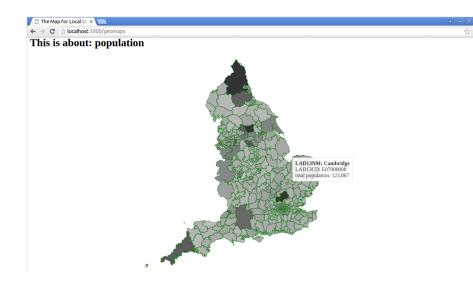


THE (FLOODED) ITALIAN JOB

Alessio Gottardo Daniel Shamoa Tommaso Benacchio

FloodHack, ECMWF, 17 January 2016

HOW IT ALL STARTED



Tools: Javascript (d3.js, node.js)

Tools: Javascript (d3.js, node.js)

LDAs plot on Mercator map of England - threshold-based color filling

Tools: Javascript (d3.js, node.js)

LDAs plot on Mercator map of England - threshold-based color filling

Tabular finer info through tooltip

Available info + model with data

Available info + model with data



Available info + model with data

 \Downarrow

Automating tool for decision and warning

Available info + model with data

 \Downarrow

Automating tool for decision and warning

SMS to flooded people e/o ONGs etc.



1. FETCHING THE DATA



Queries on rasdaman web service

Queries on rasdaman web service

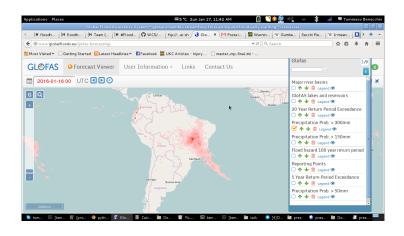
Struggling to filter available data, e.g. to take out ranges, subsets...

Queries on rasdaman web service

Struggling to filter available data, e.g. to take out ranges, subsets...

Trials with Panoply, netcdf going nowhere...

Solution: lowering expectations!



Fixed rainfall threshold all over the world!

Smaller data sets, simple measurements, no server queries

Fetching from rasdaman 2m temp and total ppn for a given time interval and a region (eg a nation)

Fetching from rasdaman 2m temp and total ppn for a given time interval and a region (eg a nation)

```
url_fmt_tp = 'http://incubator.ecmwf.int/2e/rasdaman/
ows?service=WCS&version=2.0.1&request=ProcessCoverages
&query=for c in (%s) return encode(c[Lat(%f:%f),
Long(%f:%f), ansi("%s" : "%s")], "csv") '
url_tp = url_fmt_tp % ("TP", 50.0, 51.0, 1.0, 2.0,
"2014-12-20T00:00:00+00:00", "2014-12-30T00:00:00+00:00")
```

Organize data in numpy array

```
def stuff(url_to_process):
# fetch the data
r = requests.get(url_to_process)
# clean the data
r.raise_for_status()
data = np.array(eval(r.text.replace('', '[').replace('', ']')))
print (data.shape)
# build the data structure
final = []
for col in range(data.shape[0]):
for row in range(data.shape[1]):
final.append([(np.arange(max_lat, min_lat - 0.5, -step)[row],
np.arange(min_long, max_long + 0.5, step)[col]),
data[col][row]])
matrix = np.arrav(final)
# sort and return the data structure
sorted.matrix = sorted(matrix, key=lambda x : x[0][0])
return sorted matrix
                                             4 D > 4 图 > 4 E > 4 E > E 9 Q ()
```

From Python, call R clustering function k-means

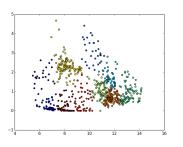
```
base = importr('base')
stats = importr('stats')
# R to py suff
from rpy2 import robjects
from rpy2.robjects import pandas2ri
from rpy2.robjects.packages import importr
R = robjects.r
KM = R.kmeans(data_t2m_tp, 10)
centers = np.array(KM.rx2('centers'))
clusters = np.array([np.array(KM.rx2('cluster'))]).T
```

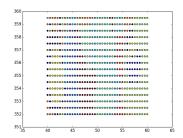
Plotting data on lat lon grid

```
complete_data = pd.concat([data_frame, pd.DataFrame(clusters,
columns=['cluster_id'])], axis=1)
#stuff = pandas2ri(data_frame)
plt.scatter(x=data_t2m_tp.ix[:,0], y=data_t2m_tp.ix[:,1],
c=complete_data.ix[:,4])
plt.show()
#R.points(data_lat_long, col = KM.rx2('cluster'))
plt.scatter(x=data_lat_long.ix[:,0], y=data_lat_long.ix[:,1],
c=complete_data.ix[:,4])
plt.show()
```

Output: new individual thresholds as means of centroids of the clusters

Output: new individual thresholds as means of centroids of the clusters





Feeding the new thresholds back to the node.js app and show data on d3.js map

Feeding the new thresholds back to the node.js app and show data on d3.js map

Interface with SMS-sending service to people, NGOs etc. for warnings