Early Adopter Report

GLOBAL HEATWAVE AND WILDFIRE ANALYSIS BASED ON THE ESDL DATA

Xikun Hu
KTH Royal Institute of Technology, Sweden
xikun@kth.se *2019-07-17*

1. Introduction

In 2018, an increase in the number of wildfires on our planet due to drought and warmer weather has been registered. 2018 was affected by drought and heatwave and it was a period of unusually hot weather without any precipitation, which led to record-breaking temperatures and wildfires in many parts of the world. According to the World Meteorological Organization, the severe heat waves across the northern hemisphere in the last summer, as well as events of extreme precipitation, are linked to climate change. The impacts were severe, even in the countries considered rich and well prepared to deal with the climate change effects, for example, several fires were out of control and Sweden requested help from neighboring countries and via the European Union's Civil Protection Mechanism [1-2].

Above all, as the dominated natural disaster caused by a heatwave, wildfires can lead to soil erosion, expansion of invasive plant species, and loss of property and life. It is urgent to analyze the relationship between heatwave and wildfire timing and location, and then understand the mechanism for the formation and transition of wildfire across time, finally to mitigate its effects. At present, we have limited knowledge about trends in wildfires worldwide and it is not easy to conclude whether fires are becoming more destructive or where the fire will advance further. Most of the research work focused on post-fire damage assessment maps more than studying the production mechanism along with wildfire like the effects of precipitation, temperature, and even carbon dioxide emission due to wildfire, which would be very significant to mitigate the impact of future emergencies. If humans are to live sustainably on flammable landscapes, we need a global system for collecting data on fires to gain a coherent picture and export wildfire fighting strategies based on the multi-source data.

ESDL provide a growing list of relevant variables for Earth System Science. Most of them have been derived from Earth Observation. At ingestion into the ESDL, all data sets are transformed in space and time to fit to the common grid of the data cube. In this report, to save the computation time, we choose the lower grid resolution 0.25 arc degree data cube as the dataset. The results could be easily redirected into higher resolution 0.083 arc degree through changing the source data cube.

2. Method

ESDL includes almost all the earth observation data for the global heatwave and wildfire analysis such as precipitation, air temperature, carbon dioxide emission and monthly burnt area, with around 20 years period [3]. The fusion of these data above could provide a comprehensive solution of the wildfire emergency forecast. Due to the anthropogenic factor, wildfire is also mainly caused by high temperature as well as events of extreme precipitation. The temperature has a direct effect on the sparking of wildfires because heat is one of the three pillars of the fire triangle. Sticks, trees, and underbrush on the ground receive radiant heat from the sun, which heats and dries potential fuels. Warmer temperatures allow for fuels to ignite and burn faster, adding to the rate at which a wildfire spreads. In this report, the following cube data variables will be considered and discussed. All the code can be accessed through the GitHub link.

Project	Name in ESDC	Description
GFED4	burnt_area	Burnt Area based on the GFED4 fire product.
	c_emissions	Carbon dioxide emissions due to natural fires expressed as carbon flux.
GPCP	precipitation	Precipitation based on the GPCP dataset.
ERAInterim	air_temperature_2m	Air temperature at 2m from the ERAInterim reanalysis product.

The analysis method consists of four steps:

1. *Analyze the burned area GFED4 data for the whole European countries*
2. *Get the suitable geometry study area of burned area*
3. *Plot the summation value of data variables across 1D time to compare the values*
4. *Systematic analysis of multi variables in 2D*

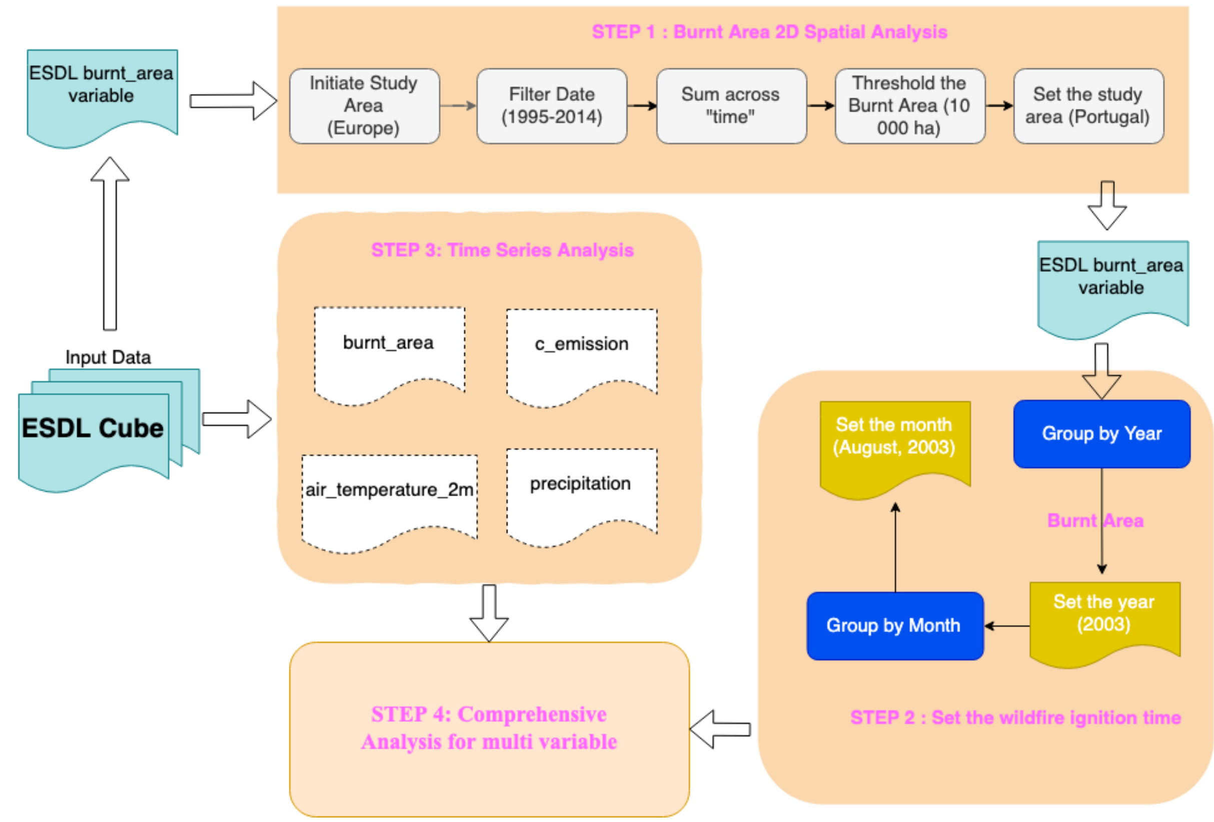

Figure 1. The methodology flowchart

3. Results

3.1 Total Burned area analysis from global to regional

This step is to find the suitable study area. ESDL provides us with more than 20 years burnt area dataset across time and through summing all of burnt area and visualize the results could help us to find the suitable study area where has dramatically historic wildfire within 20 years. The selected study area will be used for the next temporal analysis.

Coding Implementation:
https://github.com/KuntaHu/ESDL_Early_Adopter

Method A: based on the xarray with matplotlib.collections.QuadMesh

```
Europe = ESDC_img.sel(lat = slice(70.,30.), lon = slice(-20.,35.))
## time series from 1995 to 2014
Europe = Europe.sel(time = slice('1995-08-01','2014-07-30'))
%time Europe_BA_time = Europe.sum(dim='time').compute()
plt.figure()
Europe_BA_time.plot()
plt.savefig('Sum_API_time_dimension.png', dpi = 500,bbox_inches='tight')
```

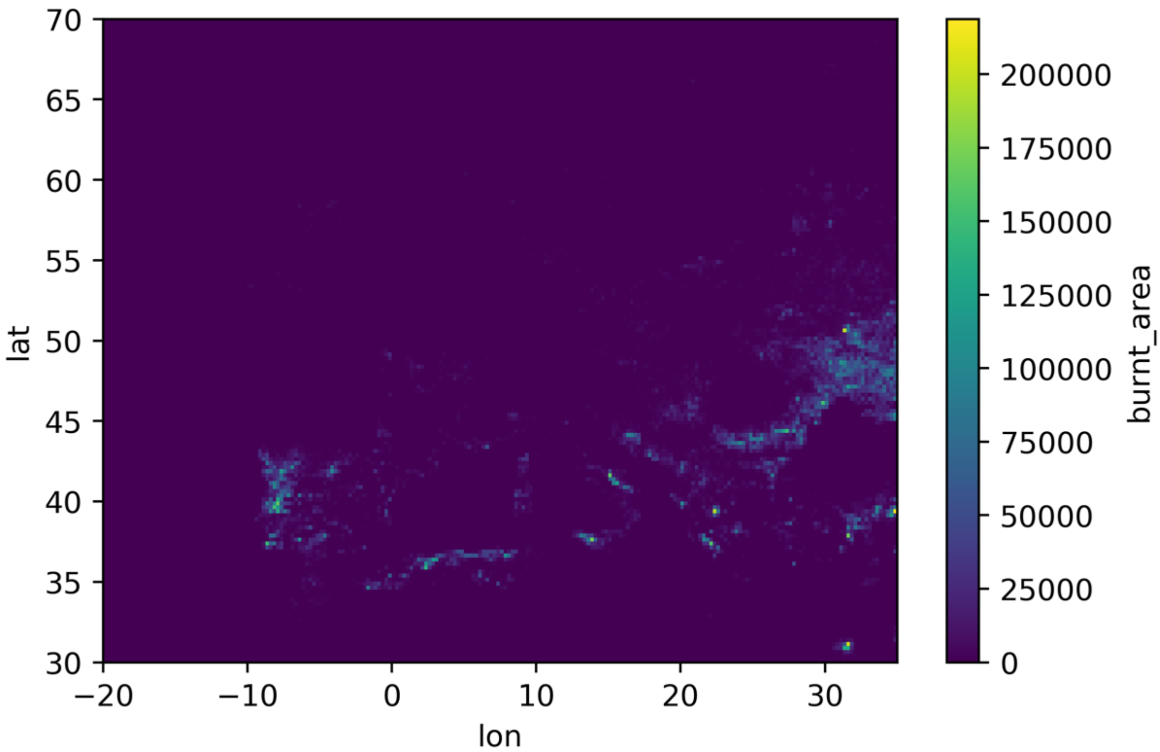

Figure 2. The burnt area in Europe based on xarray plot method

Pro: The original xarray dataset contains the lon and lat values and the variable name like burnt_area, which could be plotted directly without any more definition about the legend and colorbar as Figure 2 displayed.

Con: It cannot be easily defined the normalization for the value of sum of burnt area, which make the stretch wide and invisible for the burnt area distribution.

Method B: based on the numpy with matplotlib.image.AxesImag

Using the numpy plot methods, as Figure 3 and Figure 4, the normalization could be easily reached to highlight the burned area and unburned area, but the problem is that we need to redefine the x-axis and y-axis and make them between respective longitude and latitude range.

```
sum_axis_0_BA = np.nansum(Europe.values, axis=0) # convert the xarrat into np
from matplotlib.colors import LogNorm
plt.figure()
plt.imshow(sum_axis_0_BA, norm=LogNorm()) # show the value as logNorm
# plt.imshow(sum_axis_0_BA)
plt.colorbar()
# plt.show()
plt.savefig('BA_sum_all_years_logNorm.png', dpi=500, bbox_inches='tight')
```

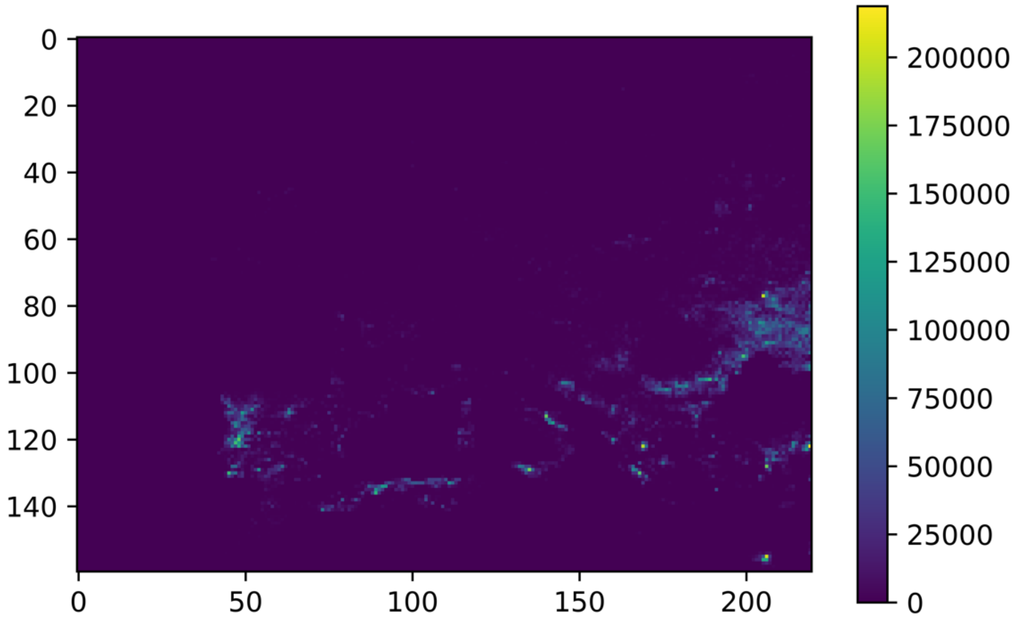

Figure 3. The burnt area in Europe based on numpy plot method

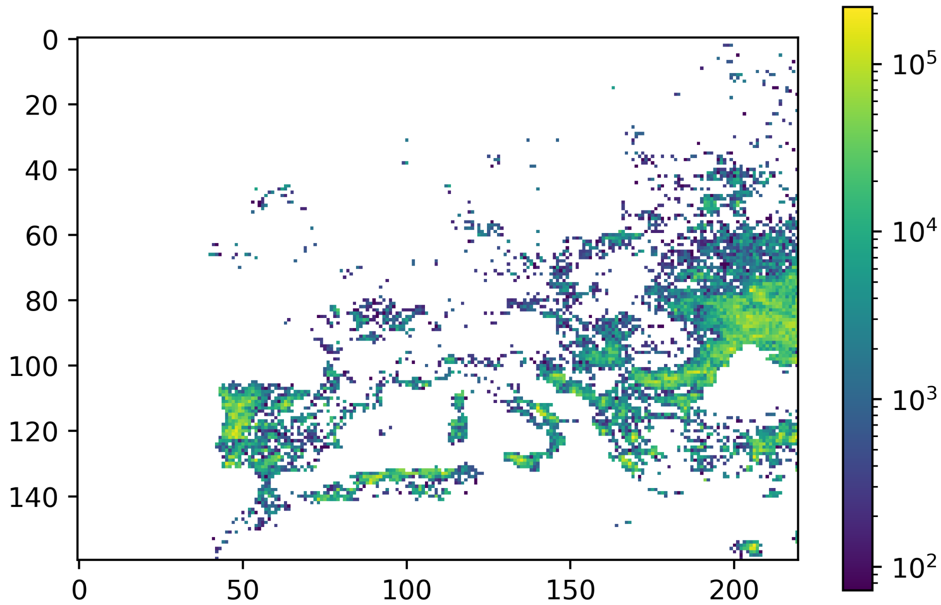

Figure 4. The burnt area in Europe based on numpy plot method in LogNorm

To show the burnt area obviously, we should stretch the burnt area into log-normal distribution and then highlight the burned area in European area as Figure 4 showed. One more drawback is that the x axis and y axis as the two dimensions are showed as the dimension shape (160, 220) which means the (lat = slice(70.,30.), lon = slice(-20.,35.)). With 40 latitude and 55 longitude, we could get the shape values as 160*220 because of the 0.25 degree spatial resolution.

Next, we need to analyze the geometry of the burned area over a specific area, within 20 year. High burnt area means frequent ignition of wildfire with dramatic burned distribution, where even more than one wildfire happened because of the probability of regrowth of vegetation within ten years.

As we know, usually the 0.25 arc degree resolution determines the grid cell has over 72 900 ha (around equator), and for the European countries (far away from equator), the cell area changes a lot but around 70 000 ha more or less. In this case we could set the burnt area threshold as 10 000 which filters out the dramatic burnt area with over 10 000 ha with 20 years. If we print the latitude values when x = 50, we could find many grids have burnt area over 10 000 ha as Figure 5 showed.

```
[221]: # get the threshold value over than 10 000
       sum_axis_0_BA_thre = (sum_axis_0_BA > 10000) * sum_axis_0_BA
       sum_axis_0_BA_thre.shape

[221]: (160, 220)

[234]: sum_axis_0_BA_thre[:,50]

[234]: array([    0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   , 41033.88 , 39578.824, 74859.86 , 21194.863,
              54380.48 , 15050.742, 39487.77 , 61309.746, 48702.57 ,     0.   ,
              39538.73 ,     0.   , 82204.41 ,     0.   , 18038.412, 10168.726,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ,     0.   ,     0.   ,
                  0.   ,     0.   ,     0.   ,     0.   ], dtype=float32)
```

Figure 5. The screenshot for the latitude grid values when the longitude grid number is 50

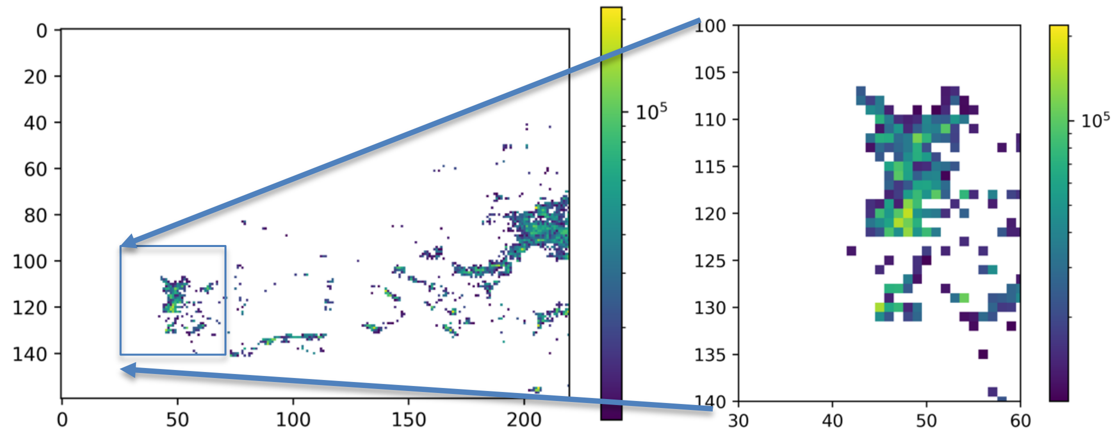

Figure 6. The study area (thresholded by 10 000ha) and confirmed study area based on the numpy polt in LogNorm

Finally, we get the study area around [35, 45] for latitude and [-12.5, -5] for longitude which corresponds to the numpy area with [100: 140, 30: 60]. The study area is located in Portugal a s Figure 6 showed. The total burnt area is 54422108.0 ha from 1995 to 2014 if we sum all the burnt area in time dimension.

3.2 Temporal Analysis for burned area in spatial 2D dimension

Coding Implementation:

Step A: Year-based temporal analysis to find the specific year with the most dramatic wildfire caused largest burnt area in history.

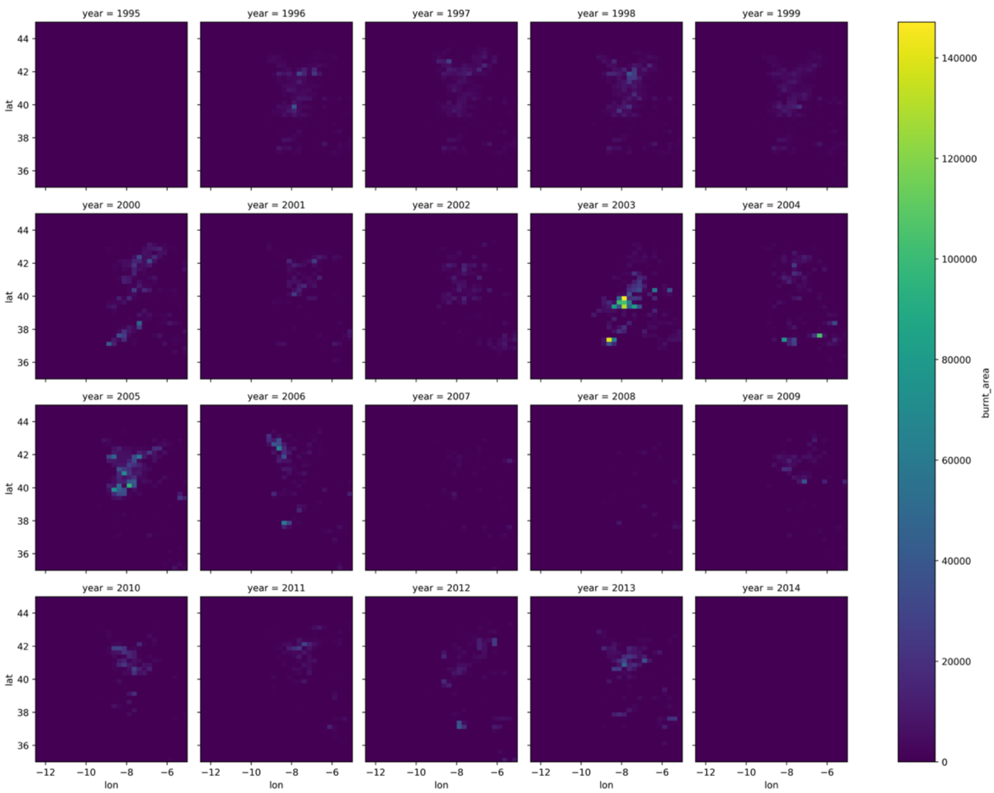

Figure 7. The year-based grouped burnt area summation map

```
Europe_study_area = ESDC_img.sel(lat = slice(45.,35.), lon = slice(-12.5,-5.))
Europe_study_area_year = Europe_study_area.sel(time = slice('1995-08-01','2014-07-30'))
BA_year = Europe_study_area_year.burnt_area.groupby('time.year').sum(dim='time')

plt.figure()
BA_year.plot.imshow(x='lon',y='lat',col='year',col_wrap=5)
plt.savefig('Portugal_BA_group_Year.png', dpi = 500,bbox_inches='tight')
```

From the grouped year with summing burnt area in Figure 7 , we could easily find the target year in 2003, which has highlighted burnt area. In this case, we also need to check the real burnt area for each year, and then we execute next script to compute the summed values for each year in OUTPUT.

```
BA_year_value = BA_year.values
BA_year_value_sum = np.sum(BA_year_value, axis = 1)
BA_year_value_sum = np.sum(BA_year_value_sum, axis = 1)
```

OUTPUT:

```
array([1.49061987e+03, 6.86568375e+05, 6.69763500e+05, 8.92133312e+05,
    6.14431562e+05, 1.03054244e+06, 4.19960656e+05, 4.81474219e+05,
    2.12947175e+06, 7.78531375e+05, 1.61631462e+06, 7.45297562e+05,
    1.25357766e+05, 8.28714844e+04, 3.47568531e+05, 5.36966500e+05,
    3.61186062e+05, 5.11173750e+05, 7.43015000e+05, 0.00000000e+00], dtype=float32)
```

Step B: Next, we will show the summation of burnt area in 2003 with monthly group and weekly group.

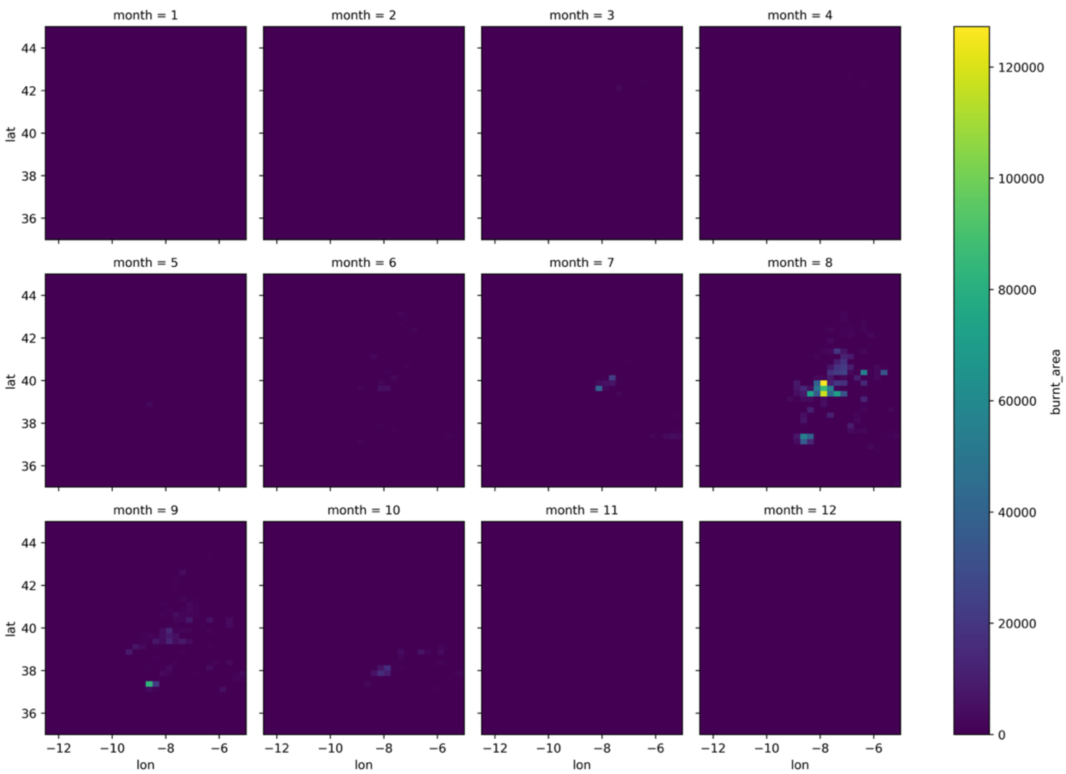

Figure 8. The month-based grouped burnt area summation map in 2003

In same way, we also could easily get the temporal change within 2003 in month even in week as Figure 8 showed. First, we could find the 8th month has the largest burned area and then we could check the burned area in August.

OUTPUT:

```
array([    0.   ,     0.   ,  5672.509 ,  4169.332 , 4908.9326,  46387.395 , 118712.73 , 1496822.4  ,
     349098.97 , 101689.914 ,   2009.6625,      0.   ], dtype=float32)
```

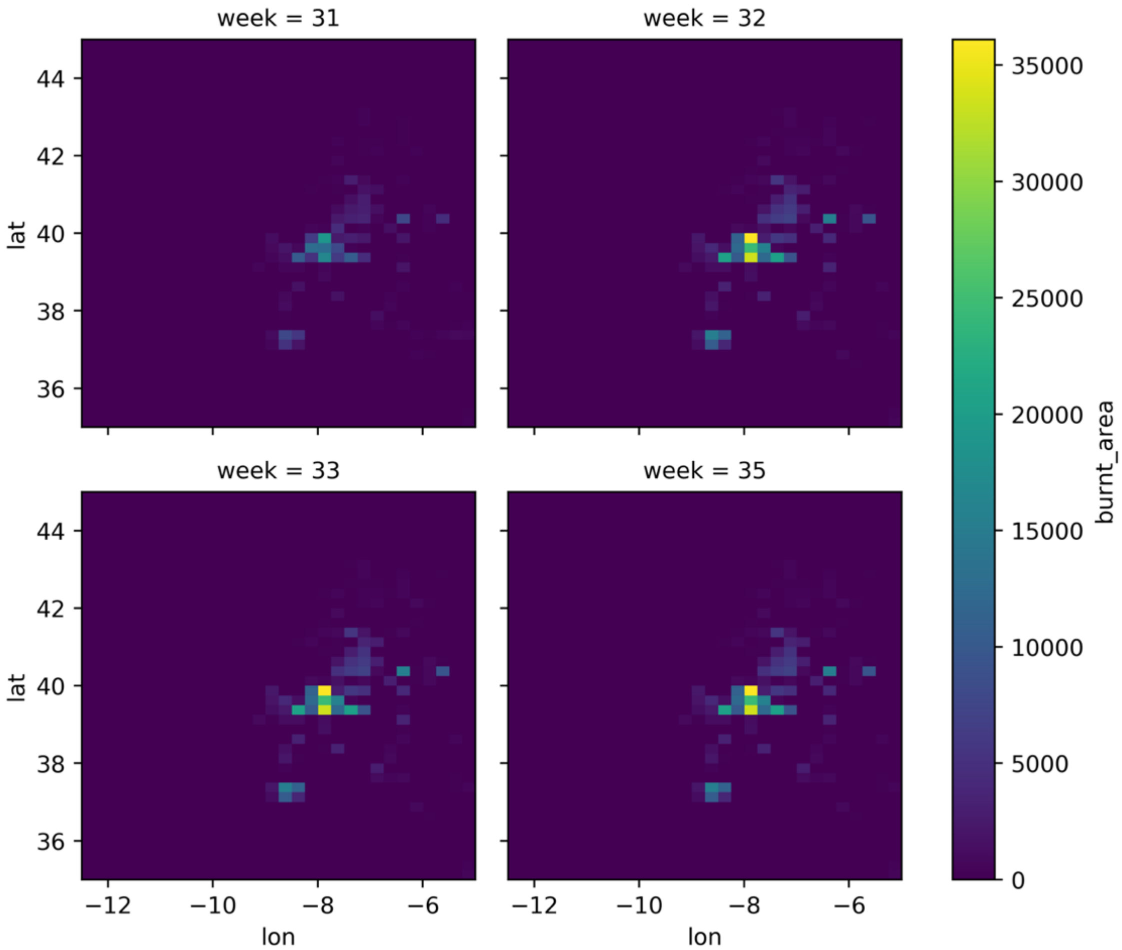

Figure 9. The week-based grouped burnt area summation map in August 2003

The real burnt area for this four weeks are 230790.73, 422010.56, 422010.56, and 422010.56 ha specifically. Base on the weekly burnt area map in Figure 9, we observe a strange thing that the last three weeks own same burnt area as 422010.56 ha which means the wildfire lasted long time even one month resulting out the replicate summations in same geolocation.

Actually, in August, the burned area should be 422010.56 ha rather than the summation burned area with around 1496822 ha. Actually, the GFED data contains the monthly burned area, in ESDL, however, it is shown as weekly data, which makes the summation results a little bit strange. In practice, we could choose the last week within this month as the final burnt area rather than summing them together as the final one to avoid the replication effect.

3.2 Time Series Analysis for burned area in 1 dimension

Next, the one dimension time series analysis for the study area should be given. Based on the ESDL coordinates ("time", "lon", and "lat"), we should compute the mean in "lon" and "lat" and then plot the array in time dimension.

```
Europe_study_area = ESDC_img.sel(lat = slice(45.,35.), lon = slice(-12.5,-5.))
Europe_study_area_year = Europe_study_area.sel(time = slice('1995-08-01','2014-07-30'))
BA_ts = Europe_study_area_year.burnt_area

BA_ts_mean = BA_ts.mean(dim='lon').mean(dim='lat')
fig, ax = plt.subplots(figsize = [14,5], ncols=2)

BA_ts_mean.plot(ax = ax[0], color ='red', marker ='.')
ax[0].set_title("Burnt Area in Portugal")
BA_ts_mean.plot.hist(ax = ax[1], color ='blue',bins = 100)
ax[1].set_xlabel("Burnt Area (hactares)")
plt.tight_layout()
plt.savefig('Portugal_BA_TS.png', dpi = 500,bbox_inches='tight')
```

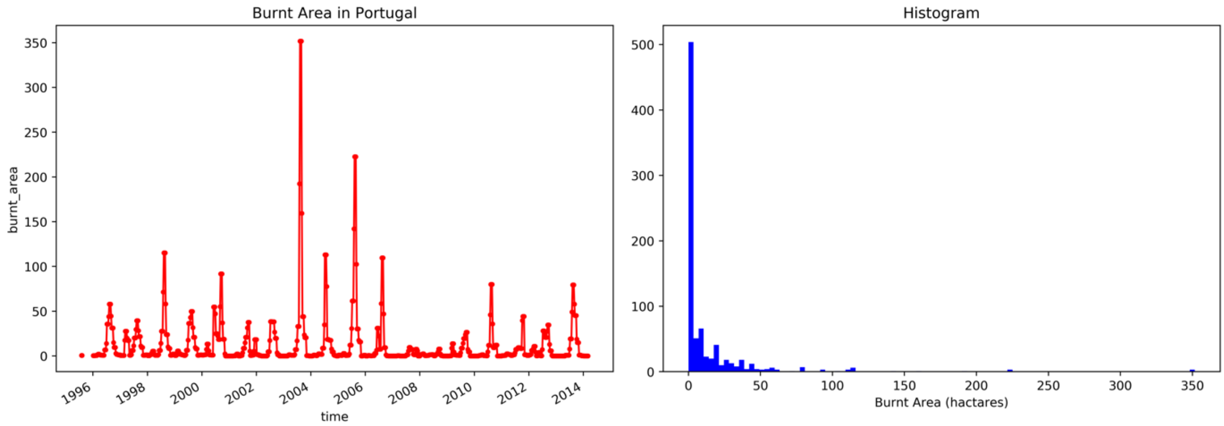

Figure 10 The time series analysis for the burned area from 1995 to 2014

Based on the time series analysis in Figure 10, it is obvious that 2003 saw the largest fire area in the history. And before and after 2003, only small wildfire started in this region which dominates from 0 to 120 ha weekly.

Next, the corresponding precipitation data across time is compared with burnt area as followed Figure 11. As the yellow circle chosen, the pretty low amount of rain makes the environment dry and biomass fuel easy to ignite, which would cause the high burned area at the same time.

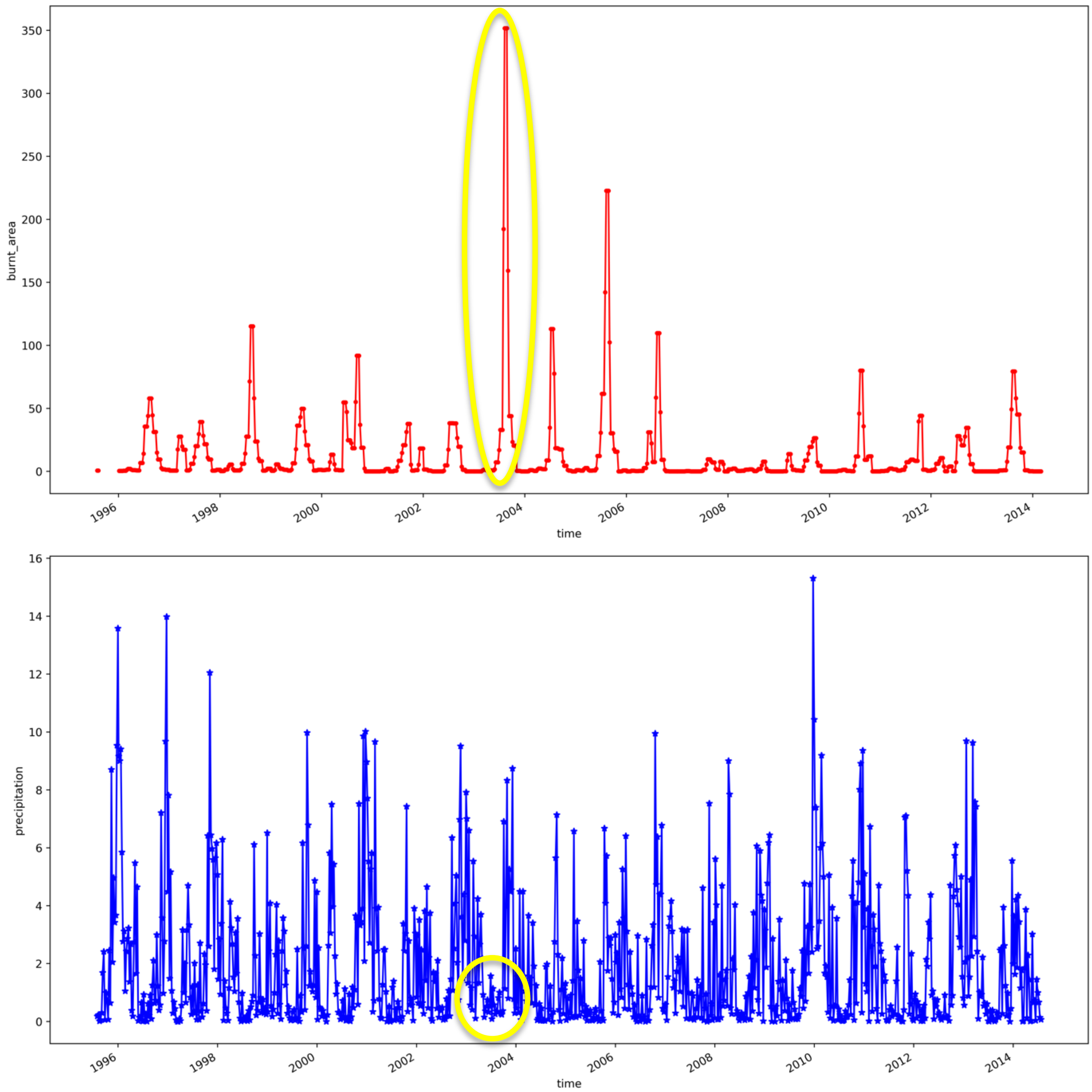

Figure 11 The time series analysis for the burned area and precipitation from 1995 to 2014

Further, we also take the temperature (air_temperature_2m) into consideration. It seems that temperature has harmonic characteristics in season as Figure 12 showed, however, in 2003, around August, we could find the anomaly temperature values which has extremely high peak around 300 K (27 °) in average which means the regional temperature would be pretty high (yellow circle). We also set a dashed line (red line) across time in 296 K, which denotes the normal temperature in average, but the 2003 late summer saw an extreme heatwave which would be a main reason for the large wildfire burnt area.

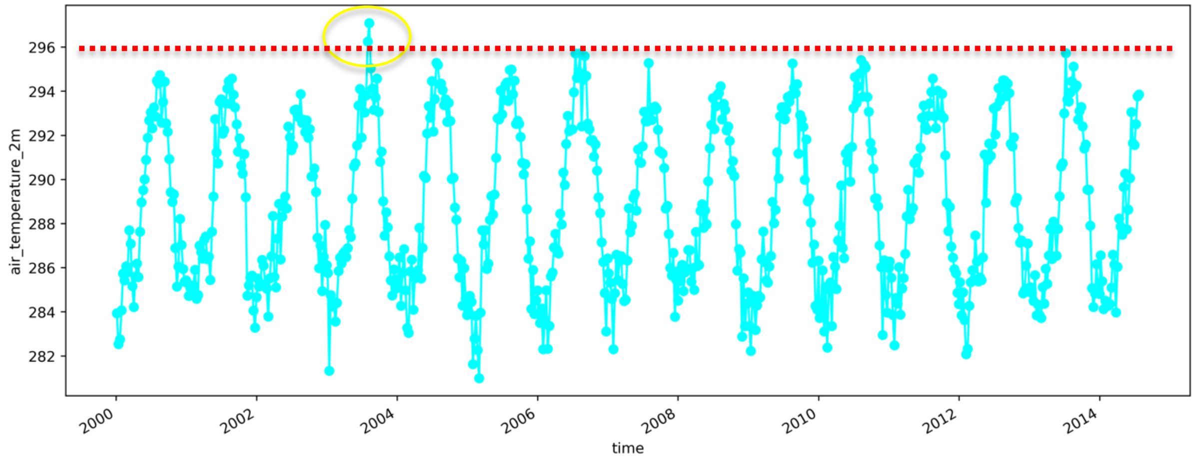

Figure 12 The time series analysis for the air temperature from 2000 to 2014

Carbon dioxide emissions due to natural fires expressed as carbon flux also is very essential during the wildfire analysis. The c_ emissions in ESDL variables could help us to analyze the emission of carbon dioxide during the fire. As we can see in the Figure 13 below, the carbon emission reaches the peak in 2003 August to around 3.5 g C m-2 month -1. At the same vision, the summer of 2005 also saw a high emission of carbon dioxide and at the same time, there was a large wildfire.

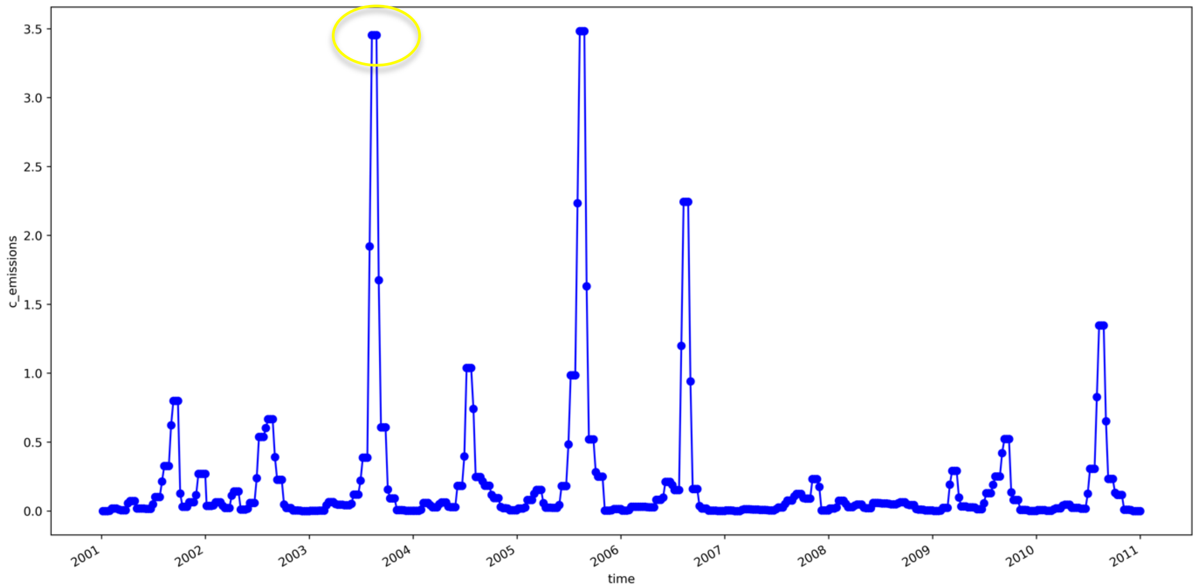

Figure 13 The time series analysis for the carbon dioxide emission from 2001 to 2011

3.3 Comprehensive Analysis for multi variables in 2D

Considering the monthly GFED burned area, actually, it is difficult to find the specific week of the fire ignition. But based on ESDL dataset, we could seek help from other variables as we analyzed in Section 3.2, like temperature, precipitation. Because c_emissions varibles distribute the monthly emissions over the days as well as the diurnal cycle, same with burnt_area variable, it cannot be used to further analyze into specific week.

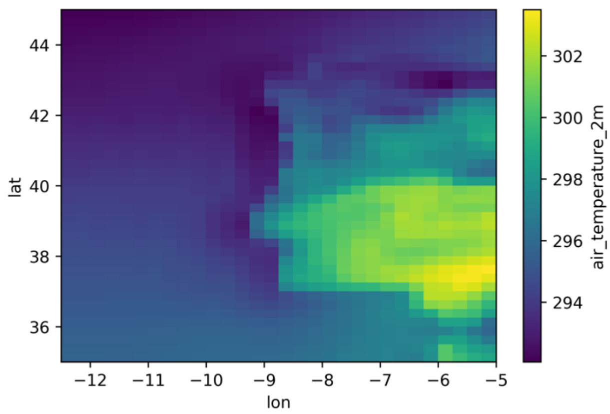

Figure 14 The average of air temperature in study area in August

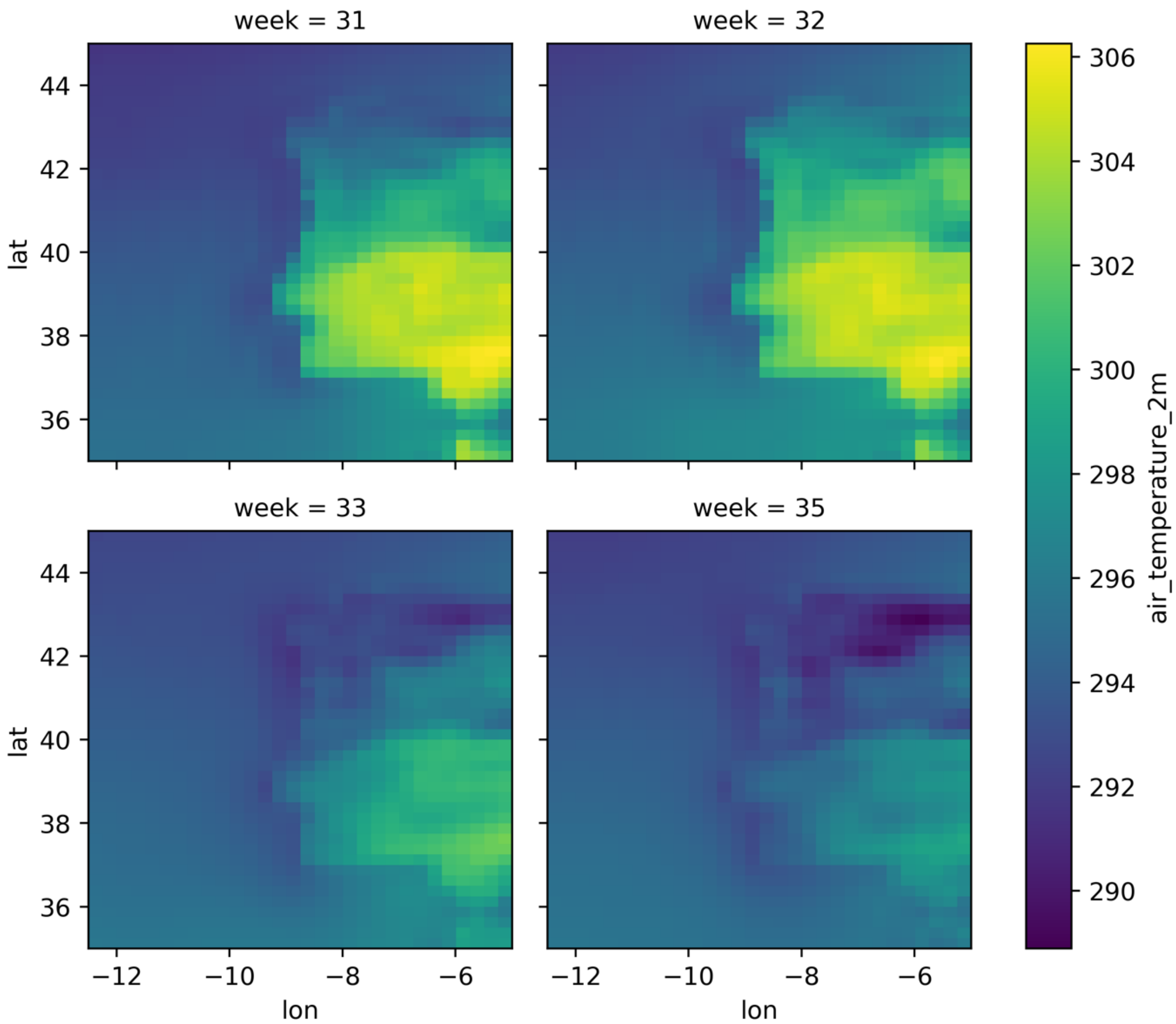

Figure 15 The average of air temperature in study area by week in August

From the temperature Figure 14 and 15 in August, we could find that the first two weeks saw a pretty high temperature over 300 K in some regions which would be the main reason of the fire ignition. Next, we will analyze the spatial distribution of temperature over 300 K. Comparing with the monthly burnt area in August, the burnt area almost is distributed inside the region where the temperature is higher than 300 K at least in one week epoch in Figure 16.

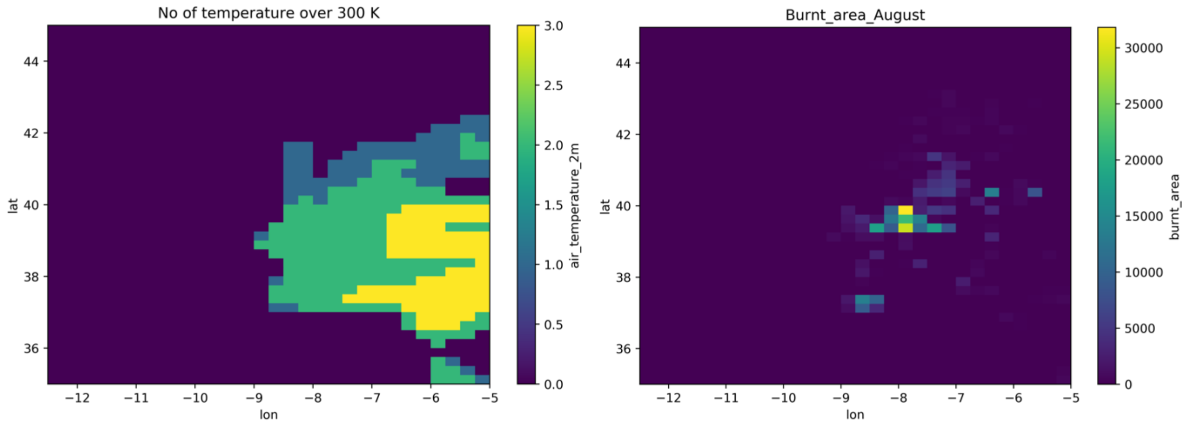

Figure 16 The Number of air temperature over 300 K and burnt area in study area in August

As for the precipitation, few rain makes the vegetation dry and easy to ignite, first, the summation of rain rate (mm/day) is given as follow and then each week will be grouped to deploy the tendency of precipitation. As we can see, the sum of precipitation only have the maximum around 10 mm in whole August as Figure 17 showed.

```
def above_Nsigma(x,Nsigma):
    return xr.ufuncs.fabs(x) > Nsigma

res = Europe.apply(above_Nsigma, Nsigma = 300)
```

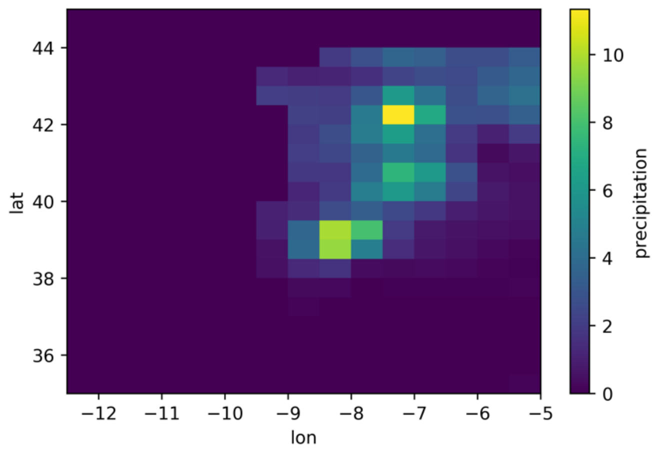

Figure 17 The summation of precipitation in study area in August

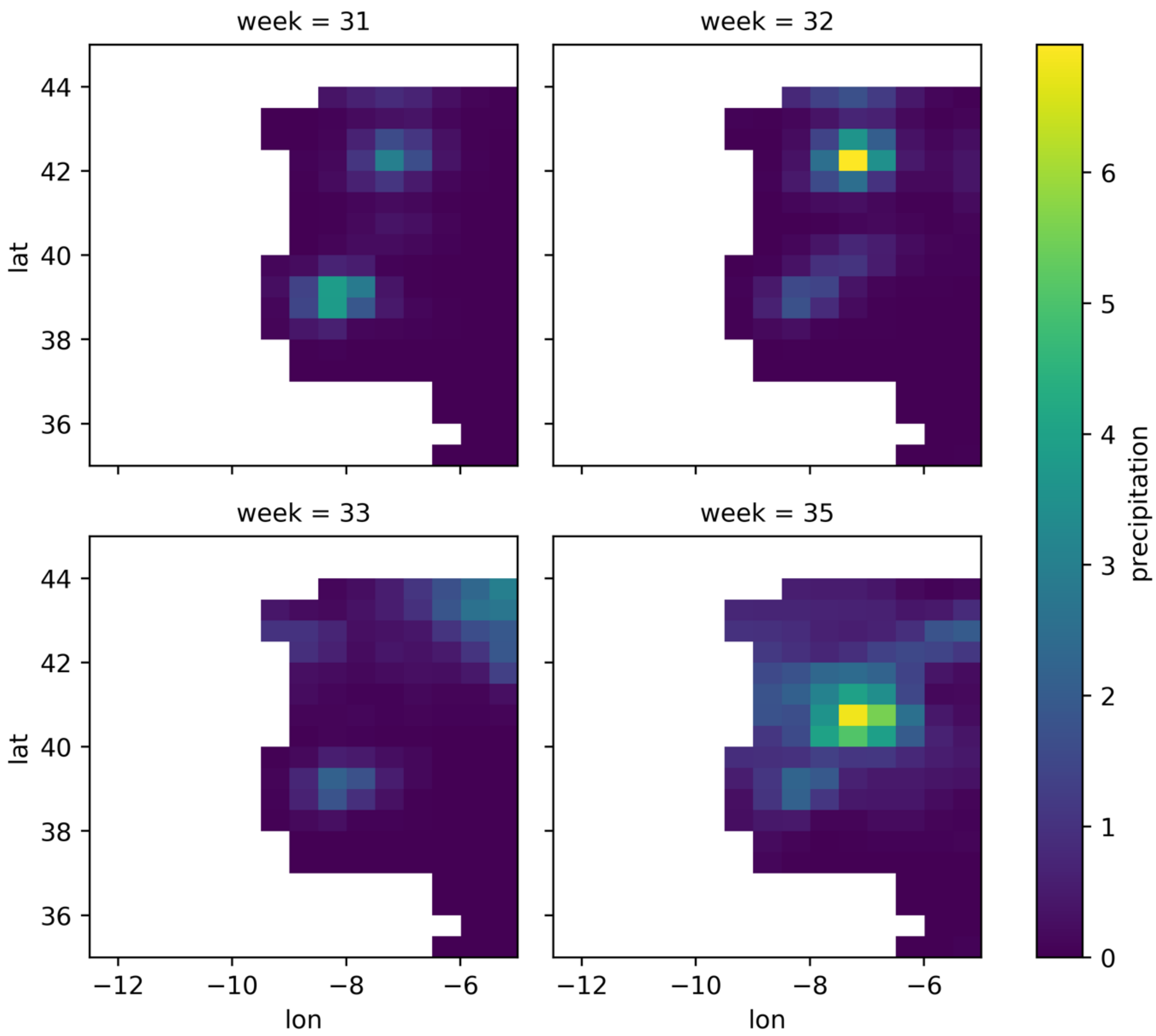

Figure 18 The sum of precipitation in study area by week in August

```
def above_Nsigma(x,Nsigma):
    return xr.ufuncs.fabs(x) < Nsigma

res = Europe.apply(above_Nsigma, Nsigma = 2)
```

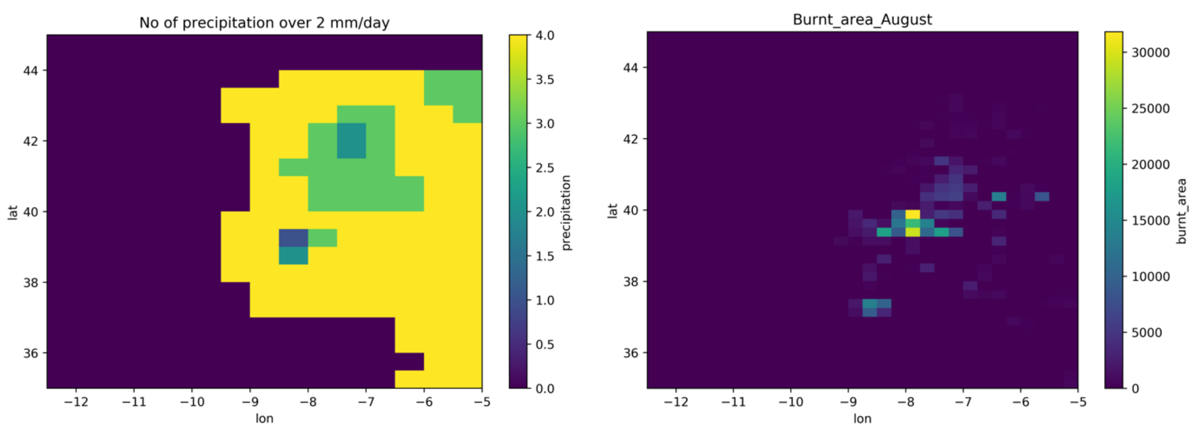

Figure 19 The Number of precipitation below 2 mm/day and burnt area in study area in August

From the Figure 18 and 19, we could conclude that the burnt region saw a drought in August because most of the study area got less than 2 mm precipitation per day and the burnt area is located in the extreme drought area.

4. Discussion

In the help of ESDL data cube, it becomes much easier to reach the climate data derived from satellite data such as GFED, precipitation, temperature and carbon emission. Based on these long-term data, first we search regionally the Europe for the most dramatic burnt area in the history from 1995 to 2014. And then zoom in the vision into specific area located in Portugal as the study area. Next, the 2D spatial visualization is applied to search the study area by Year, Month and Week. The default functions in ESDL makes the spatial analysis by specific grouped time dimension accessible easily. Finally, the study area in specific ignition time are confirmed then used for the next step.

In the second part, multi-variables time series analysis are implemented in order to verify the spatial analysis conclusion and also provide the comparison results across time from multi-source like precipitation, temperature and carbon emission. The relationship between these four variables are strong. Low precipitation and high air temperature are always aligned before and during the wildfire which can be regarded as the main causes of the wildfire. And then the carbon emission also enlarges dramatically during the fire since the biomass fuel consumes the oxygen, and produced mass of carbon dioxide. Based on the time series analysis, we could make it clear about the complex relationship between different impact variables. And finally, we return to the spatial analysis in 2D and draw the distribution of extreme temperature and low precipitation in study area, which could provide an spatial vision for the burnt area with temperature and precipitation. Another reason is that the smaller scale in time with temperature and precipitation by week could further help us to analyze the short-term consequence caused by them, since GFED only provides the monthly burnt area and carbon emission.

In the future, the higher grid resolution will be applied to get more reliable and accurate results. And other fire-related variables also will be taken into account. In addition, the fire model should be applied for numerical system analysis regionally and even globally [4].

5. Acknowledge

Thanks ESDL team for provision of computing server and guidance. And thanks the ESDL Forum for answering the questions about programming. Very grateful to all data owners for kindly providing the data sets free of charge from the owners, the ESDL team or ESA.

6. Reference

[1] https://en.wikipedia.org/wiki/2018_European_heat_wave
[2] "French soldiers land in Sweden to battle wildfire inferno". thelocal.se. 23 July 2018. Retrieved 25 July 2018.
[3] ESDC webpage: http://www.earthsystemdatacube.net/
[4] Krasovskii, A., Khabarov, N., Pirker, J., Kraxner, F., Yowargana, P., Schepaschenko, D., … Obersteiner, M. (2018). Modeling burned areas in Indonesia: The FLAM approach. *Forests*, *9*(7), 437. https://doi.org/10.3390/f9070437