Water footprint of cotton production

Susan Härtig

Overview and Motivation

After visiting Kazakhstan and the Aral sea region in 2011 I learned about the impacts intensive cotton grow and water regulation can have on a ecosystem. The continuous cultivation of cotton in the Aral Sea basin of Uzbekistan has caused a tremendous decrease in the surface area of the sea - the water volume has been reduced to 10%. The reason is that two of the rivers that formerly fed the Aral Sea (the Amu Darya River and Syr Darya River) were diverted for cotton production. Cotton uses a huge amount of water both to produce and process. It can take 2,700 litres to produce the cotton needed to make a single t-shirt.

Cotton production and processing are also a major source of pollution of fresh water. Globally, freshwater resources are becoming scarcer due to a growing population, increasing wealth and consumption and hence increasing water withdrawals for human use. The impact of consumption can be quantified with the concept of the "water footprint", a concept introduced by Hoekstra and Hung (2002), and further elaborated by Chapagain and Hoekstra (2004). The water footprint of a nation has been defined as the total volume of freshwater that is used to produce the goods and services consumed by the inhabitants of the nation (Chapagain et al., 2006).

With my visualization I want to show the world wide cotton production and compare this data with the consumptive water use. The water consumption can be divided into green water (water from rainwater stored in soil) and blue water (fresh water in lakes, rivers, aquifers). Blue water is stress-weighted because it is consumed at rates faster than its short-term replacement and many cotton producing countries have limited blue water resources. I want to compare this 2 types per country and show the result in relative (per capita) and absolute values.

Related Work

Papers and other resources:

- Everything started with a trip to Aral sea in 2011
- A paper about Changes in water volume of the Aral Sea after 1960: http://download.springer.com/static/pdf/754/art%253A10.1007%252Fs13201-012-0048-z.pdf?auth66=1397339085_93e0e7dd55e4e091beeff6aece14688a&ext=.pdf
- NASA pictures of Aral sea: http://earthobservatory.nasa.gov/Features/WorldOfChange/aral_sea.php?all=y
- Background report on Water footprints of crops and derived crop products (1996-2005): http://www.waterfootprint.org/Reports/Mekonnen-Hoekstra-2011-WaterFootprintCrops.pdf
- Paper on Water footprints of crops and derived crop products (1996-2005): http://www.waterfootprint.org/Reports/Report47-WaterFootprintCrops-Vol1.pdf

Visualizations that inspired me or were helpful regarding code:

- http://viewer.phildow.net/world-gdp-growth/
- Fresh water supply worldwide: http://www.josephbergen.com/viz/water/
- http://mbostock.github.io/protovis/ex/symbol.html
- http://data.karmi.cz/d3/life-expectancy-map/
- http://bl.ocks.org/KoGor/5685876
- http://chartsbin.com/view/1455

Questions

1. Which countries are the main cotton producers?
2. How much do the produce?
3. How much did this production per country change since 1961?
4. How much green and blue water is used for cotton production in relation to each other?
5. How much green and blue water is used for cotton production in total? After my Exploratory Data Analysis I discovered that India is using huge amounts of blue water due to its big production. But relatively it is using more green water
6. How many percent of total renewable water resources per capita is used for cotton production
7. Finally see which countries are suitable for cotton production

Data

Data sources and Calculation

- **Cotton production quantity (cotton lint, cotton seed):** FAO. 2013. FAOSTAT database, Food and Agriculture Organization of the United Nations (FAO). Website accessed [03/2014]
- **Total Population - Both sexes (1000):** FAO. 2013. FAOSTAT database, Food and Agriculture Organization of the United Nations (FAO). Website accessed [03/2014]
- **Total renewable water resources per capita (actual):** FAO. 2013. AQUASTAT database, Food and Agriculture Organization of the United Nations (FAO). Website accessed [04/2014]
- **Water footprints of crops and derived crop products (1996-2005):** WaterStat, Water Footprint Network, Enschede, the Netherlands, Mekonnen, M.M. and Hoekstra, A.Y. (2011) The green, blue and grey water footprint of crops and derived crop products, Hydrology and Earth System Sciences, 15(5): 1577-1600. Website accessed [03/2014]
- **Country codes:** http://www.geonames.org/countries/ and http://wits.worldbank.org/WITS/wits/WITSHELP/Content/Codes/Country_Codes.htm. Website accessed [03/2014]
- **Geoinformation (Lat and Lon):** https://opendata.socrata.com/dataset/Country-List-ISO-3166-Codes-Latitude-Longitude/mnkm-8ram and http://itouchmap.com/latlong.html. Website accessed [03/2014]
- **Topojson "world-50m.json":** https://github.com/mbostock/topojson/blob/master/examples/world-50m.json. Website accessed [03/2014]

A friend helped me to calculate the total water consumption for cotton products and the per capita values of cotton production and water consumption (Data/raw data/cotton_water.xlsx).

We have used data on the production of cotton seed and cotton lint (1st Level only to avoid double counting), and export data for all 1st and 2nd level (see figure 1) cotton products from FAOSTAT to gather the annual domestic production from 1961 – 2012 on a country level. The data was available in 1000tons/yr. For the water consumption, we have downloaded data from waterfootprint.org (Appendix II. Water footprint per ton of crop or derived crop product at national and sub-national level (m3/ton) (1996-2005)). Data was only available for 1996-2005 averages for green/blue/grey water consumption, for all 1st level and 2nd level cotton products. We have used these average values for the whole time period which maybe lead to an over- or underestimation for earlier years. For easier readability, we have only calculated green/blue water use for total production.

In order to compare the water consumption of countries we have normalised total values by population (data from FAOSTATS population database). The per capita requirements of water for cotton production can show important links to food security and water shortages, especially in dry regions that are large producers of cotton and cotton derived products.

The data for total renewable water resources per capita (actual) from AQUASTAT was only available for every 5 years. So we calculated the missing years via linear interpolation. We used this dataset to calculate the percent of the total renewable water resources per capita withdrawn for cotton production.

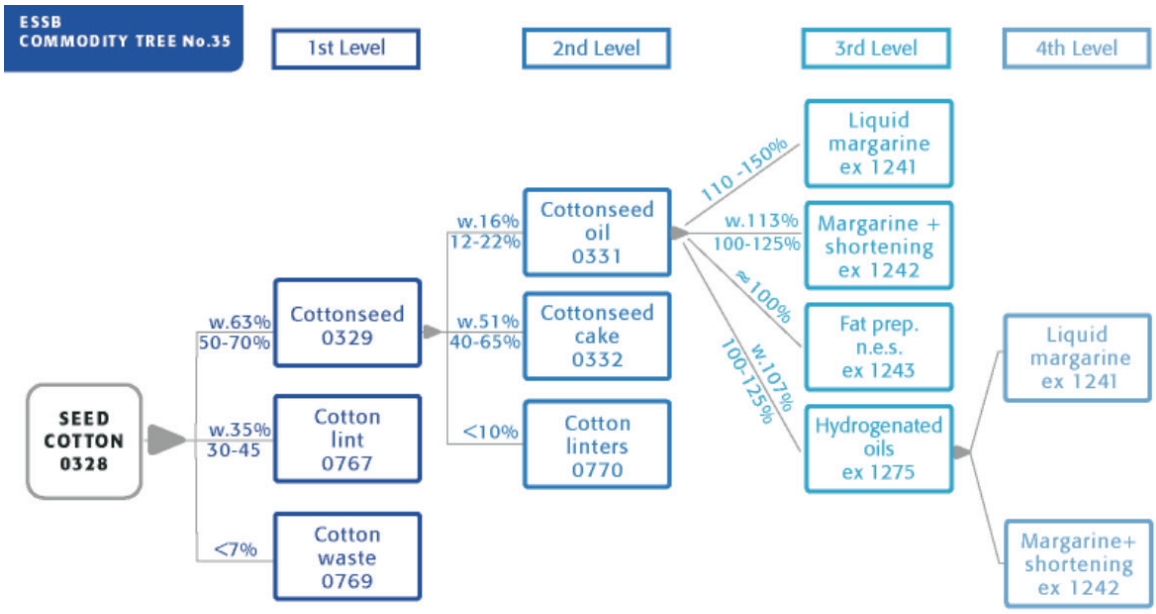


Figure 1: Cotton commodity tree (source: Faostat)

Data clean-up

We calculated averages for cotton producing countries that don't have water consumption data or split up since 1961:

- **Uruguay**: Average of Argentina, Brazil, Paraguay
- **Sri Lanka**: Values of India
- **Saint Vincent and Grenadines, Saint Kitts and Nevis, Montserrat, Haiti, Guadeloupe, Grenada, Dominican Republic, Cuba, Antigua and Barbuda**: Average of Nicaragua, Venezuela, Costa Rica, Colombia, Guatemala, Honduras
- **Rwanda**: Congo, Tanzania, Uganda, Burundi
- **Philippines**: Average of Indonesia, Viet-Nam
- **Nigeria**: Average of Benin, Cameroon, Niger
- **Korea**: Values of Korea, Democratic People's Republic of
- **Italy**: Values of Spain
- **Cyprus**: Average of Greece, Turkey
- **Bangladesh**: Average of India, Myanmar
- **USSR**: Average of all cotton producing USSR states
- **Yugoslav SFR**: Average of all cotton producing Yugoslav SFR states

We used linear Interpolation for the missing years in the dataset "Total renewable water resources per capita (actual)" because there was only data available every 5 years.

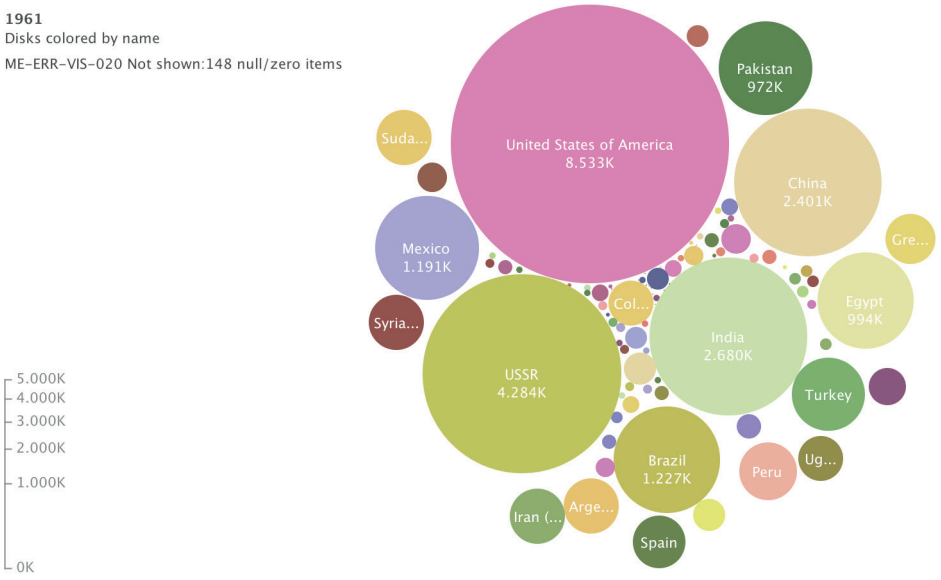
Exploratory Data Analysis

I explored my data with “Many Eyes” http://www-958.ibm.com/software/data/cognos/many-eyes/

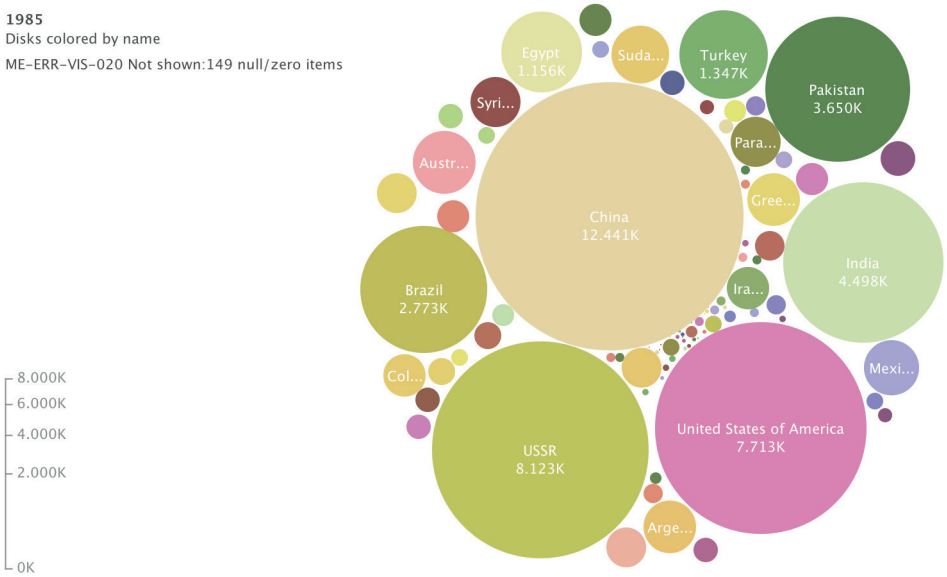
After testing different views with my data I decided on a choropleth map as foundation for my visualization because it’s important to see the geographical data in connection to cotton production and the related water consumption. The blue and green water data for each country will be implemented as a pie or stacked bar chart on top of each country to see the relationship of all three values at a glance. The bubble view helped my to investigate the total and per capita blue and green water use and supported me in my decision to also show the total values. India for example uses relatively more green water but the blue water consumption is still immense due to tremendous cotton production. The line graphs helped me to discover peaks for certain countries. By looking at the choropleth map I discovered that I need to find experiment with different colour scales for my dataset because the differences between the values of the countries can be huge.

Total cotton production in tons over years

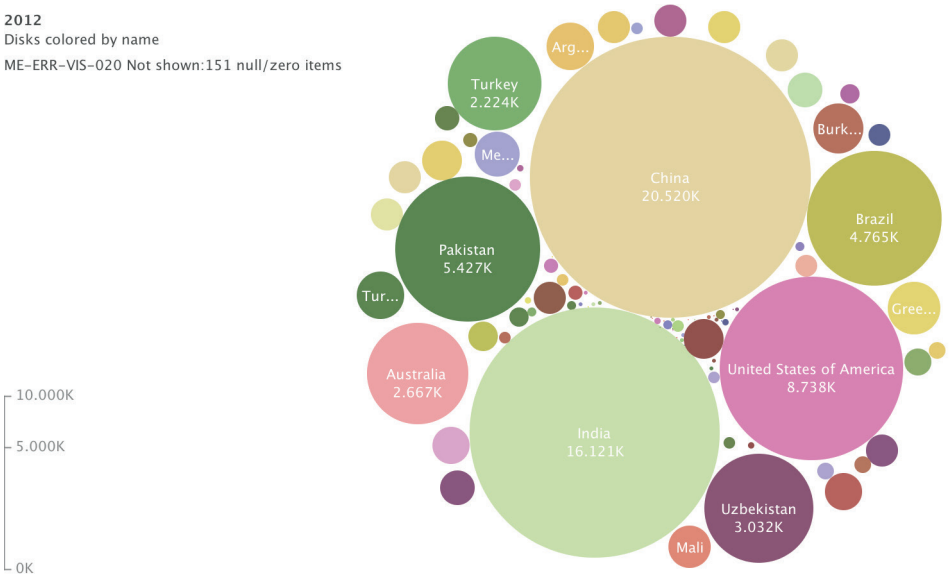
1961
Disks colored by name
ME-ERR-VIS-020 Not shown:148 null/zero items



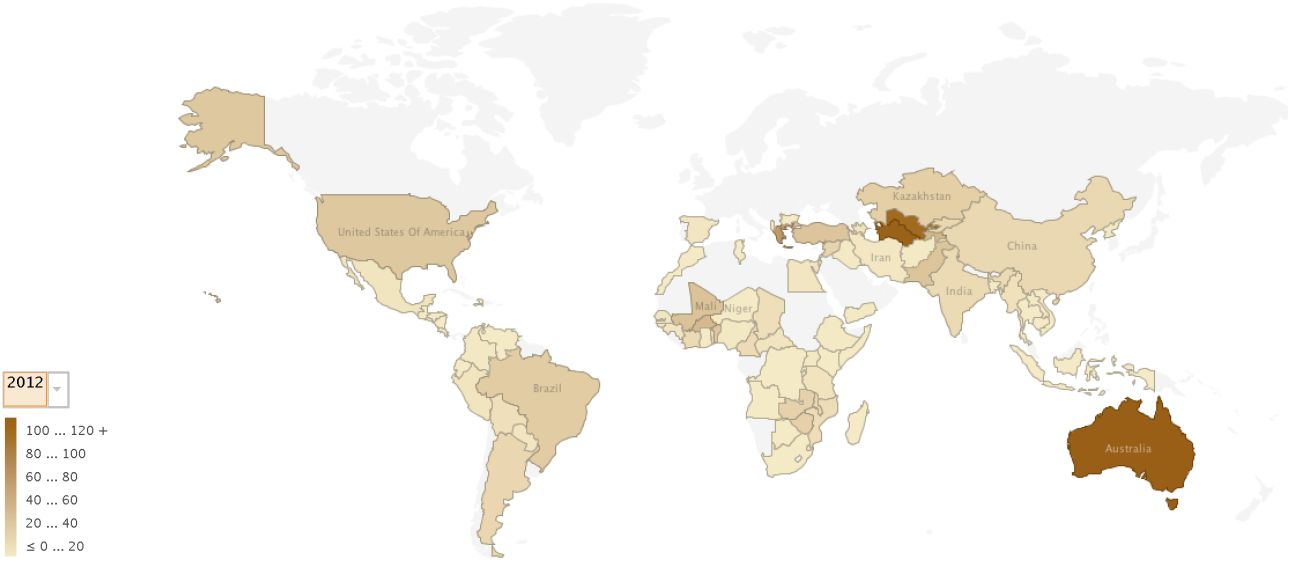
1985
Disks colored by name
ME-ERR-VIS-020 Not shown:149 null/zero items



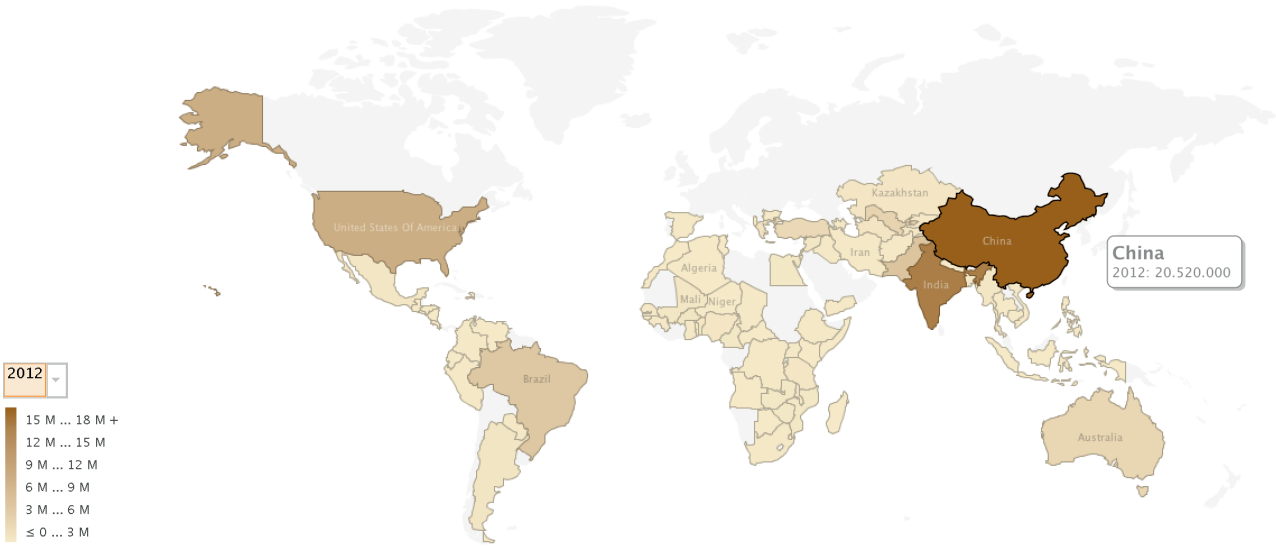
2012
Disks colored by name
ME-ERR-VIS-020 Not shown:151 null/zero items



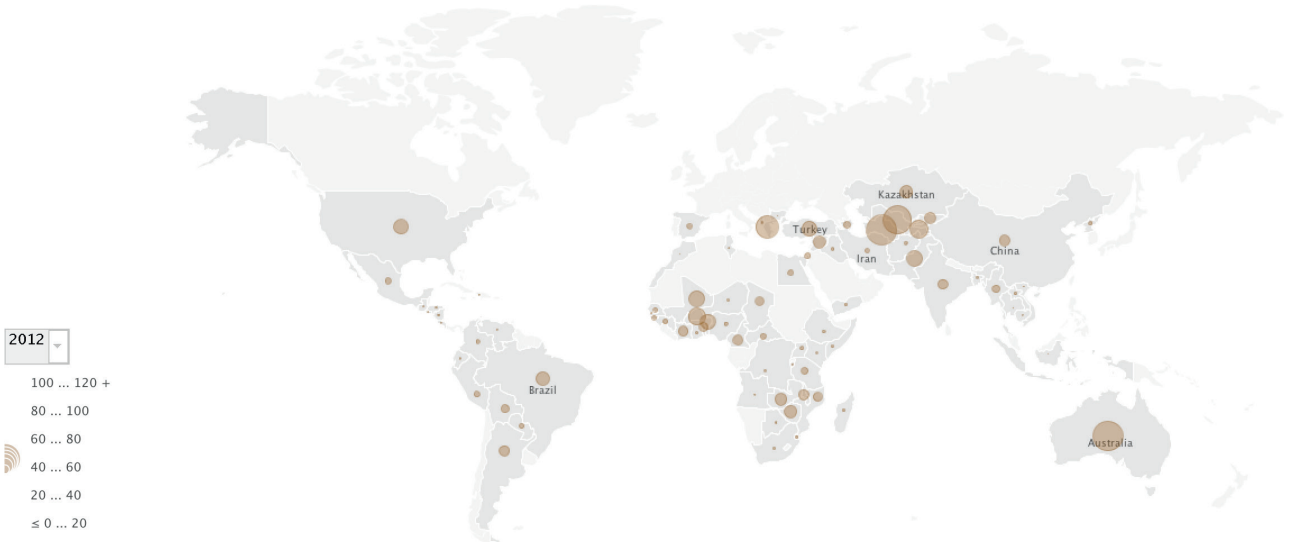
cotton production in tons/per capita 2012



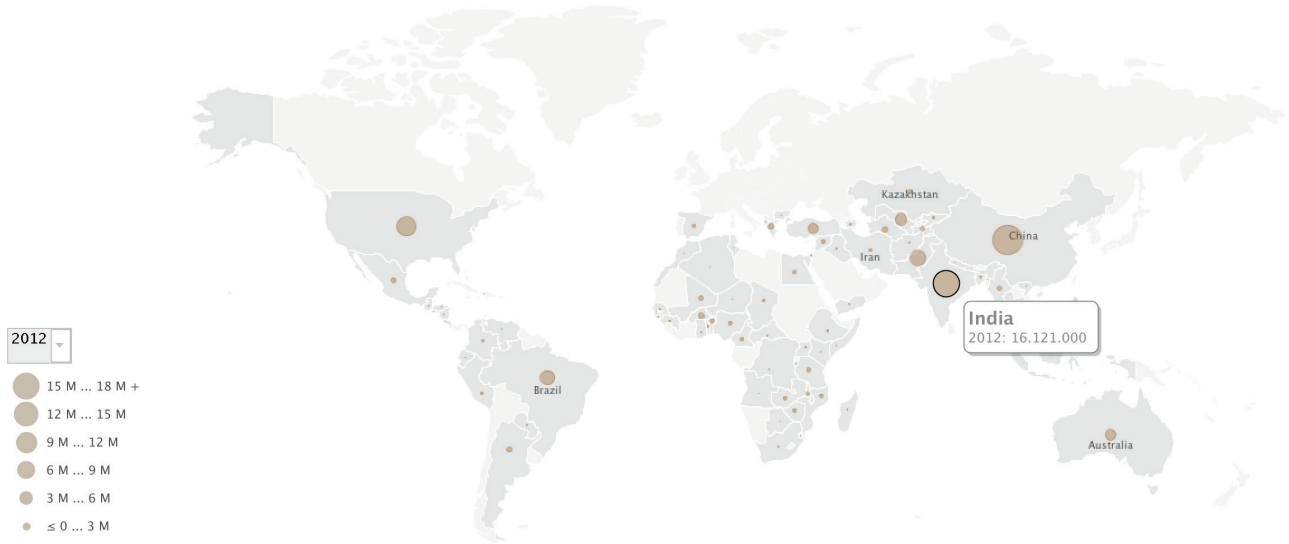
total cotton production in tons 2012



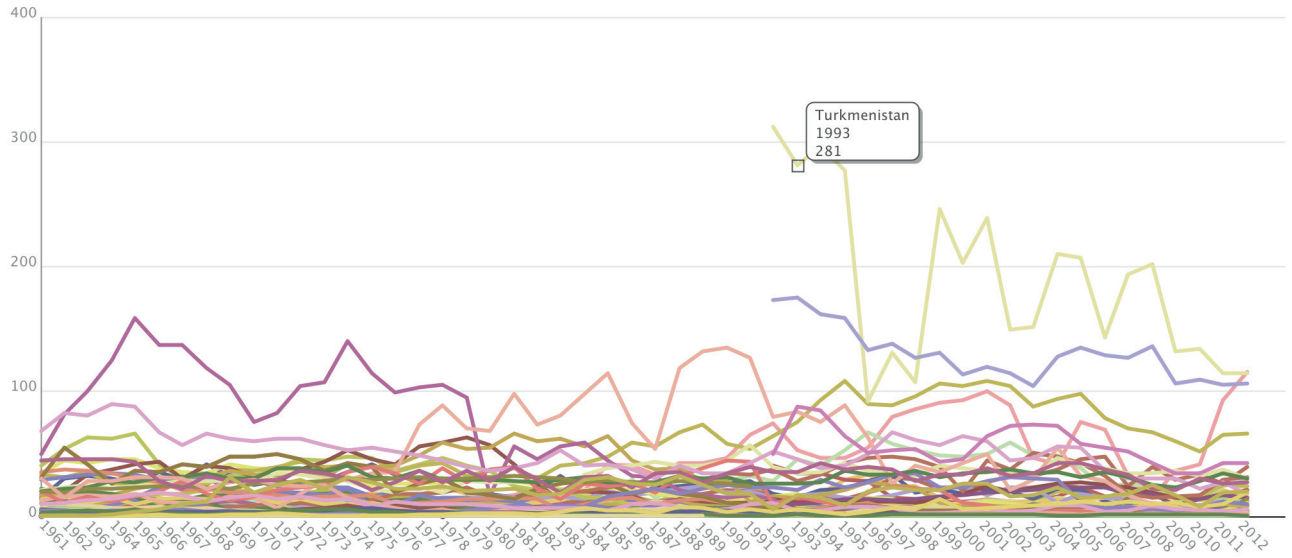
cotton production in tons/per capita 2012



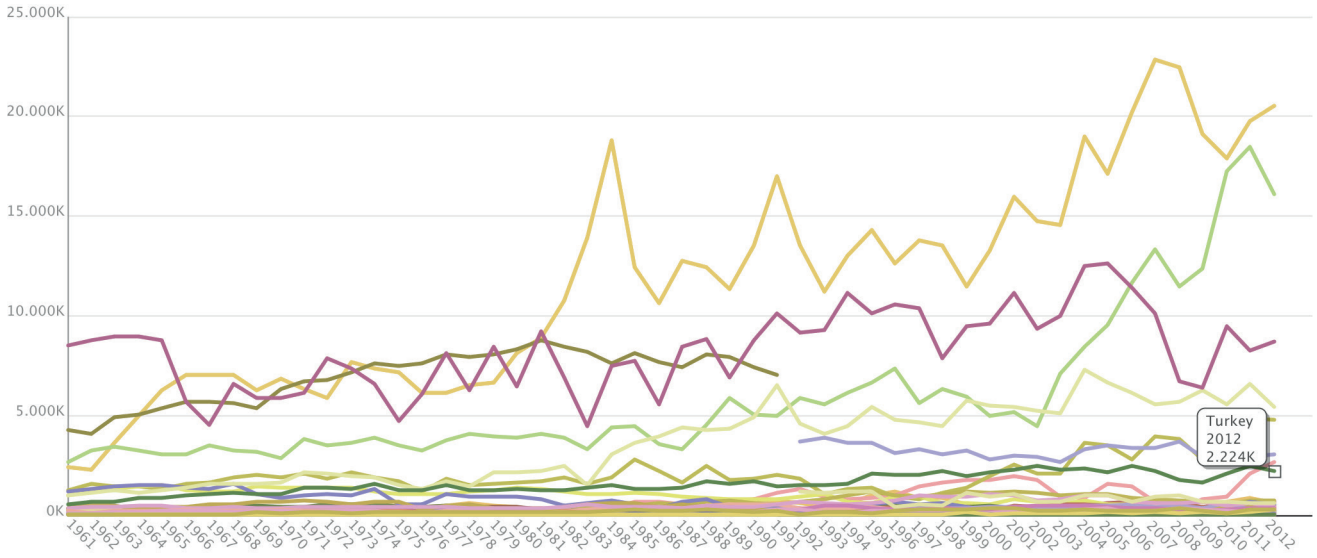
total cotton production in tons 2012



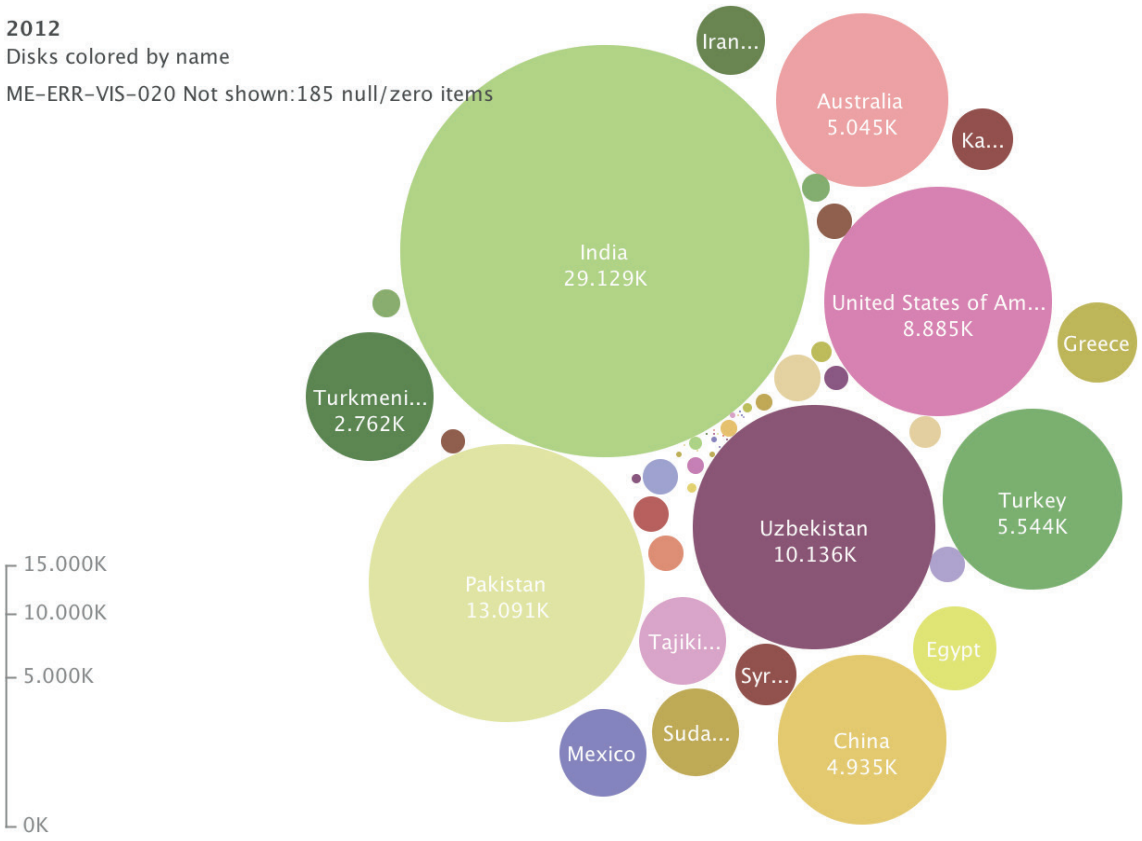
cotton production in tons/per capita



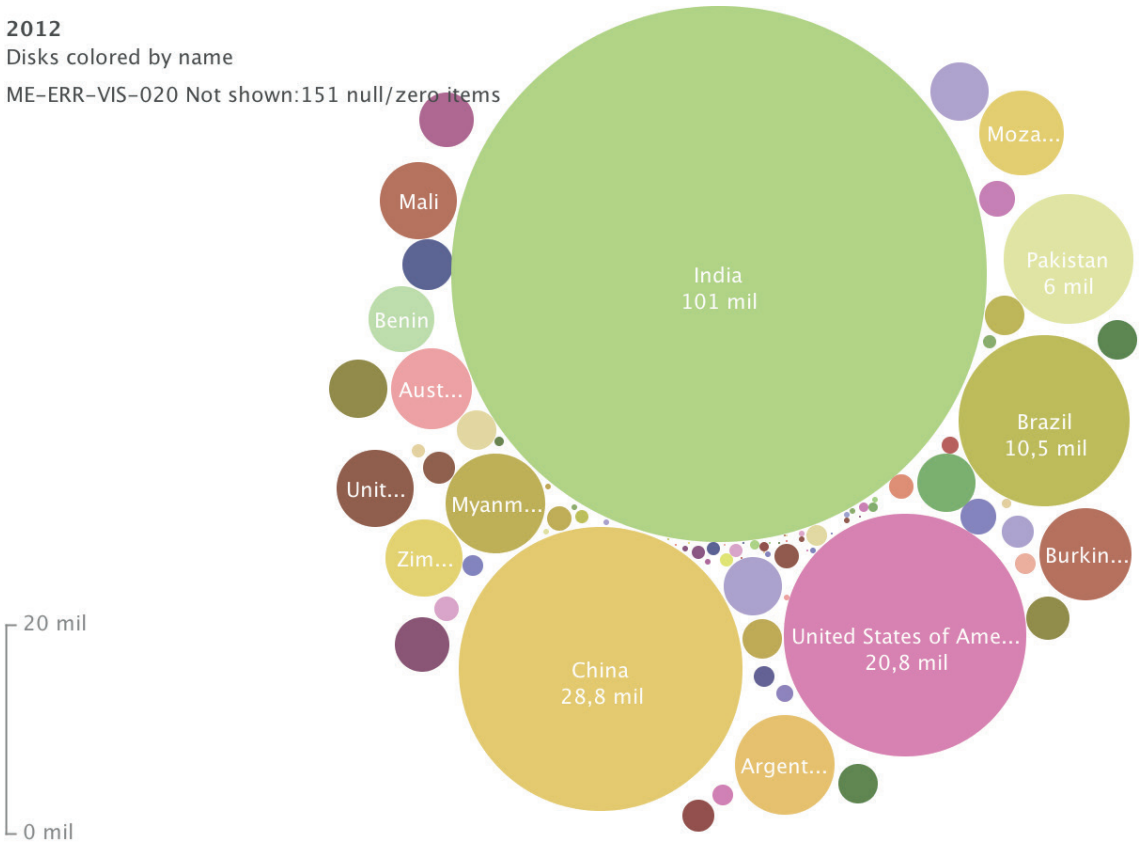
total cotton production in tons



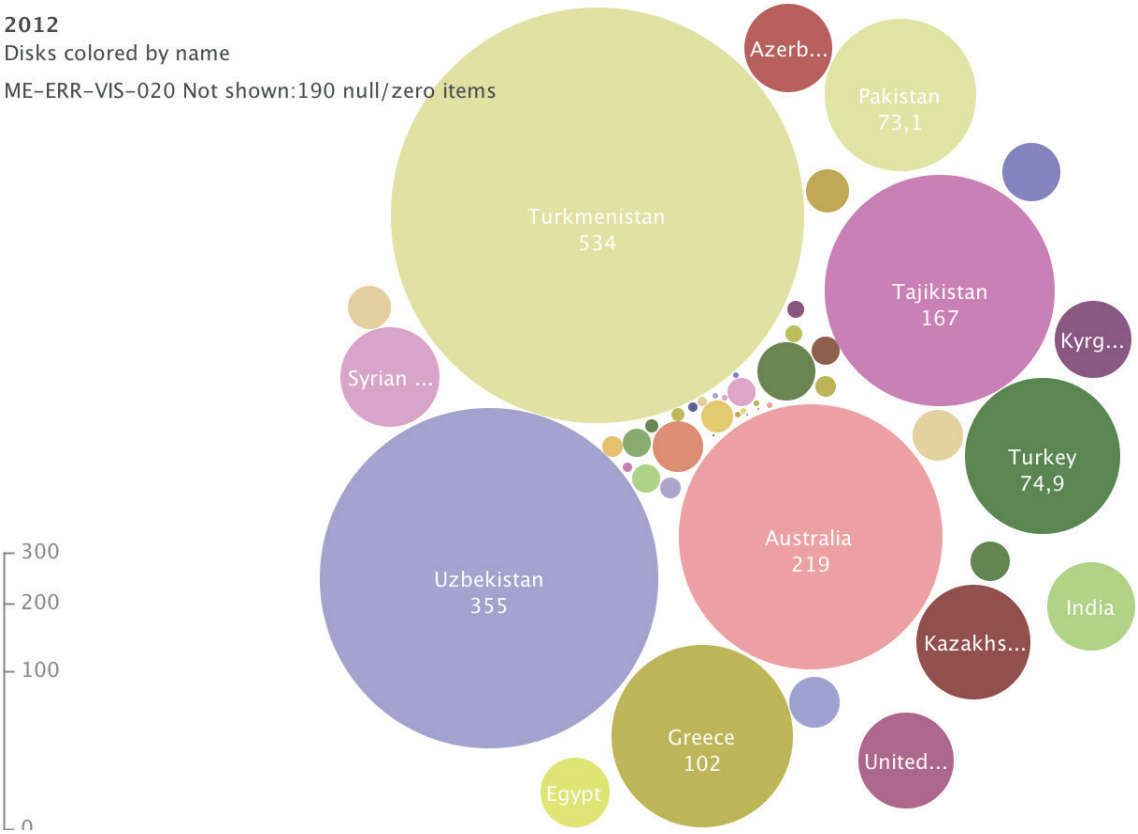
blue water consumption total (1000 m^3)



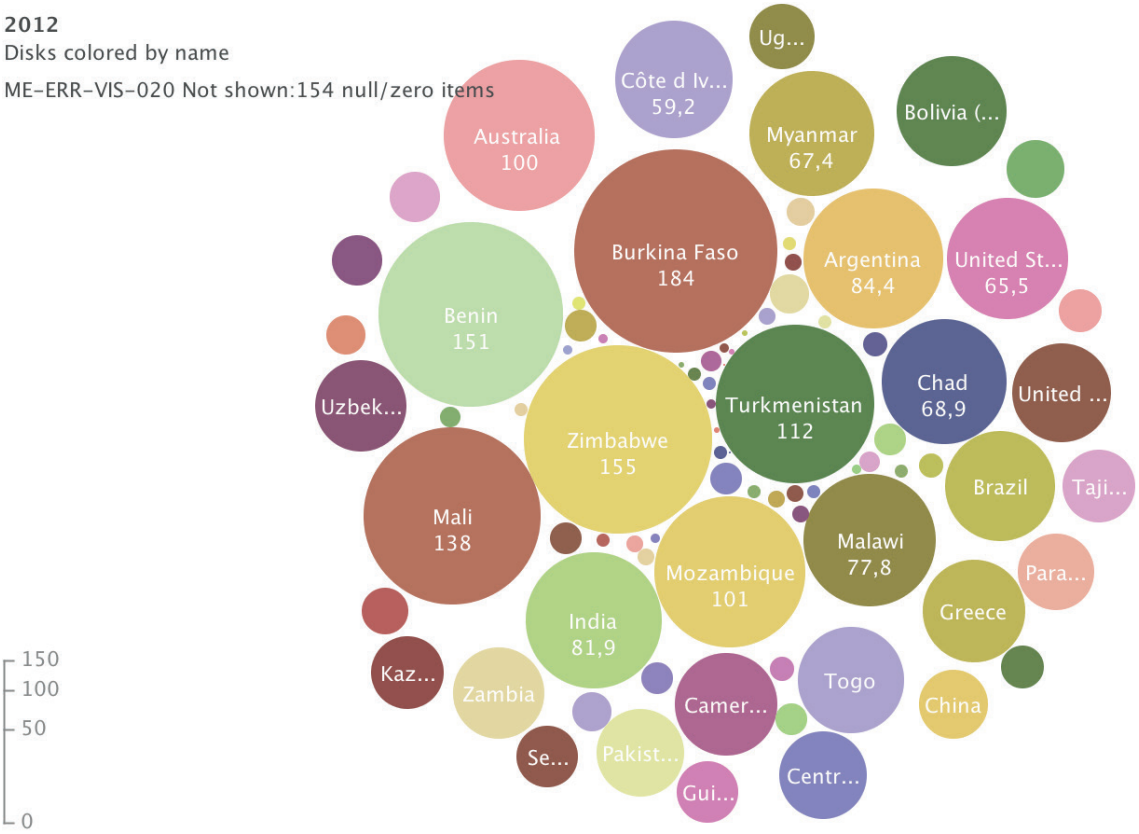
green water consumption total (1000 m^3)



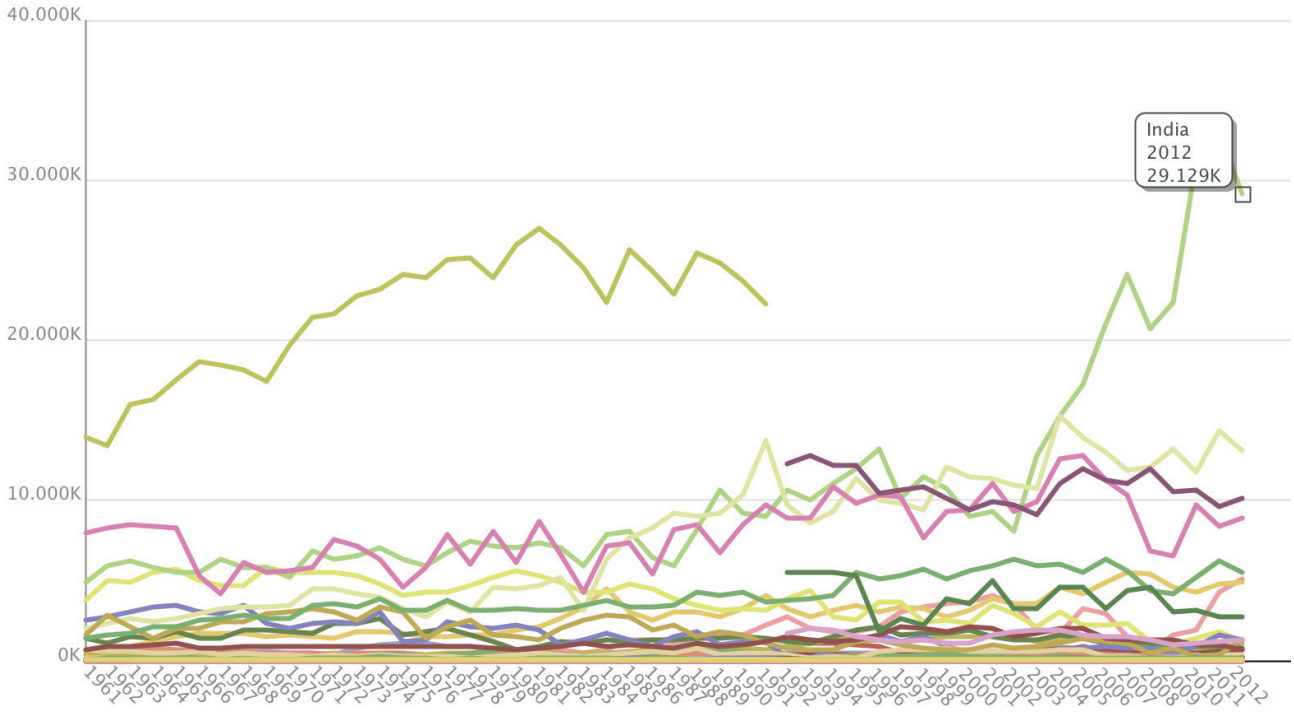
blue water consumption (m^3/capita)



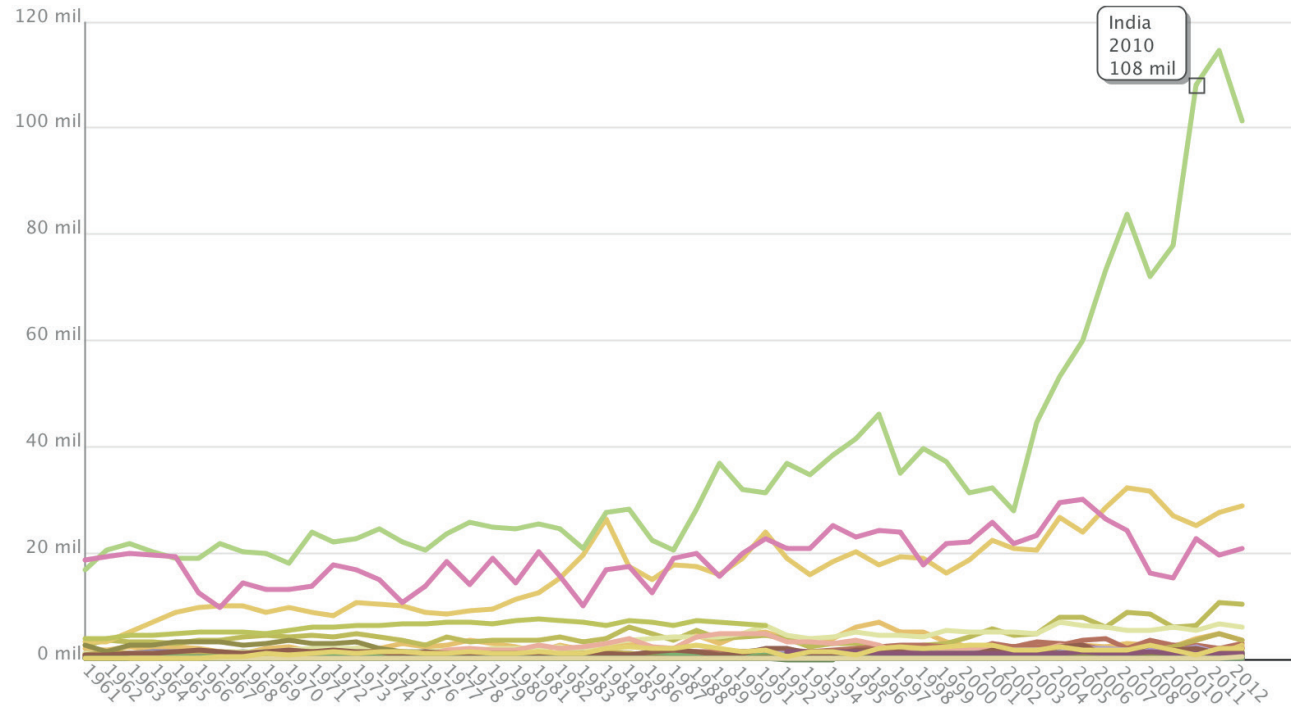
green water consumption (m^3/capita)



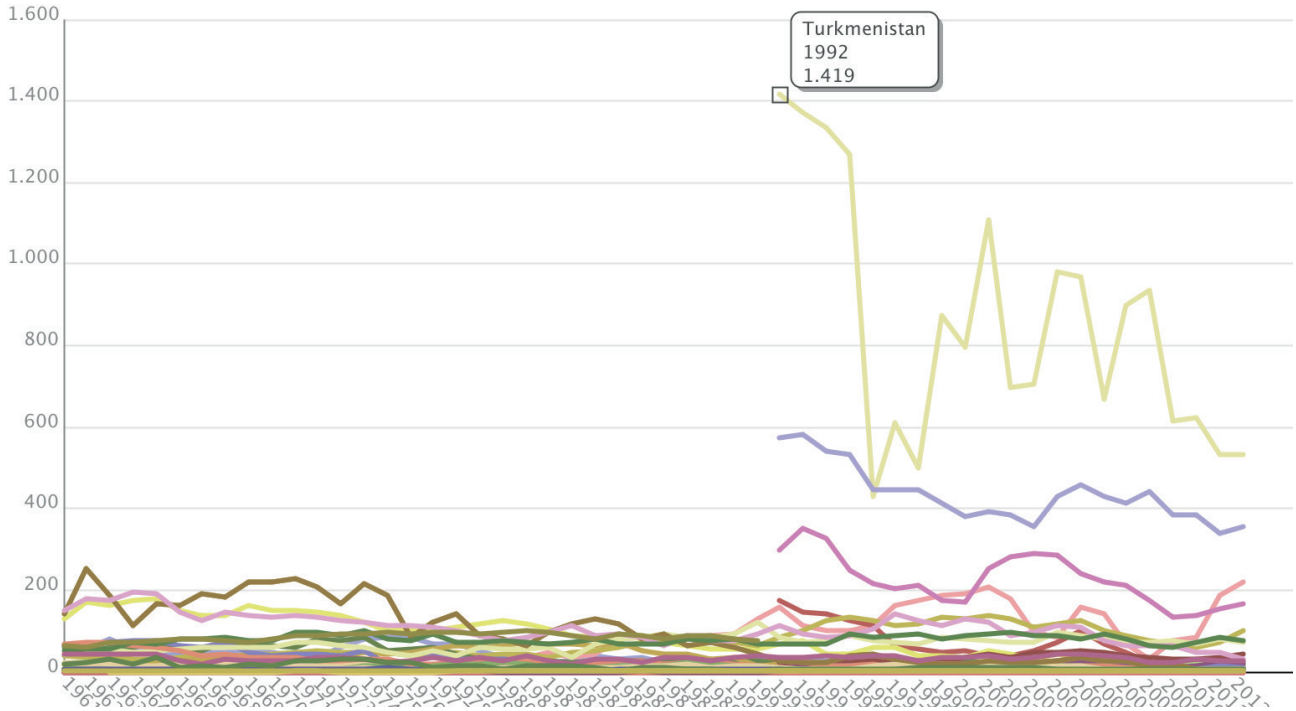
blue water consumption total (1000 m^3)



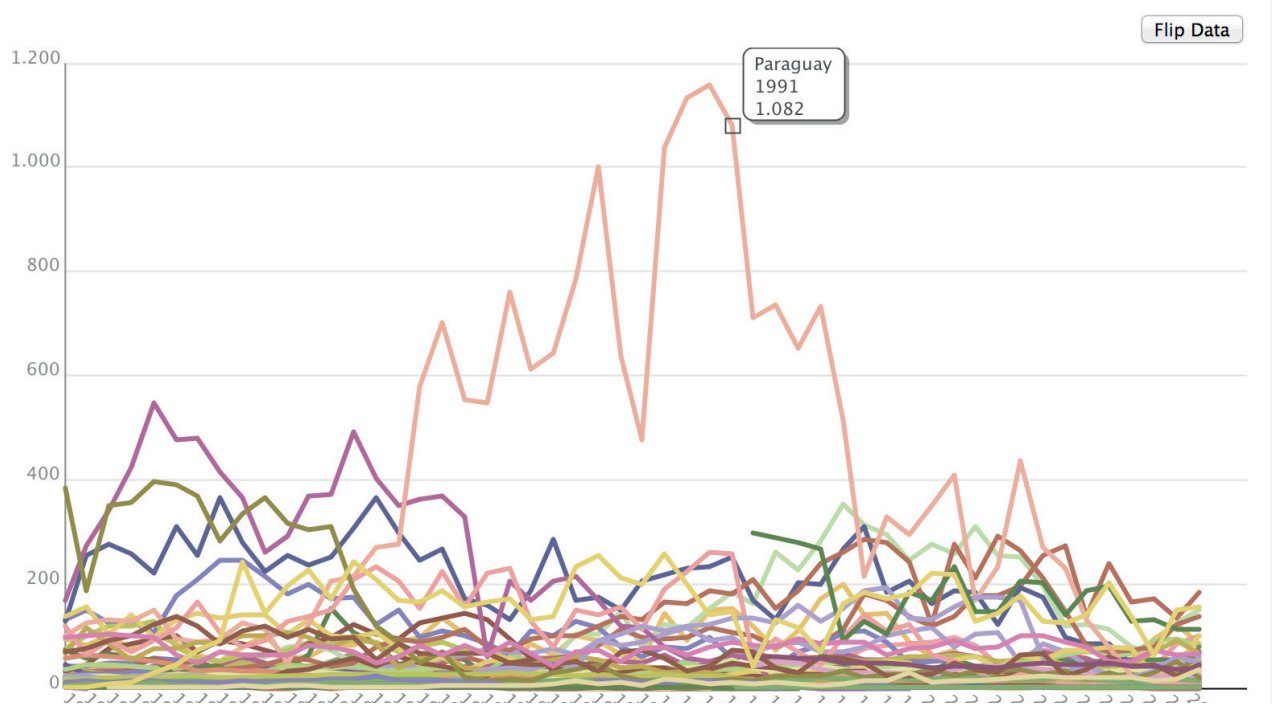
green water consumption total (1000 m^3)



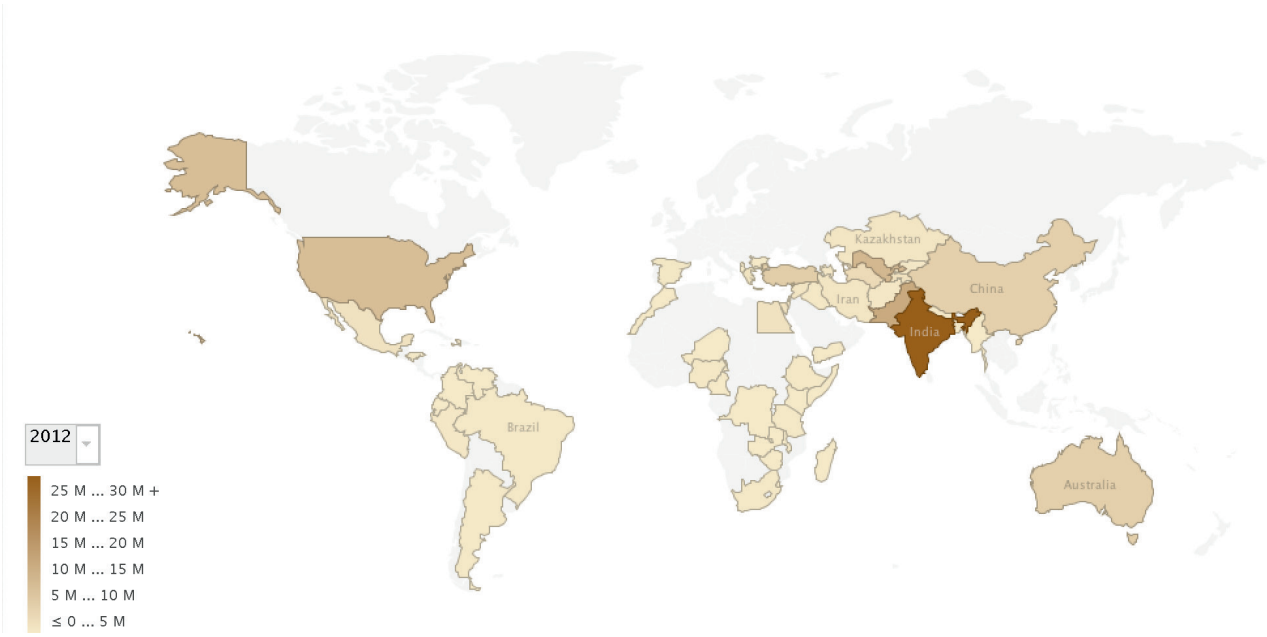
blue water consumption (m^3/capita)



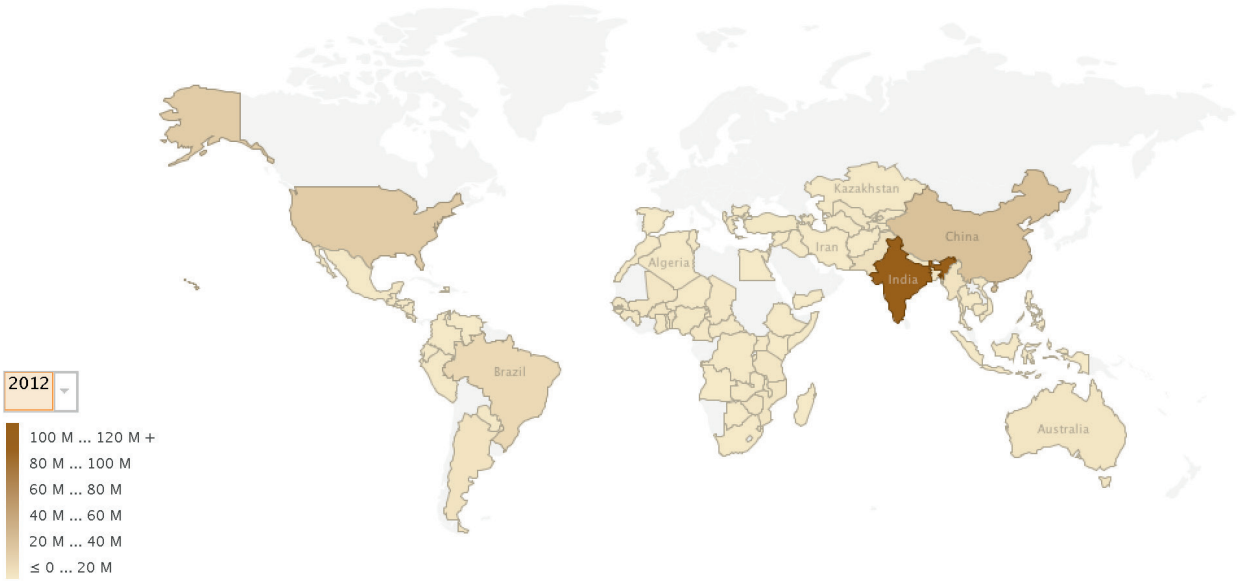
green water consumption (m^3/capita)



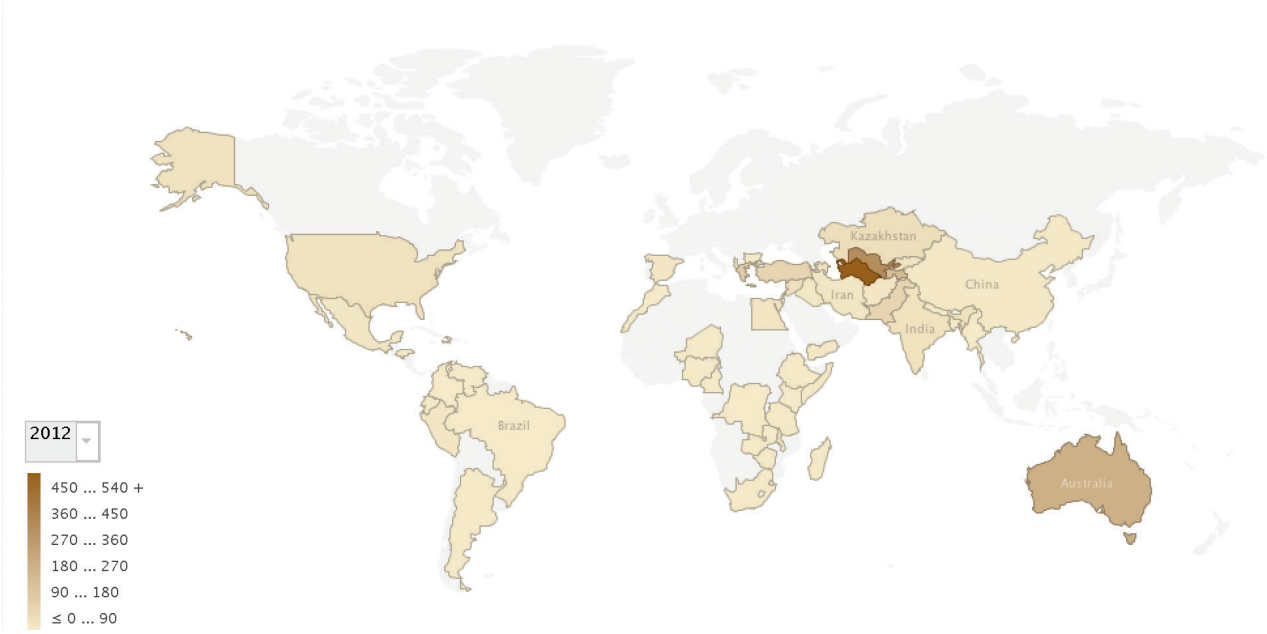
blue water consumption total (1000 m^3)



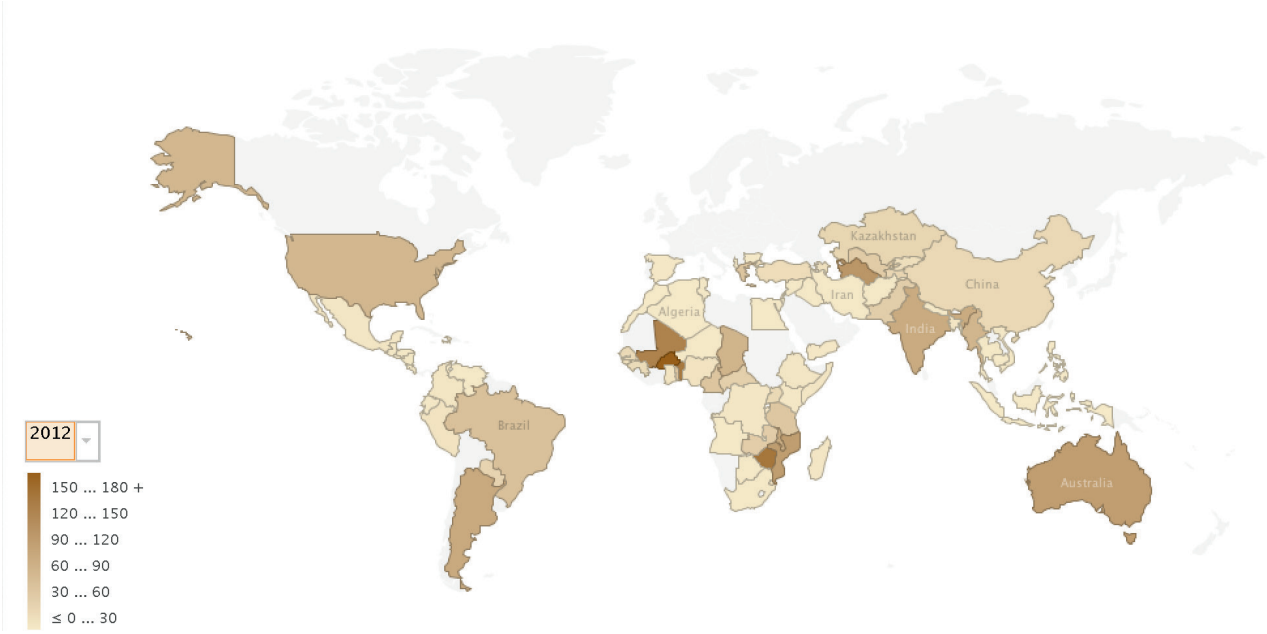
green water consumption total (1000 m^3)



blue water consumption (m^3/capita)

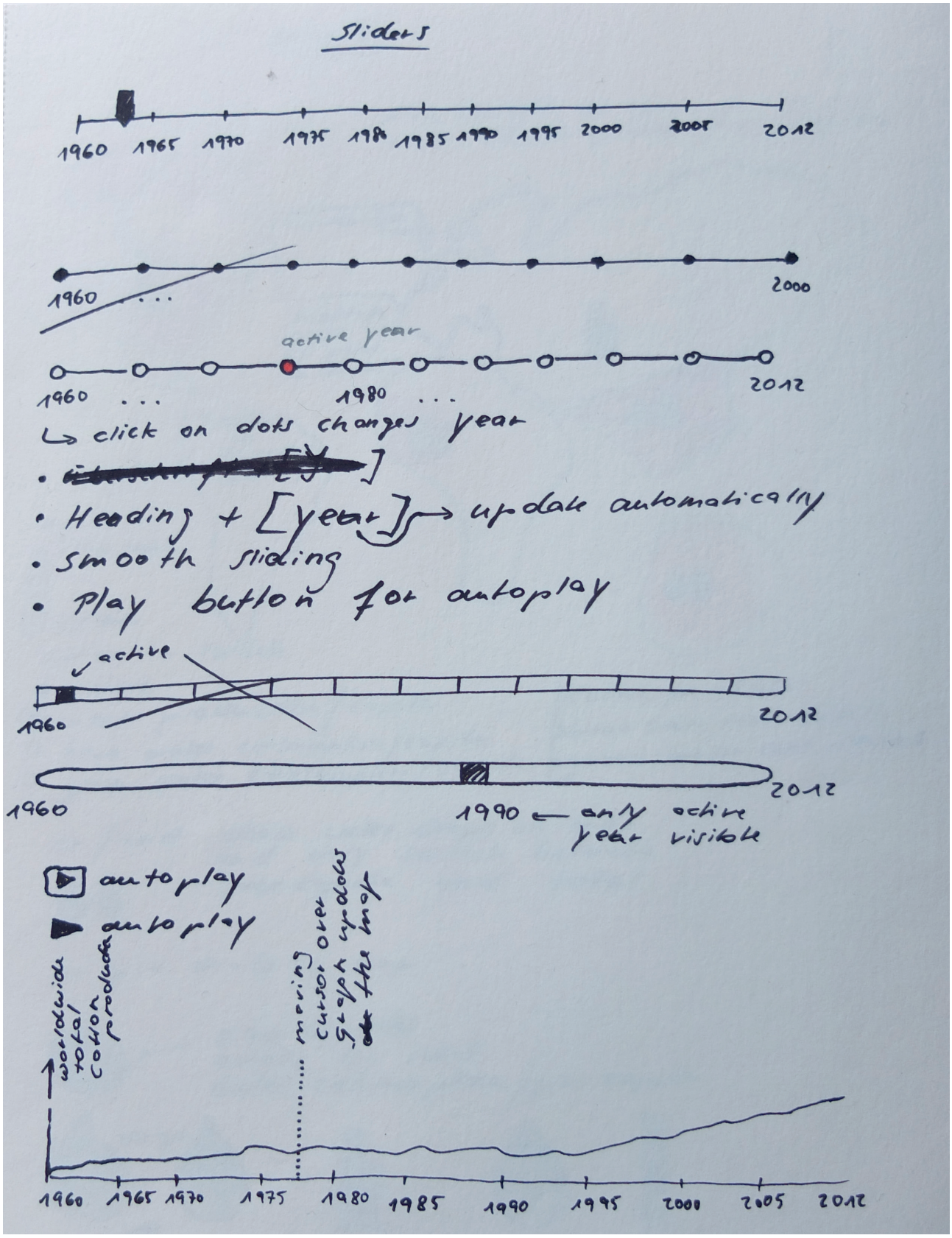


green water consumption (m^3/capita)



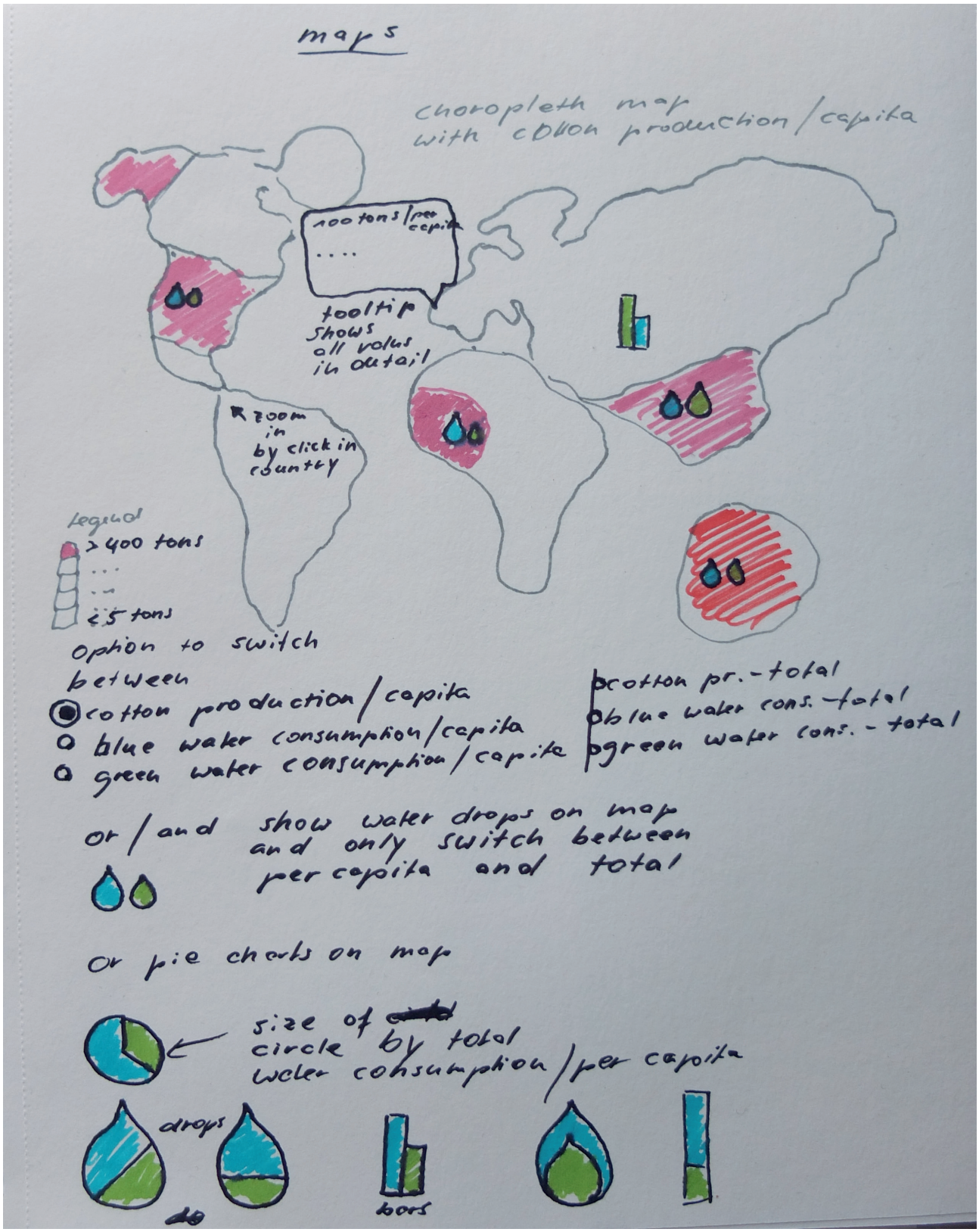
Design Evolution

Sketches of possible time sliders

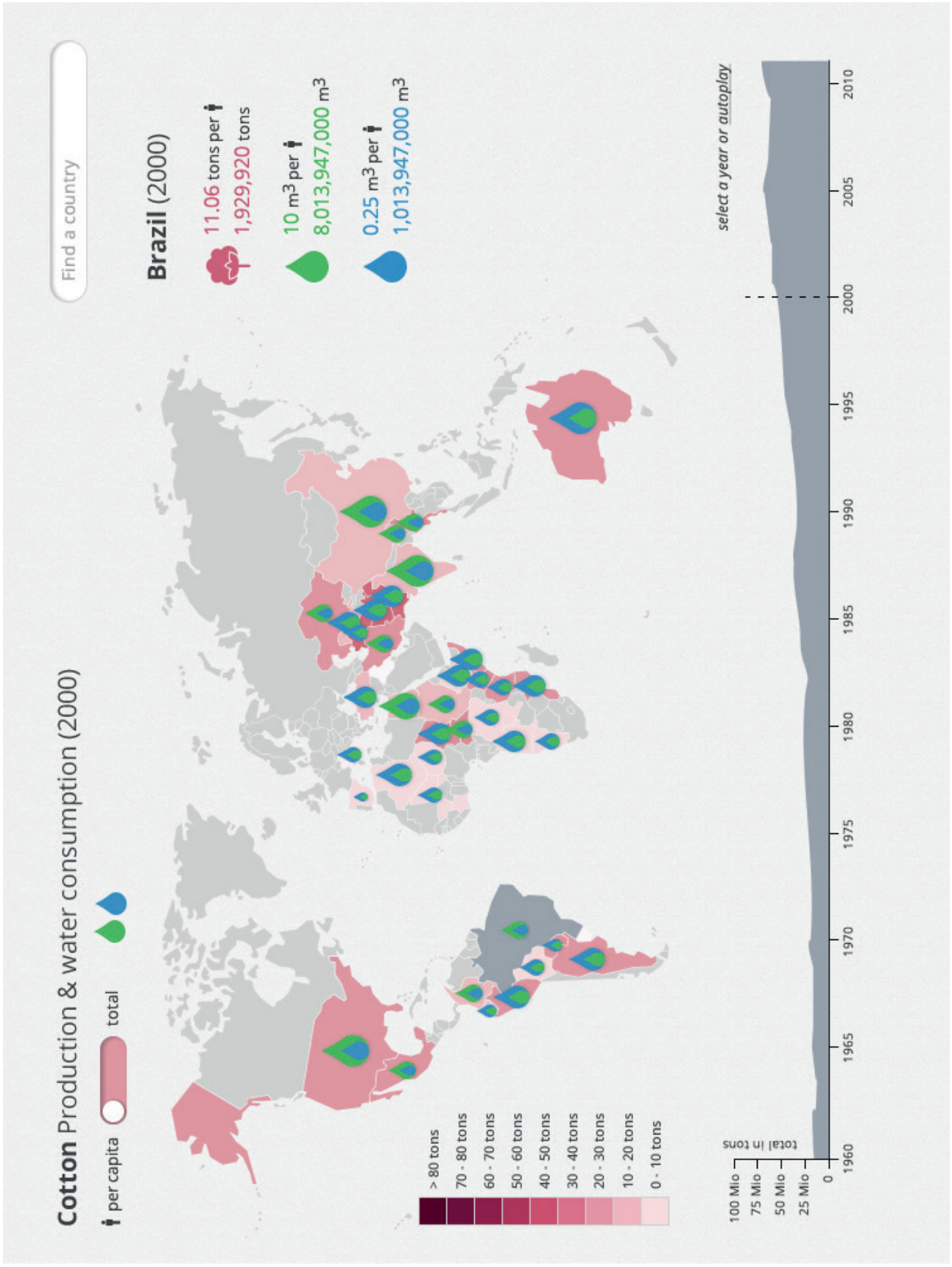


choropleth map sketch and display of water values

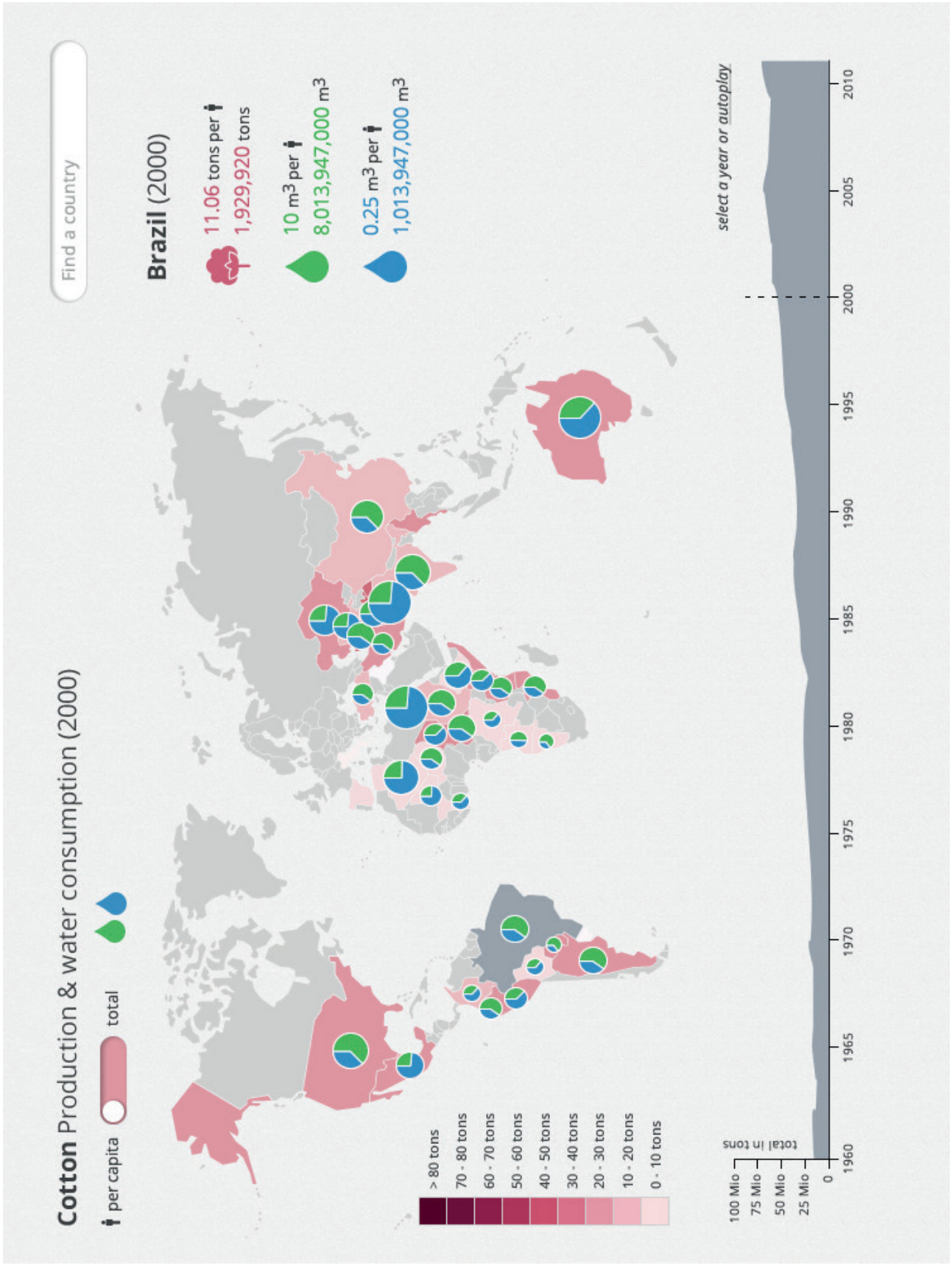
possible integration of filters and switches



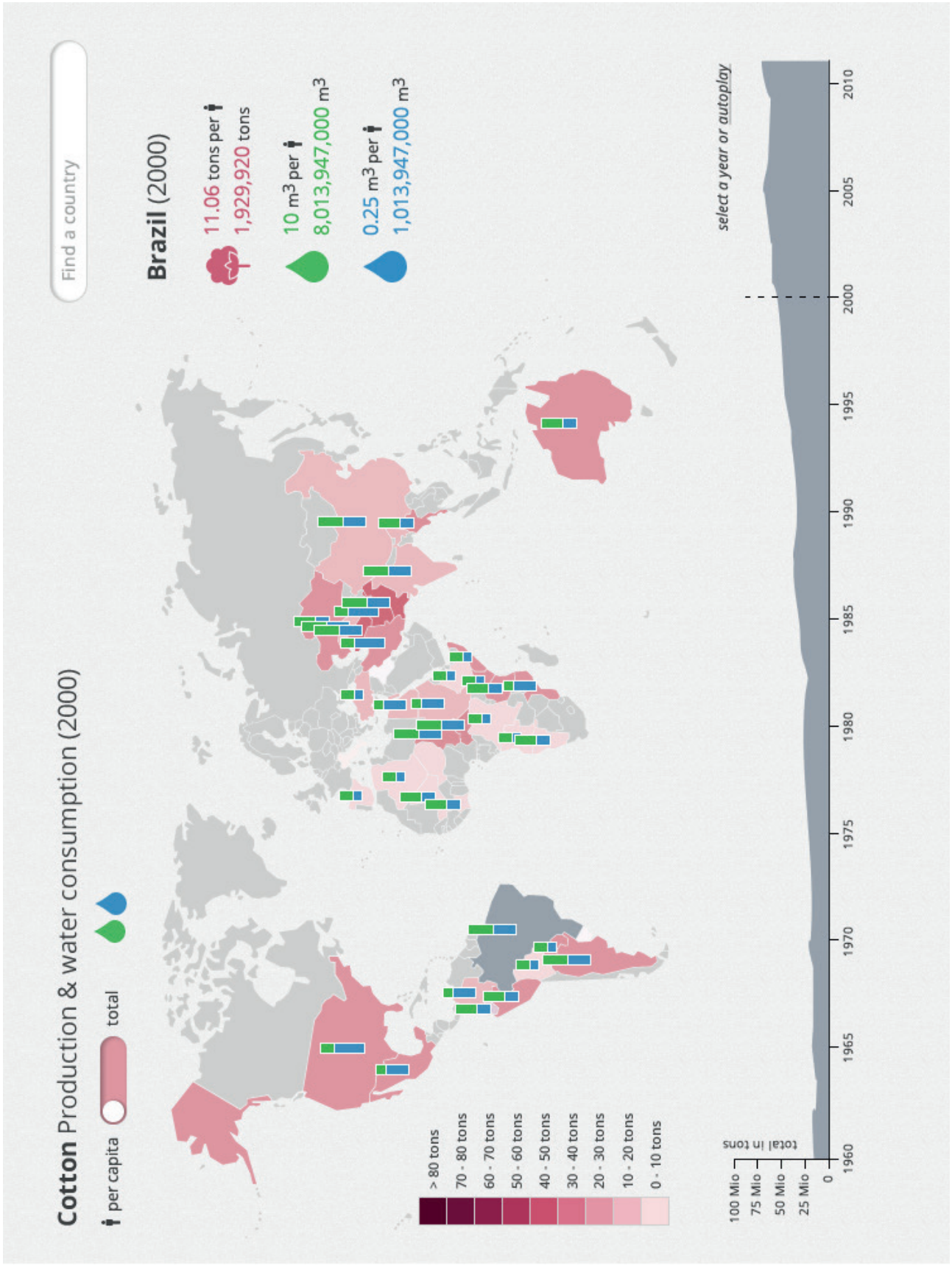
final design with drops



final design with pie charts



final design with stacked bar charts



Design explanation

I will implement a choropleth map in combination with symbol map to show all 3 values in relation to each other. I will choose between pie charts and stacked bar charts because with drops it will be difficult to compare the values for blue and green water. At the bottom of the map there is a area or line chart with the total production worldwide. This chart is also the time slider but there is also the possibility to navigate via arrow keys or cycle through years via auto-play. The year values in the heading and the right side will update automatically. At the top of the map you can choose between per capita and total values to compare the differences for each country in each year. I will implement a zoom via scroll and click. By hovering over a country you can see all values for that country on the right side of the map and the bottom graph will only display the values for that country. I will implement a country search that instantly gives me the values of the country and highlights it on the map. By clicking the green or blue drop on the top you can disable/enable the display of that values on the map.

Update

I integrated a help functionality that is always accessible but hidden by default to give the visualization more space and only show important elements. I decided to keep the design clean and not overload the screen with too many elements. I worked a lot with icons to highlight important functions and make the visualization more memorable. I chose not to implement the country search functionality because a map in combination with a zoom should make it easy to find a country. I worked with bright colours so the important coloured elements can stand out. I set the opacity of the pie charts to 80% so the underlying countries are still visible. I decided to use bar charts instead of stacked bar charts because they harmonize better with the map and show the differences in the water values clearly.

Implementation

I started the implementation of my visualization with the choropleth map. I used topojson to merge countries to USSR, Yugoslavia, Ethiopia,.. for certain time intervals. This example helped me: http://bl.ocks.org/mbostock/5416440. I implemented a first basic time slider to navigate through the years following this example: http://data.karmi.cz/d3/life-expectancy-map/. I tried different colour scales and implemented a zoom using my code from HW4. I tested the display of bubbles on my map and will continue to add the pie charts or stacked bar graphs on top of the map.

Update

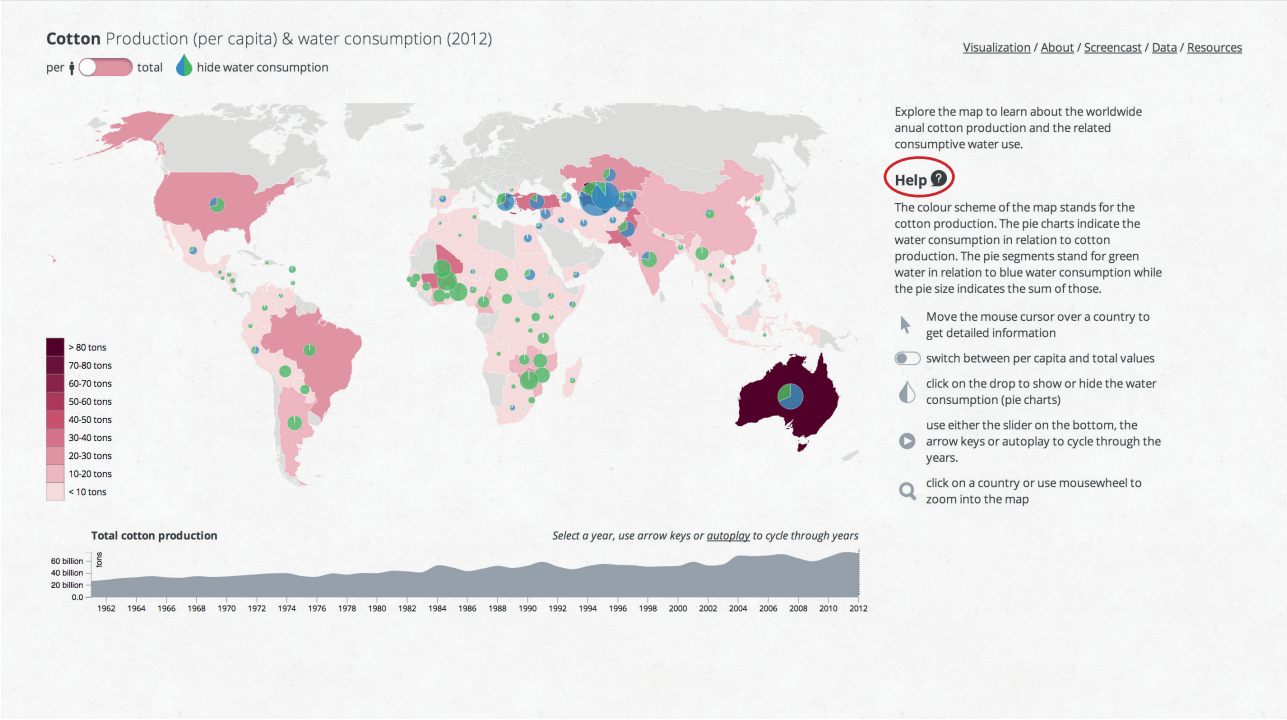
I implemented the pie charts and had to find a workaround for the case that both, blue and green water values can be zero in certain years. I that case I got a lot of error messages because the ankles couldn't be calculated. In my opinion that's a bug in d3. So I set the zero values to a very low number to avoid this error messages.

I used an area chart in combination with d3.js slider (https://github.com/turban/d3.slider). I had to modify the code to get my chart time slider.

The interaction elements of my visualization are explained on the next pages.

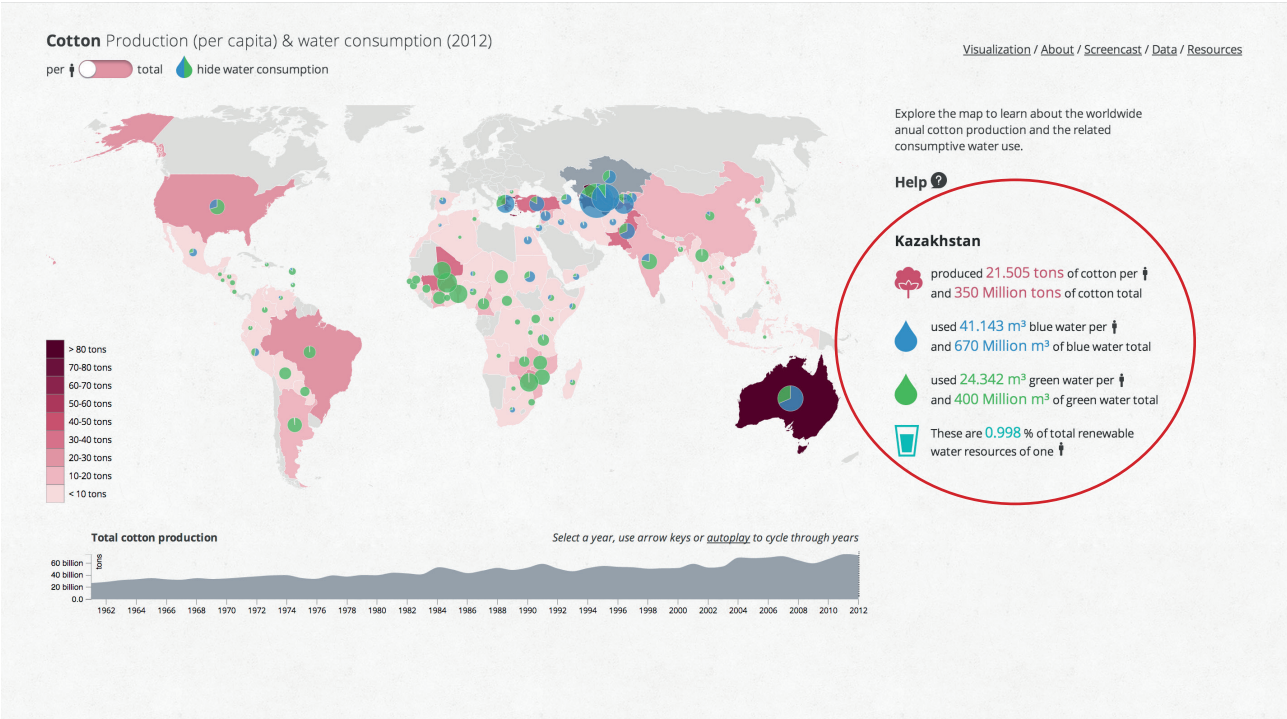
Help

The user can get help about the interactive features of this visualization by moving the mouse over the “Help” text:



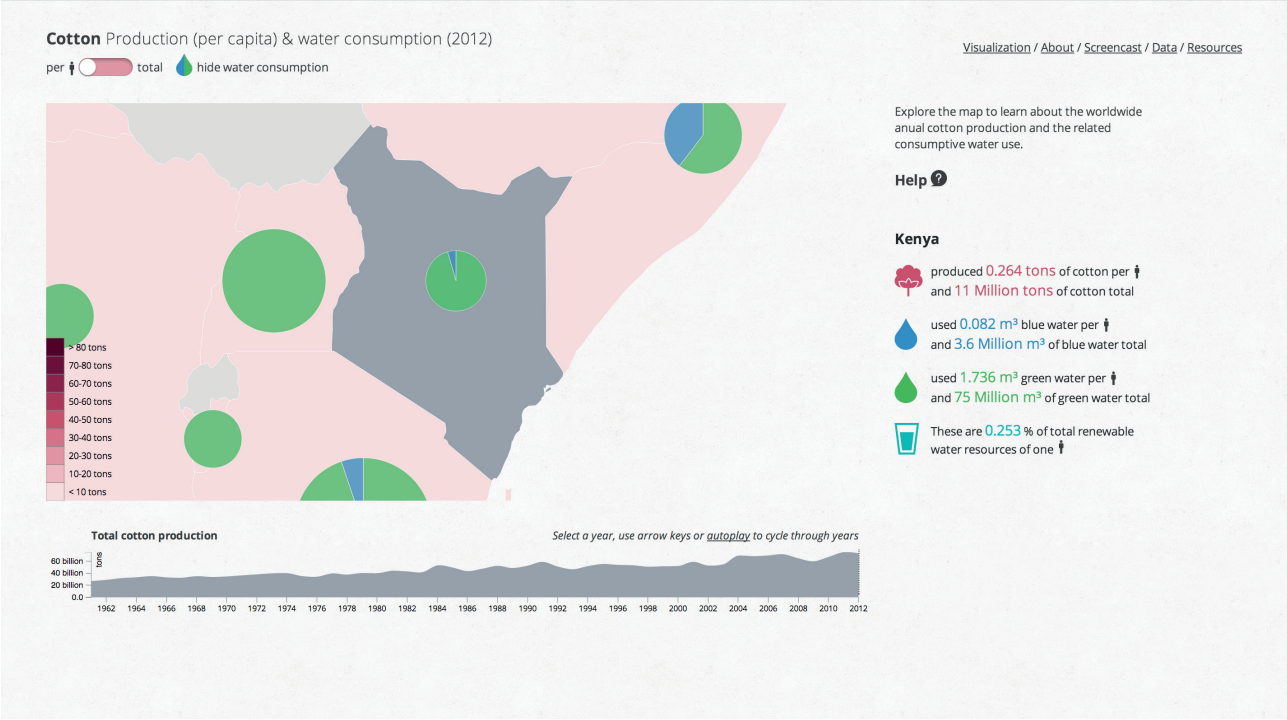
Detailed Information about a country

By moving the cursor over a country it will be highlighted and detailed information about that country in that specific year is showing up in the sidebar. This includes a calculation of the percentage value of total renewable water resources per capita used for cotton production. :



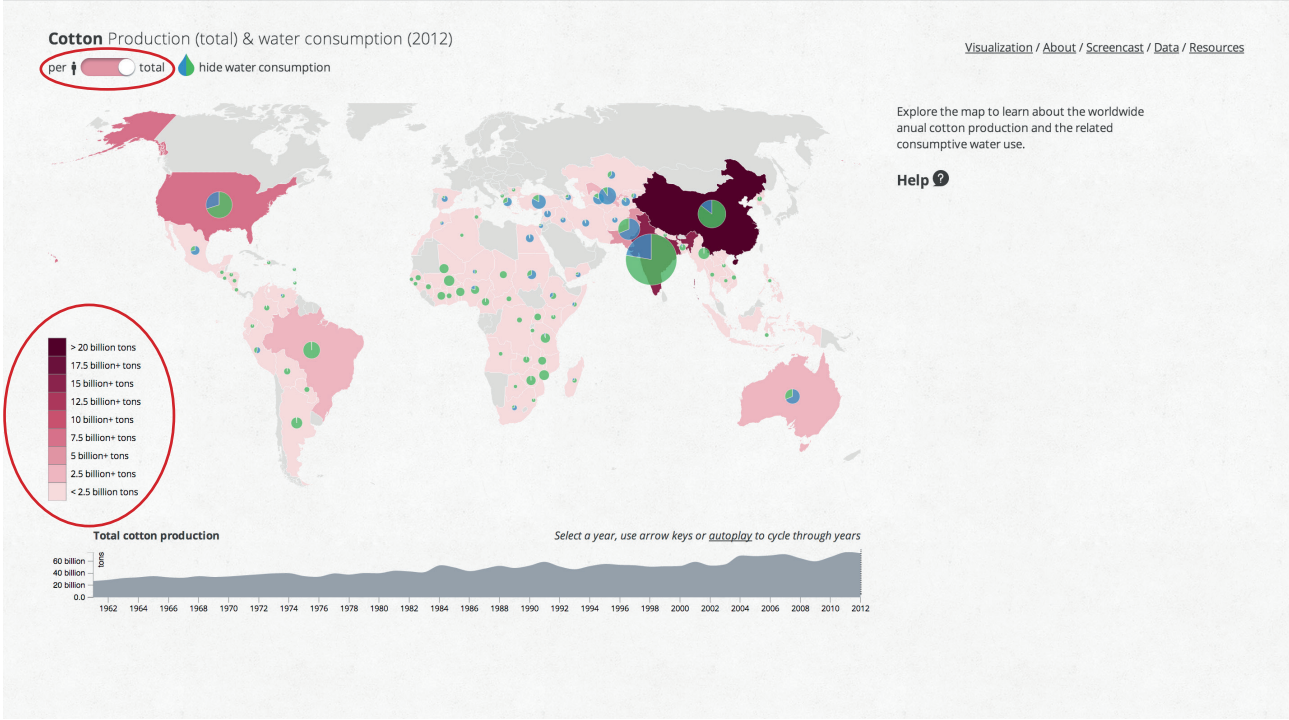
Zoom

There are 2 different ways to zoom into the map: clicking on a country or using the mouse wheel. So it's easier to see details of smaller pie charts.



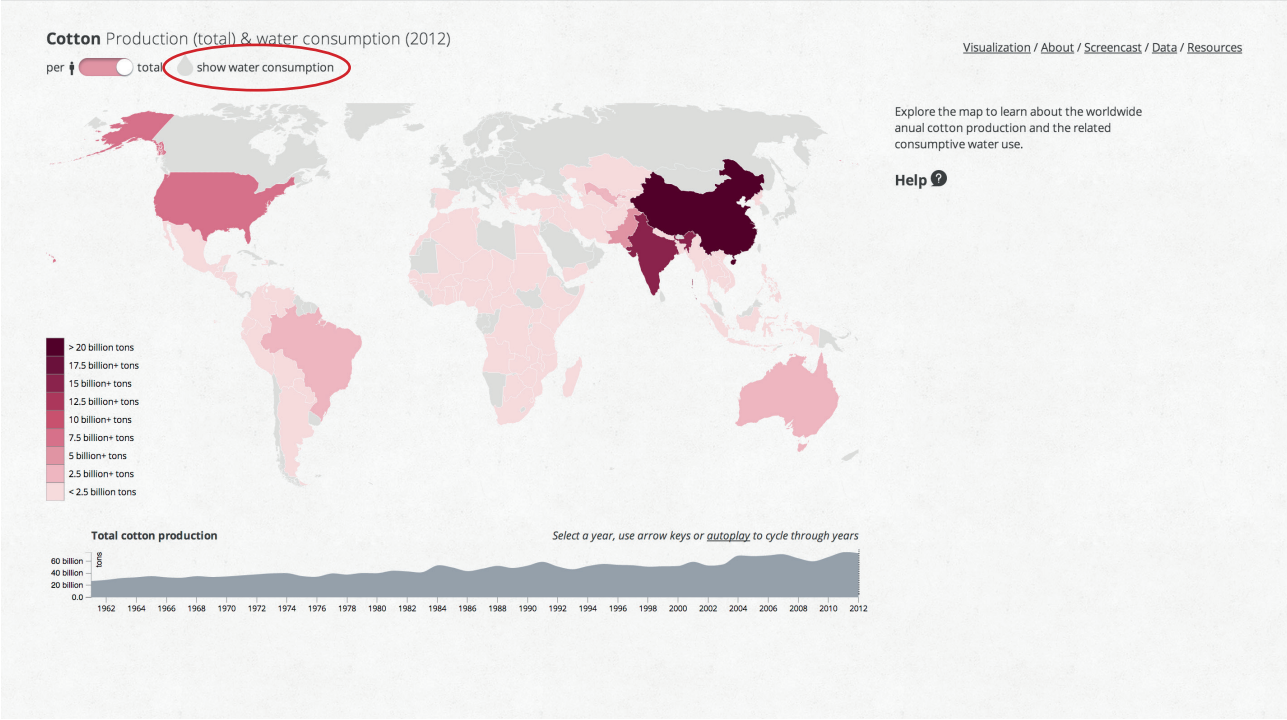
Switch

By default per capita values are selected. The user can switch to total values and the map and the legend change:



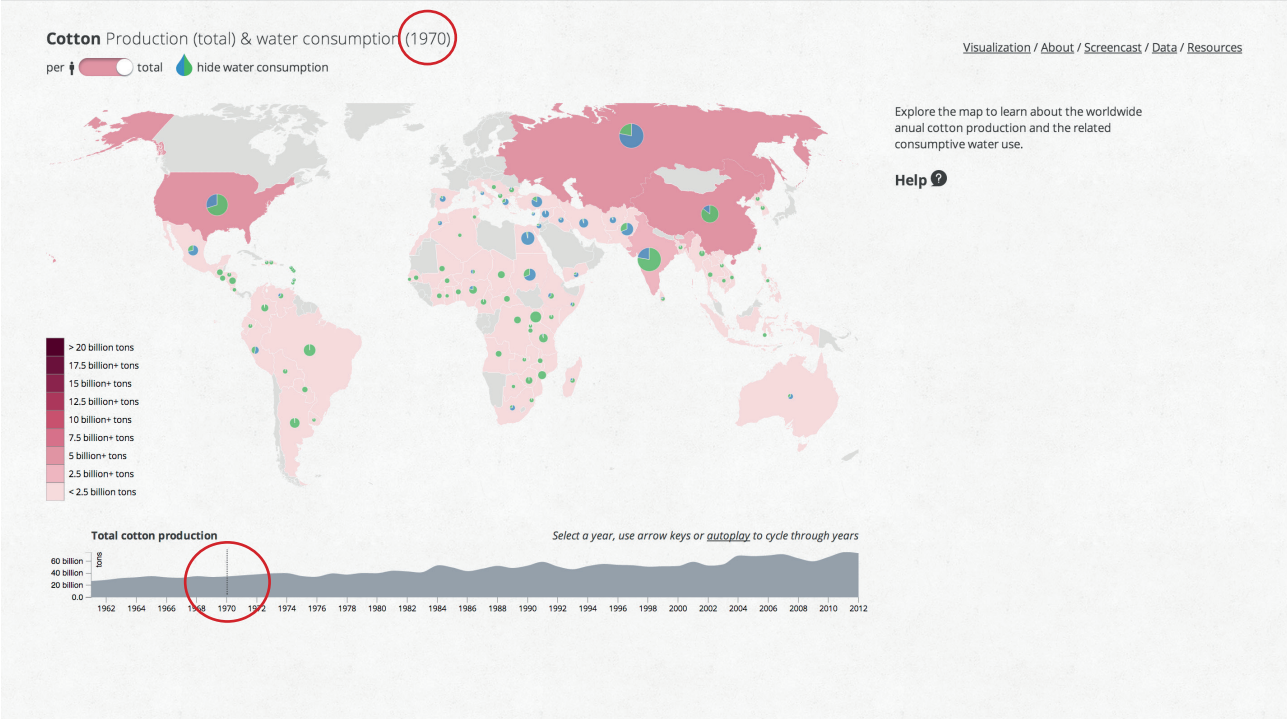
Toggle display of pie charts

The user can toggle the display of the pie charts to concentrate only on cotton production and see countries that are usually hidden by big pies:



Time slider

By default the year 2012 is selected. The user can either use the time slider to select another year or the arrow keys or the auto-play function. The chart behind the slider shows the total worldwide cotton production over the years. To actual year the user navigates to is displayed in the heading:



Evaluation

I learned a lot about my data while implementing the visualisation. It was shocking to see that Uzbekistan is using 22.5 % of total renewable water resources per capita just for cotton production. It was interesting to see that India is using so much blue water in total, even if the share of green water is bigger. I feel like my questions are more then answered.

Further improvements

When changing the year via auto-play or the arrow keys it would be nice to see the time slider move as well. It would be better to find an easier more dynamically way to load the data into d3. But there is just too much data clean-up to do when working with annual data and data from different sources.

Possible extensions

I would like to extend the map with import and export data to see how much water is actually exported in form of cotton. There is also the idea to integrate the whole manufacturing process because textile production uses huge amounts of water. This map can also be used for other crops. It would be interesting to see that values as well, maybe provide a select box and show the 10 most thirsty crops.