Playing with Uttarakhand rainfall data

Last year we received an amazing dataset from David Hopkins in Uttarakhand , who compiled daily weather readings from his weather station and shared the data with us. At the time we were trying to find more sources of rainfall data that was more granular then monthly averages that the IMD provides. Since then we have gotten Rajasthan's daily rainfall and some data from IMD's automatic weather stations.

Daily rainfall information at weather station level is actually the most useful rainfall data, it can be used for more analysis and planning efforts so we were very grateful for the contribution from Mr. Hopkins. It is always possible to aggregate up for trends but deaggrating is impossible especially for places like Uttarakhand where the climates can change from one area to the other. A district average can't help people plan or monitor a situation in a real time manner.

After the Uttarakhand floods last year there was no better reason to look at the shortfalls of only providing aggregated rainfall data. We believe that the IMD should be more proactive and share the daily rainfall data at the weather station level, in a more open and accessible way. This will help many people in the sector with planning water management projects from rain water harvesting to groundwater management.

We decided to showcase the major problems with district and monthly data, especially for a place like Uttarkhand where the mountains create a lot of different climates and multiple weather stations are needed to really get information for the area that you are working in.

We first looked at comparing the daily weather station data to the district average data to try to see how different they are. Mr. Hopkins lives in Almora district so we took that data and graphed it with his. You can see full analysis and yearly charts here.
If you look at 2002 you can see how the daily rainfall makes shows outliers and rain intense days. You can see that a district average doesn't well represent what is happening in Kausani.
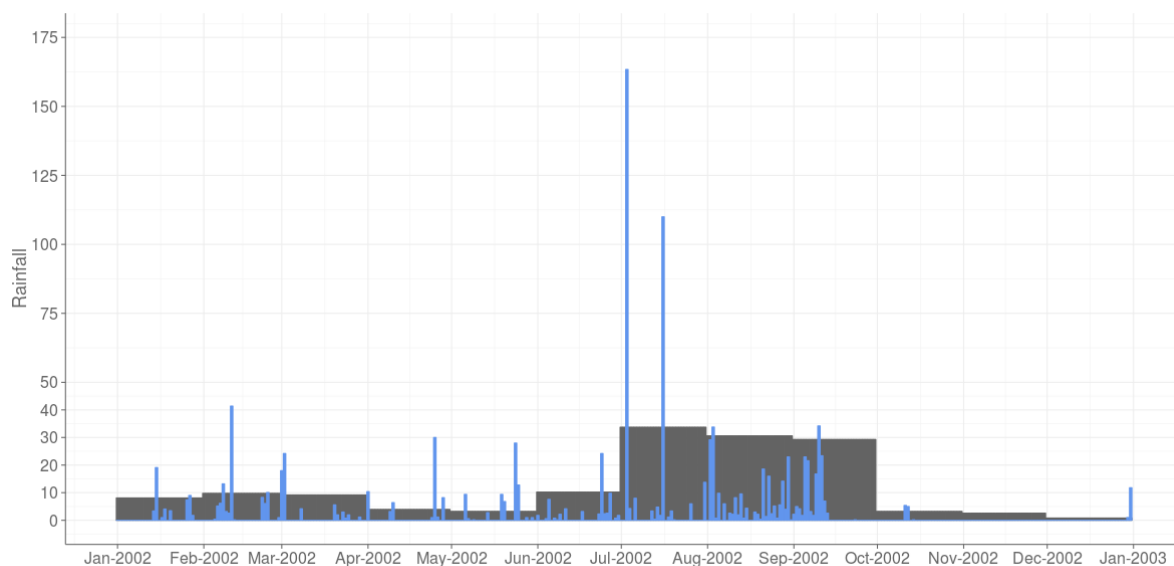
Then we looked at rainfall days and intensity and put htem on a heat map. With the daily rainfall you can see how the number of rainy days and their intensity has changed over a given time period. With the average you can't get a sense of that change.
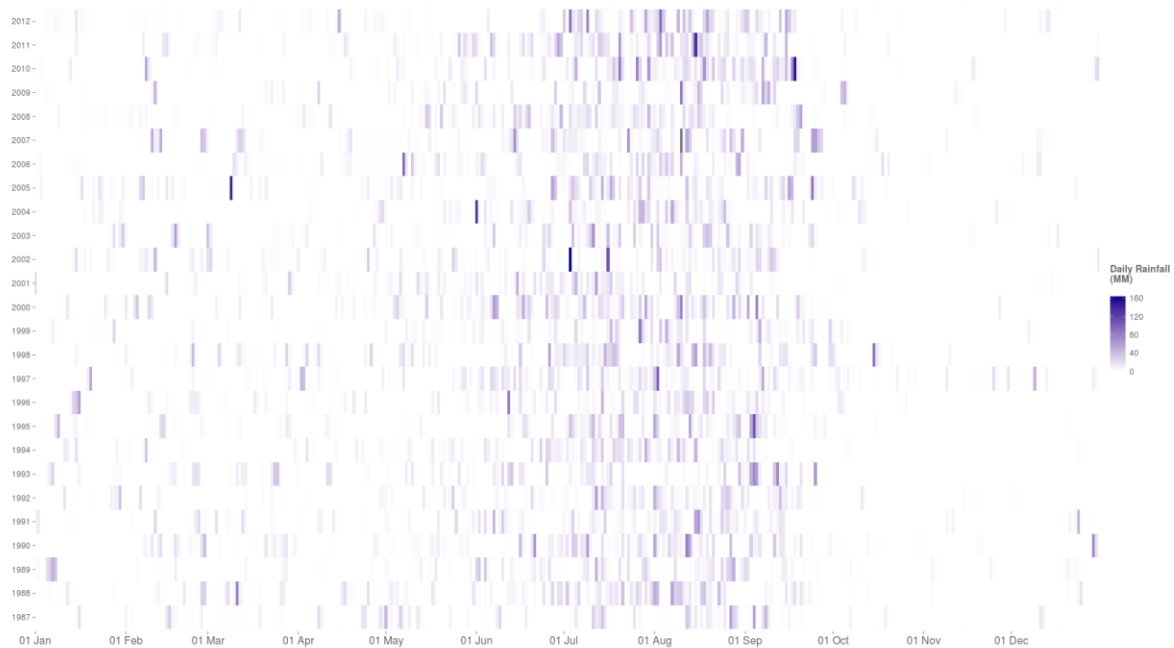
Then we looked at last June's rainfall data from around the time of the Uttarkhand floods. The IMD has not put up the 2013 monthly averages yet. So let's use the daily hourly data we got from IMD Automatic Weather stations. Even though they put the automatic weather station data out they only archive it for 2 weeks and make it difficult to download a large amount of data. Also there are gaps in where the weather stations are, they also don't. If the IMD releases this data according to more open formats and with a robust public archive it will be even more useful.
Let's look at Dehradun District where there was a great deal of rainfall.

This is the hourly data that the Dehradun weather station was reporting on June 15th, 16th, and 17th.
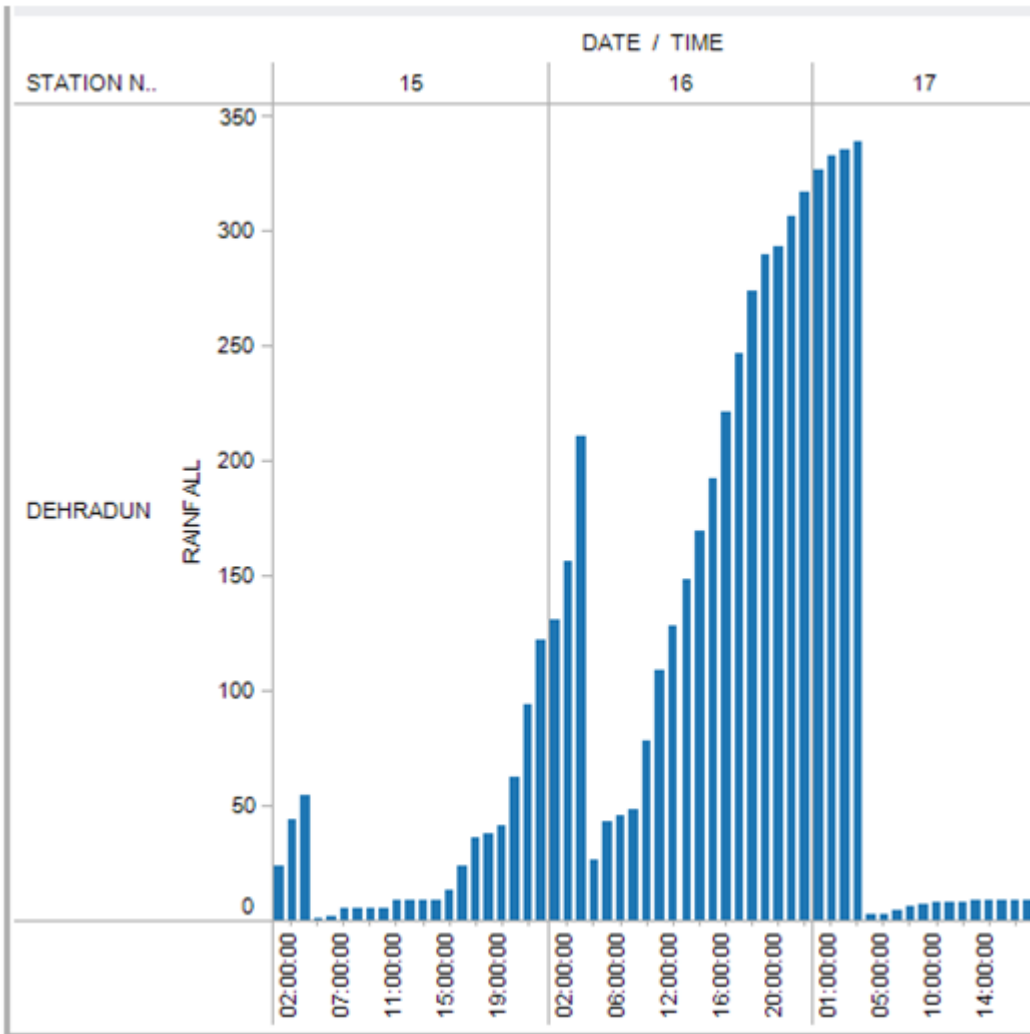
You can see very clearly see when the bulk of the rain is falling and when the cloudburst happened. For understanding the flooding this is much more useful then say the weekly average and definitely more than the monthly/district average.

The IMD collects an incredible amount of useful information, and while it is good they are doing some proactive disclosure, it would be useful if they let real time, weather station level data open so people everyone can better utilize it.