



# **Du local au global aborder le problème des échelles et des données multidimensionnelles** *(Illustration à travers l'exemple des flux RSS des journaux)*

GRASLAND C.  
Université Paris Diderot

« Les données de l'Anthropocène : disponibilité, interopérabilité, visualisation - Journée d'étude du 7 mars 2014 »



**1ere Partie : Quelles données ?**

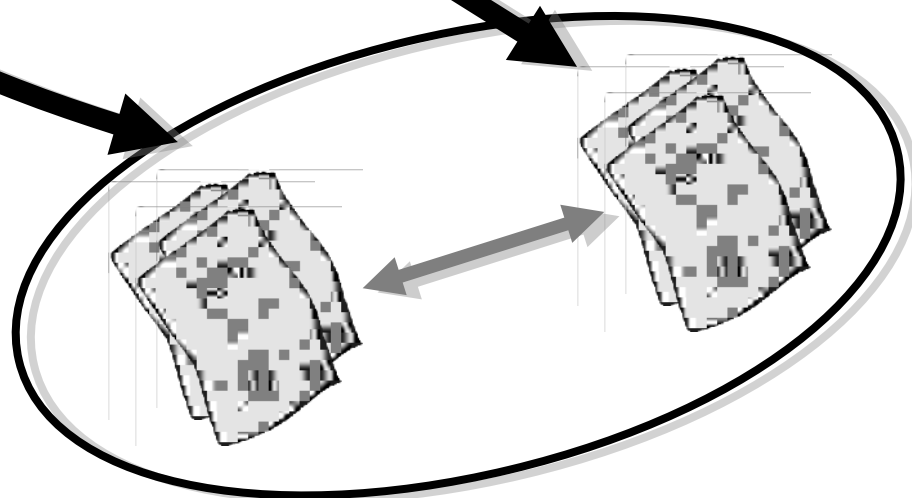
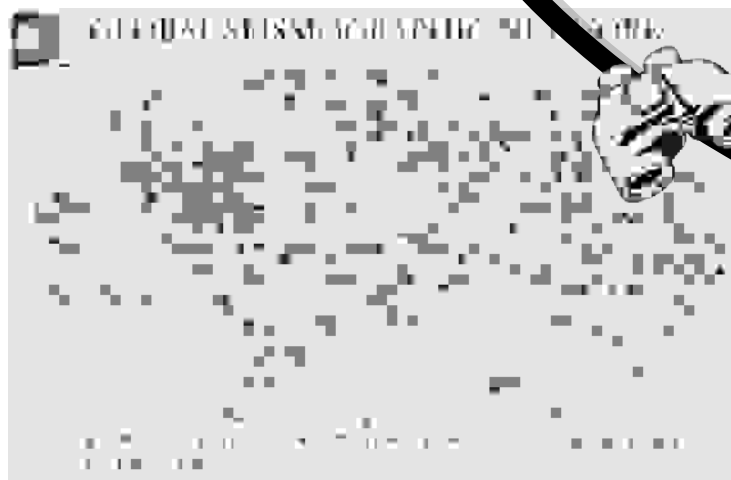
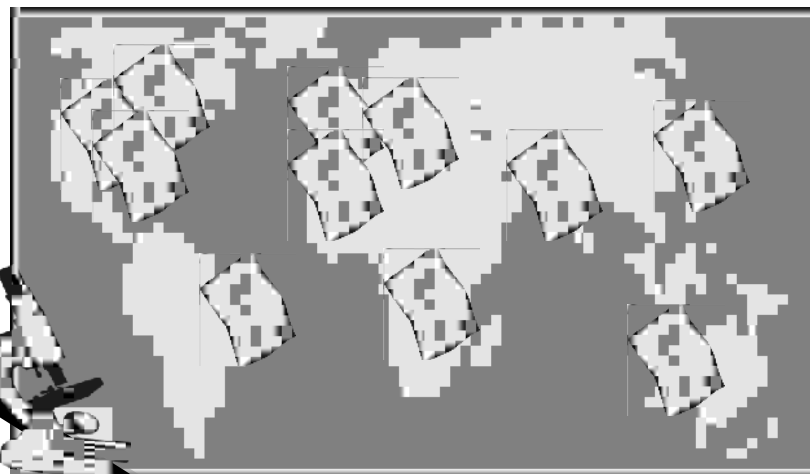
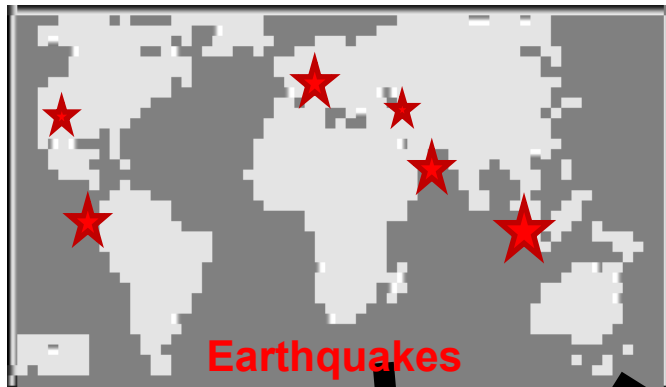
**L'EXEMPLE DE LA MEDIATISATION DES  
CATASTROPHES NATURELLES**

GRASLAND C.

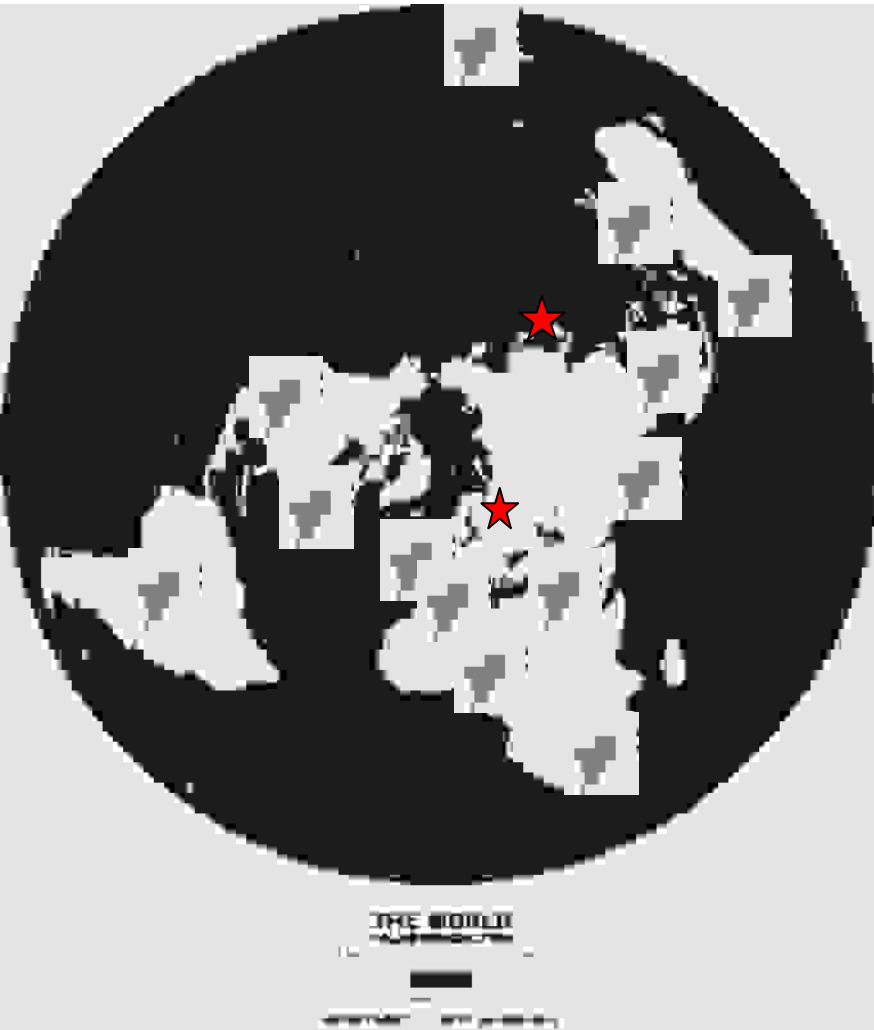
Université Paris Diderot

# INTRODUCTION





# Hypothèse du capteur global



Agrégation de tous les signaux d'un grand nombre de journaux en un signal unique. (ex. presse FACTIVA)

# Test sur Factiva

**Type de médias** : « Journaux »

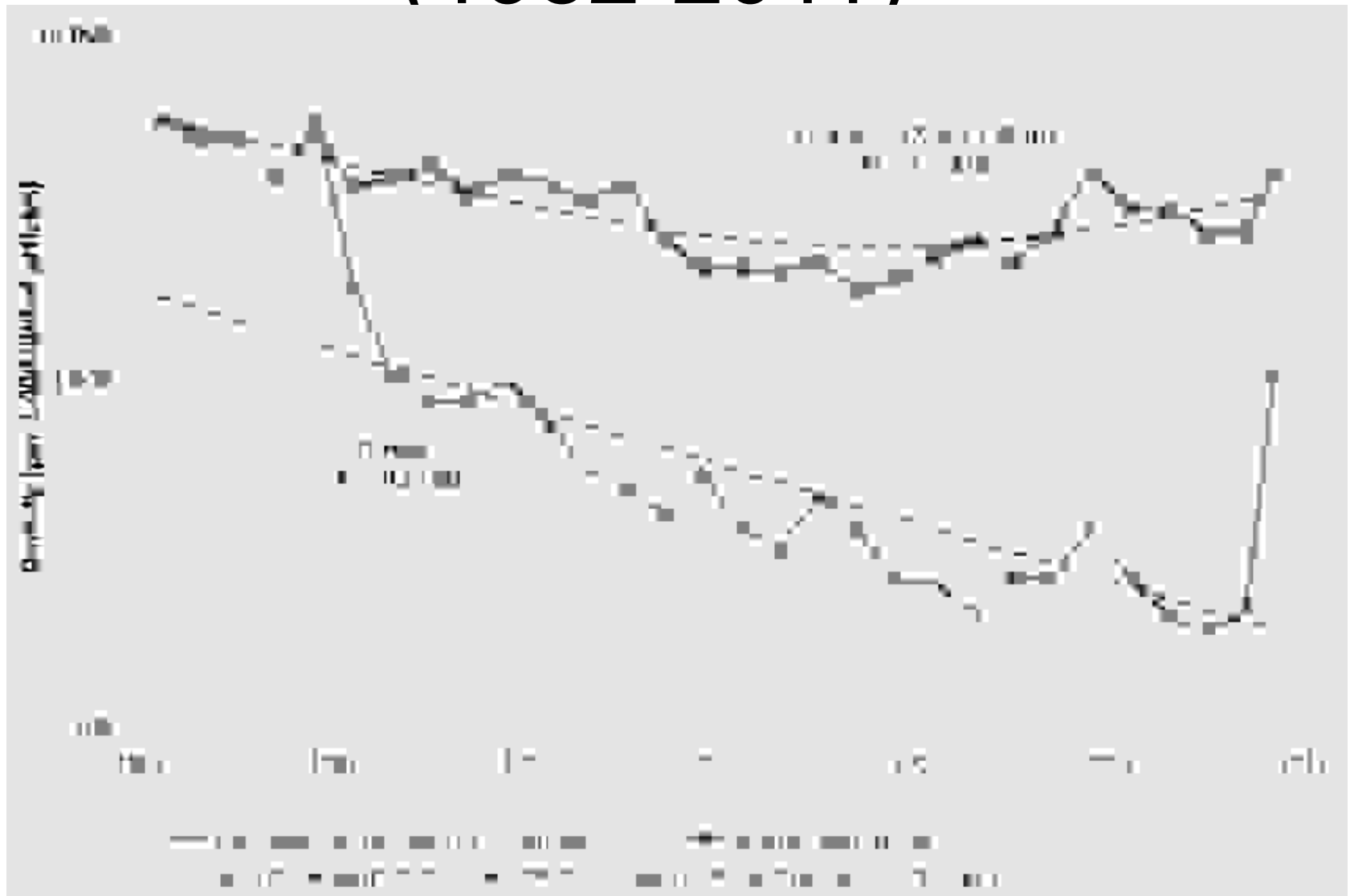
**Langue** : « Anglais »

**Thème 1** : « *Chernobyl* » OR « *Tchernobyl* »  
OR « *Cernobyl* »

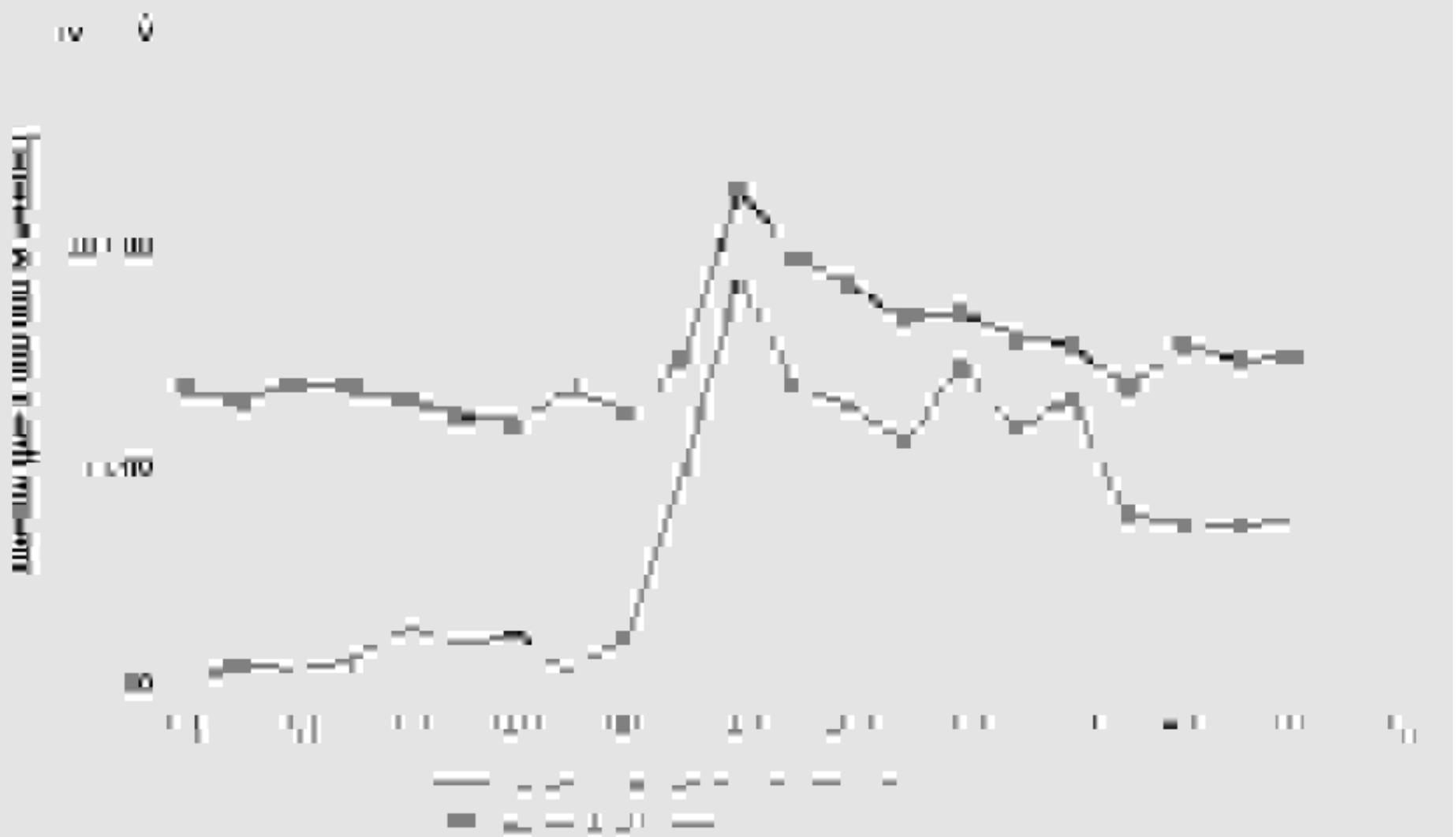
**Thème 2** : « *Energy* » AND « *Nuclear* »

**Contrôle** : « *the* »

# Densités annuelles (1982-2011)

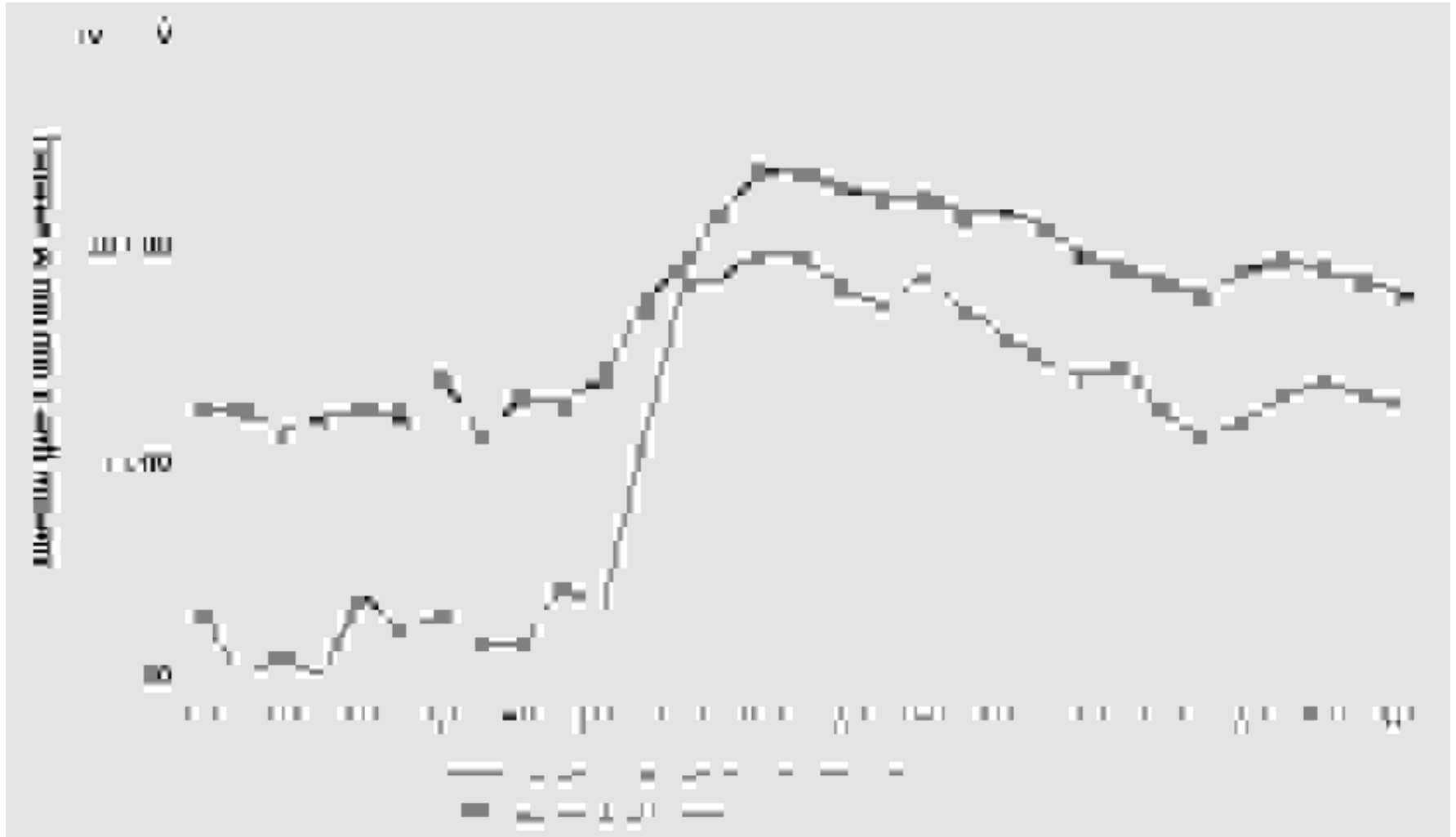


# Densités hebdomadaires (3 janvier au 23 Mai 2011)

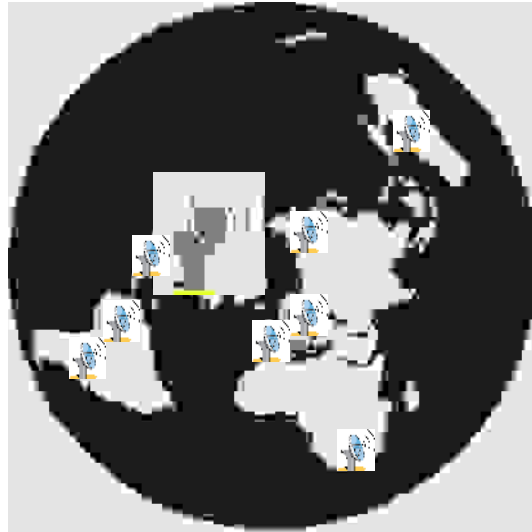




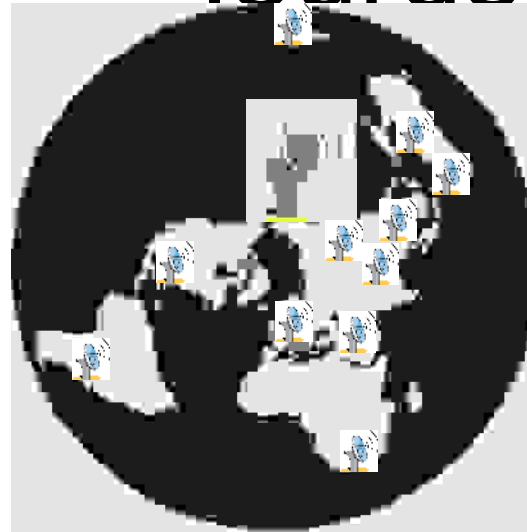
# Densités quotidiennes (1<sup>er</sup> au 31 Mars 2011)



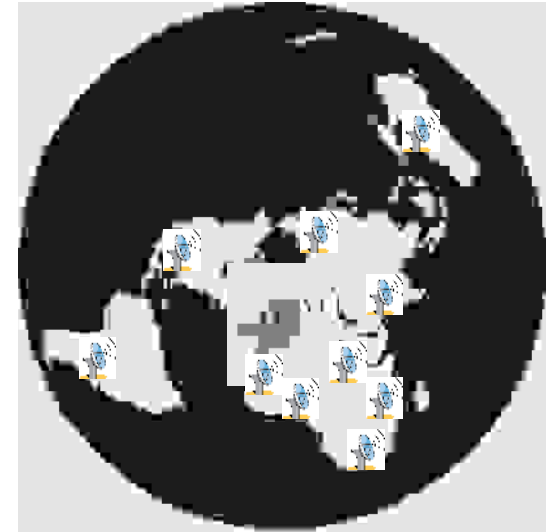
# Hypothèse de capteurs parallèles lourds



Associated Press



Reuters



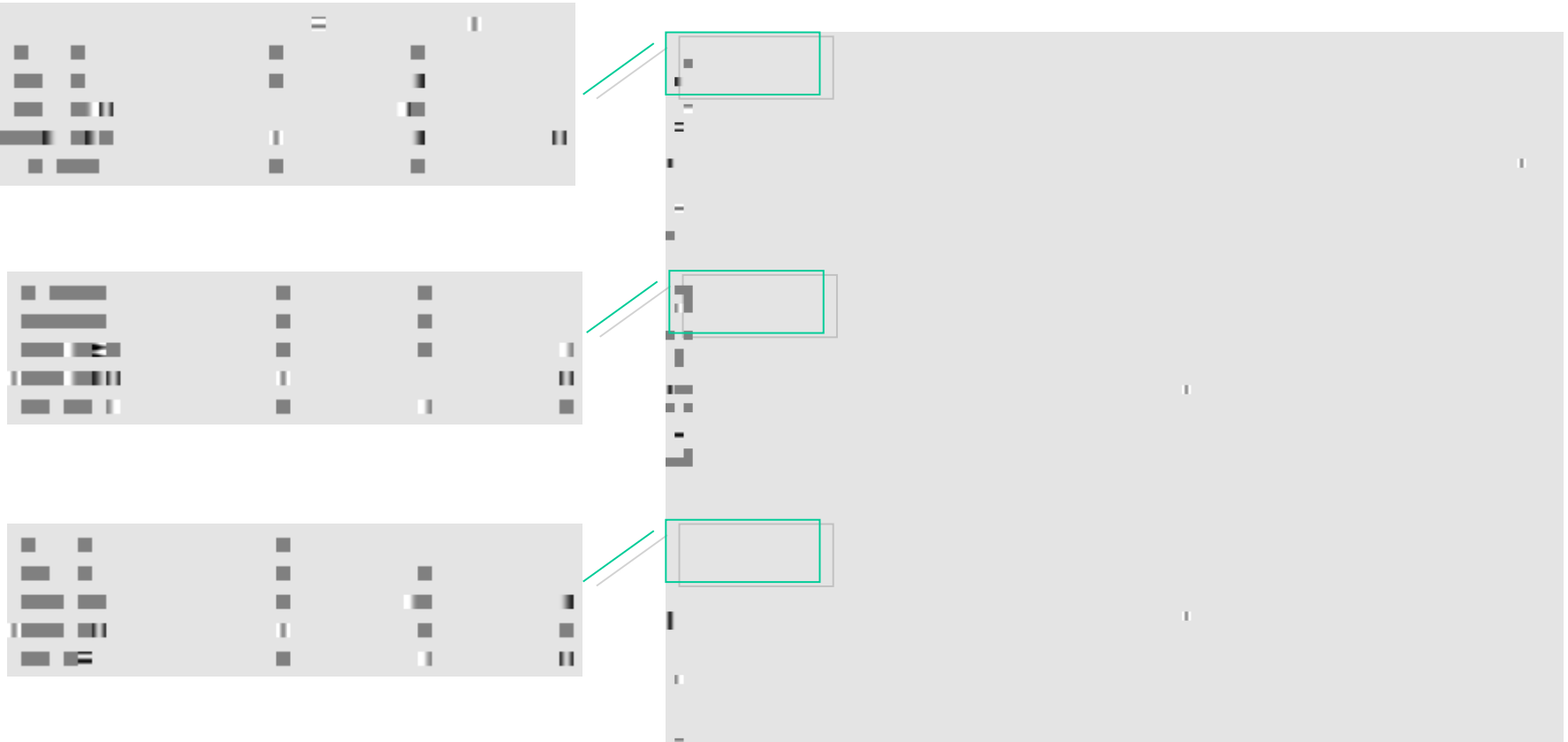
Agence France Presse

Comparaison du signal de trois réseaux de collecte d'information à vocation mondiale, dont les dépêches sont ensuite reprises par les journaux

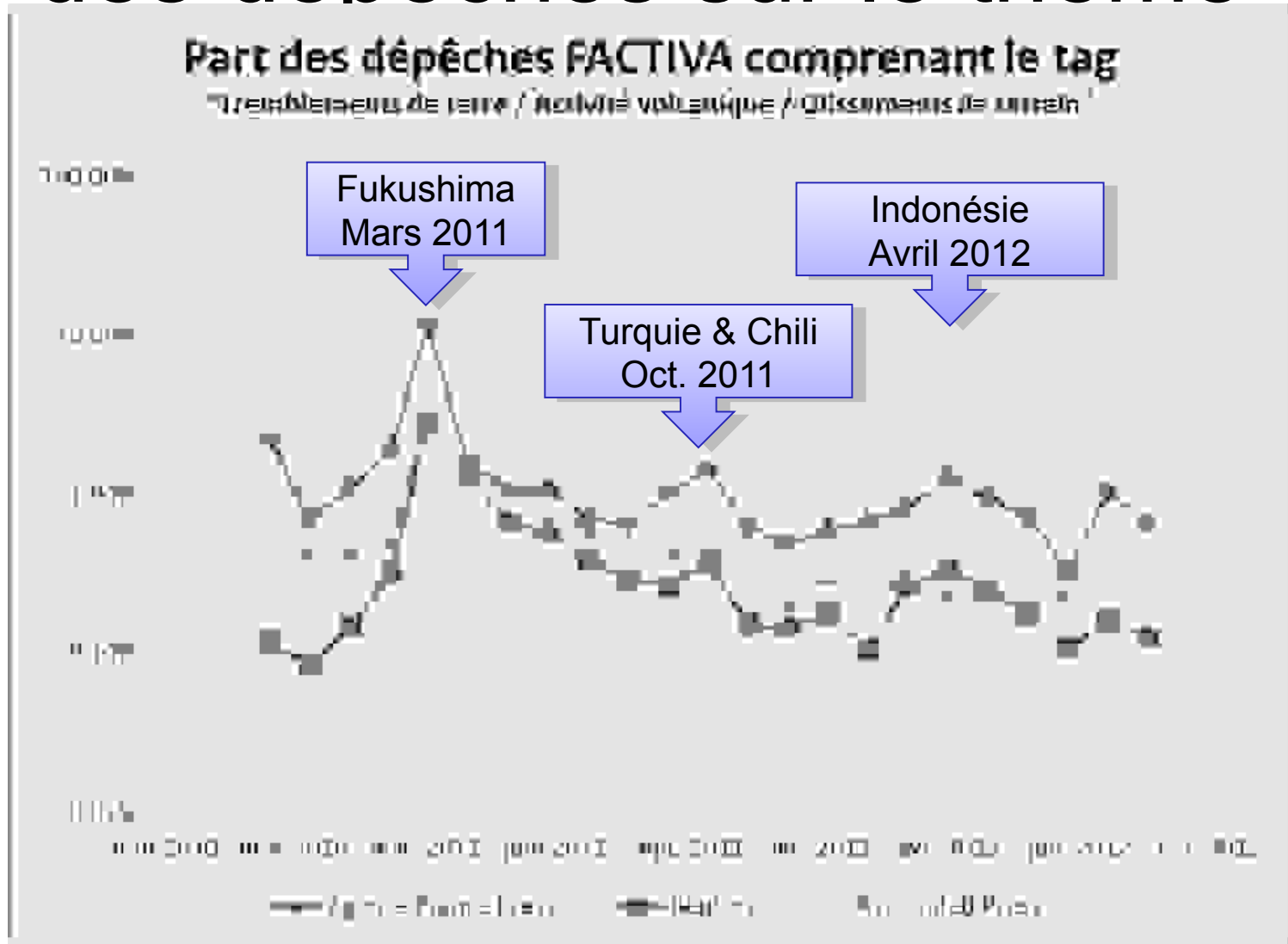
23/11/11

# Tableau Agence\*Pays\*Temps

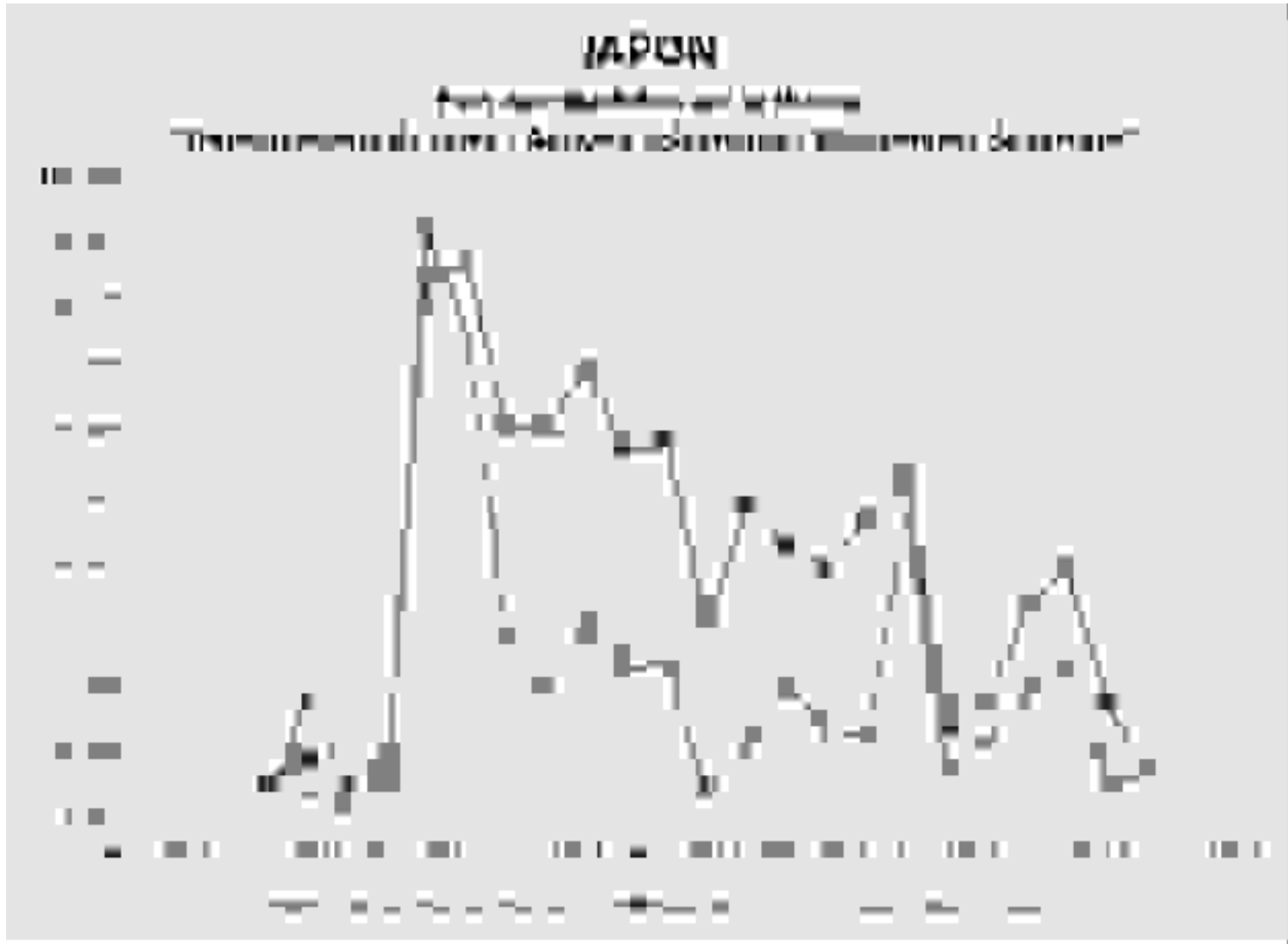
Source FACTIVA – Tag « tremblement de terre »



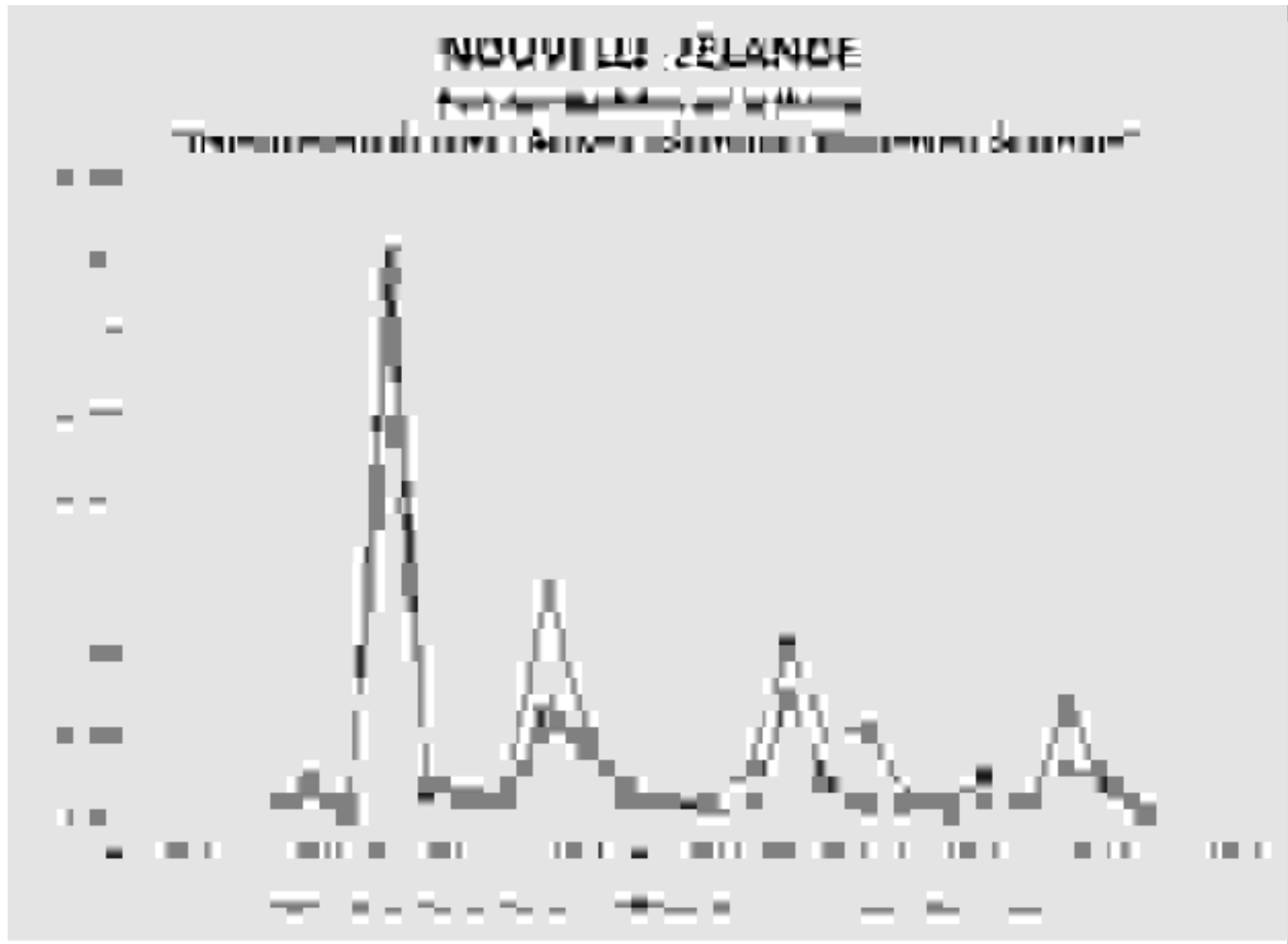
# Variation mensuelle de la part des dépêches sur le thème



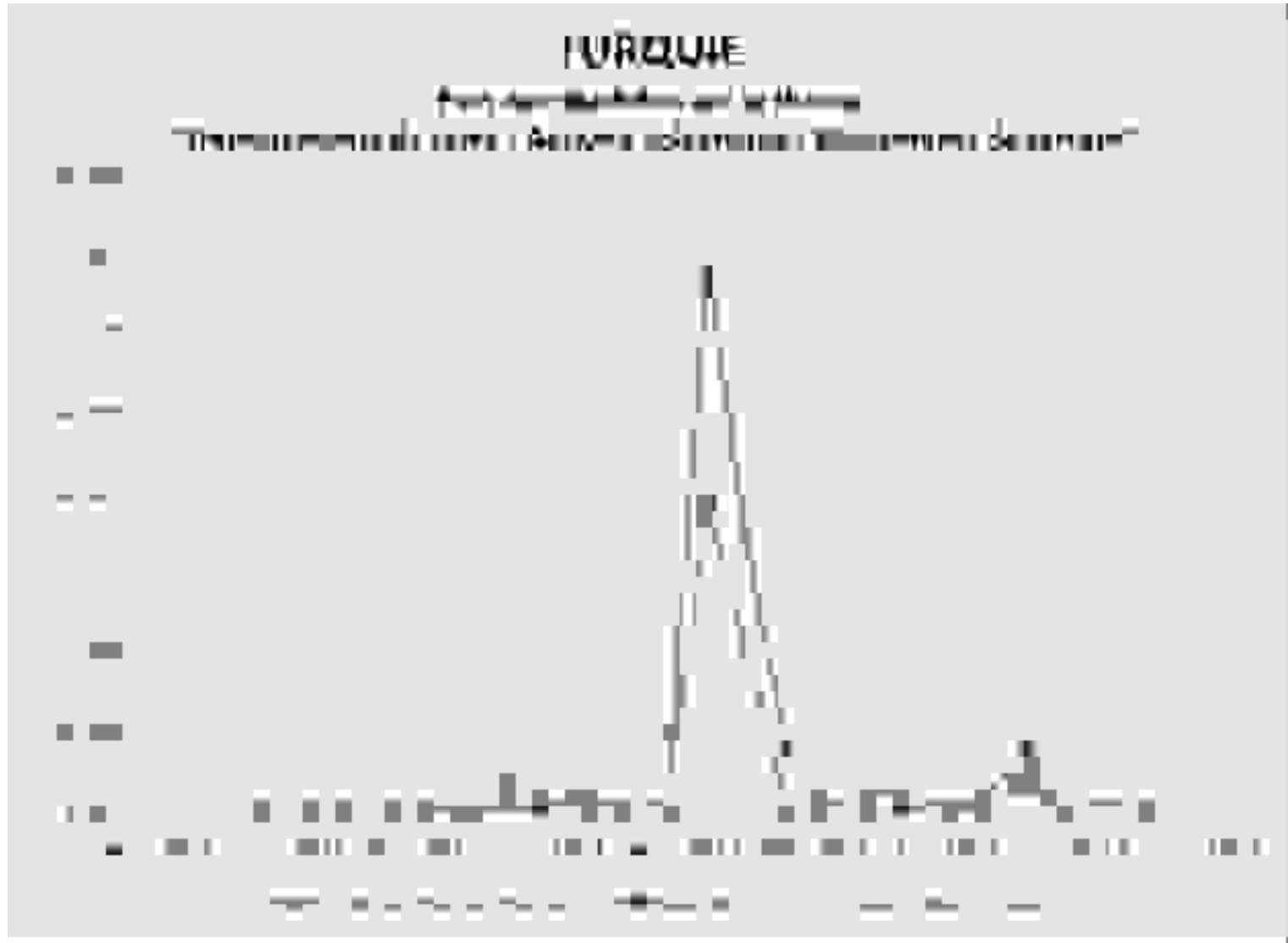
# Analyse par pays



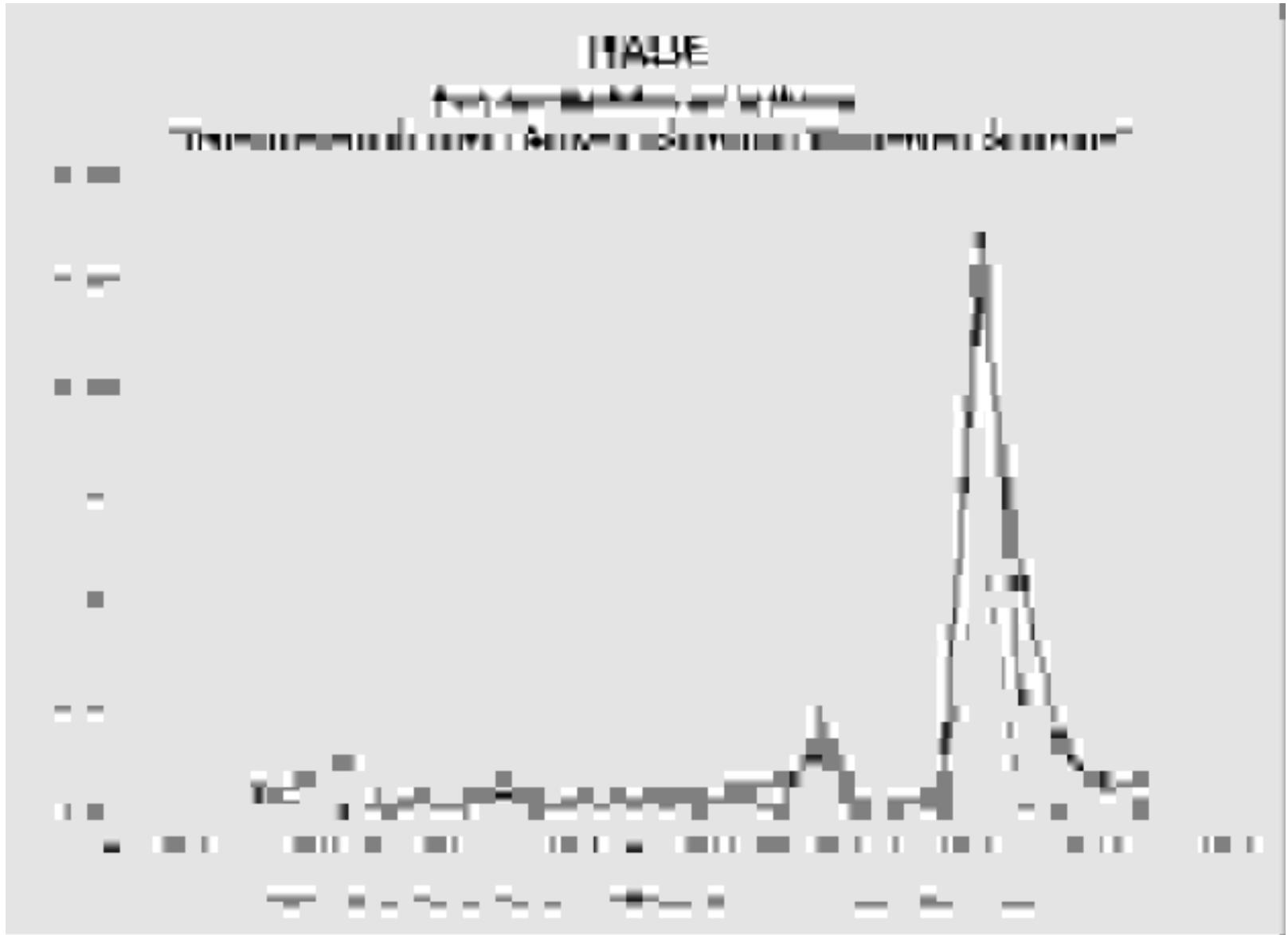
# Analyse par pays



# Analyse par pays

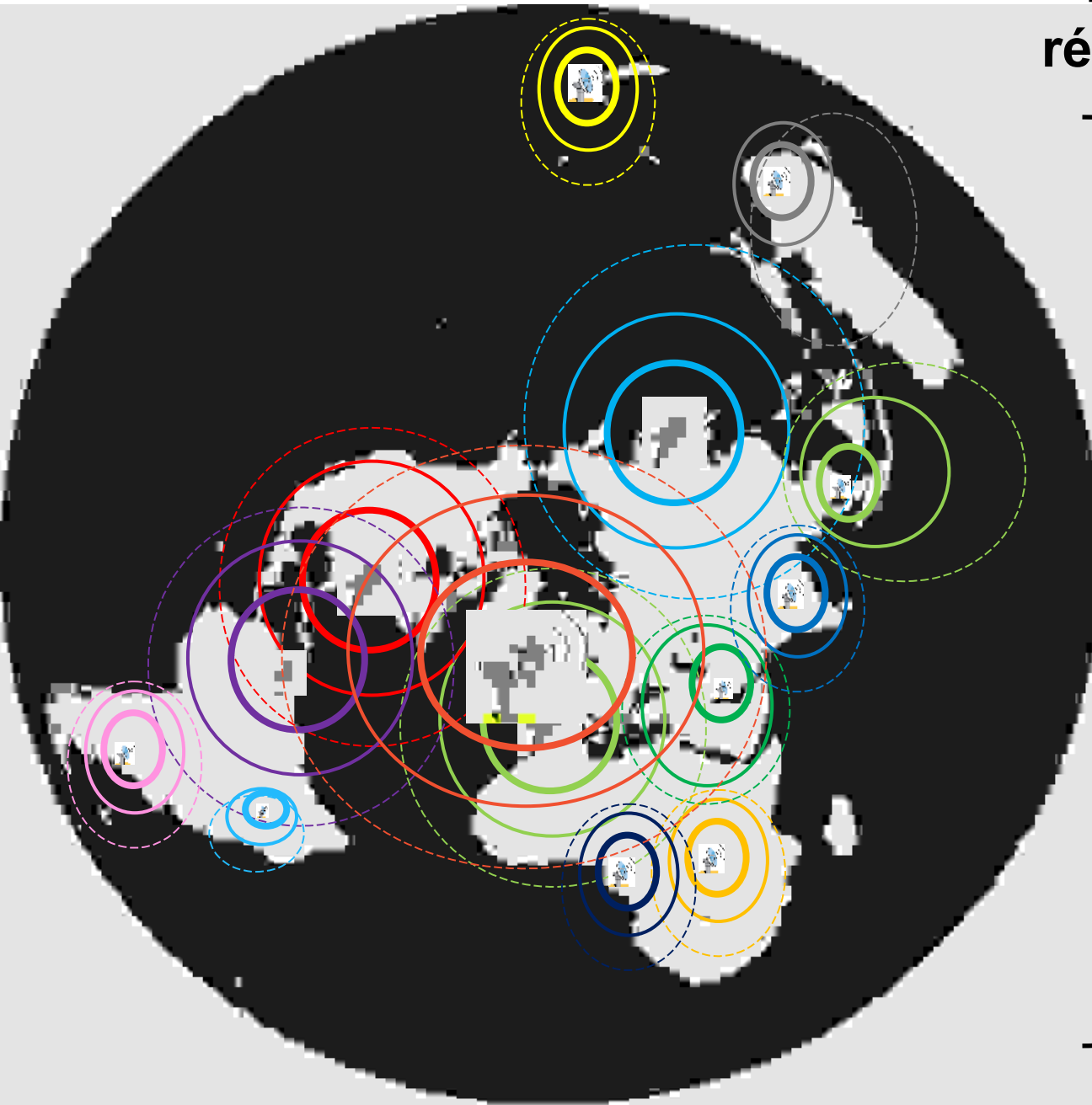


# Analyse par pays





## Hypothèse d'un réseau de capteurs légers



Comparaison du signal de journaux ayant chacun une implantation locale mais des portées spatiales plus ou moins grandes.

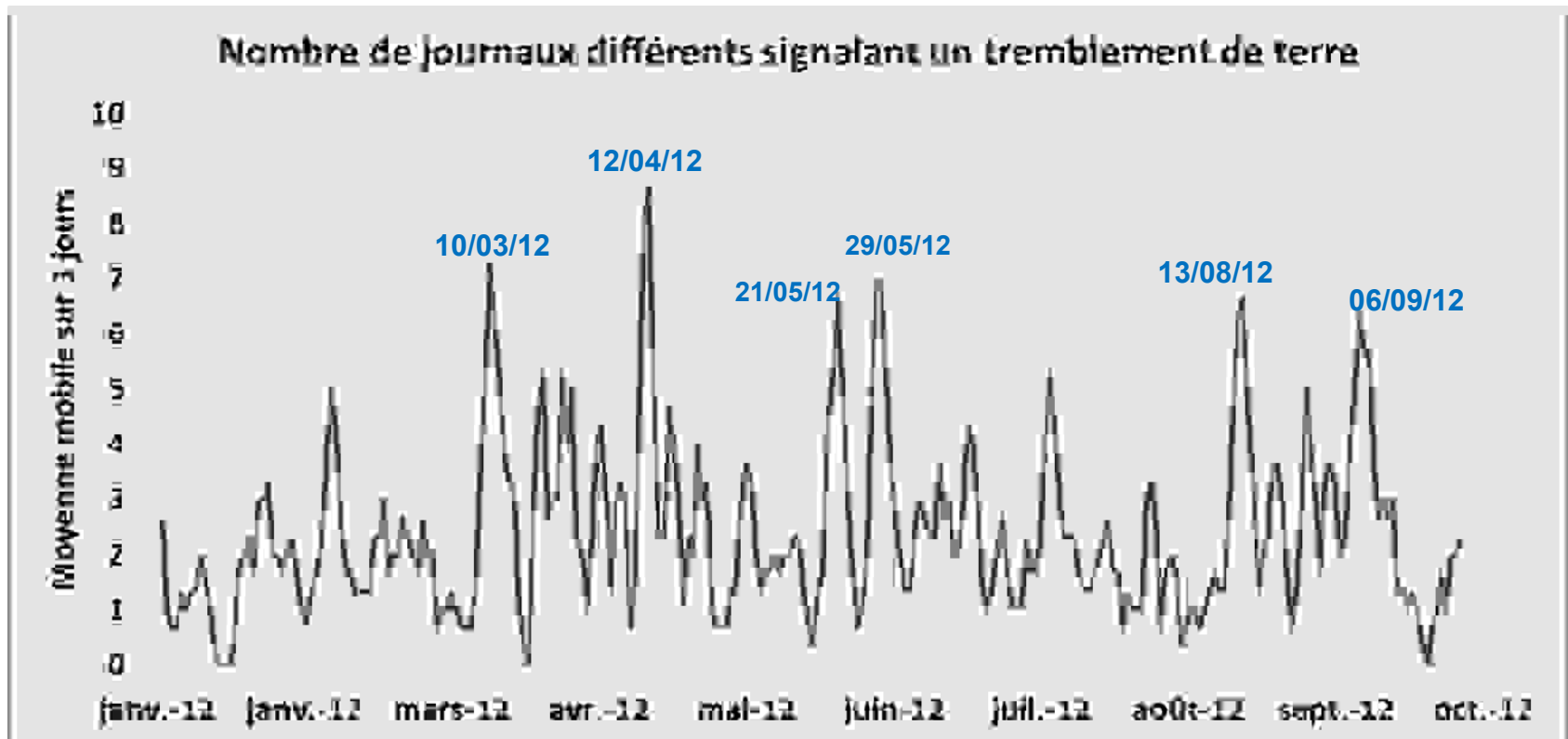
Utilisation des flux RSS pour constituer une base libre de droits.

# Exemple de recherche dans les flux RSS 11 journaux

Flux RSS contenant "earthquake" ou "tsunamis"	01/01/2012	02/01/2012	03/01/2012	04/01/2012	05/01/2012	06/01/2012	07/01/2012	08/01/2012	09/01/2012	10/01/2012	11/01/2012	12/01/2012
China Daily	0	0	1	0	0	0	0	0	0	0	0	0
New York Times world	1	0	0	0	0	0	0	0	0	0	0	0
New Zealand Herald	0	0	0	0	0	0	0	0	0	0	0	0
Philippine Daily Inquirer	0	0	0	0	0	0	0	0	0	0	0	0
The Australian	0	0	0	0	0	0	0	0	0	1	0	0
The Times	0	0	0	0	0	0	0	0	0	0	0	0
The times of India	1	1	0	0	0	0	1	0	1	0	1	1
The times of Malta	0	0	0	0	0	0	0	0	0	0	0	0
The Vancouver Sun	1	0	0	0	0	0	0	0	0	0	0	0
The Washington Post	1	1	0	0	1	1	1	0	1	1	1	0
Toronto Star	1	0	0	0	0	0	0	0	0	0	0	0
Total	5	2	1	0	1	1	2	0	2	2	2	1
Moyenne mobile 3 jours		2.7	1.0	0.7	0.7	1.3	1.0	1.3	1.3	2.0	1.7	1.5

N.B. : on ne tient pas compte du nombre de RSS envoyé par jour mais uniquement de l'existence d'au moins un flux portant sur les tremblements de terre ou tsunamis

# Repérage des pics d'activité maximale



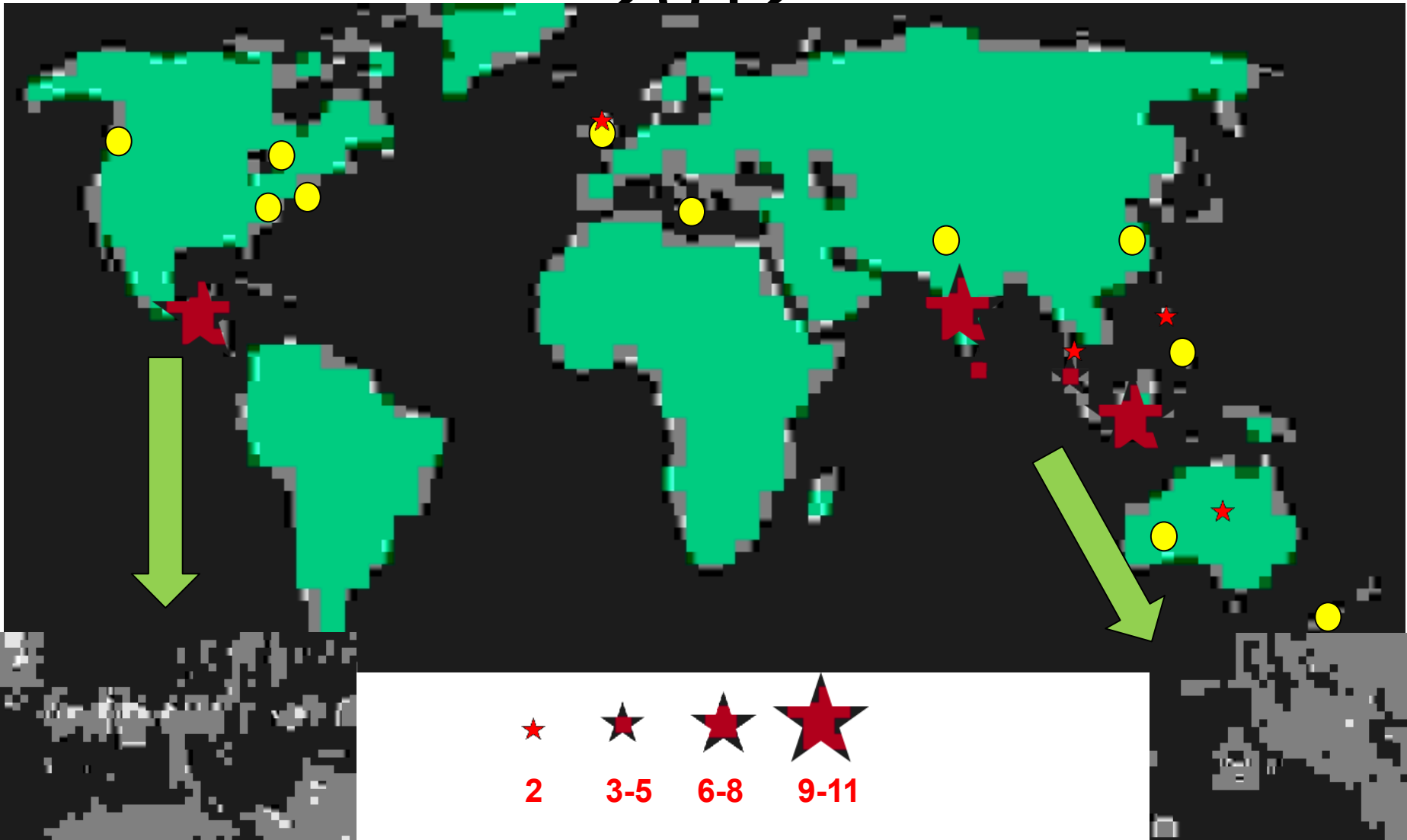
# Analyse des pays les plus cites dans les flux RSS du 9-11 Mars 2012



23/11/11



# Analyse des pays les plus cités dans les flux RSS du 11-13 Avril 2012



# Analyse des pays les plus cités dans les flux RSS du 20-22 Mai 2012



23/11/11



# Analyse des pays les plus cités dans les flux RSS du 28-30 Mai 2012



23/11/11



# Analyse des pays les plus cites dans les flux RSS du 12-14 Août 2012



23/11/11







**2e Partie : Quelles méthodes ?**  
**LES FLUX RSS ET L'ANALYSE DES**  
**RELATIONS INTERNATIONALES**

GRASLAND C.  
Université Paris Diderot

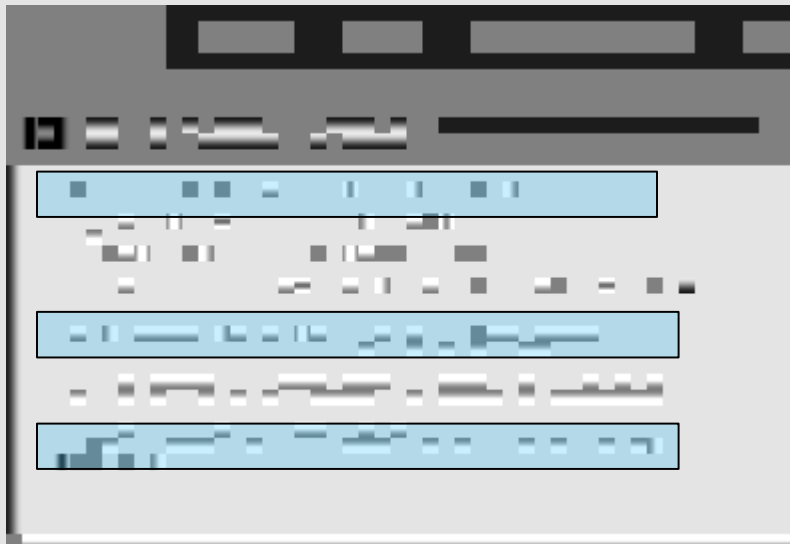
# INTRODUCTION



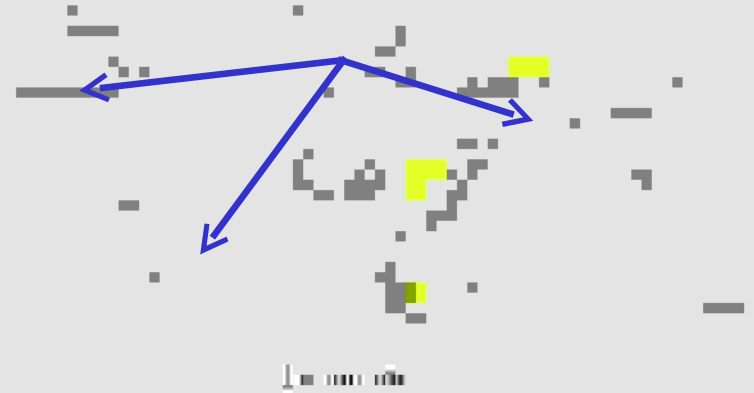
# Flux médiatiques et puissance internationale

Flux médiatiques directs et paradigme réaliste (*Carr, Morgenthau, Waltz; ...*)

**Puissance des Etats = nombre de citations de médias étrangers**



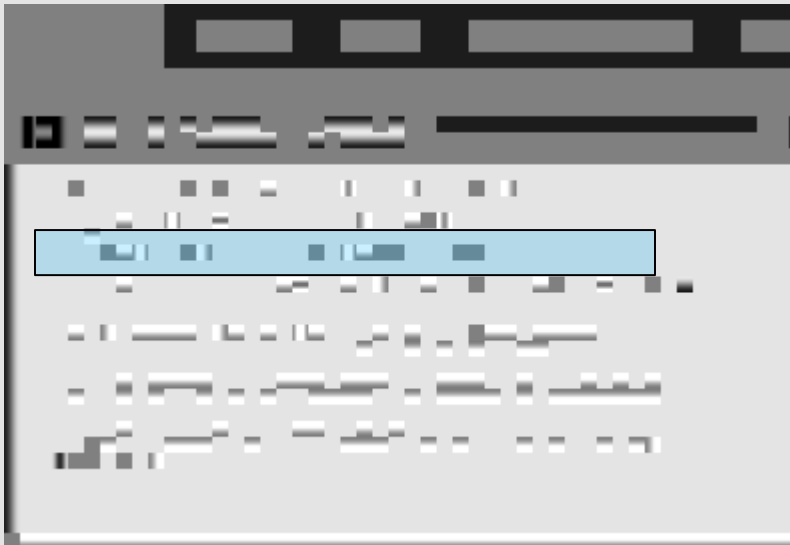
Lecteur de flux RSS



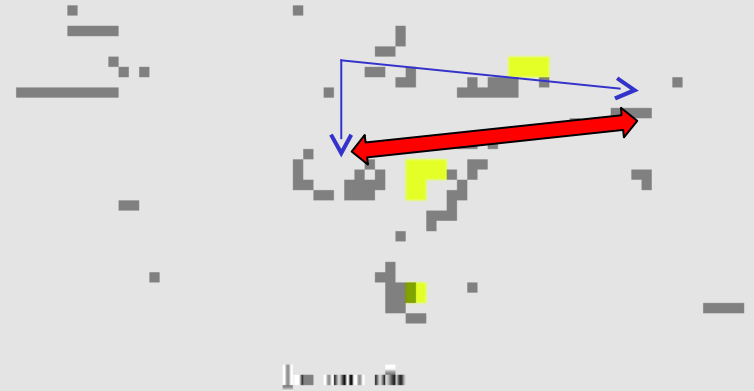
Flux médiatiques directs

# Flux médiatiques et puissance internationale

Flux médiatiques indirect et paradigme transnationaliste (Keohane & Nye)  
**Puissance des Etats = nombre de cocitations avec d'autres pays**



Lecteur de flux RSS

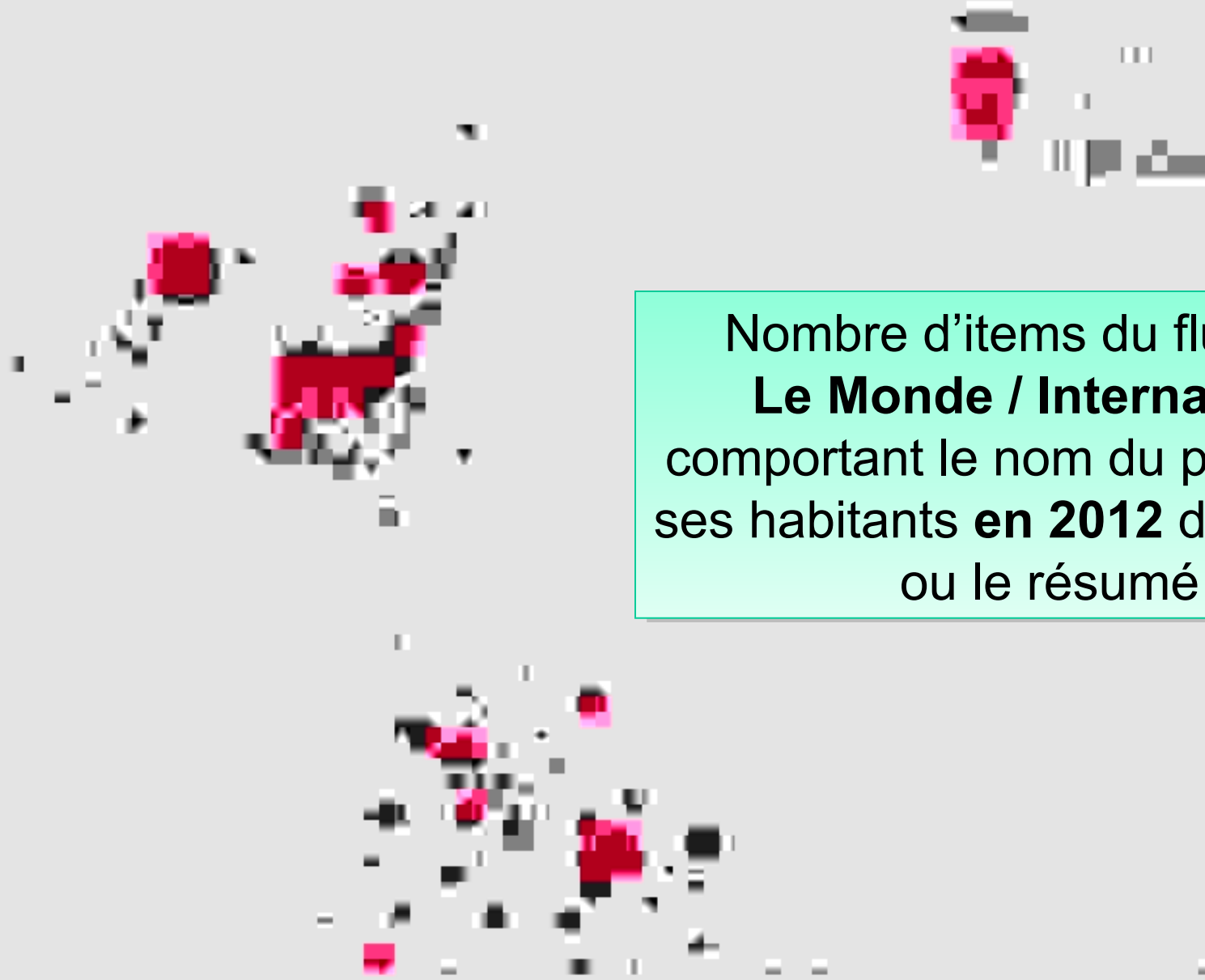


Flux médiatiques indirects

# Puissance des états et relations médiatiques de premier ordre *(paradigme réaliste)*

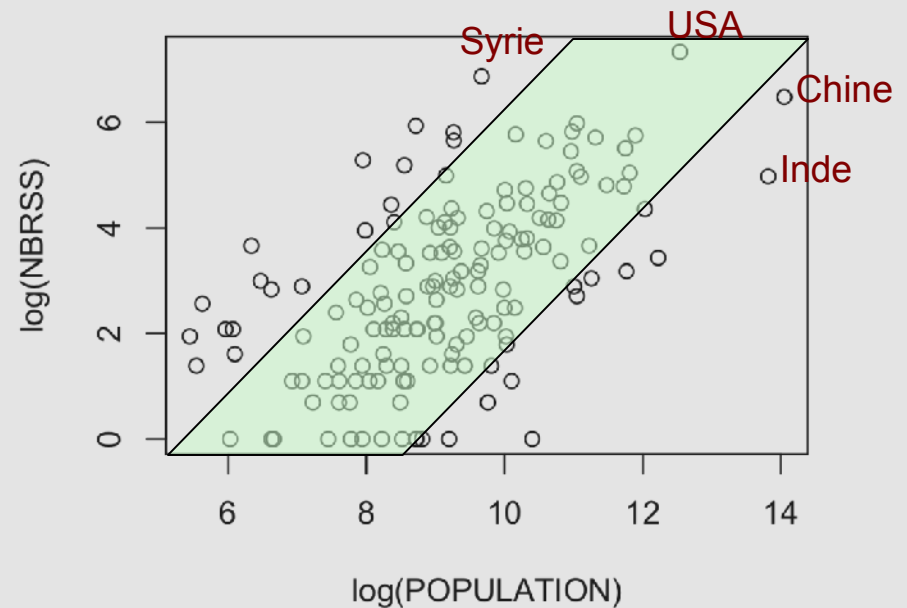
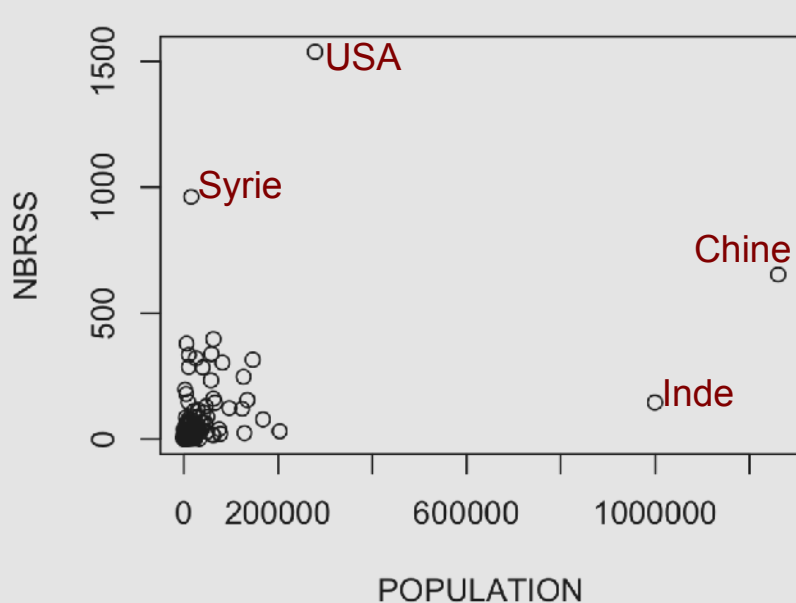


# « *Le MONDE* » (2012)



Nombre d'items du flux RSS  
**Le Monde / International**  
comportant le nom du pays ou de  
ses habitants **en 2012** dans le titre  
ou le résumé

# H1 : L'importance médiatique d'un pays est proportionnelle à sa population



R- Pearson	POPULATION	log.POPULATION.
NBRSS	0,39	0,41
log.NBRSS.	0,33	0,58

# H1 : L'importance médiatique d'un pays est proportionnelle à sa population

$$\text{NBRSS} = 0.25 (\text{POPULATION})^{0.57}$$

*L'importance d'un pays est environ proportionnelle à la racine carré de sa population –  
Les habitants des petits pays sont plus médiatisés que ceux des grands*

```
1st-Order Spl. B-spl. = Value P<Value>
(Intercept) -1.382017 0.057819 -24.07 <2e-16 ***
logPOPULATION 0.509822 0.005115 107.12 <2e-16 ***
---
Signif. codes:  '0' '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for poisson family taken to be 1)
```

```
Null deviance: 27393 on 154 degrees of freedom
Residual deviance: 16838 on 153 degrees of freedom
AIC: 17598
```

*Ce modèle explique (27393 – 16838)/27393 de la déviance totale  
**39% expliqué – 61% résiduel***



# H1 : L'importance médiatique d'un pays est proportionnelle à sa population

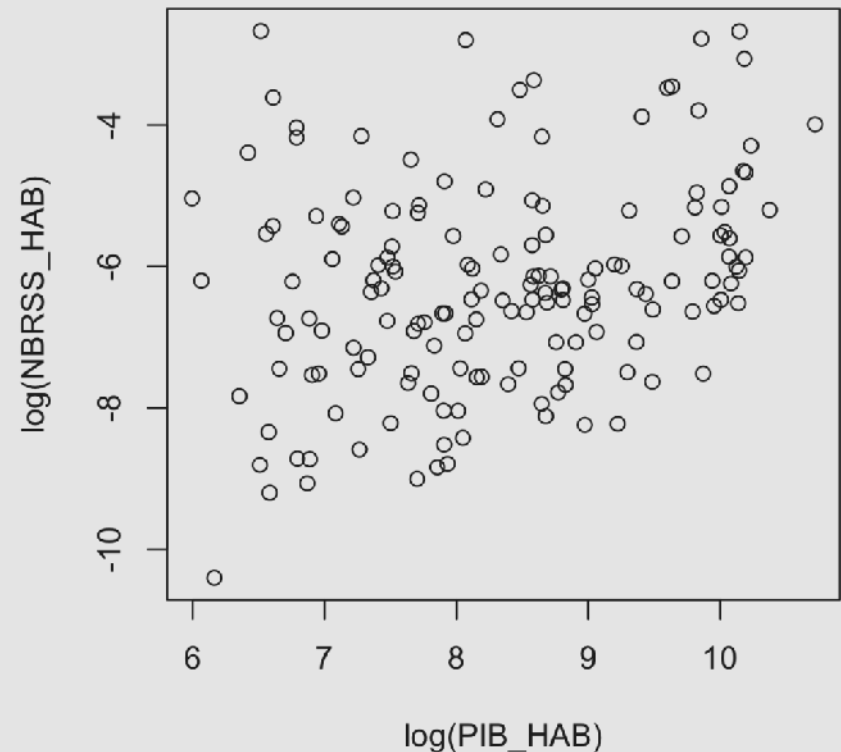
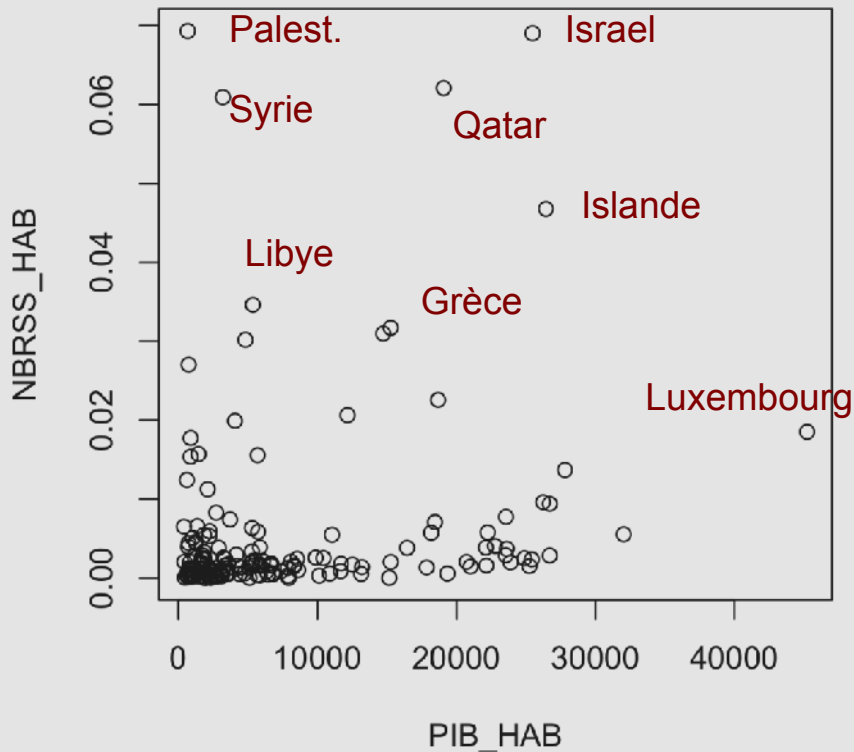


**Résidus relatifs :**  
**Obs / Est**



**Résidus absolus :**  
**Obs - Est**

## H2 : A population égale, on parle plus des pays riches que des pays pauvres.



## H2: A population égale, on parle plus des pays riches que des pays pauvres.

$$\text{NBRSS} = 0.01 \cdot (\text{POPULATION})^{0.57} \cdot (\text{PIB/hab})^{0.37}$$

*L'importance d'un pays est proportionnelle à sa population mais aussi à la richesse de ses habitants.*

```

Coefficients:
(Intercept)      -0.004802    0.087130    -0.7300    <2e-10 ***
logPOPULATION    0.571535    0.005405    106.29    <2e-16 ***
logPIB_HAB      0.371644    0.008000    47.18    <2e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

```

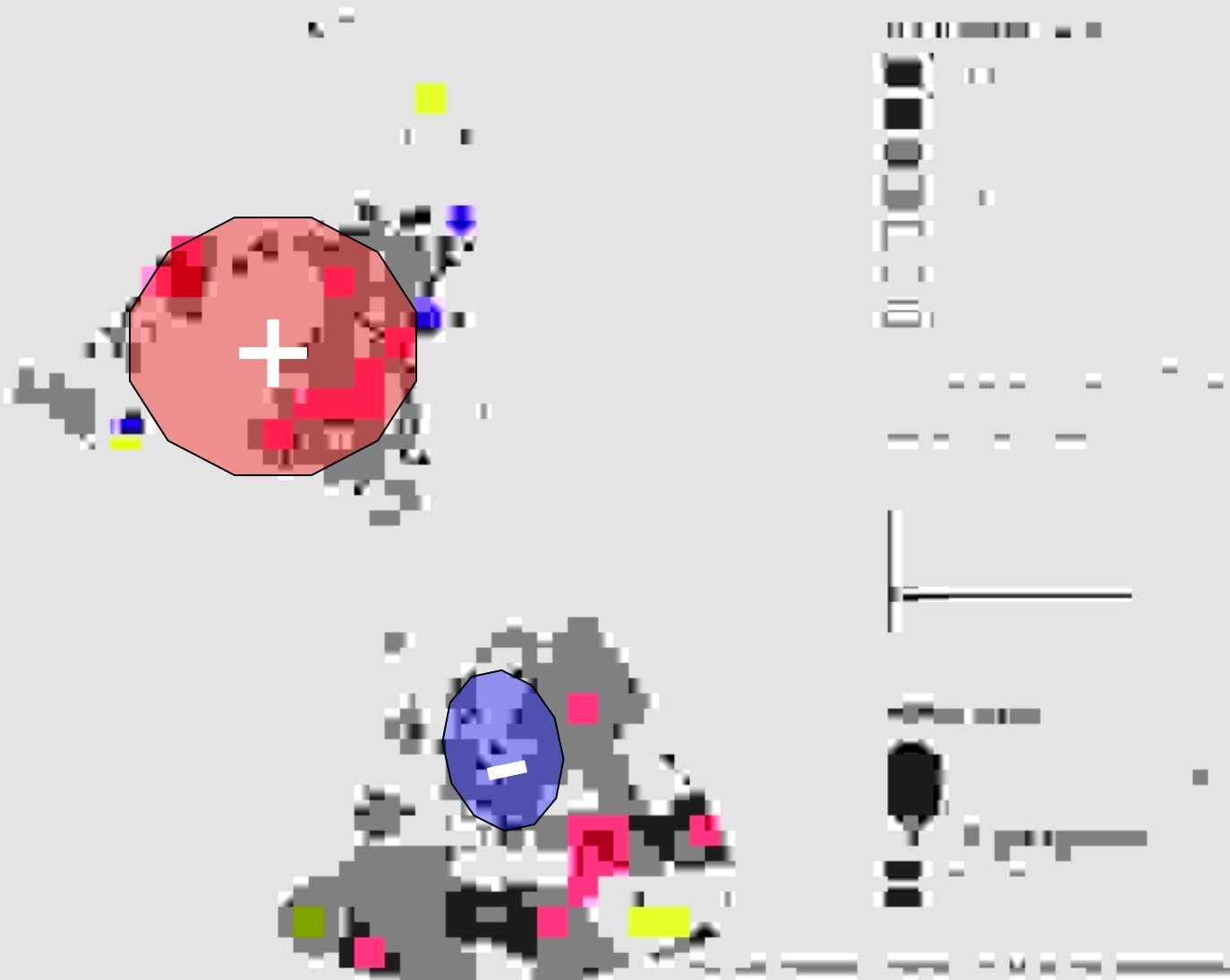
```

Null deviance: 27393 on 158 degrees of freedom
Residual deviance: 14945 on 107 degrees of freedom
AIC: 15718

```

*Ce modèle explique (27393 – 14945)/27393 de la déviance totale  
**45% expliqué – 55% résiduel***

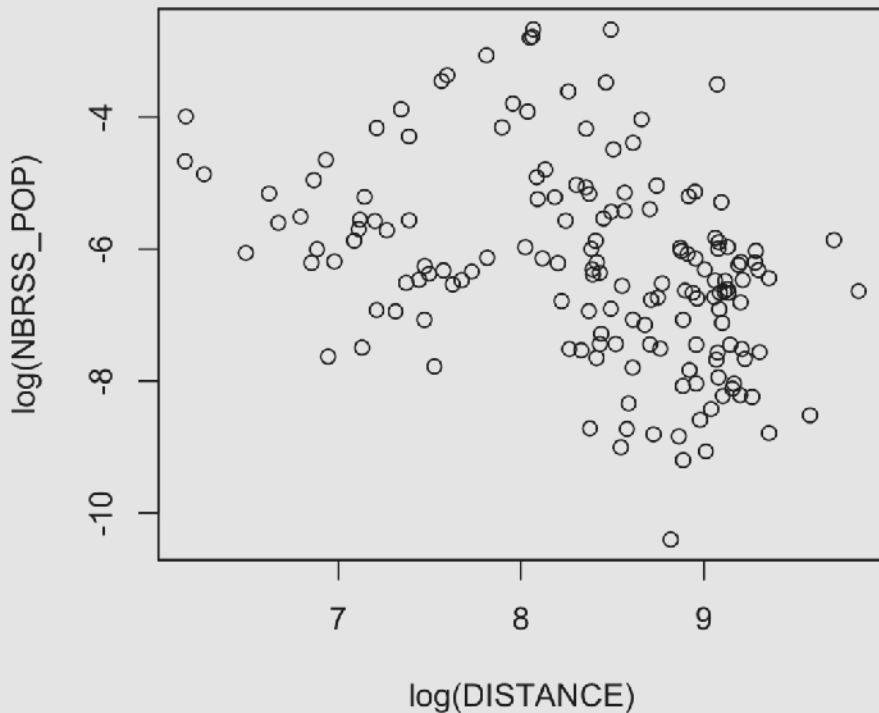
## H2: A population égale, on parle plus des pays riches que des pays pauvres.



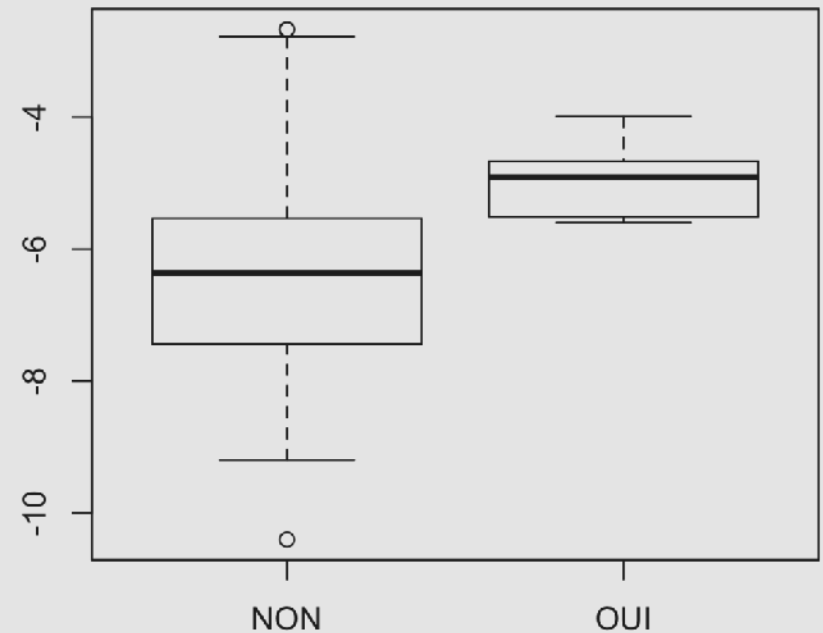
Résidus  
relatifs :  
 $\text{Obs} / \text{Est}$

Résidus  
absolus :  
 $\text{Obs} - \text{Est}$

### H3 : A taille égale (pop., richesse, ...), on parle plus des pays proches que des pays éloignés.



**Proximité continue**  
*Distance euclidienne*



**Proximité discrète**  
*Frontière commune*

# H3 : A taille égale (pop., richesse, ...), on parle plus des pays proches que des pays éloignés

$$\text{NBRSS} = 0.16 \cdot (\text{POPULATION})^{0.63} \cdot (\text{PIB/hab})^{0.30} \cdot (\text{DISTANCE})^{-0.32} \cdot (0.70)^{\text{CONTIGUITE}}$$

*Le nombre d'items diminue avec la distance, mais toutes choses égales on parle moins (-30%) des pays ayant une frontière commune que des autres.*

	Estimate	Std. Error	t value	Pr(> z )
(Intercept)	-1.813299	0.152426	-11.978	<2e-16 ***
logPOPULATION	0.027469	0.003986	104.629	<2e-16 ***
logPIB_HAB	0.304668	0.009611	31.699	<2e-16 ***
logDISTANCE	-0.329582	0.013403	-24.591	<2e-16 ***
CONTIGUITE	-0.354681	0.040670	-8.719	<2e-16 ***

```
Null deviance: 27393 on 164 degrees of freedom
Residual deviance: 14329 on 100 degrees of freedom
AIC: 15098
```

*Ce modèle explique (27393 – 14945)/27393 de la déviance totale  
**48% expliqué – 52% résiduel***

# TIMES OF INDIA (2012)

Rank	DOM	IMPRESS
1	ETATS UNIS	1160
2	INDIE	714
3	CHINE	687
4	SYARIE	610
5	ROYAUME UNI	547
6	AFRIQUE DU SUD	486
7	INDON	473
8	EGYPTE	463
9	ITALIE	400
10	ISRAEL	375
11	FRANCE	333
12	JAPON	313
13	ESPAGNE	292
14	ROYAUME DES NEUCHES	270
15	RUSSIE	260
16	AUSTRALIE	257
17	PAKISTAN	250
18	THAIE	230
19	ITALIE	220
20	CHINE	210

# TIMES OF INDIA - 2012

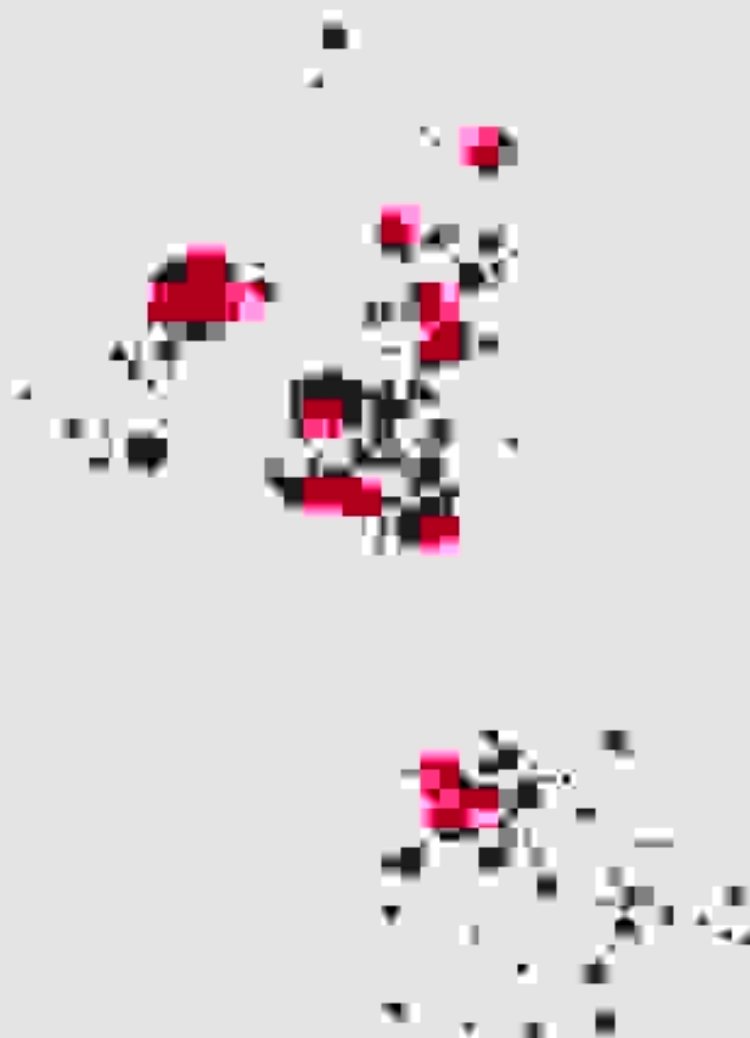
Variable	Estimate	SE	t-stat	p-value
Distance	-0.111	0.009	-12.22	<.0001
Contiguity	0.190	0.011	17.18	<.0001
Langue	0.270	0.012	22.50	<.0001
Colonisation	0.160	0.013	12.31	<.0001
Population	0.000	0.000	1.12	0.261
PIB/hab	0.000	0.000	0.00	1.000

- **Population** : augmente fortement les flux.
- **PIB/hab** : augmente fortement les flux
- **Distance** : Influence très négative (-1.11)
- **Contiguïté** : divise les flux par 1.9
- **Langue** : multiplie les flux par 2.7
- **Colonisation** : multiplie les flux par 1.6

**66 % expliqué – 34% résiduel**



# The SOWETAN - 2012



Rang	NOM	NBRÉS
1	ETATS UNIS	171
2	AFRIQUE	91
3	CHINE	90
4	INDE	88
5	ROYAUME-UNI	82
6	ETHIOPIE	76
7	NOBENT	73
8	BRESIL	71
9	FRANCE	60
10	ESPAGNE	51
11	JAPON	47
12	Canada	45
13	MEXIQUE	43
14	ALLEMAGNE	43
15	ITALIE	37
16	THAILANDE	36
17	HONGRIE	32
18	EGYPTE	31
19	RUSSIE FEDERATION DE	29
20	NOUVELLE ZELANDE	28
21	BRANA	26

# The SOWETAN - 2012

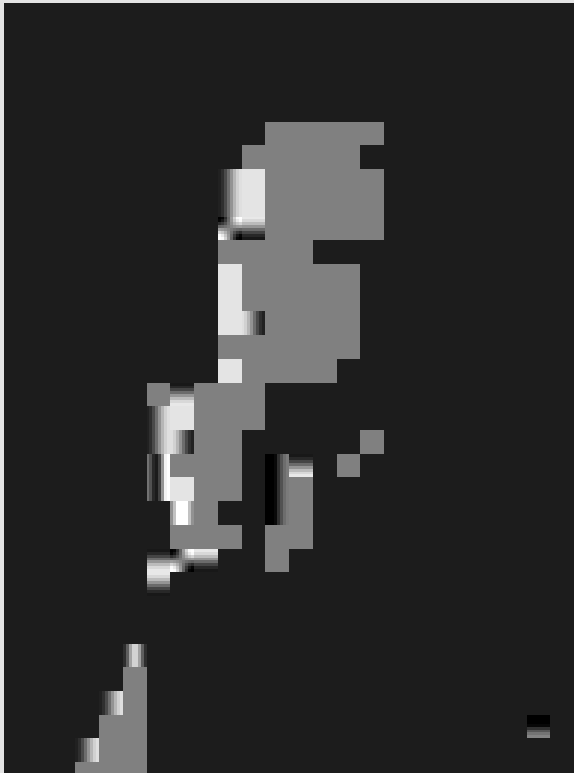
Variable	Coef	Std. Error	t-Statistic	Prob >  t
Interception	1.000000	0.000000	1.000000	1.000000
Distance	-0.740000	0.000000	-0.740000	0.450000
Contiguïté	0.000000	0.000000	0.000000	1.000000
Langue	2.300000	0.000000	2.300000	0.020000
Colonisation	0.320000	0.000000	0.320000	0.750000

- **Population** : augmente les flux.
- **PIB/hab** : augmente les flux
- **Distance** : Forte influence négative (-0.74)
- **Contiguïté** : pas d'effet
- **Langue** : multiplie les flux par 2.3
- **Colonisation** : divise les flux par 3.2

**73 % expliqué – 27% résiduel**



# Puissance des Etats et relations médiatiques de second ordre *(paradigme transnationaliste)*

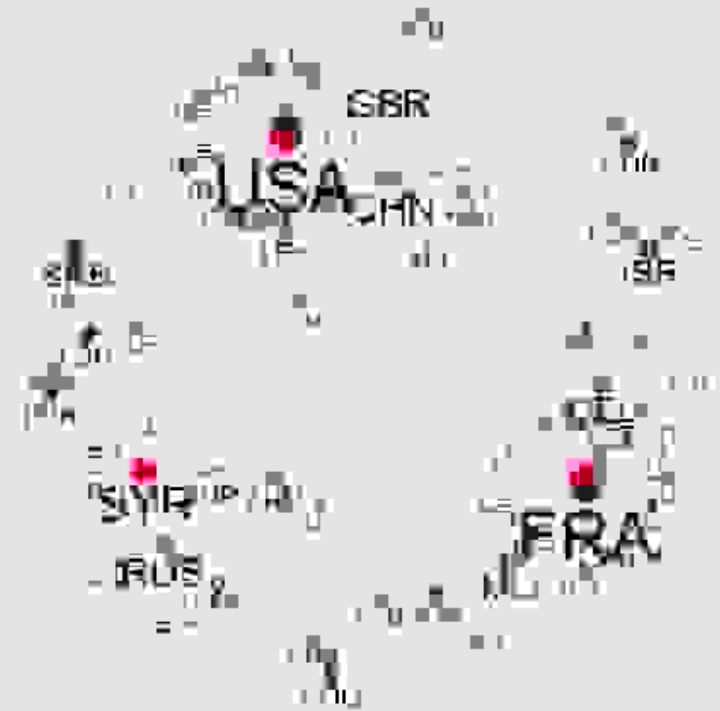
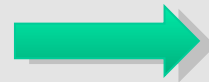


# Transformation des flux de co-citations en graphe de domination (ex. *Le Monde*, 2012)

Le pays  $i$  est « dominé » par le pays  $j$  si  $i$  envoie son plus grand flux vers  $j$  &  $j$  reçoit plus de flux que  $i$



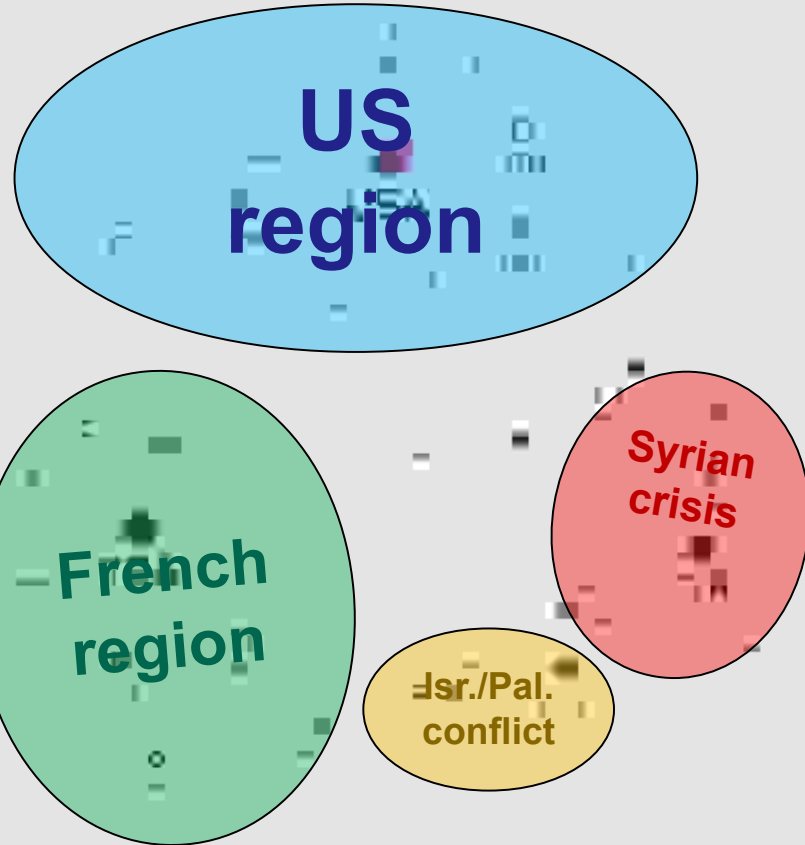
Flux bruts de co-citation



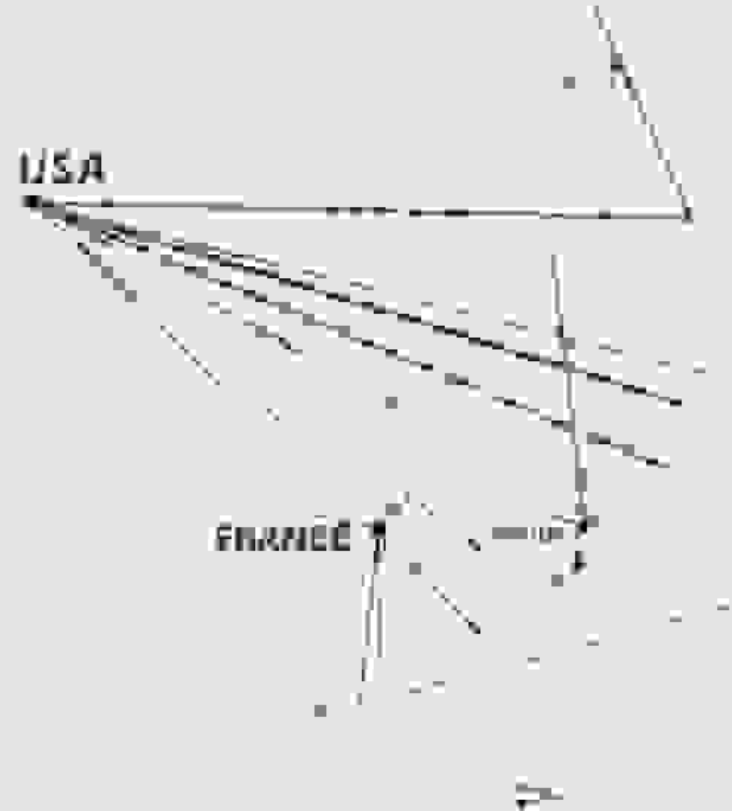
Flux dominants de co-citation

# Dominant links of **Le Monde** (May 2013 – Dec. 2012)

Network



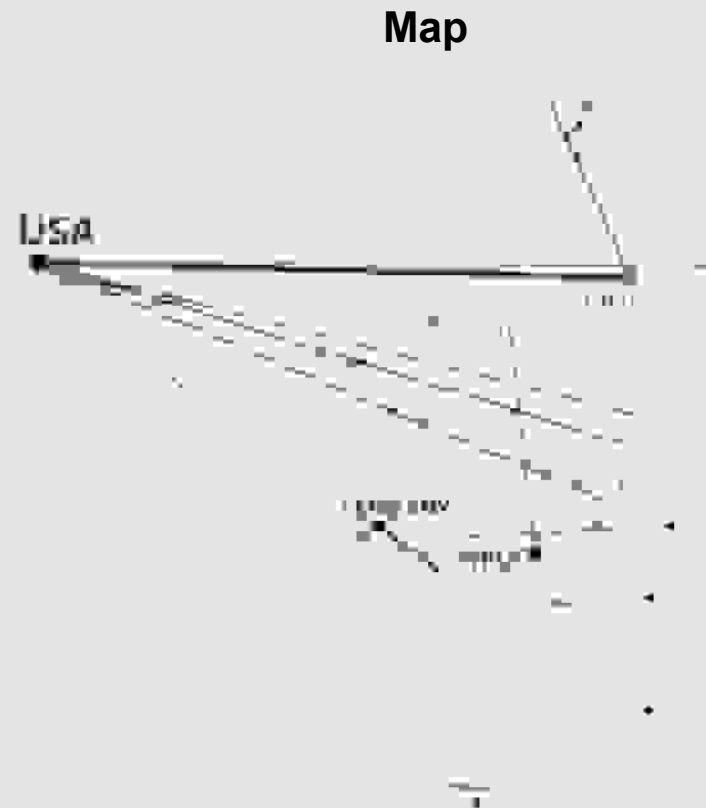
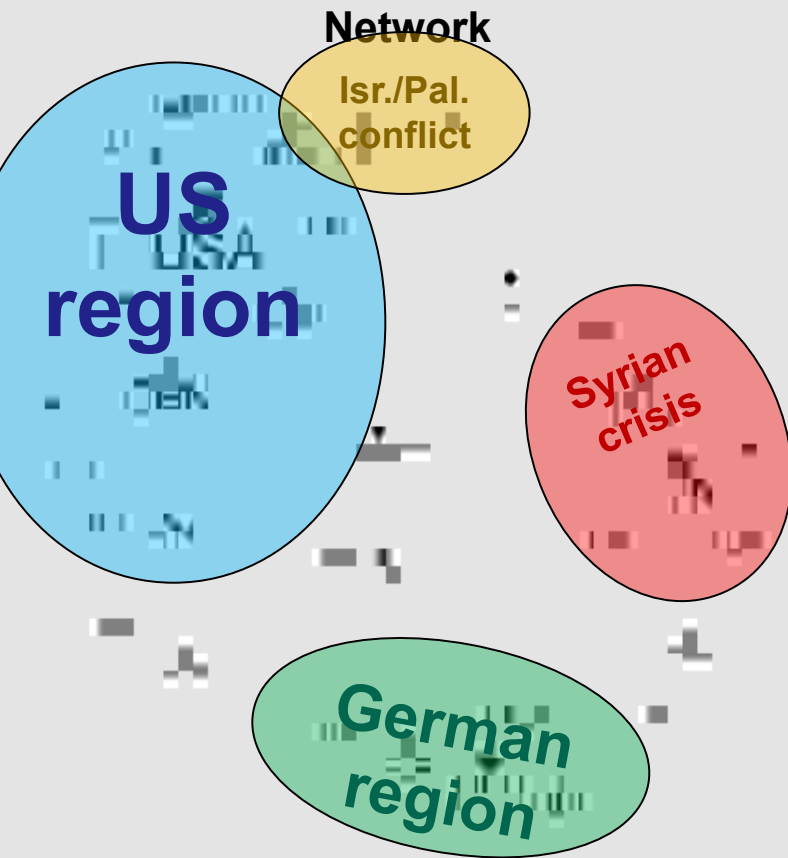
Map



## Comment

- Division of the World between major French & US regions
- Syrian conflict identified as specific region (with Russia)
- Israel-Palestinian conflict as autonomous region

# Dominant links of *Financial Times* (May 2013 – Dec. 2012)

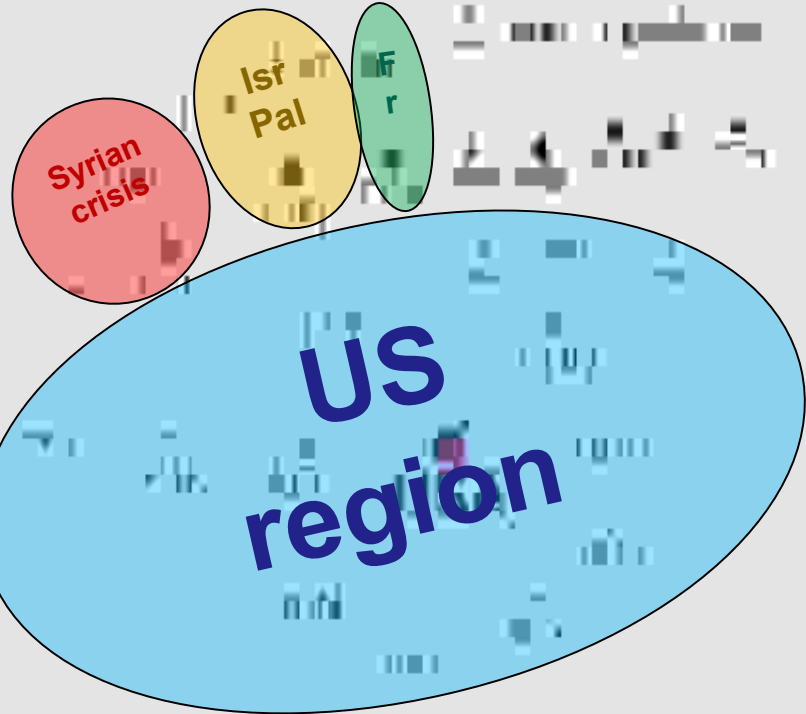


## Comment

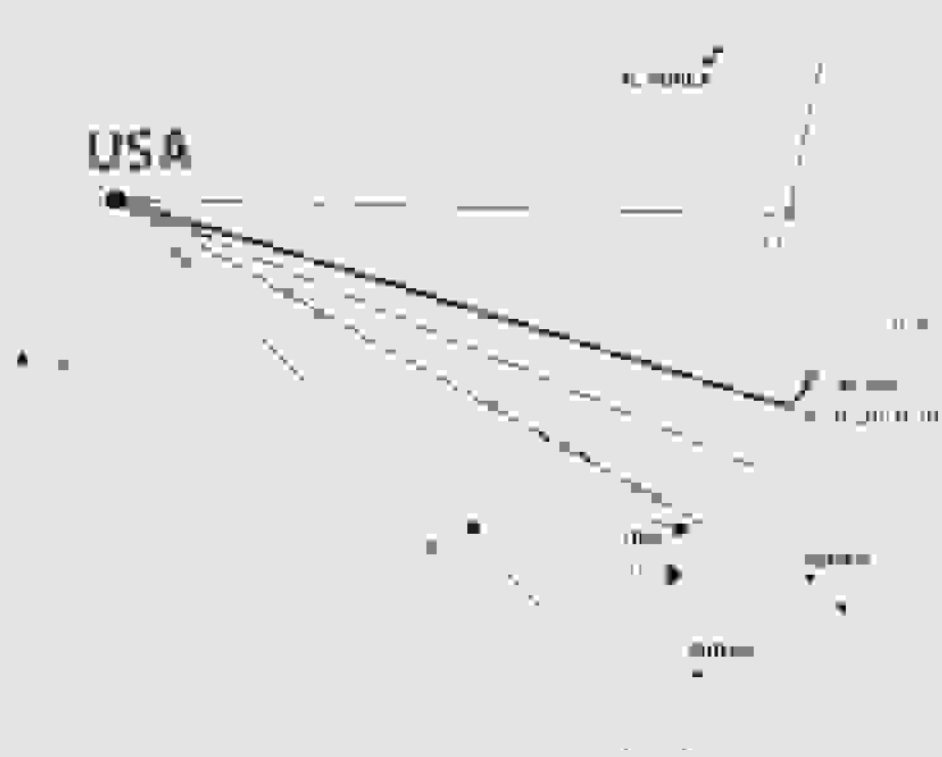
- US as dominant world region with Chinese subregion
- Syrian conflict identified as specific region (with Russia)
- EU region dominated by Germany

# Dominant links of *The Washington Post* (May 2013 – Dec. 2012)

Network



Map

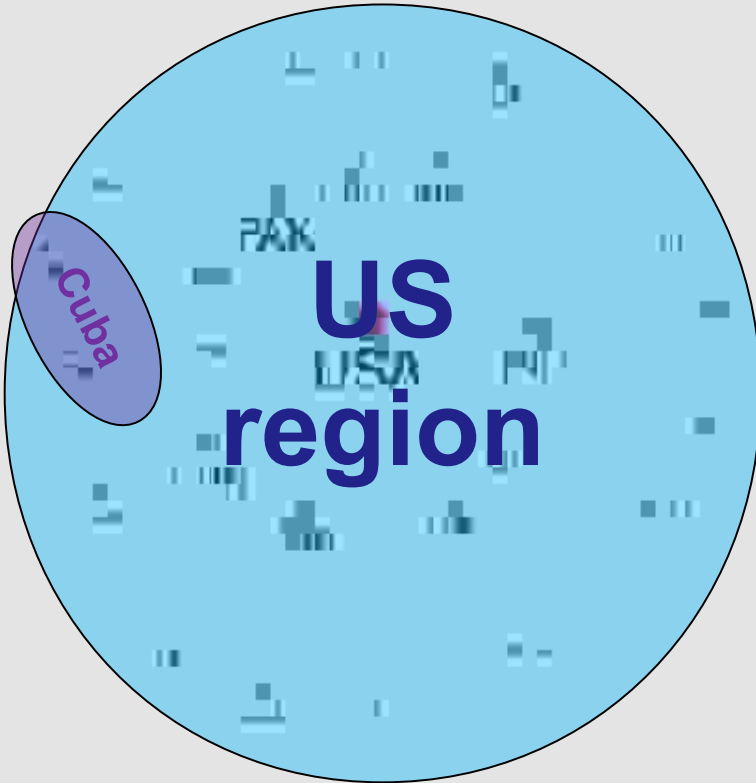


### Comment

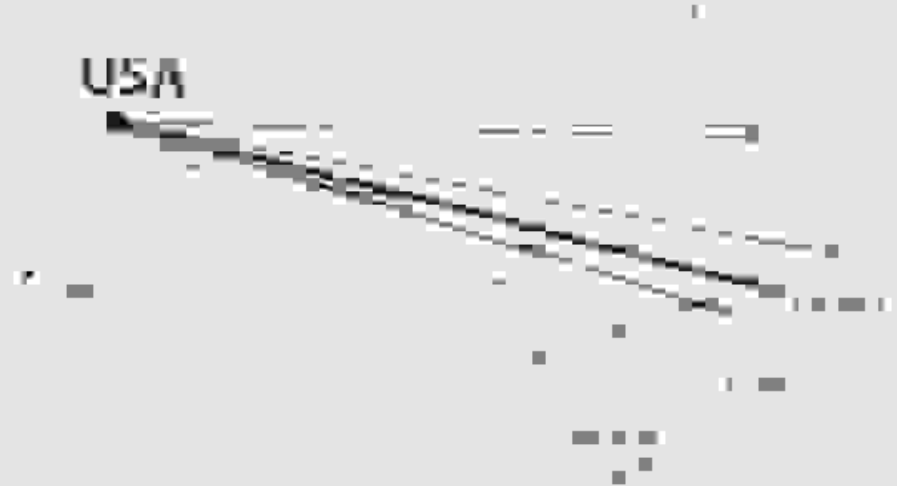
- Major organisation of the world around US (and China)
- Syrian conflict associated to neighbours (not Russia)
- Israel-Palestinian conflict as autonomous region

# Dominant links of *Times of India* (May 2013 – Dec. 2012)

Network



Map

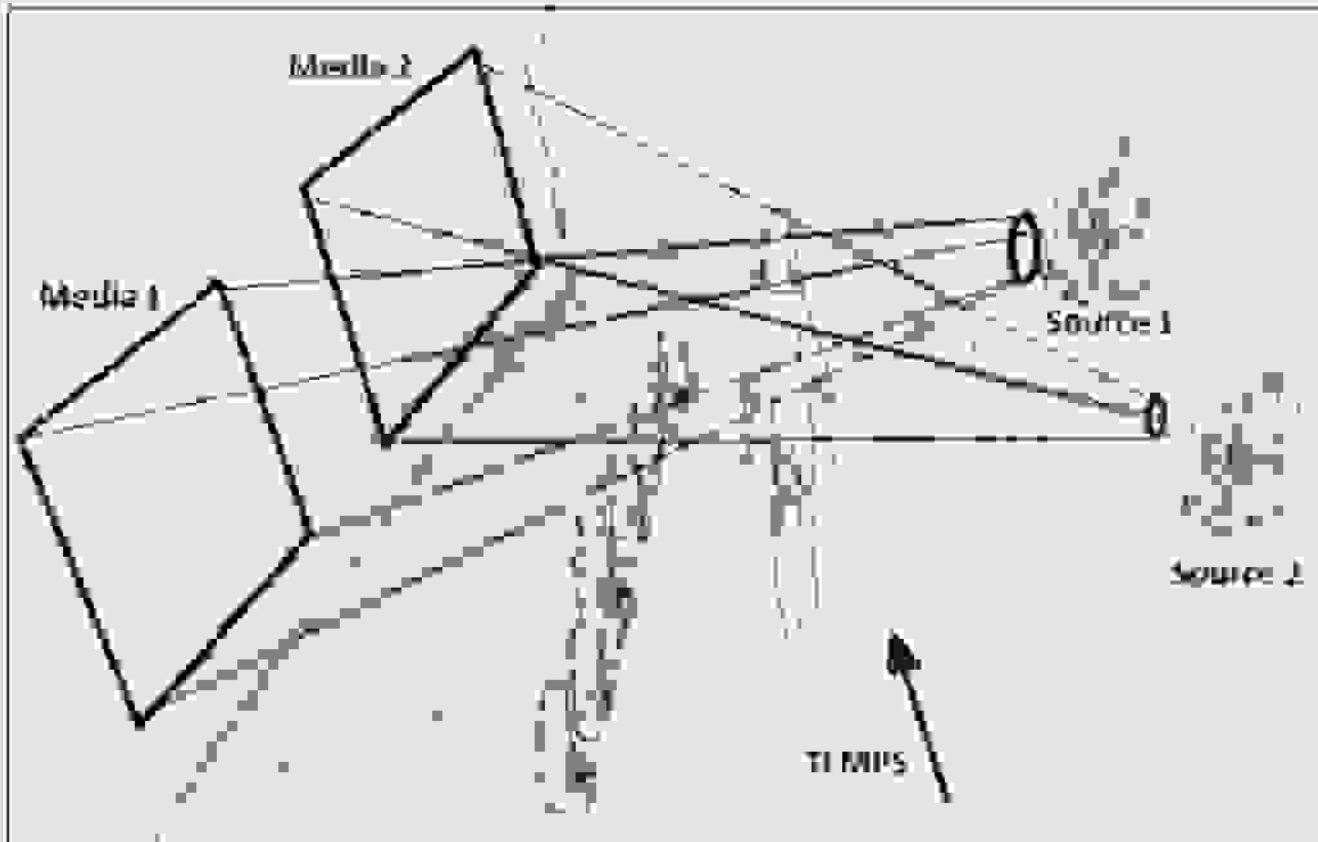


## Comment

- Vision of the world fully dominated by the US
- Syrian conflict as sub region (Russia, Turkey, France)
- Only Cuba-Venezuela escape from full US domination ...



# CONCLUSION



L'Allegorie de la Caverne revisitée

# Validation de l'intérêt des flux RSS pour l'étude des relations internationales ?

1. **Relations directes** : Confirmation de lois connues sur les flux médiatiques.
2. **Relations indirectes** : Mise en évidence de bassins médiatiques récurrents dans les médias.

Distinguer les structures pérennes et les événements conjoncturels.

Merci de votre attention

