Precipitation Visualization

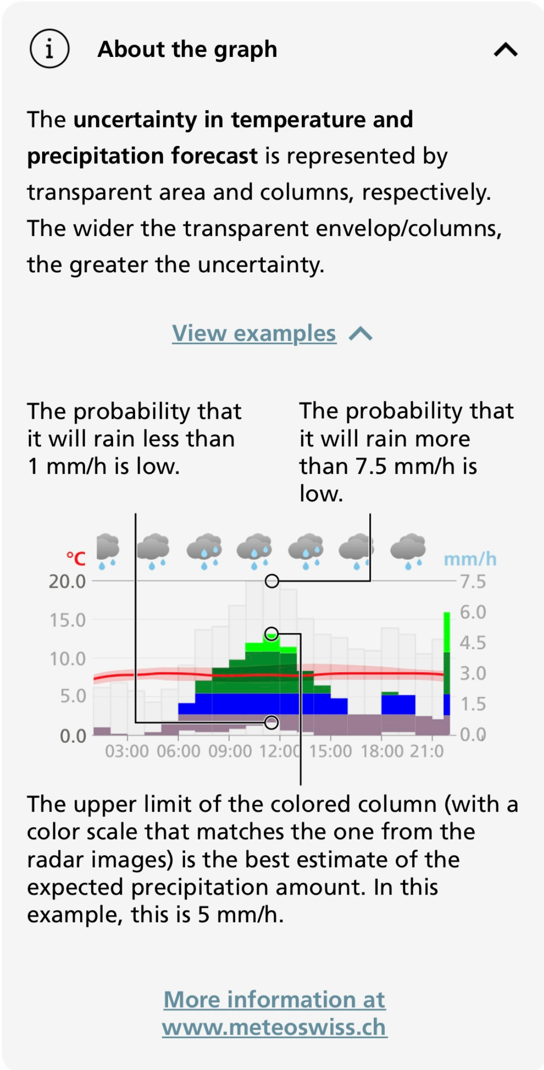
Elisa Baux, Justine Stoll, Tibor Vaughan

/ Data Visualization / COM-480

Contents

1 PROCESS BREAKDOWN

1.1 MOTIVATION



The three of us, share a deep passion for climate related issues and were eager to incorporate it into our project in some way. While looking for a real-world project, we were fortunate enough to connect with Lea Beush, an expert in forecasting and analysis at Meteo Swiss. With Lea's guidance we decided to work with a dataset that represents the forecast of precipitations over a week. Motivated by the opportunity to work on a project grounded in concrete data, and that interested people working in a company outside of EPFL, we set out to tackle the challenge of visualizing uncertainties in precipitation forecasting. Our main goal was to find efficient and engaging methods to communicate the inherent variability and uncertainty in weather predictions. When looking at the current way that meteo swiss displays uncertainty (see image above), we found that it lacked clarity and intuitiveness. Which is why we wanted to focus on visualizing these uncertainties. To provide a more comprehensive understanding of weather forecasts, enabling users to make informed decisions based on a nuanced view of the data.

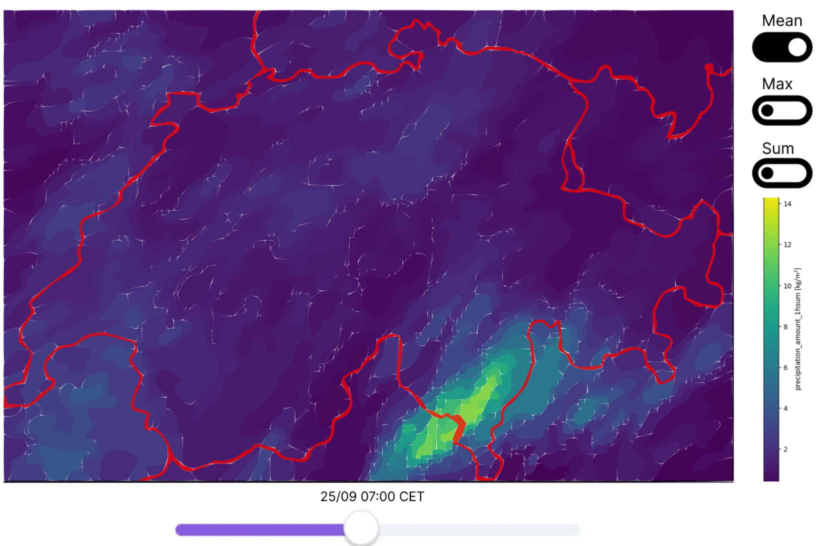
1.2 STARTING THE PROJECT

1.2.1 DATA EXPLORATION

The great value of this dataset is its level of precision when it comes to spatial predictions, and the fact that for each point in time and space, we have predictions from multiple ensemble members. That is, instead of having one predicted outcome, we have several. This is obviously extremely powerful to assess for example the uncertainty when it comes to predicting future events. However, it has the drawback of making the dataset difficult to handle, and it limits the visualization opportunities, as it is rather intended for computational studies. We played around with different ways of aggregating the different outcome predictions, as we explained in previous milestones.

1.2.2 FIRST IDEAS

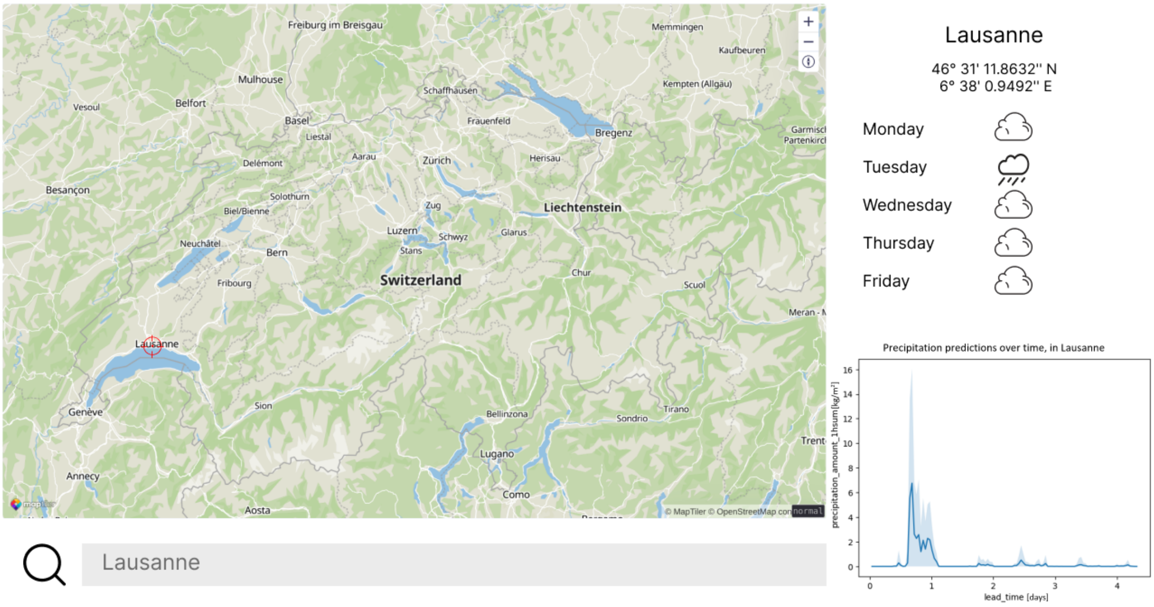
After exploring different ideas, we narrowed down our focus onto two main visualisations. The first one, whose aim would be to improve the user's understanding of MeteoSwiss's prediction system, provides insights into how different ensemble members represent possible realizations of weather behavior over five days. This visualization not only showcases the variability and changes over time but also allows users to explore variations accross different locations within Switzerland.



We wanted to let the user switch between the different modes of interpretation of data, letting them explore the spatial and temporal variability of the precipitations in different ways. As you can see there is a toggle letting the user glide through the different times of the predictions, and three toggles to decide how the values are presented.

From our data exploration we found 3 ways to group the predictions that seemed to give the most insights into the data. The mean, which would give an understanding of the overall forecast trend, the max which would let the user view the places with the most rain possible or the locations with no risk of rain at all, the sum which we have found highlights the areas with a greater contrast in weather patterns.

The second visualisation that we thought of presents the weather predictions aggregated by location. The goal being to allow the user to view predictions for a given point in space over the next five days.



This second visualization adopts a more familiar weather forecast map format but with enhanced interactivity and a clearer representation of the prediction and its associated uncertainty. To avoid overwhelming the viewer with information, we would provide the prediction when a location is clicked on or searched for. The map staying in view to help the user gain a better sense the whole.

Once we had a clear vision of our desired visualizations, we began implementing the second one. To build this visualization, we utilized Maplibre for the map rendering and React for the overall structure. This combination allowing us to create an interactive experience without relying on a specialized data visualization library.

1.2.3 TRYING A CHANGE OF DIRECTION

After getting some feedback on our project, we realised that our visualisations might not meet the expected level of complexity and storytelling. We were constrained by our dataset selection and found it difficult to make significant changes. We decided to explore additional data sources to try and enhance our visualisations.

We considered comparing the predicted precipitation data with transportation datasets to gain insights into population trends in different modes of transportation influenced by rainfall. We discovered a promising dataset provided by opentransportdata.swiss. However, we encountered a hurdle as the dataset contained string-based location data, requiring a seperate database to cross-reference location names.

We searched for alternative datasets but encountered similar challenges, either with string-based locations or incompatible date ranges with our predictions dataset. While we explored various possibilities, we were careful not to fall into the trap of relying on unrelated correlations, such as linking drowning incidents with ice cream consumption, which further reduced our choices when it came to finding a good match.

Considering the work we had already invested in our first visualizations, we made the decision to focus on that aspect, accepting that our final visualizations might not be as compelling as initially envisioned.

So we continued to work on delivering accurate and informative representations of the weather forecasts.

2 CHALLENGES AND DESIGN DECISIONS

2.1 COMPLEXITY OF THE DATASET

The dataset containing precipitation forecasts allows for great computational work. However, we realized after receiving some feedback, that this aspect was predominant with respect to the visualization potential. In fact, we spent more effort exploring and testing ways to represent for example the uncertainty, rather than complex simulations.

It was also particularly challenging to process a dataset with a lot of information, and simplify it in a way that would allow for meaningful visualizations, that would not overwhelm the user.

2.2 INTERACTION WITH THE USER

Another challenge we faced was to create a meteo app that would allow an interaction with the user. As we explained previously, our primary goal was to make the data accessible to anyone, and easily explorable. This challenge was particularly relevant, given the complexity of the dataset. Our intention was to allow people to have access to weather forecasts, in a more useful way than what is currently achieved on the Meteo swiss app. For example, our approach allows for an easy selections of a particular city, and access to its predictions.

2.3 ADDING VALUE

After discussions with the person from MeteoSwiss, as well as feedback from milestones 1 and 2, we realized that it was extremely complicated to come up with original visualizations that would effectively add value, compared to the data that is already accessible on the MeteoSwiss app. In fact, even though we added the possibility for the user to interact with our app, the information contained in our visualizations is not significantly higher than that of MeteoSwiss. We tried to face this challenge by exploring a second dataset, trying to bring forward correlations between the precipitations forecast, and the use of public transportation.

2.4 COMBINING DATASETS

The additional dataset we wanted to integrate in our work was more difficult to process than what we thought, and ultimately we were not able to obtain significant results. Indeed, while we found a dataset on the usage of public transportation over the same time interval as our original dataset, it was significantly difficult to work with it. In a next step, we would spend more time to make the data numerical instead of natural language.

3 FINAL RESULTS

We implemented the app using React as it allows us to provide a very interactive experience to the user, we also used rechart [1] (an abstraction layer over D3.js) to provide dynamic charts and Maplibre [2] for the maps. We decided to go for a static-only setup as it allows us to easily host in on Github Pages while keeping great performances, it is available at https://com-480-data-visualization.github.io/project-2023-meteo/.

In order to avoid overwhelming the website during the initial load we split the data in tiny chunks that are lazy loaded and cached using a custom cache system, however we did not pre-process the data specifically for the graph as it is completely dependent on a single point, this come with the downside that in order to display the graph all the data have to be loaded first, this problem could be alleviated using a server-side service that process the data only for a single point, it's however not possible to do so with our current static-only setup.

From a less technical standpoint, we also decided to merge our to visualizations to present a simpler interface, for an overall streamlined experience for the user.

4 PEER ASSESSMENT

The following are the main tasks that our project implied. (Tibor: **T**, Elisa: **E**, Justine: **J**)

- Reaching out to people of interest to find a suitable dataset. (**T**, **E**, **J**)
- Literature review. (Lead: **E**, Support: **J, T**,)
- Preliminary data exploration. (Lead: **J**, Support: **T**)
- Computational tasks (Lead: **J**, Support: **E**)
  - Dimensionality reduction of original dataset
  - Multiple aggregation techniques for different ensemble members
- Secondary data exploration: transport dataset. (Lead: **T, E** Support **J**)
- Visualizations (Lead: **T**, Support: **E, J**)
  - Drafts on Figma
  - Actual implementation of visualizations
- Milestone: writing. (Lead: **E**, Support: **J, T**)
- Technical implementation. (Lead: **T**, Support: **E, J**)

[1] https://recharts.org
[2] https://maplibre.org