

Poor harvest in India in 2002 shown with CDC wheather data

Tilman Hinnerichs, Anja Reusch, Veronika Soloviova

Bertelsmann Data Science Scholarship Program

July 21, 2018

What question would we like to answer?

On the search for good data we stumbled accross the CDC (Climate Data Center) of the DWD (German Wheather Center). They are providing data from all kinds of different wheather stations all around the world as we visualized in Figure 1.



Figure 1: Wheather stations provided by the CDC

What question would we like to answer?

The CDC is providing all different kinds of data for those stations. Their database holds the statistics on snowfall, precipitation and sunshine hours and literally everything. The data is provided in a raw heap of data which we tried to classify.

But what to do with that giant pile of data?

Hence, we found another source of data: The **Food and Agriculture Organization of the United Nation**, which provides data on quantitative measurements of the yield of harvest. There we found a peak in the yield of paddled Rice of India in 2002, seen in Figure 2, where more than 22% of the production got lost.

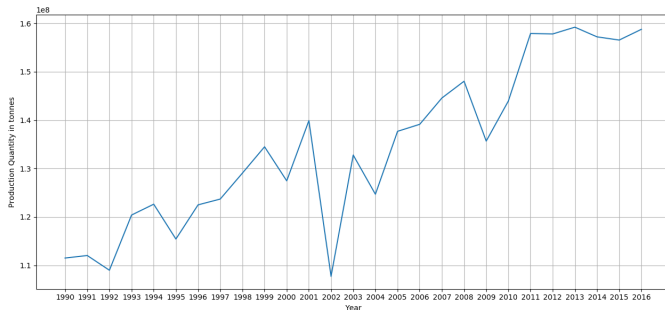


Figure 2: We can see a harsh peak at 2002

What question would we like to answer?

So we would like to know:

- ▶ Can we see possible reasons for this peak in our wheather data?

We therefor had a look at the rice production on the world shown below in Figure 6. Even though it is showing the productivity, it is nevertheless a good indicator for the absolute production.

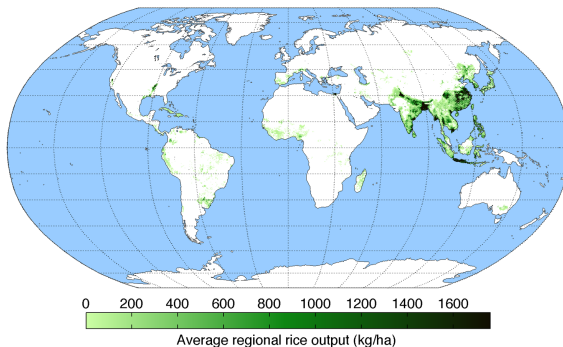


Figure 3: World rice production highlighted

So first have a look at the common rice crops in India and their growth and harvest time intervals. Here we can see the Kharif and the Rabi crop displayed for India.

RICE: planting and harvesting calendar



Now lets have a look at our wheather data for some city located in the productice are of India concerning rice cultivation.

We therefore considered the following cities, while rejecting cities not including the crucial data:

- ▶ Guwahati,
- ▶ Cherrapunji,
- ▶ Daltonganj,
- ▶ Jagdalpur,
- ▶ Chennai Minambakkam, and
- ▶ Bangalore

Now concerning the crucial diagrams (You can get the data and the other diagrams on request):

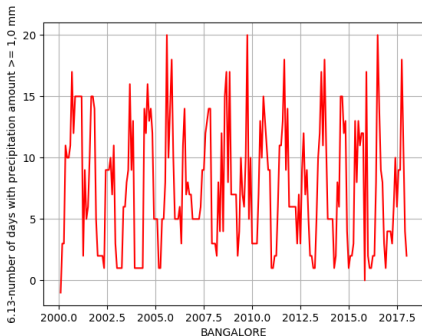


Figure 4: Precipitation for Bangalore

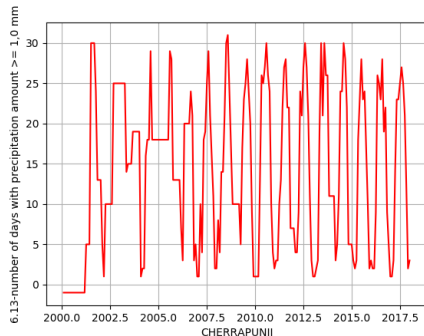


Figure 5: Precipitation for Cherrapunji

Considering those we can see the cut in the middle of year 2002.

- ▶ Hence, we can see that the Kharif Crop, which is decisive for the production quantity, had to suffer a lot under the weather conditions given in 2002

Sources

- ▶ This presentation was build using the data of the previously mentioned sites
- ▶ Plots and analysing has been done with several python libraries, e.g. Matplotlib, Pandas, NumPy, etc., and Vega

Other sources:

RICE: planting and harvesting calendar: AMIS Crop Calendar
Figure 3: Wikipedia