# USER CENTERED AUDIO INTERFACE FOR CLIMATE SCIENCE

*Visda Goudarzi, Hanns Holger Rutz, Katharina Vogt*

Institute of Electronic Music and Acoustics
University of Music and Performing Arts
Inffeldg. 10/3, 8010 Graz, Austria
{goudarzi,rutz,vogt}@iem.at

## ABSTRACT

This paper presents a user centred design approach to create an audio interface in the context of climate science. Contextual inquiry including think-aloud protocols was used to gather data about scientists' workflows in a first round of interviews. Furthermore, focus groups were used to gather information about the specific use of language by climate experts. The interviews have been analysed for their language content as well. Two goals are envisaged with this basic assessment. First, a climate terminology will help realising a domain-specific description of the sonifications that are understandable in the field. Second, identifying metaphors can help building a metaphoric sound identity for the sonification. An audio interface shall enrich their perceptualisation possibilities, based on the language metaphors derived from the interviews. Later, in a separate set of experiments, the participants were asked to pair sound stimuli with climate terms extracted from the first interviews and evaluate the sound samples aesthetically. They were asked to choose sound textures (from a set of sounds given to them) that best express the specific climate parameter and rate the relevance of the sound to the metaphor. Correlations between climate terminology and sound stimuli for the sonification tool is assessed to improve the sound design. Intuitiveness, learnability, memorability, and aesthetic preference of the sounds is measured by evaluations.

## 1. INTRODUCTION

As sensors increase resolution and models become more complex, the amount of data being processed today is steadily increasing, and both scientists and society need new ways to understand scientific data and their implications. Sonification is especially suited to the preliminary exploration of complex, dynamic, and multidimensional data sets. These kind of large and multivariate time-based data sets appear for example in climate science, coming both from empirical satellite provided sources as well as models. Examples of sonification in the context of climate research are given by Halim et al [1], who present a rain prediction auditory icon, and by Bearman [2], who uses sound to represent uncertainty in UK climate projections data. A. Polli sonified storm data from weather models [3].

In our research project (SysSon), we apply a systematic approach to design sonifications of climate data. In collaboration with the (Wegener Center for Climate and Global Change) climate research group, we assessed the parameters climate scientists use and their typical workflows. This background has been used to design and develop a multi modal interface (our sonification tool), which is integrated with the visualisation tools the scientists use already for data analysis. A sonification prototype is built and will be evaluated according to its functionality and usability for climate scientists, as well as under aesthetic criteria. In the current stage of the project, conceptual links between climate science and sound have been elaborated and first sonification designs have been developed.

## 2. NEEDS ASSESSMENTS

In order to assess the needs of climate scientists with regard to their data analysis methods we investigated their research context applying contextual inquiry and observed focus groups. Based on the collected data, we applied different evaluation techniques: from simple quantitative analysis and a reflection of the workflows to experimental qualitative analysis of the terms and metaphors used in communication between climate scientists. Interviews have been conducted in German, the native language of all participants, audio-recorded and partly transcribed for analysis. All participants received headphones as an acknowledgement for their participation (meanwhile encouraging the research lab with additional audio infrastructure.)

### 2.1. Contextual Inquiry and Focus Groups

Contextual task analysis is an established technique in HCI [4] therefore we decided to explore different data analysis tasks that climate scientists are regularly involved in. The approach is challenging because each scientist uses an individual set of programs and performs different tasks, due to different habits and background. Therefore we conducted a usability study in a non-classical sense, following Karat et al [5]. In a field study an observer and an interviewer visited climate scientists in their workplace to capture their workflows, and the environmental factors while analysing data. Following a questionnaire they assessed the general questions and marked if all relevant topics have been covered during the open task. Interviews took about an hour and consisted of three parts. The central part of the individual interview session consisted of a walk-through of a self-chosen data analysis task. Finally, expectations about an auditory display were collected, including a recording of what the data in the task would sound like, which data sets would be most useful for the participants to sonify, and how and if they would use sound at their work. Focus groups were conducted to observe more specific information about the communication between the experts. Participants belong to three different research groups. Participants brought their own task results, as had been demonstrated in the contextual inquiry, and were asked to briefly present and discuss them with the other members of the group. The focus groups took about one hour each and were

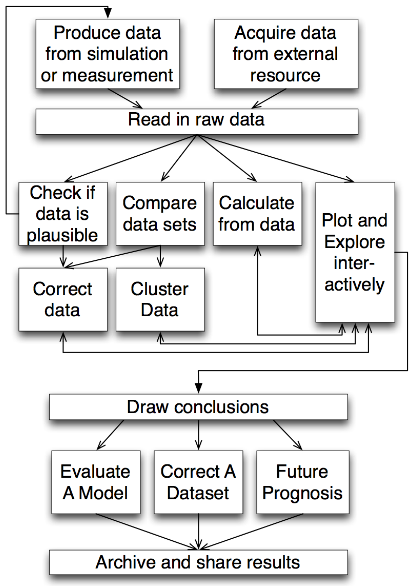


Figure 1: Workflow from the observed tasks.

observed by the authors of this paper without interfering.

### 2.2. Work Flow Analysis

Fig. 1 shows a common workflow summarising the data analysis process in all three user groups. The task of data analysis is very similar and can probably be generalised to other scientific disciplines as well. The first step is the acquisition of data, either from external research institutions or from their own simulations. What usually follows in a next step is a quick check on the plausibility of the data, e.g., by plotting or, sometimes, scanning through the numbers by hand. Then, in many cases some secondary data are derived from the first raw data by calculations following a hypothesis. Following the results of these steps and the tentative plots of data, the original data are corrected or clustered; results at this stage are always plotted and/or explored interactively. Conclusions from this process consist of either the evaluation of a model, the correction of a data set, and/or some future prognosis, such as climate predictions. Finally, results are archived and shared; here the plots usually serve as a basis in discussions and publications.

The workflow analysis shows that climate scientists depend heavily on the visual display of their data. At the same time, the amount and multi-variance of the data makes them hard to visualise, therefore many scientists expressed their dissatisfaction with existing visualisation techniques or their knowledge thereof. For a further analysis, we assessed and analysed the visualisations that were used in the exemplary tasks during the interviews. We were surprised to find that the average number of data sets that the scientists wanted to compare with each other in one task was as high as 47, with single tasks demanding up to 400 sets (25 different colour-coded climate models, provided for four different altitudes of the atmosphere and for four different regions, i.e. each having 16 sequential plots.) About half of the visualisations are more or less self-explanatory, assuming a basic understanding of the field, but a few of them were either difficult to understand or, in the case of the 400 data slices, even confusing. The visualisation methods chosen mostly employed standard methodology, e.g. line charts and maps. A few researchers developed their own visualisations. Fig. 2 shows typical visualisations in climate science.

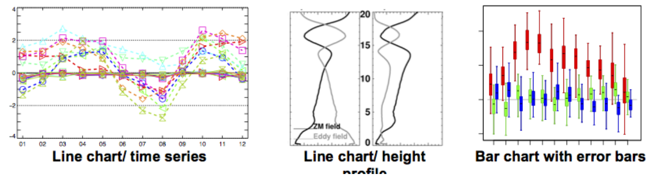


Figure 2: Overview of typical visualisations of climate data.

### 2.3. Contextual Analysis

As a qualitative analysis, both the interviews from the contextual inquiry and the focus groups have been analysed for their language content. A climate terminology will help realising a domain-specific description of the sonifications that are understandable in the field and identifying metaphors can help building a metaphoric sound identity for the sonification. In a quick check on the correlation of mentions of words, a small trend to using similar vocabulary within the same research group could be seen. The difference of the focus of the research groups is reflected in the language. The richness in vocabulary, i.e., the number of different words mentioned by each person, does not correlate with his/her experience in the field.

In the next step of analysis, the words have been grouped. The categories for the groups have been determined iteratively, where final categories emerged while trying to group the data as far as possible. The categories most often cited in the interviews are *data analysis*, *simulation*, *description of climate phenomena*, and *data properties*, which is not surprising because of the task the participants have been asked to show. Comparing the master-master communication in the focus groups and the master-layman communication in the personal interviews, it turned out that in the latter condition the scientists talked more about general phenomena and less about data analysis. The top 20 sub-categories used by the subjects in interviews and focus groups were analysed and showed the following:

- Climate scientists use visualisation as a basis of their work (e.g., *look at* and *plot*);
- Temperature is the most important climate parameter they are interested in;
- In terms of working style, programming is the daily job of most of the scientists (e.g., using words such as *climate model*, *program structure*);
- The mathematics used is often rather basic, e.g., *difference* is still in the top 10, the most important basic method when comparing data sets amongst each other.

Regarding the generalised categories *data* and *climate phenomenon*, it turned out that for data analysis the most important method is correlating or finding relations between two data sets. Also visual analysis is often used. Next, preparatory steps are important, including for instance data acquisition, listing, simple calculations, calibration, and transformation of grids, sorting and retrieval. When describing phenomena, subjects mostly use comparisons; followed by logical, emotional (*good/bad*, *interesting*), and aesthetic statements (*beautiful/not*).

## 3. PARAMETER MAPPING

In general, few metaphors have been found in the collected words. The participants used the standard vocabulary of science. In the contextual analysis these terms cannot be interpreted as metaphors, but become metaphoric when shifted to the auditory domain. Therefore we attempted to collect such *metaphoric* climate terms.

- Climate data is inherently **dynamic**: climate scientists *run a simulation* or collect time series data; Therefore, in general, the time axis can be used as direction of reading for the playback independently of further processing, filtering, amplification, etc., that depend on the specific sonification design.
- **Periodicities** and any associated type of wave phenomena play an important role in climate science and can directly be linked to sound oscillation and rhythmic phenomena.
- **Resolution** is a big topic in climate science, when comparing different data sets with each other or trying to find phenomena at a certain range; resolution in audio is given by the sampling rate. It can be changed by interpolation, an approach the scientists are used to as well, e.g., when fitting a plot.
- **Missing data** plays a large role in climate science; an obvious analogy is making it hearable as breaks, which can be used for a quick scanning of the completeness of a data set.
- The **ensemble** in climate science is a group of data sets resulting from different runs of a simulation. Because a single outcome is always the product of random processes, only the ensemble of many simulations can be regarded as trustworthy; in music, an ensemble is a group of different instruments—the metaphor can be used by mapping, e.g., different climate models to different sound timbres although the concept of ensemble in the two domains has a different impact.
- **Noise**—climate scientists who work with measurement data or with simulation data both know about the signal-to-noise ratio; one participant called the atmosphere *noisy*, when a high amount of greenhouse gases was to be found there; the scientists search for long-term trends within the noisy/random behaviour of everyday weather. Although noise in climate data has a different meaning than noise in sound, it could be a useful metaphor in sonification.
- **Obvious mapping strategies** comprise the height dimension in climate data (altitude) to the height in sound (pitch), but also temperature has a very tight association to mapping to pitch; the geographical spread can be used for spatial rendering of audio.
- **Weather phenomena** are linked to typical sounds and can be used, e.g., rain or wind sounds.
- On a more **conceptual level**, terms as for instance *extreme*, *dramatic* or *beautiful* will have to be transferred to the sound design and evaluated in listening tests by the future users.
- Furthermore, the control of the audio interface will involve actions that climate scientists are used to anyway, e.g., **calibrating** or **filtering** data/sound.

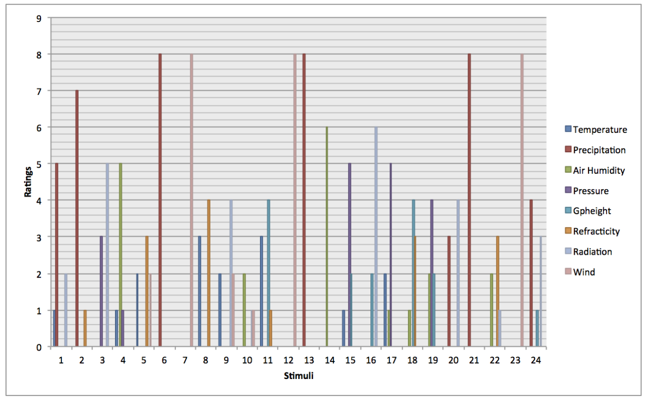


Figure 3: Frequency with which each parameter was mapped to each stimuli by the EG.

## 4. SOUND PREFERENCES

Aesthetical preference of the climate scientists and the intuitive mapping of climate characteristics to sound parameters are crucial and are explored in a set of experiments. For this study 24 sound samples of 10 seconds duration each were used. We chose these sounds from freesound database so that each three would constitute a group thematically or metaphorically connected to one of eight climate parameters determined in workflow analysis of climate scientists: *Temperature, Precipitation, Air Humidity, Pressure, Geopotential height, Refractivity, Radiation, Wind*. The reason for this selection was to provide a broad range of sounds which can be used to elucidate whether the climate scientists will be able to associate these sounds to parameters of their domain, and whether this association is unanimous.

Each experiment was divided into two sections; the purpose of the first stage was mainly to evaluate the sound samples (stimuli) aesthetically, and the second part for mapping the stimuli to the climate parameters. Altogether each experiment took between 35 and 45 minutes. The participants were given identical settings, listening to the stimuli via the same type of headphones. In order to control the effect of auditory experience and music knowledge on evaluating the aesthetics of sounds, the experiment was repeated on two different groups of participants. The first group are the domain scientists (the climate scientists), and the second group are all sound experts from IEM. The first group is taken as our experimental group (EG) and the sound experts as control group (CG). Each group consists of 8 participants. The EG is a subset of the group we interviewed in the contextual inquiry and focus groups.

The experiments show that not all eight parameters provided perceived links to the sound samples. (Fig. 3.) Especially challenging were the parameters that are more abstract such as **temperature** or **refractivity** or **geopotential height**. The results for non-abstract parameters such as **wind** were very clear over both experiments and over both participant groups. As a result from the two experiments, we decided to use sounds in the sonification tool that satisfy either one of these criteria:

- A sound was mapped to the same parameter by EG and CG.
- A sound was mapped to the same parameter by EG as by the hypothesis.

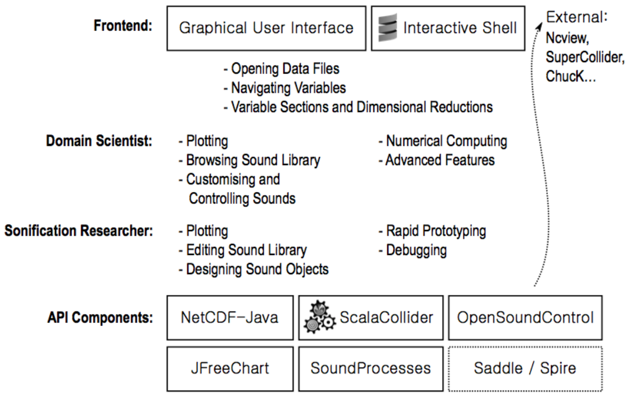


Figure 4: Sonification software framework.

- A sound was rated highest by EG or CG, and mapped to the hypothesised parameter by EG or CG.

## 5. TECHNICAL IMPLEMENTATION

Fig. 4 shows the components of the framework we are developing. It provides a rich application programming interface (API) which interlinks data I/O, visual and auditory display. Both measured and simulated climate data is provided by the collaborating institution in the Network Common Data Form (NetCDF [6]), an open standard for multi dimensional data. A standalone application provides both a graphical user interface and a text based shell for the Scala language. Scala was chosen because it can easily incorporate libraries running on the Java Virtual Machine, such as our sound synthesis layer and the NetCDF interface, while providing a succinct syntax not unlike dynamic languages commonly used in scientific computing, such as Python. Also, third party numerical computing libraries such as Saddle can be easily integrated. Communication with external clients such as Ncview—a commonly used plotter—is possible through the OSC protocol.

The framework dual functions as **analysis tool** for the climate scientists and development environment for the sonification design. It will be subject to **user-based testing** and will drive a sound installation. Real-time sound synthesis is provided through the ScalaCollider library, and we are evaluating the use of the (SysSon) framework which allows us to record a historical trace of the sound design process in order to better understand and to formalise it.

## 6. CONCLUSIONS AND FUTURE WORK

A main outcome of the extensive analysis of the contextual inquiry is improving our understanding of the way climate scientists work, communicate, and, to a certain extend, how they think. The audio tool has to be designed in a way that scientists integrate it naturally to their workflows, and allow them to be creative with using it. Therefore we are working to reach this goal by creating a sound space of intuitive sounds. The final sonification designs will be implemented in the audio tool and tested at each iteration. Furthermore, the experiments discussed in this paper evaluated our primary sound design which leads to a more advanced soundscape and improvement of the auditory display. The next steps are to evaluate the dynamics of sounds and see how and if they correlate with related climate phenomena. Those experiments should be designed within the tool to give the participants the option to interact with the user interface and to adjust the sound dynamically while analysing data. Sonification does not aim at replacing visual displays. Rather, the specific characteristics of visual and acoustic perception shall be used in an optimal way to complement each other. Questions for further research on sonification include which techniques are best suited for analysing climate data.

In terms of the systematic design process, we also consider tracking both the sound design process from our side as well as the interaction of the climate scientists with the framework. Several aspects found in visualisation, such as overlaying graphs, highlighting regions, showing differences and error boundaries, adding threshold guides, adding labels to particular parts of a plot, and so forth, will be studied in terms of potential analogues in the auditory domain.

## 7. ACKNOWLEDGEMENT

We would like to thank the volunteers at Wegener Center and IEM for participating in our interviews.

## 8. REFERENCES

[1] Z. Halim, "Sonification: A novel approach towards data mining," in *Proc. International Conference on Engineering Technologies.* 2006, pp. 548–553, IEEE Publishing.

[2] N. Bearman, "Using sound to represent uncertainty in future climate projections for the united kingdom," in *Proc. of the International Conference on Auditory Display*. 2011, ICAD Publishing.

[3] A. Polli, "Atmospherics/weather works: A multi-channel storm sonification project," in *Proc. of the International Conference on Auditory Display*. 2004, ICAD Publishing.

[4] Deborah J. Mayhew, "The usability engineering lifecycle," in *CHI '99 Extended Abstracts on Human Factors in Computing Systems*, New York, NY, USA, 1999, CHI EA '99, pp. 147–148, ACM.

[5] Claire-Marie Karat, Robert Campbell, and Tarra Fiegel, "Comparison of empirical testing and walkthrough methods in user interface evaluation," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 1992, CHI '92, pp. 397–404, ACM.

[6] Harry L Jenter and Richard P Signell, "Netcdf: A public-domainsoftware solution to data-access problems for numerical modelers," in *Proceedings of the 2nd International Conference on Estuarine and Coastal Modeling*. American Society of Civil Engineers (ASCE), 1992, vol. 72, pp. 72–82.