian Köberl. Validation of the 3d vision and visualization frameworks provip and pro3d for the mars2020 and exomars stereo panoramic camera systems. In Louise Prockter, Eileen Stansbery, David Draper, and Walter Kiefer, editors, *49th Lunar and Planetary Science Conference*. Lunar and Planetary Institute, 2018.

[PMT+13] Gerhard Paar, Jan-Peter Müller, Y. Tao, T. Pajdla, J. G. Morley, K. Willner, I. Karatchevtseva, Christoph Traxler, Gerd Hesina, L. Tyler, D. Barnes, and Sanjeev Gupta. Provide: Planetary probes' mass vision data processing. In *Proceedings of European Planetary Science Congress (EPSC 2013)*, volume 8, pages 289–2. European Planetary Science Congress, 2013.

[PMT+15] Gerhard Paar, Jan-Peter Müller, Y Tao, T Pajdla, M Giordano, E Tasdelen, et al. Provide: Planetary robotics vision data processing and fusion. In *European Planetary Science Congress*, 2015.

[PTN+20] Gerhard Paar, Christoph Traxler, Rebecca Nowak, Filippo Garolla, Andreas Bechtold, Christian Köberl, Miguel Yuste Fernandez Alonso, and Oliver Sidla. Mars-dl: Demonstrating feasibility of a simulation-based training approach for autonomous planetary science target selection. In n.n., editor, *Europlanet Science Congress (EPSC) 2020*, 2020.

[RvG18] Angelo Pio Rossi and Stephan van Gasselt, editors. *Planetary Geology*. Springer, Cham, 2018.

[RVLH+14] Frank Rarity, Xavier M.T. Van Lanen, David Hodgetts, Robert L. Gawthorpe, P. Wilson, Ivan Fabuel-Perez, and J. Redfern. Lidar-based digital outcrops for sedimentological analysis: workflows and techniques. *Geological Society, London, Special Publications*, 387(1):153–183, 2014.

[SBC+17] Jürgen Schieber, David Bish, Max Coleman, Mark Reed, Elisabeth M. Hausrath, John Cosgrove, Sanjeev Gupta, Michelle E. Minitti, Kenneth S. Edgett, and Mike Malin. Encounters with an unearthly mudstone: Understanding the first mudstone found on mars. *Sedimentology*, 64(2):311–358, 2017.

[SMM12] Michael Sedlmair, Miriah Meyer, and Tamara Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE transactions on visualization and computer graphics*, 18(12):2431–2440, 2012.

[TFN+20] Christoph Traxler, Laura Fritz, Rebecca Nowak, Gerhard Paar, Christian Köberl, Andreas Bechtold, Filippo Garolla, and Oliver Sidla. Simulating rover imagery to train deep learning systems for scientific target selection. In *Europlanet Science Congress (EPSC) 2020*. Europlanet Science Congress, 2020.

[Tuc88] Maurice Tucker, editor. *Techniques in Sedimentology*. Blackwell Scientific Publications, Osney Mead, Oxford OX2 OEL, 1988.

[Tuc03] Maurice Tucker. *Sedimentary Rocks in the Field, 3rd Edition*. John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, 2003.

[VRVa] VRVis. The aardvark platform. https://aardvarkians.com/. Accessed: 2020-05-06.

[VRVb] VRVis. Pro3d. http://pro3d.space. Accessed: 2021-03-24.

[Wla18] Scott Wlaschin. *Domain Modeling Made Functional: Tackle Software Complexity with Domain-Driven Design and F.* Pragmatic Bookshelf, 2018.

[WSFL18] Xinming Wu, Yunzhi Shi, Sergey Fomel, and Fangyu Li. Incremental correlation of multiple well logs following geologically optimal neighbors. *Interpretation*, 6(3):T713–T722, 2018.

[ZS09] Peng Zhou and Zhicai Shang. 2D molecular graphics: a flattened world of chemistry and biology. *Briefings in Bioinformatics*, 10(3):247–258, 03 2009.