



USING TEXT MINING TO ASSESS THE IMPACT OF EXTREME EVENTS ON BUSINESSES : TWO USECASES

STEPHANIE HIMPENS & EDITH STOJANOVIC
BUSINESS SURVEYS UNIT

CENTRAL BANK BUSINESS SURVEYS CONFERENCE

ROMA 2024





SUMMARY

- 1 AN INTRODUCTION TO BANQUE DE FRANCE'S MONTHLY BUSINESS SURVEY
- 2 USE CASE 1 : IMPACT OF STRIKES ON BUSINESSES' ACTIVITY
- 3 USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME
 - A. INITIAL WORK OF NOVEMBER 2023 : SEVERE WEATHER EVENTS BASED ON KEYWORDS SEARCH
 - B. TEXT MINING METHODS TO EXPAND KEYWORDS SEARCH
- 4 CONCLUSION

1

AN INTRODUCTION TO BANQUE DE FRANCE'S MONTHLY BUSINESS SURVEY

1. AN INTRODUCTION TO BANQUE DE FRANCE'S MONTHLY BUSINESS SURVEY

Survey sample

- Scope: manufacturing sector, market services, construction sector
- 8 500 firms, divided in 13 regional samples
- Participation of interviewees is voluntary

Questionnaire

- Semi-directed phone interview led by Banque de France's interviewer
- Afterwards coded into opinion marks...
 - **7 levels ordered by positivity (scaled from -200 to 200)**
 - **Both level and evolution (wrt last month), past and forward looking**
- ... supplemented with a free, non compulsory, textual field allowing to check the concordance of certain items declared in a company's interview

1. AN INTRODUCTION TO BANQUE DE FRANCE'S MONTHLY BUSINESS SURVEY

Example of additionnal textual responses :

“ After a very good August, business in September was weak, as customers were uncertain and had no visibility. Inventories were therefore particularly high in September (+800 KE in value), as production had anticipated a normal month. Resin purchase prices are stable overall (very slightly down) after the increases seen at the beginning of the year. Sales prices are indexed to the resin price. The energy contract is a summer/winter contract. ”

“ The company has seen a few decreases in the price of materials, but is maintaining its selling prices. Customer payment terms are deteriorating, a sign that the sector is suffering. It is worth noting the development of the rental business to the detriment of the sales business.”

Over the last 5 years, we have developed some tools to make better use of those comments in order to :

- Quantify *ad hoc* events
- Produce counterfactual of the activity
- Analyze the spread of extreme events through the economy

2

USE CASE 1 : IMPACT OF STRIKES ON BUSINESSES' ACTIVITY

2. USE CASE 1 : IMPACT OF STRIKES ON BUSINESSES' ACTIVITY

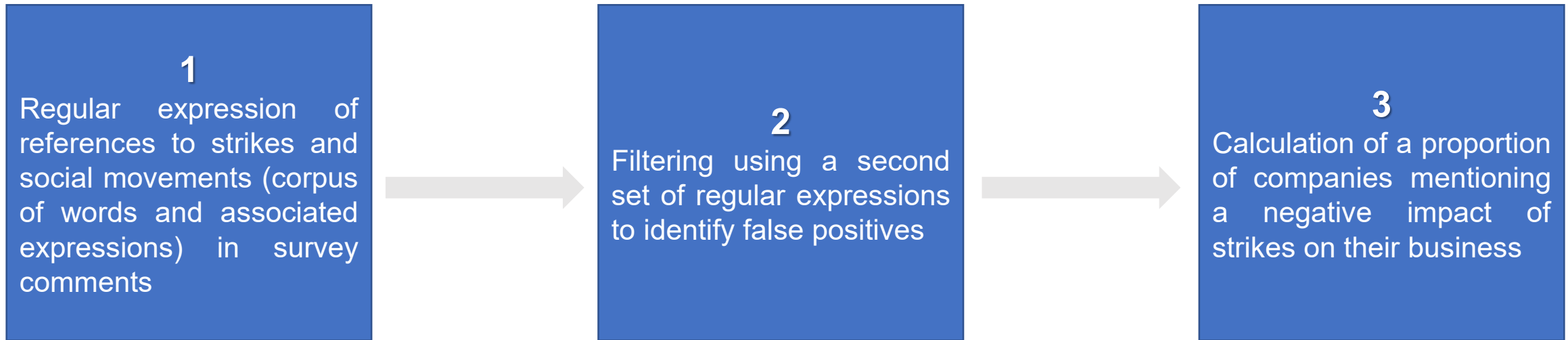


February and March 2023 were marked by numerous strikes and social movements linked to the pension system reform.

- Is this phenomenon visible in the monthly business survey?
- Is it possible to put a figure on the impact of strikes based on observation of monthly business survey data?

2. USE CASE 1 : IMPACT OF STRIKES ON BUSINESSES' ACTIVITY

In order to assess the impact of strikes :

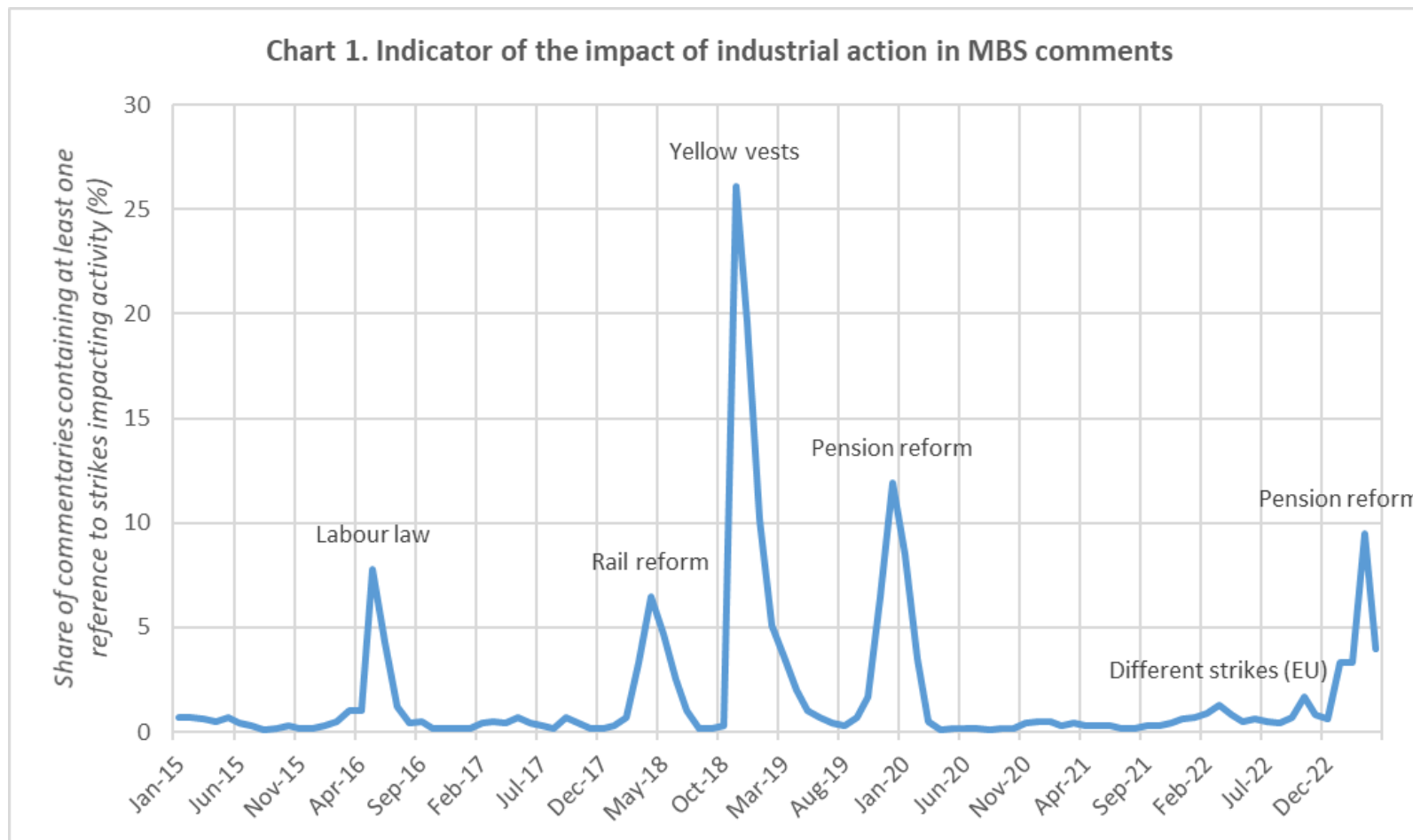


Construction of a dummy variable :

- 1 = observations which comment has a phrase that mentions confirmed an impact of strike on the activity
- 0 = all other observations with a filled comment

2. USE CASE 1 : IMPACT OF STRIKES ON BUSINESSES' ACTIVITY

For March 2023, 8.9% of companies in industry, 12.8% of companies in market services, and 2.3% of companies in construction mentioned strikes in their business survey : **9.5% of companies in total** (unweighted).



2. USE CASE 1 : IMPACT OF STRIKES ON BUSINESSES' ACTIVITY

Based on the proportions established, we calculated a counterfactual of the balances of opinion on activity, excluding companies that mentioned an effect of strikes in their comments: i.e. attributing to companies reporting a negative effect of strikes on their business the average balance observed within their sector of activity.

Sector	Proportion of companies whose test does not mention strikes	Proportion of companies whose test mentions strikes	Balance of production calculated for companies not reporting strikes	Survey balance	Difference
Chemical industry	83%	17%	19,8	7,0	12,8
Pharmaceutical industry	87%	13%	21,8	19,3	2,5
Rubber, plastic, and other products	93%	7%	32,9	31,7	1,2
Metallurgy	91%	9%	25,9	22,9	3,0
Other Manufacturing	94%	6%	23,4	21,8	1,6
Agricultural and food industry	91%	9%	16,6	12,5	4,1
Electrical and electronic equipment	88%	12%	32,4	31,1	1,2
Total Industry	91%	9%	23,5	21,0	2,5
Car repair	97%	3%	31,2	30,1	1,1
Transportation and storage	78%	22%	0,4	-4,9	5,3
Other services	95%	5%	34,7	33,6	1,1
Total Services	87%	13%	25,2	24,2	1,0
Total Construction	98%	2%	16,9	17,2	-0,3

2. USE CASE 1 : IMPACT OF STRIKES ON BUSINESSES' ACTIVITY

- Thus, we note a negative effect of the mention of strikes on EMC activity balances, albeit in moderate proportions.
- Covid point of comparison: -111 points for industry, -123 for market services.
- Calculation repeated for April 2023: the effects on balances largely dissipate, in line with a much lower proportion of strikes mentioned.
- Estimation with the growth prediction team* : overall, the social unrest of March and April 2023 would have reduced GDP growth by around -0.06 pp in Q1, but would have been followed by a catch-up effect of around +0.06 pp in Q2, with the rebound effect in April and May offsetting the drop in activity observed in March.

* : details in [*Measuring the impact of industrial action on activity*](#), Coueffe, Jousselin, Penalva & Stojanovic, Dec. 2023

2. USE CASE 1 : IMPACT OF STRIKES ON BUSINESSES' ACTIVITY

Some methodological issues:

1. **Sample size for recalculating activity balances:** the new balances are calculated on a slightly smaller sample (10% fewer respondents).
2. **Representativeness of the new sample:** composition similar to the previous sample, and behavior of the two groups (mention of strike yes/no) is substantially similar during 2022 on questions relating to production or activity : the two groups have similar characteristics in terms of sectors, sizes and response behavior.
3. **Method is dependent on Banque de France interviewer behavior:** we believe that this bias is controlled by the previous point and the large number of testers spread across the country. However, this point is regularly monitored by our Unit via the monthly instructions distributed to the network for the benefit of interviewers, as well as yearly training programs.

2. USE CASE 1 : IMPACT OF STRIKES ON BUSINESSES' ACTIVITY

Key take-aways at this point:

- Our textual data can be used to match some punctual events : strikes or other events can be monitored in an *ad hoc* way (e.g. Olympics in Paris).
- We create some tools to derive the effect of those events on business activity and growth.
- National coverage of Banque de France interviewers + sample size can alleviate some methodological issues arising.

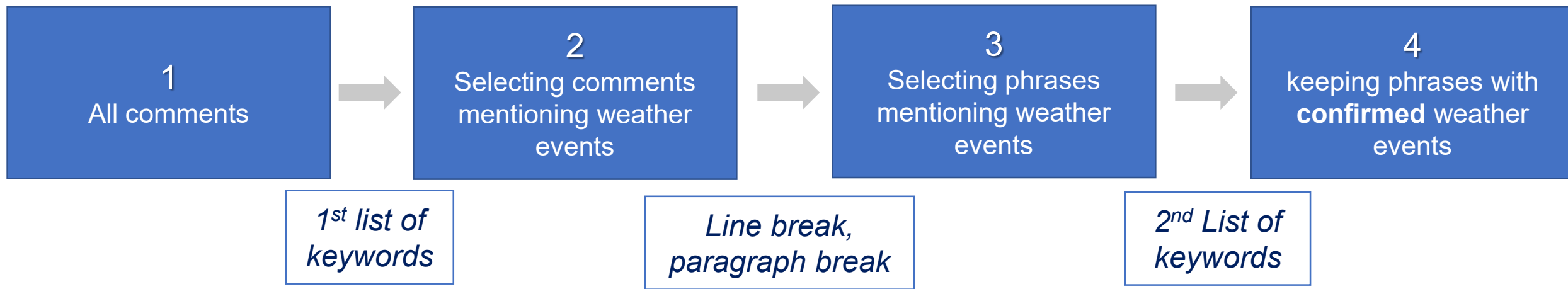
Moving forward, we now try to improve and use this method with different kind of events:

- Some events are not binary (impacted/not impacted)
- Some events are linked to several and different corpus of vocabulary
- How to take into account more complex, changing through times events?

3

USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME

2. USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME



Construction of a dummy variable

- 1 = Observations which comment has a phrase that mentions confirmed weather events
- 0 = All other observations with a filled comment

2. USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME

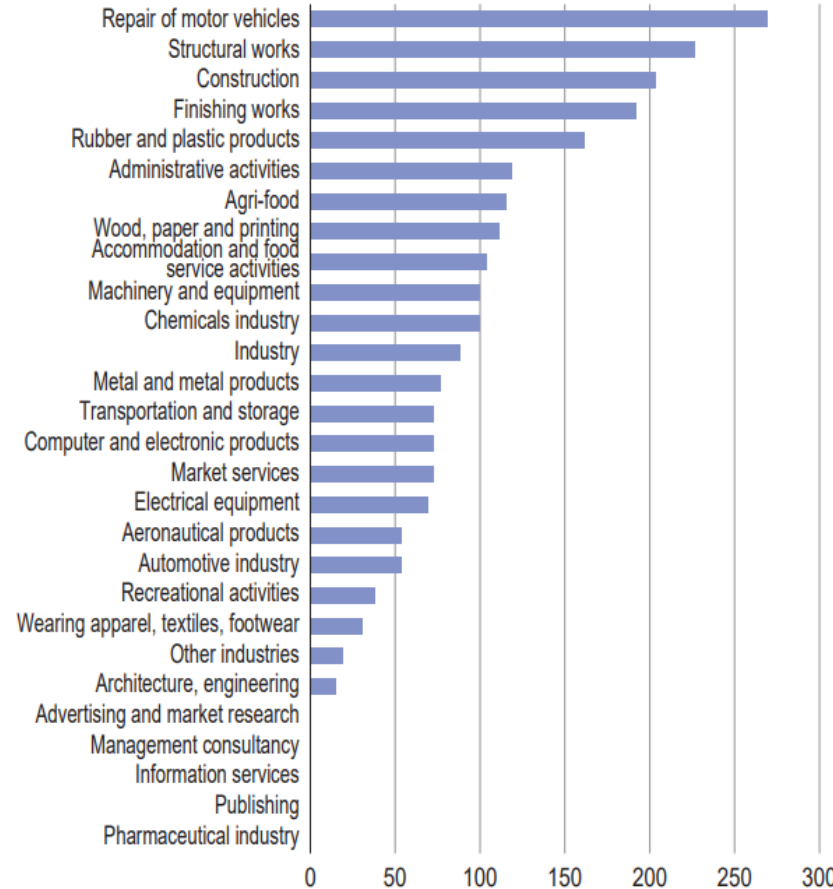
A. INITIAL WORK OF NOVEMBER 2023 : SEVERE WEATHER EVENTS BASED ON KEYWORDS SEARCH

Number of events relative to a base of 100 across all sectors.

For example : the proportion of comments mentioning weather events in construction was almost twice that of all the economy.

INCLUDED IN THE COMMENTS

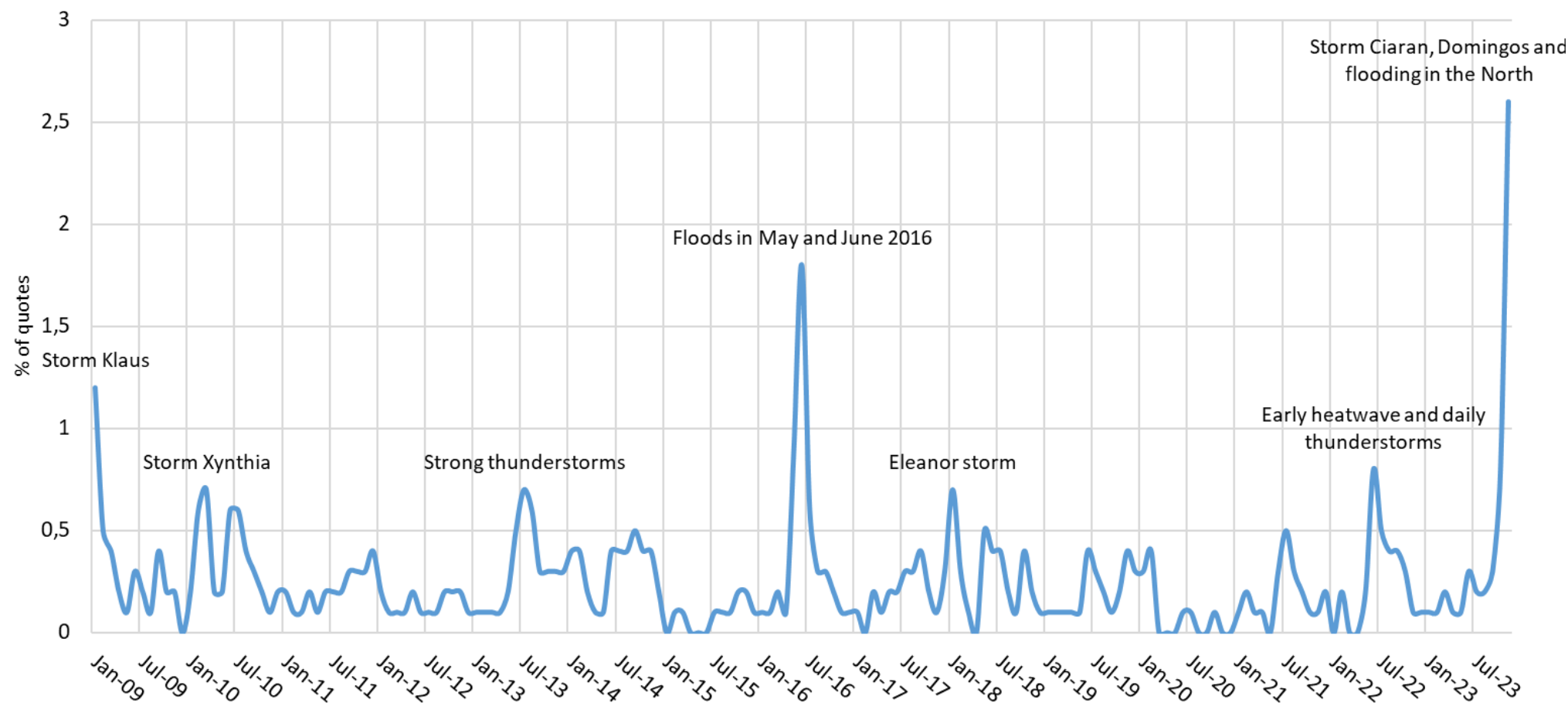
(total = 100)



Monthly Business Survey –
Start of December 2023 |
Banque de France (banque-
france.fr)

2. USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME

A. INITIAL WORK OF NOVEMBER 2023 : SEVERE WEATHER EVENTS BASED ON KEYWORDS SEARCH



2. USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME

B. TEXT MINING METHODS TO EXPAND KEYWORDS SEARCH

External sources

- Find word in weather reports using word counts
- Uses of lexicons of weather terms available on the web

Based on survey comments

- On a sample of comments containing at least one weather event keyword
- bigram search (succession of 2 or 3 words)
- Using semantic proximities (word2vec embedding)

2. USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME

B. TEXT MINING METHODS TO EXPAND KEYWORDS SEARCH

Search successions of 2 or 3 words to qualify weather events

rank	successions of words	Frequency	RAKE score
1	high added value	11	4,2
	poor weather		
2	conditions	7	4
	poor climatic		
3	condition	8	3,9
	unfavorable weather		
4	conditions	3	3,9
5	low added value	3	3,9
	unfavorable weather		
6	conditions	3	3,8
	photovoltaic panel		
7	installed	9	3,6
	small multi-brand		
8	garage	3	3,6
	large geographical area		
9	large geographical area	3	3,6
	favorable climatic		
10	conditions	8	3,4

RAKE score

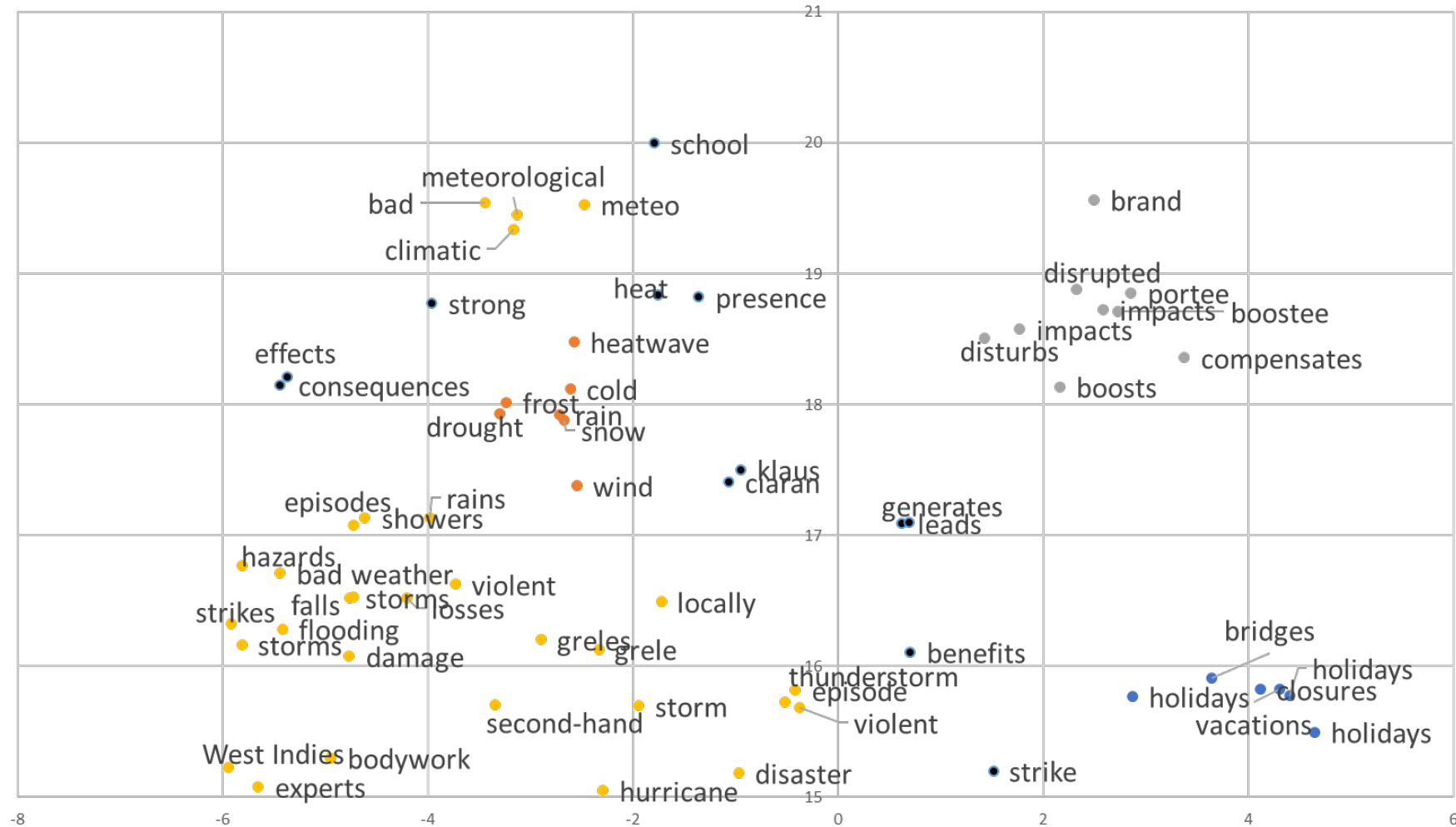
$$= \frac{\text{Frequency of the word sequence}}{\text{Frequency of the first word}} + \frac{\text{Frequency of the word sequence}}{\text{Frequency of the second word}}$$

Calculation were made on a sample where comments previously detected are over-represented.

It is also possible to search for bigramm where a specific word is present (such as meteo or climatic).

2. USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME

B. TEXT MINING METHODS TO EXPAND KEYWORDS SEARCH

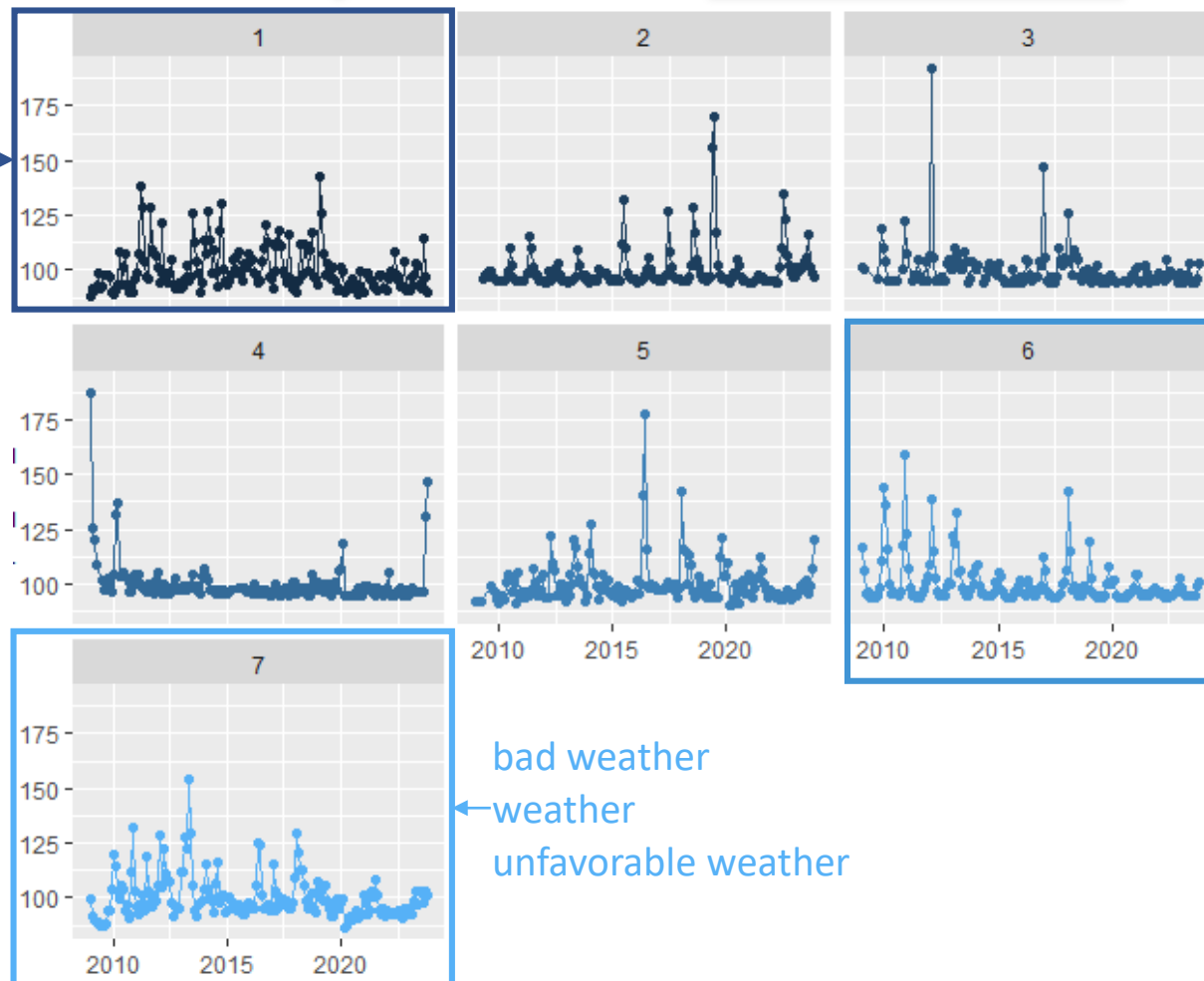


2. USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME

B. TEXT MINING METHODS TO EXPAND KEYWORDS SEARCH

Common vocabulary to describe an everyday weather situations

fine weather
sunny
exceptional
weather
favorable weather



inclement weather
snow

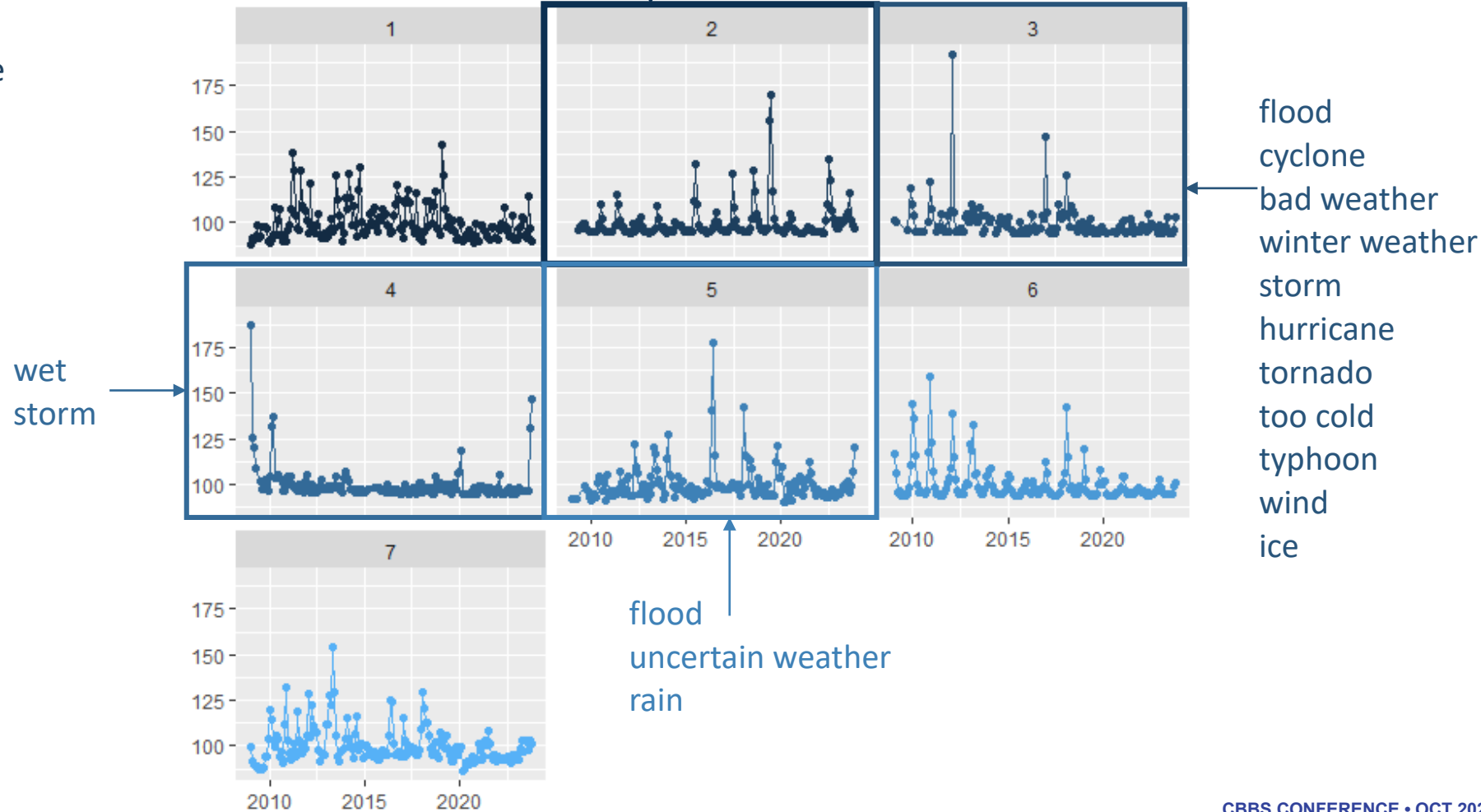
bad weather
weather
unfavorable weather

2. USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME

B. TEXT MINING METHODS TO EXPAND KEYWORDS SEARCH

Common vocabulary to describe an extreme weather situations

heatwave
hail
drought
too hot

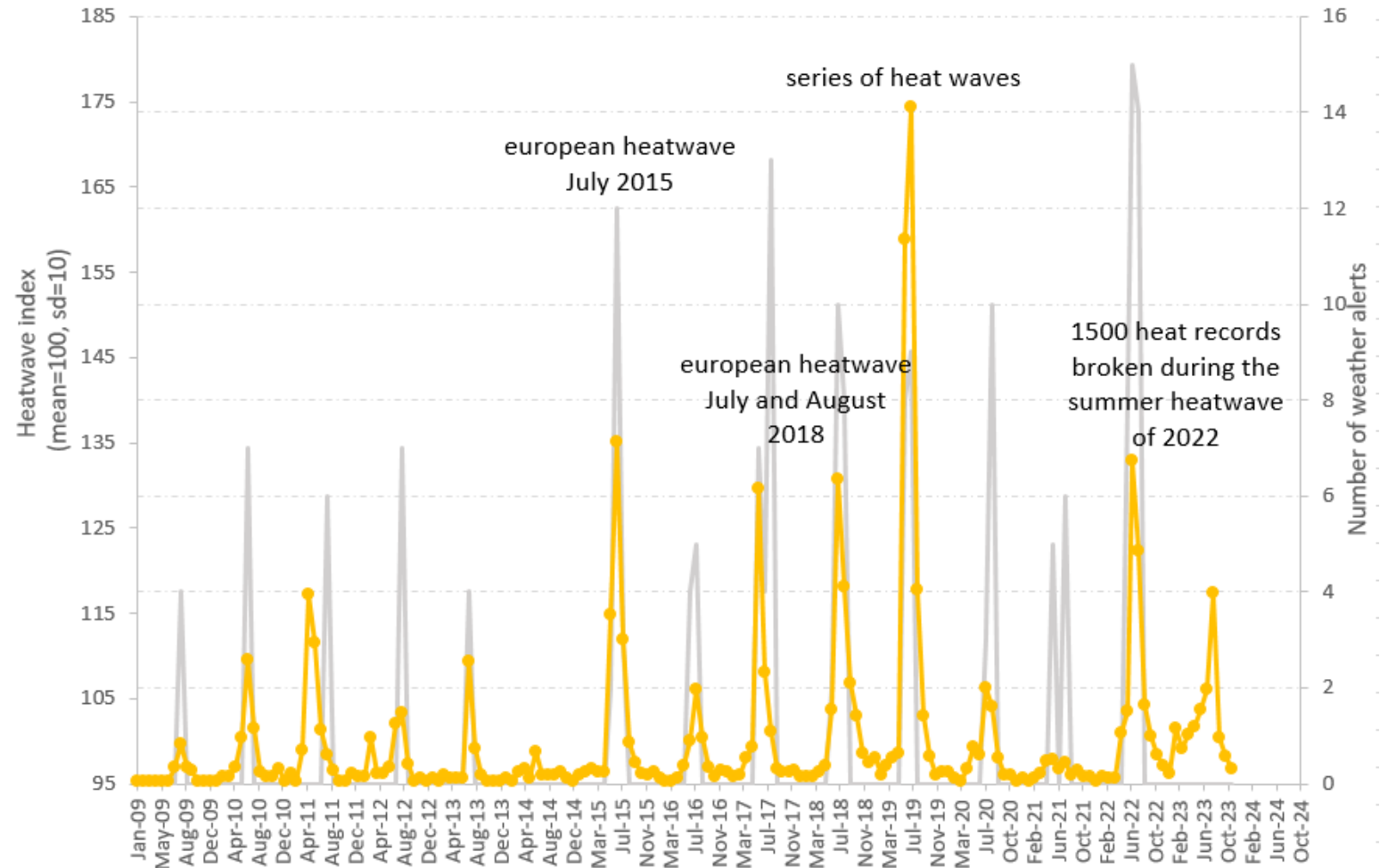


2. USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME

B. TEXT MINING METHODS TO EXPAND KEYWORDS SEARCH

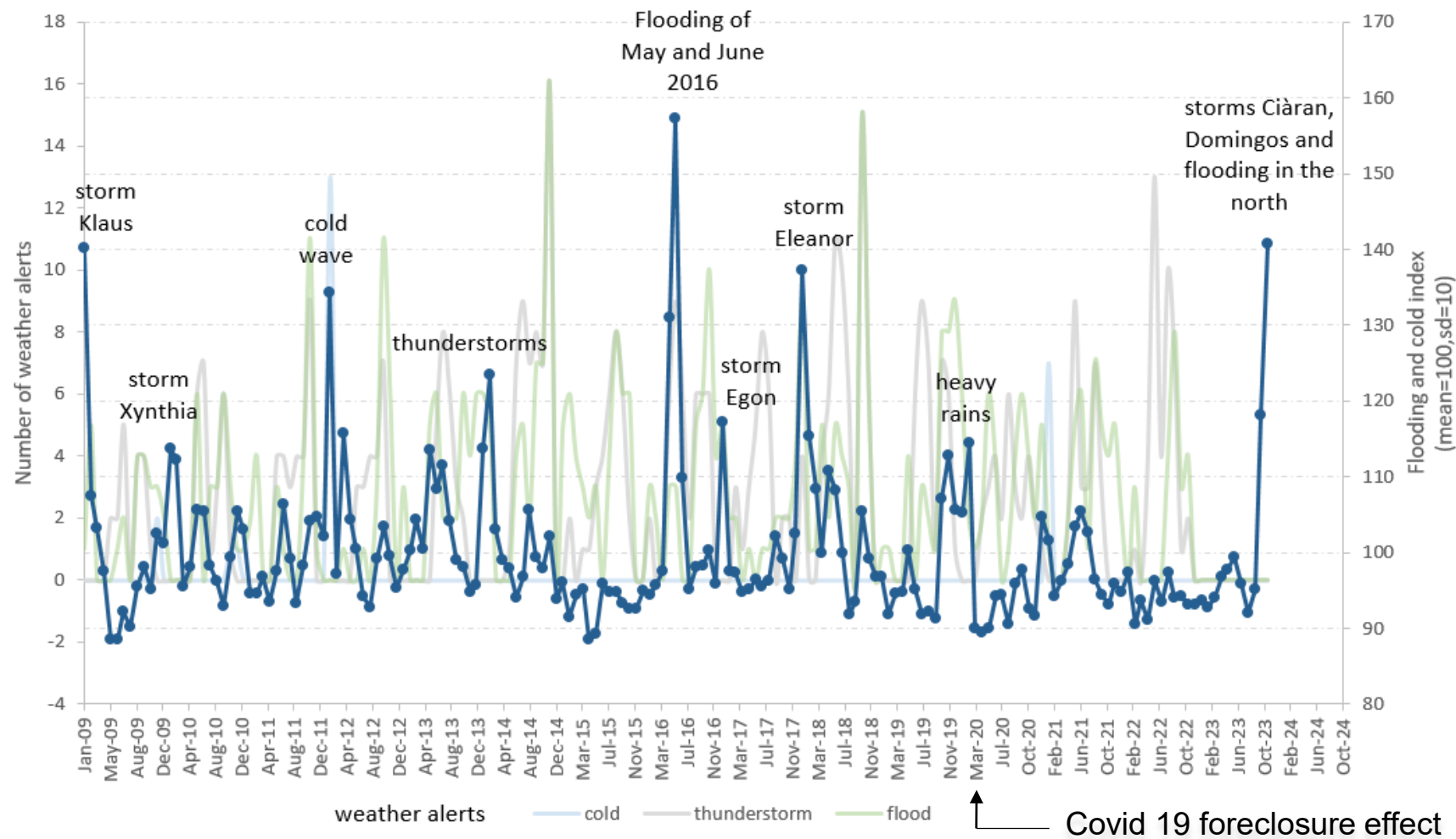
Heatwave index
(mean=100,sd=10)
from word counts

Number of weather
alerts



2. USE CASE 2 : ASSESSING THE IMPACT OF SEVERE WEATHER EVENTS ON BUSINESSES OVERTIME

B. TEXT MINING METHODS TO EXPAND KEYWORDS SEARCH



CONCLUSION

1. POSITIVE ASPECTS

Three conclusions can be drawn from this work (still in progress).

- The monthly business survey is sensitive to weather events (extreme events as well as casual events), through the comments made by companies.
- It is possible to establish categories of weather events.
- The computed series are consistent with data from other sources (such as weather alerts from the national weather agency).

CONCLUSION

2. CHALLENGES

Some challenges are identified moving forward :

- lot of hand work to set up a list of weather keywords,
- non-binary events and evolving effects are harder to take into account,
- some difficulties to make a link with production data,
- the number of comments with keyword is sometimes very low (comments are optionnal),
- hypothesis of stability of response behaviour over time (changing instructions and interviewers, eviction effects).

