The association between precipitation and the spread of shiga toxin-producing Escherichia coli infection in Lombardy, North of Italy

A. Novaro
A.Baronetti
G. Ardissino
F. Acquaotta

The disease

The HUS (Hemolytic-Uremic Syndrome) is a complication of an infection by shiga toxin produced by *Escherichia coli* (STEC) infection.

- affect children
- fairly **rare**
- endemic in central Europe and North America
- usually **not epidemic**
- **vector unknown**
- cases appear in **time related clusters** -> can suggests the role of the weather conditions as a hot gun

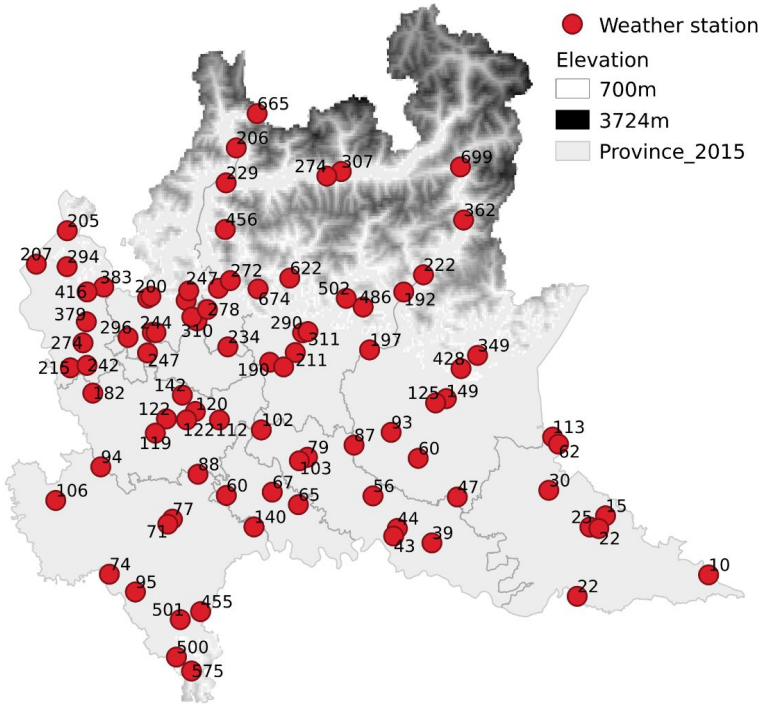
Study area: Lombardy (Italy)

Population density: 418.3 ab./km^2

Population under 18: over 1.8 million

Home of the national HUS study centre: Centro SEU del Policlinico di Milano

Elevation max: 700 m



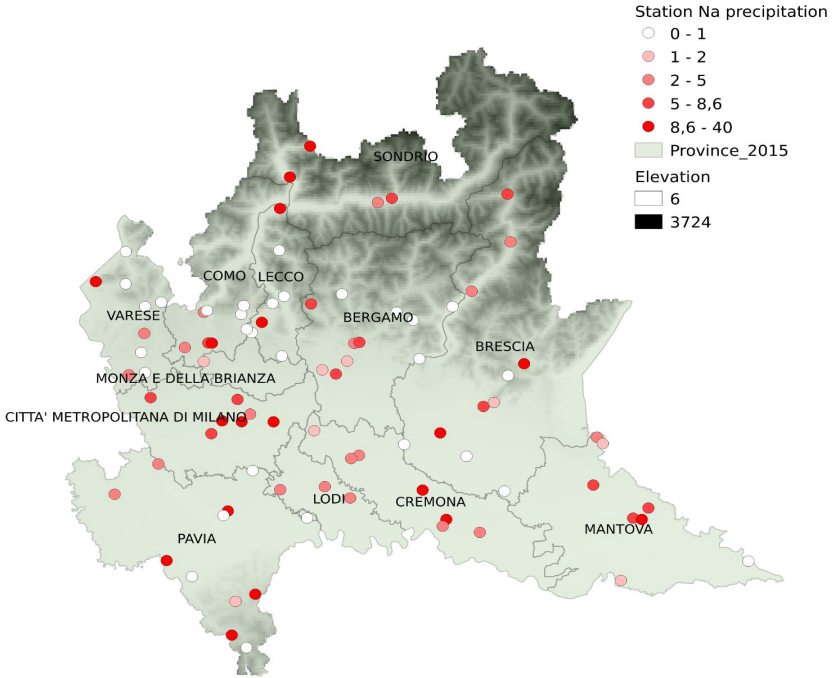
Climate database

Weather station: **88** from ARPA Lombardy, under 700 m altitude

daily rainfall sum

time series: **2010 -2021**

high quality: 93% annually



HUS cases

Dataset: 4113 test-> **249 positive**

3864 negative

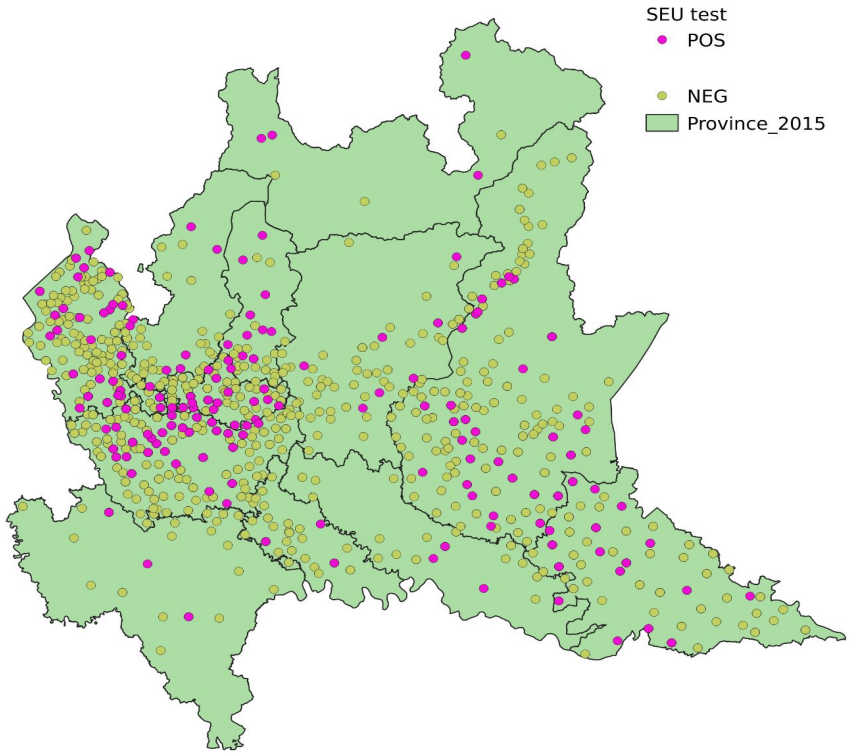
location: Lombardy

time: **2010:2020**

age: <19 years old

Date of positive test is set 3 days before (incubation period)

Analyzed distribution: Annual, Monthly and Weekly (“selected weeks”>3 cases)



Spatial distribution of SEU (*Biggeri et al*)

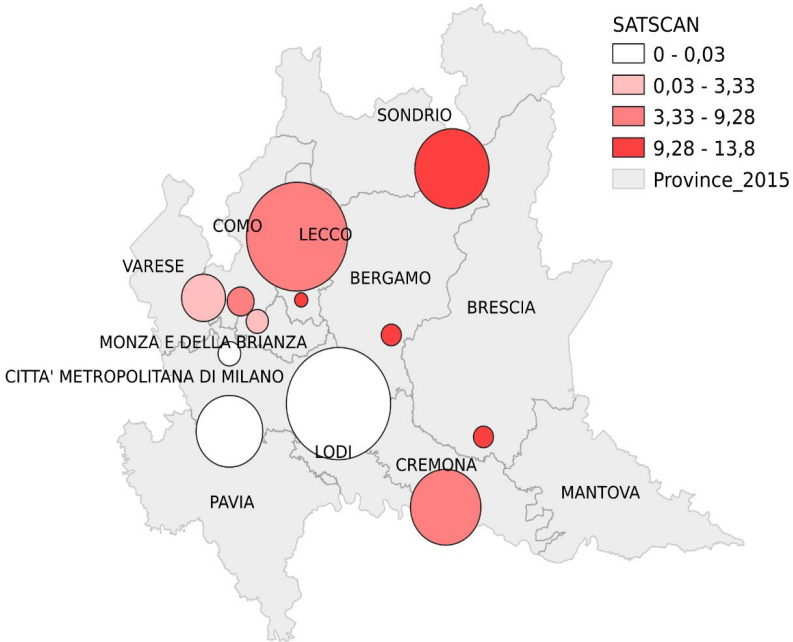
Spatial Scan Statistic (SatScan) identify the **spatial aggregation** to localize high frequency areas. Probability is calculated by a MonteCarlo simulation conducted assigning cases in the region due to population density.

radius: 20 km

seasonality correction: April-October

high-risk zone –> red

low-risk zone –> white



Methodology: rainfall

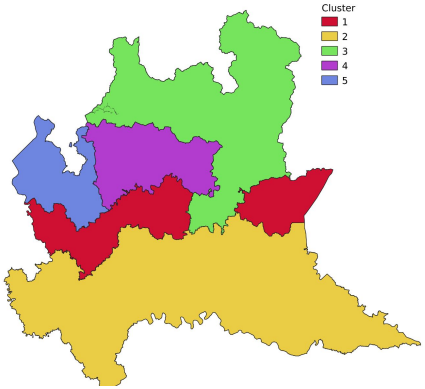
Cluster analysis (top-down, method *ward.d2*) on the monthly rainfall sum

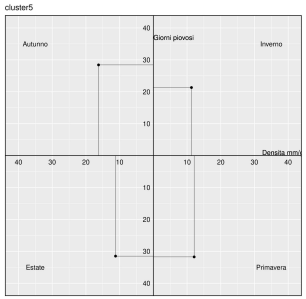
Rappresentative series for each cluster

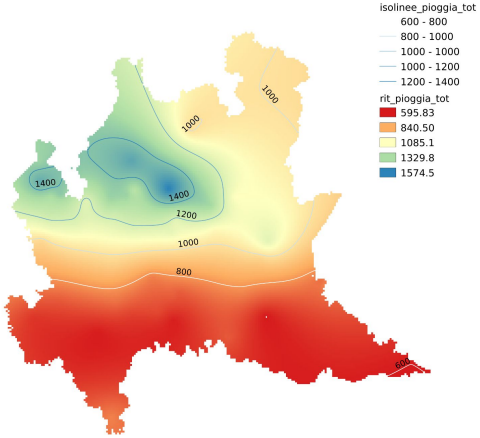
Balseinte polygons

Simple kriging

Precipitation **indices**







Indices used for precipitation pattern & analysis

Code	Name	Index	Unit
CDD	Consecutive Dry Days	Maximum number of consecutive dry days (when PR < 1.0 mm)	days
R20mm/ R10mm	Number of very rain days	Number of days when PR >= 20 mm & PR >= 10 mm	days
R95p & R99p	Total PR from heavy rain days & from very heavy rain days	Annual sum of daily PR > 95th percentile & Annual sum of daily PR > 99th percentile	mm
Rx1day & Rx5day	Max 1-day PR & Max 5-day PR	Maximum 1-day PR total & Maximum 5-day PR total	mm
q05, q10, q50, q90, q95, q99	Percentiles	percentiles of daily sum rainfall distribution	mm

Methodology: disease & rainfall

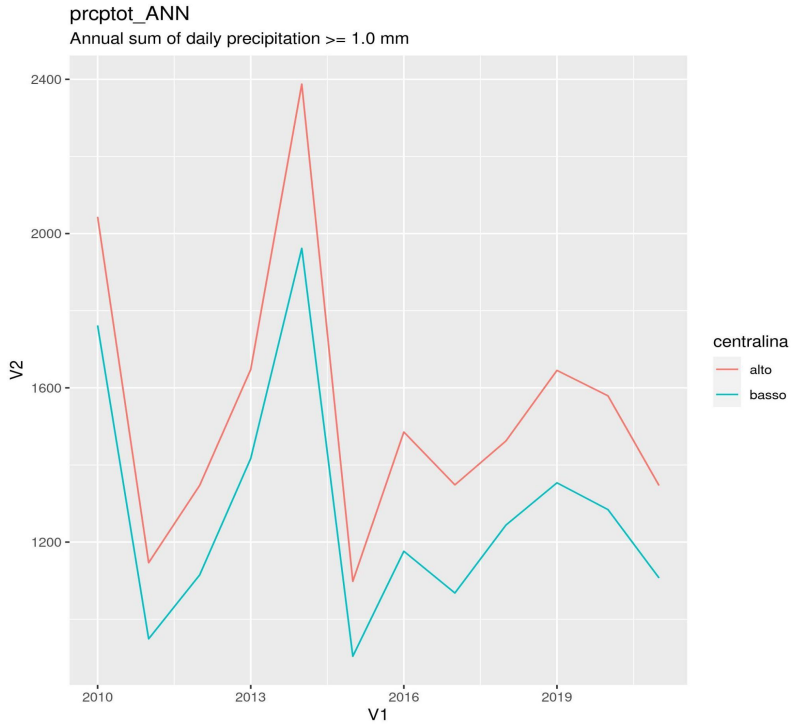
Identificate weather stations in **high risk zone** and **low risk zone**

Create two type series throughout the mean of daily value

Indices **plotting** and **comparison** (Wilcoxon Mann Whitney test)

Plotting “selected weeks” cases

N. of days between “selected weeks” cases and precipitation over a threshold



Rainfall

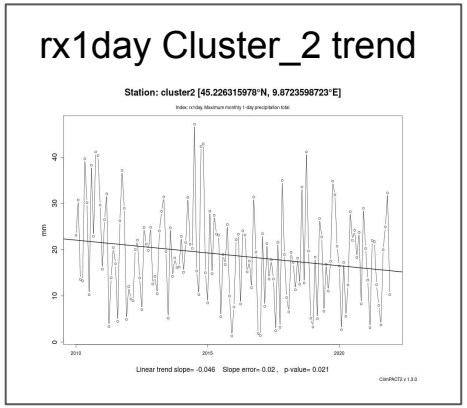
Annual rainfall sum descendent in NW-SE direction

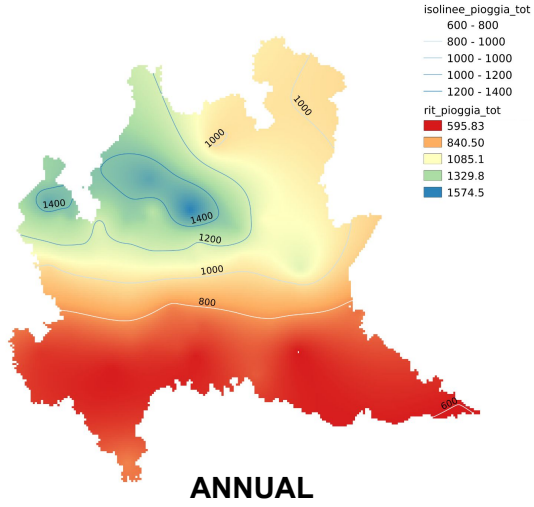
minimum in **winter**

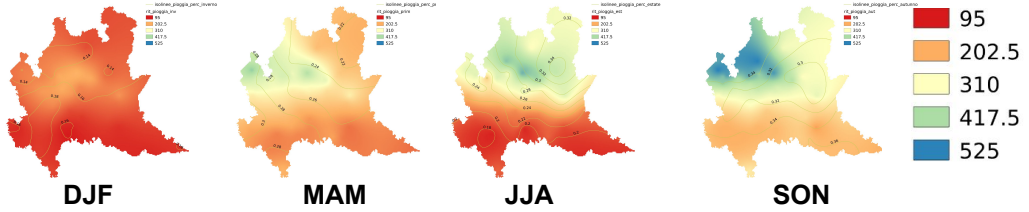
maximum in **autumn**

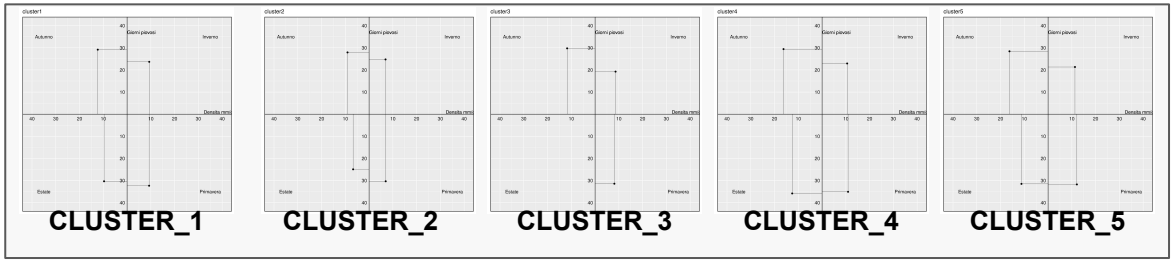
secondary peak in summer (NE) or spring (S and NW)

Descendent trends for: maximum 1-5 day precipitation amount of rainfall from very wet days (only in the southern cluster)







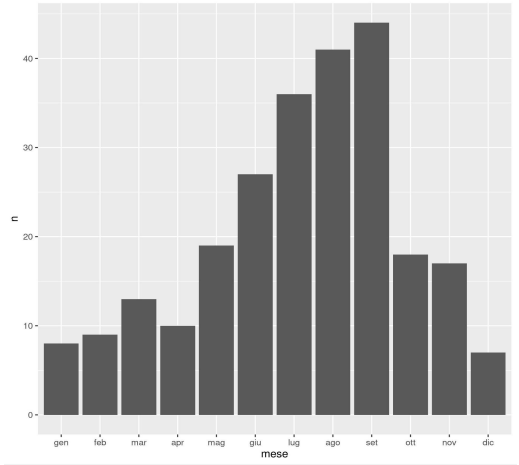


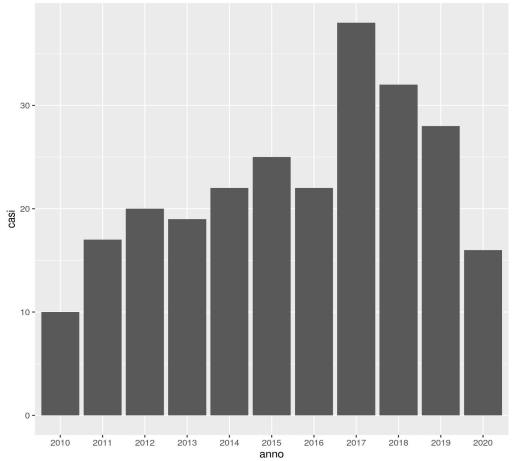
Disease distribution

Most of positive cases are on **september** (44 cases), august (41), july (36) and june (27).

It suggests a **seasonal nature**

Most of cases are in **2017** (38), **2018** (32), **2019** (28), **2015** (25)

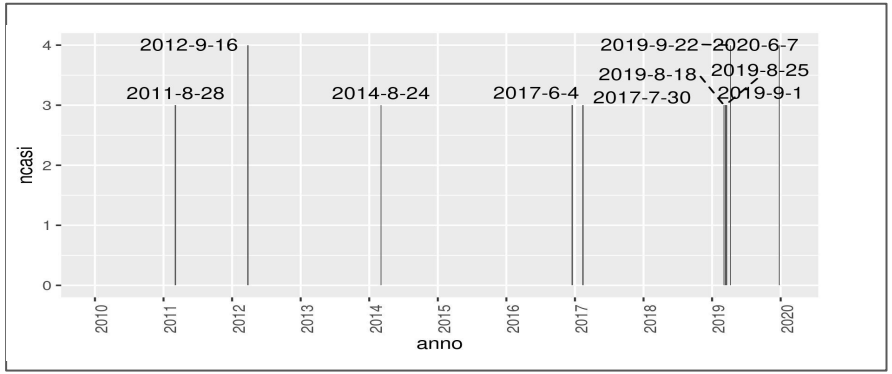


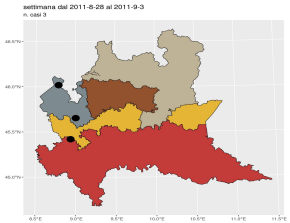


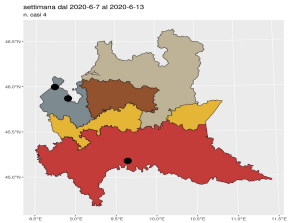
Weekly distribution

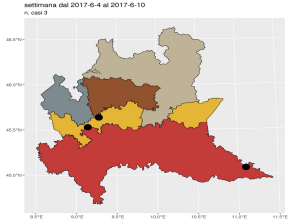
3 or more positive tests in a week -> 10 cases

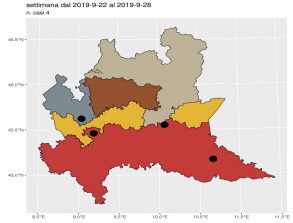
Maps suggest that they are **not space related**

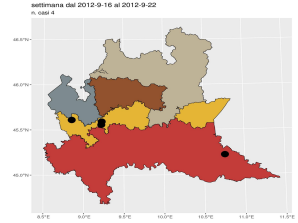


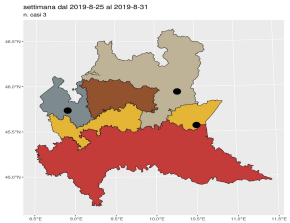


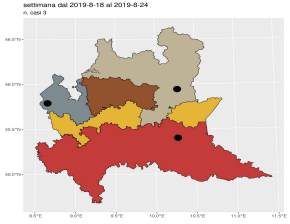


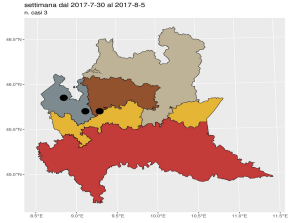


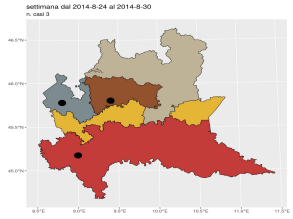


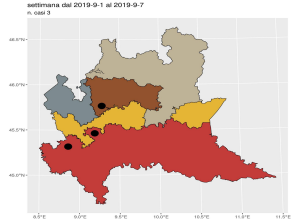








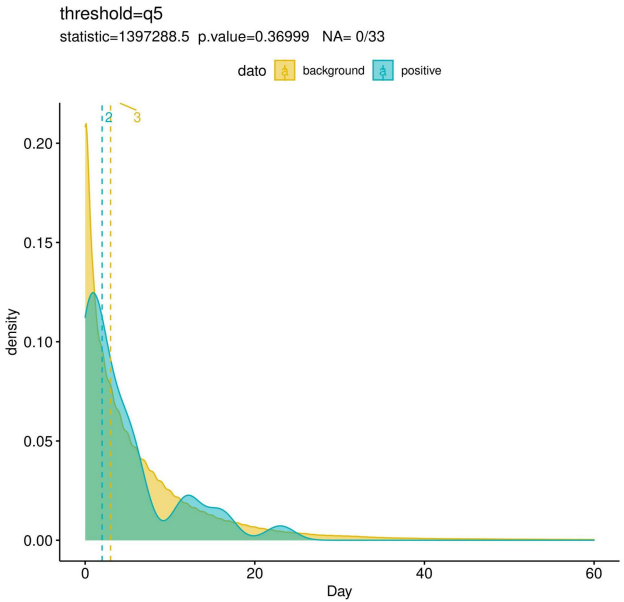


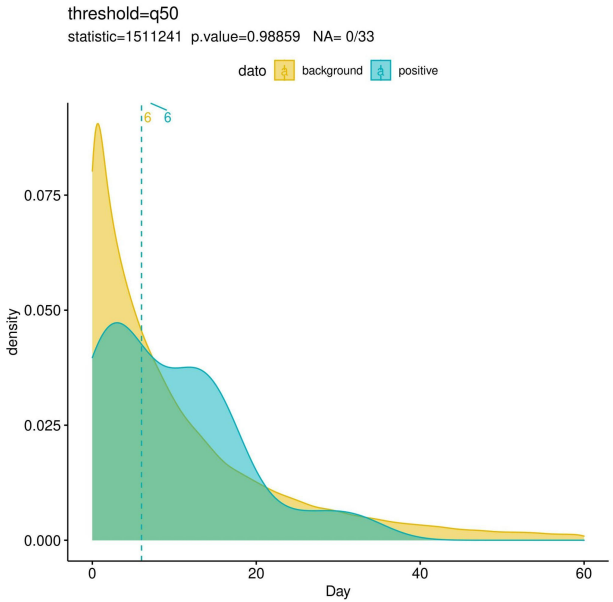


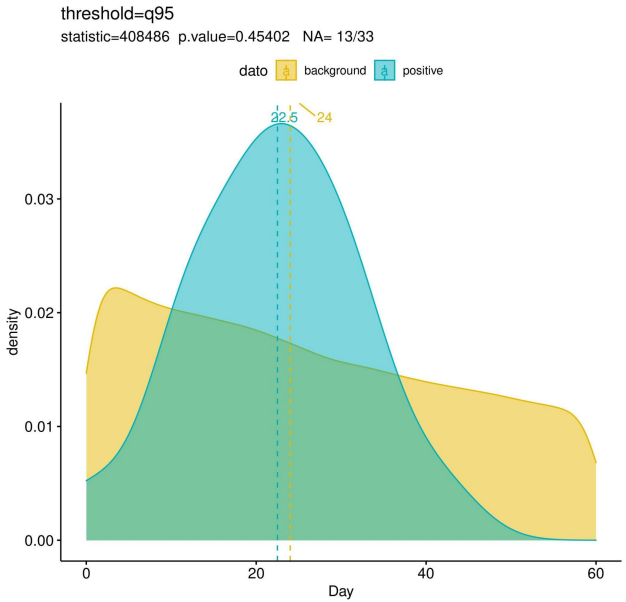
Lags between days of “selected weeks” and rainfall over thresholds

N. days between positive cases day and daily precipitation > threshold (percentile: **5**, 10, **50**, 90, **95**, 99)

No statistical differences between control and positive tests



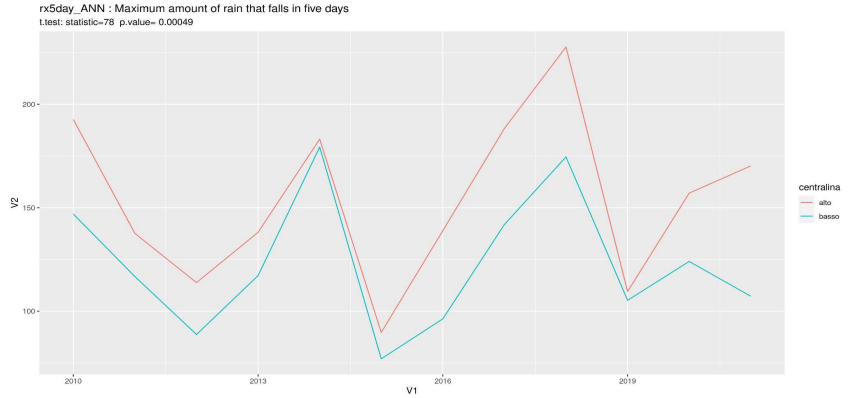


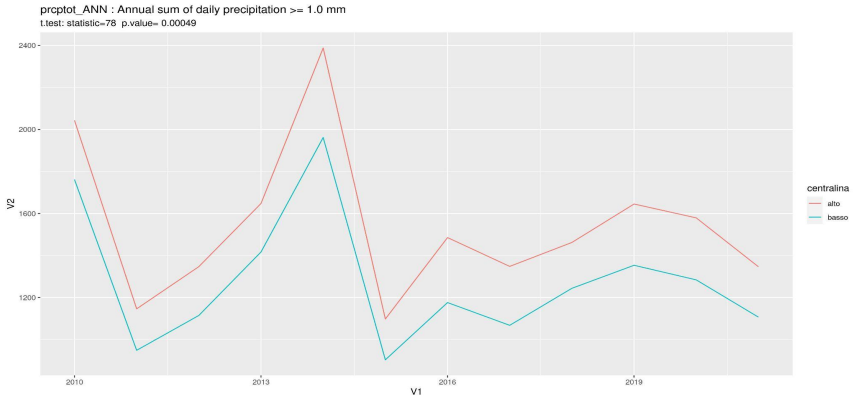


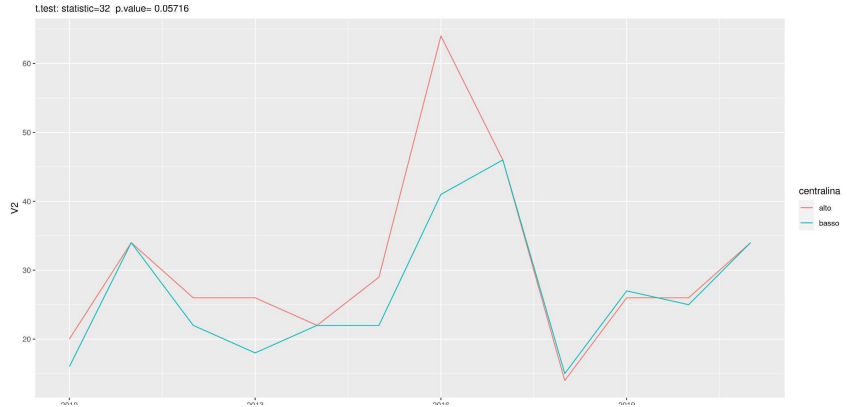
High risk and low risk zone analysis

High risk zone shows higher value in: **prcptot,** r10mm, r20mm, **r95p**, **rx1day**, rx5day.

No statistical differences: CDD, r99p







Conclusion

- Lombardy **precipitation pattern is not uniform**, it presents a first peak in autumn and a second in spring or summer
- **Five homogeneous areas** have been identified
- **Minimal or no trend emerge**, otherwise there is a tendency to registrate less rainfall annual sum, less wet days and less very rain days but without a statistical confirmation
- Spread of disease show different value in the whole year (**seasonality**): especially in august, july and june with the max value in **september**
- **Time related cluster** (weekly scale), 10 cluster with 3 or 4 cases each
- **Different precipitation pattern over high risk zone and low risk zone**: high risk zone show up a higher annual sum precipitation and higher value in intense rainfall indices
- **No link between low/normal/heavy daily rain and positive tests**

Thank you for your attention

Andrea Novaro
andrea.novaro@edu.unito.it