



# Un modèle auto-régressif multivarié et multi-niveaux de données ESM émotionnelles

Jean-Luc Kop<sup>1</sup>, Bruno Dauvier<sup>2</sup>, Sarah Le Vigouroux<sup>2</sup>,  
Jean-Baptiste Pavani<sup>2</sup> et Anne Congard<sup>2</sup>

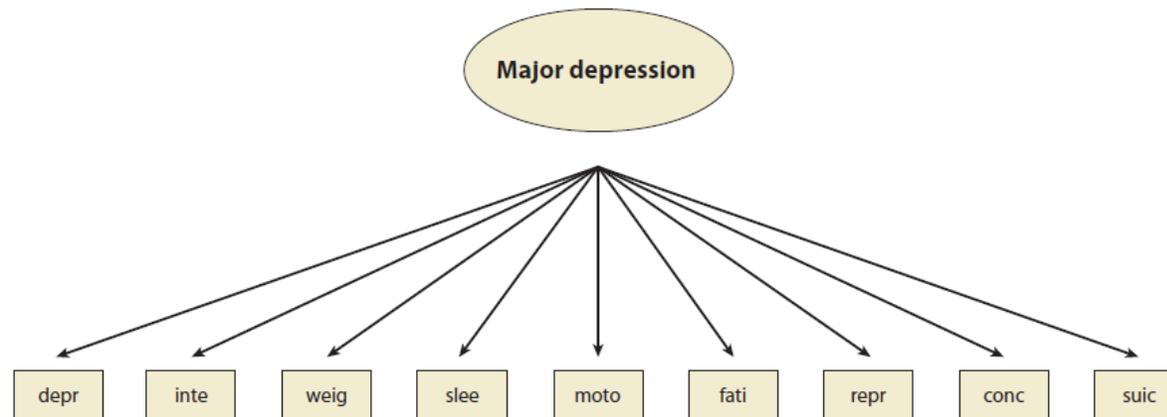
<sup>1</sup> Université de Lorraine

<sup>2</sup> Aix-Marseille Université

# En préambule : Mais où sont les construits ?

## D'une approche en variables latentes vers une approche en réseaux

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9, 91-121.



depr : Depressed mood  
inte : Loss of interest  
weig : Weight problems  
slee : Sleep problems  
moto : Psychomotor disturbances  
fati : Fatigue  
repr : Self-reproach  
conc : Concentration problems  
suic : Suicidal ideation

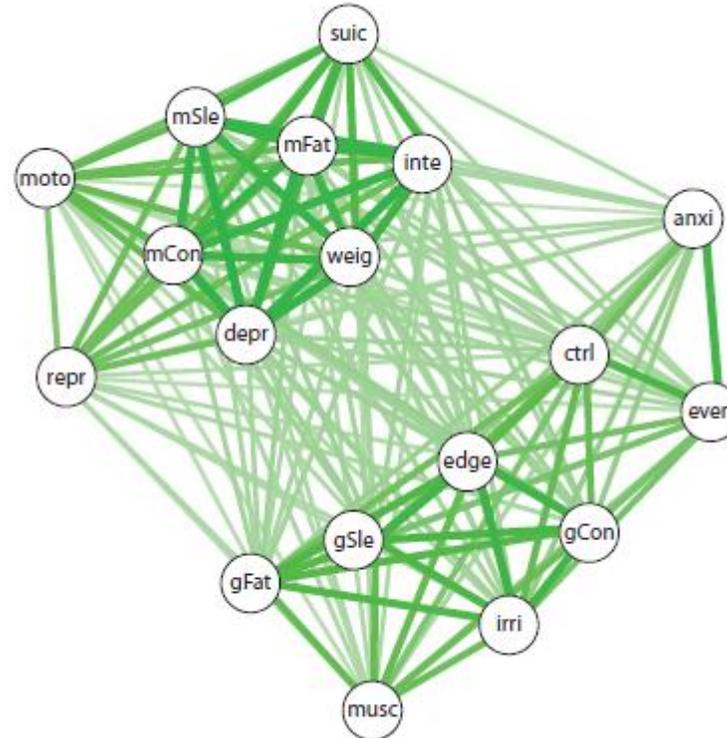
“But where is the university? I have seen where the members of the Colleges live, where the Registrar works, where the scientists experiment, and the rest. But I have not yet seen the University. . .” [Ryle’s (1949, p. 16)]

# En préambule : Mais où sont les construits ?

## D'une approche en variables latentes vers une approche en réseaux

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology, 9*, 91-121.

*In such approaches, disorders are conceptualized as systems of causally connected symptoms rather than as effects of a latent disorder. Network modeling has the philosophical advantage of dropping the unrealistic idea that symptoms of a single disorder share a single causal background, while it simultaneously avoids the relativistic consequence that disorders are merely labels for an arbitrary set of symptoms: It provides a middle ground in which **disorders exist as systems, rather than as entities***



Network for symptoms of major depression (MD) and generalized anxiety disorder (GAD) based on correlations based on the National Comorbidity Survey Replication data.

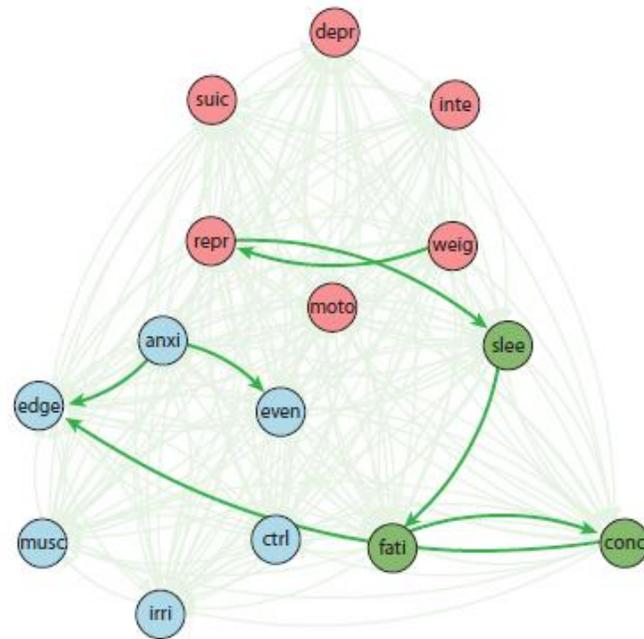
Depr : Depressed mood MD  
Inte: Loss of interest MD  
Weig : Weight problems MD  
Moto : Psychomotor disturbances MD  
Repr : Self-reproach MD  
Suic : Suicidal ideation MD  
mSle : Sleep problems MD  
mFat : Fatigue MD  
mCon : Concentration problems MD  
Slee : Sleep problems MD/GAD  
Conc : Concentration problems MD/GAD  
Fati : Fatigue MD/GAD  
gSle : Sleep problems GAD  
gFat : Fatigue GAD  
gCon : Concentration problems GAD  
Anxi : Chronic anxiety/worry GAD  
Even : Anxiety about >1 event GAD  
Ctrl : No control over anxiety GAD  
Irri : Irritable GAD  
Musc : Muscle tension GAD  
Edge : Feeling on edge GAD

# En préambule : Mais où sont les construits ?

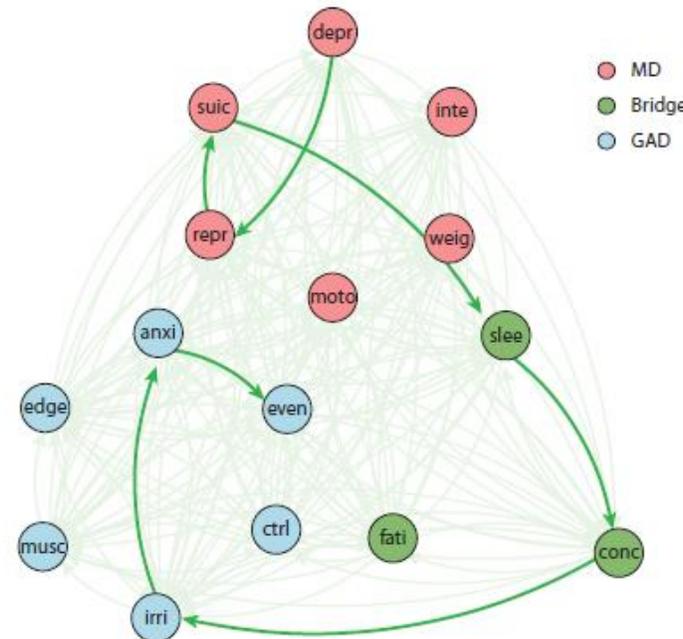
## D'une approche en variables latentes vers une approche en réseaux

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology, 9*, 91-121.

Les interactions entre éléments peuvent être différentes d'une période à l'autre et d'un individu à l'autre.



Bob



Alice

Depr : Depressed mood MD  
Inte: Loss of interest MD  
Weig : Weight problems MD  
Moto : Psychomotor disturbances MD  
Repr : Self-reproach MD  
Suic : Suicidal ideation MD  
mSle : Sleep problems MD  
mFat : Fatigue MD  
mCon : Concentration problems MD  
Slee : Sleep problems MD/GAD  
Conc : Concentration problems MD/GAD  
Fati : Fatigue MD/GAD  
gSle : Sleep problems GAD  
gFat : Fatigue GAD  
gCon : Concentration problems GAD  
Anxi : Chronic anxiety/worry GAD  
Even : Anxiety about >1 event GAD  
Ctrl : No control over anxiety GAD  
Irri : Irritable GAD  
Musc : Muscle tension GAD  
Edge : Feeling on edge GAD

# Comment étudier des réseaux dynamiques ?

Des techniques statistiques existent depuis longtemps pour inférer des réseaux à partir de données empiriques mais elles ne sont pas forcément adaptées aux données psychologiques :

- Les séries temporelles sont souvent courtes
- Les données sont structurées hiérarchiquement (observations emboîtées dans des personnes)

# Un modèle auto-régressif multivarié et multi-niveaux (*multilevel vector autoregression (VAR) model*)

OPEN ACCESS Freely available online



## A Network Approach to Psychopathology: New Insights into Clinical Longitudinal Data

**Laura F. Bringmann<sup>1\*</sup>, Nathalie Vissers<sup>1</sup>, Marieke Wichers<sup>2</sup>, Nicole Geschwind<sup>3</sup>, Peter Kuppens<sup>1</sup>, Frenk Peeters<sup>2</sup>, Denny Borsboom<sup>4</sup>, Francis Tuerlinckx<sup>1</sup>**

<sup>1</sup> Department of Psychology, University of Leuven, Leuven, Belgium, <sup>2</sup> Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, The Netherlands, <sup>3</sup> Department of Clinical Psychological Science, Maastricht University, Maastricht, The Netherlands, <sup>4</sup> Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

# Un modèle auto-régressif multivarié et multi-niveaux (*multilevel vector autoregression (VAR) model*)

AR : modèle auto-régressif (variable au temps  $t$  dépend de la même variable au temps  $t-1$ ,  $t-2$ , etc...)

VAR : modèle auto-régressif multivarié (variable au temps  $t$  dépend en plus d'autres variables au temps  $t-1$ ,  $t-2$ , etc.)

→ Modélisation intra-individuelle

Modèle multi-niveaux permet de modéliser les différences inter-individuelles (i.e. les coefficients de modèles auto-régressifs peuvent varier d'un individu à l'autre [effets aléatoires])

→ Modélisation inter-individuelle

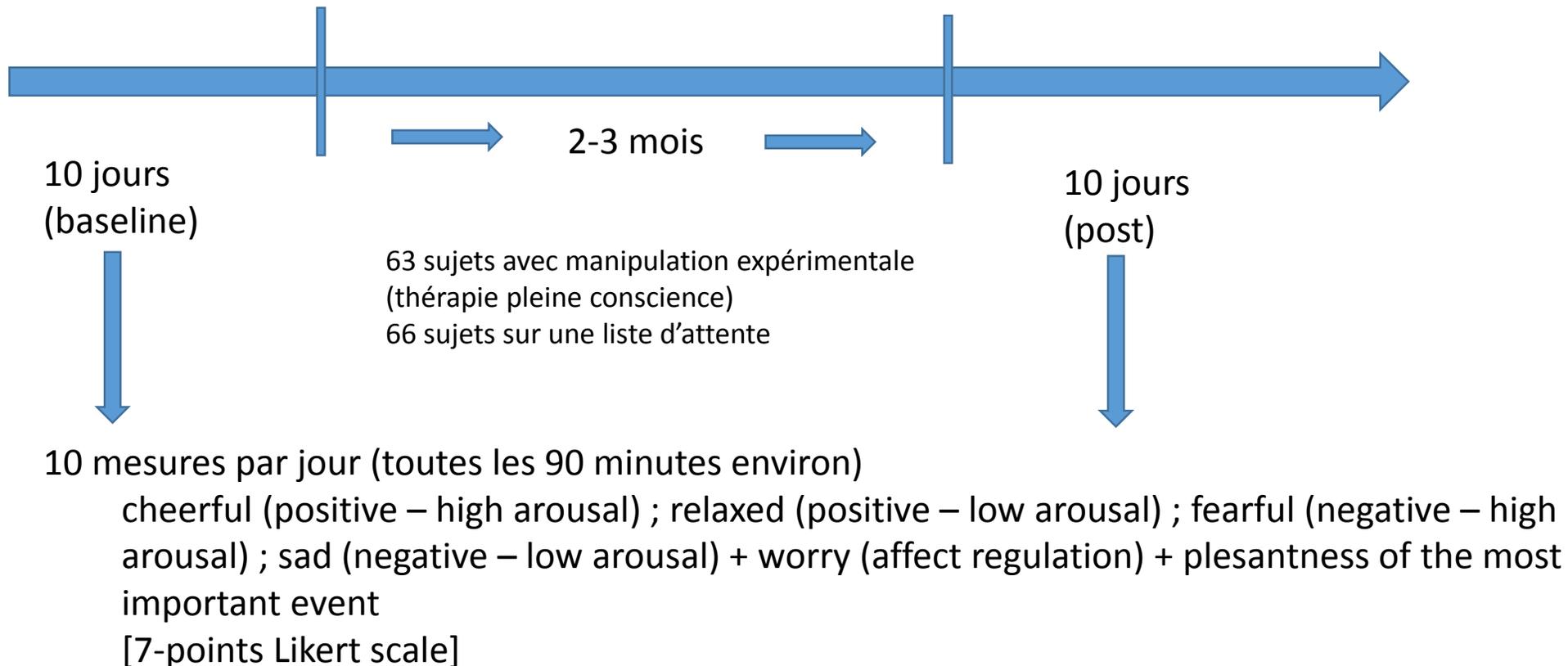


# EXEMPLE

## A Network Approach to Psychopathology: New Insights into Clinical Longitudinal Data

Laura F. Bringmann<sup>1\*</sup>, Nathalie Vissers<sup>1</sup>, Marieke Wichers<sup>2</sup>, Nicole Geschwind<sup>3</sup>, Peter Kuppens<sup>1</sup>,  
Frenk Peeters<sup>2</sup>, Denny Borsboom<sup>4</sup>, Francis Tuerlinckx<sup>1</sup>

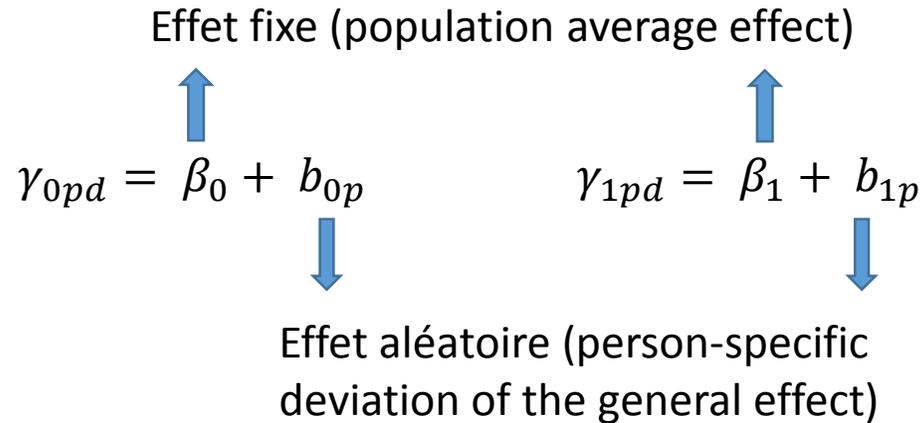
<sup>1</sup>Department of Psychology, University of Leuven, Leuven, Belgium, <sup>2</sup>Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, The Netherlands, <sup>3</sup>Department of Clinical Psychological Science, Maastricht University, Maastricht, The Netherlands, <sup>4</sup>Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands



# Le modèle

$$cheerful_{pdt} = \gamma_{0pd} + \gamma_{1pd} \times cheerful_{p,d,t-1} + \gamma_{2pd} \times sad_{p,d,t-1} + \gamma_{3pd} \times worry_{p,d,t-1} + \gamma_{4pd} \times fear_{p,d,t-1} + \gamma_{5pd} \times event_{p,d,t-1} + \gamma_{6pd} \times relaxed_{p,d,t-1} + \epsilon_{pdt}$$

p : personne  
d : jour  
t : moment



**MATRICE B**  
des effets fixes

	cheerful (L1)	pleasant (L1)	worry (L1)	fearful (L1)	sad (L1)	relaxed (L1)	
cheerful	$\beta_{11}$	$\beta_{12}$	$\beta_{13}$	...			] population average of the network structure
pleasant	...						
worry							
fearful							
sad							
relaxed							

# Estimation

## MATRICE B

des effets fixes

	cheerful (L1)	pleasant (L1)	worry (L1)	fearful (L1)	sad (L1)	relaxed (L1)	
cheerful	$\beta_{11}$	$\beta_{12}$	$\beta_{13}$	...			] population average of the network structure
pleasant	...						
worry							
fearful							
sad							
relaxed							

- Les coefficients sont estimés à partir de 6 modèles univariés (chaque variable est prise tour à tour comme variable dépendante dans 6 modèles différents) → *pseudo-likelihood method*
- La plupart des paramètres sont estimés directement ; les autres (corrélations entre les termes d'erreur des différentes régressions univariées) ne peuvent être estimés qu'indirectement [calculées à partir des corrélations entre valeurs prédites ???].

# Illustration

## Le tableau de données

sujet	jour	beep	group	period	cheerful	pleasant	worry	fearful	sad	relaxed	nevroisme
10720	1	1	NA	0	NA	NA	NA	NA	NA	NA	NA
10720	1	2	NA	0	NA	NA	NA	NA	NA	NA	NA
10720	1	3	NA	0	NA	NA	NA	NA	NA	NA	NA
10720	1	4	0	0	4	3	6	5	4	2	34
10720	1	5	0	0	2	1	5	4	4	2	34
10720	1	6	0	0	3	2	5	3	5	5	34
10720	1	7	NA	0	NA	NA	NA	NA	NA	NA	NA
10720	1	8	0	0	2	2	6	5	3	2	34
10720	1	9	NA	0	NA	NA	NA	NA	NA	NA	NA
10720	1	10	NA	0	NA	NA	NA	NA	NA	NA	NA
10720	1	NA	NA	0	NA	NA	NA	NA	NA	NA	NA
10720	2	1	0	0	5	2	5	3	5	5	34
10720	2	2	0	0	5	-3	6	2	3	5	34
10720	2	3	0	0	3	1	6	3	5	3	34
10720	2	4	0	0	5	-2	5	3	4	4	34
10720	2	5	NA	0	NA	NA	NA	NA	NA	NA	NA
10720	2	6	0	0	4	2	4	2	3	5	34
10720	2	7	NA	0	NA	NA	NA	NA	NA	NA	NA
10720	2	8	0	0	4	2	5	2	2	5	34
10720	2	9	0	0	5	3	2	1	2	6	34
10720	2	10	0	0	6	3	2	1	2	5	34
10720	2	NA	NA	0	NA	NA	NA	NA	NA	NA	NA
10720	3	1	NA	0	NA	NA	NA	NA	NA	NA	NA
10720	3	2	0	0	3	2	5	2	4	3	34
10720	3	3	0	0	3	-2	5	2	3	2	34

On ne tient pas  
compte des  
observations entre  
10<sup>e</sup> mesure du jour J  
et 1<sup>ère</sup> mesure du  
jour J+1

129 sujets \* 10  
jours \* (10 + 1)  
beep \* 2 périodes  
= 28380 lignes

# Illustration

Le tableau de données après ajout des variables décalées (1 seul décalage ici)

sujet	jour	beep	groupe	period	cheerful	pleasant	worry	fearful	sad	relaxed	nevrosisme	cheerfulL1	pleasantL1	worryL1	fearfulL1	sadL1	relaxedL1
10720	1	1	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
10720	1	2	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
10720	1	3	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
10720	1	4	0	0	4	3	6	5	4	2	34	NA	NA	NA	NA	NA	NA
10720	1	5	0	0	2	1	5	4	4	2	34	4	3	6	5	4	2
10720	1	6	0	0	3	2	5	3	5	5	34	2	1	5	4	4	2
10720	1	7	NA	0	NA	NA	NA	NA	NA	NA	NA	3	2	5	3	5	5
10720	1	8	0	0	2	2	6	5	3	2	34	NA	NA	NA	NA	NA	NA
10720	1	9	NA	0	NA	NA	NA	NA	NA	NA	NA	2	2	6	5	3	2
10720	1	10	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
10720	1	NA	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
10720	2	1	0	0	5	2	5	3	5	5	34	NA	NA	NA	NA	NA	NA
10720	2	2	0	0	5	-3	6	2	3	5	34	5	2	5	3	5	5
10720	2	3	0	0	3	1	6	3	5	3	34	5	-3	6	2	3	5
10720	2	4	0	0	5	-2	5	3	4	4	34	3	1	6	3	5	3
10720	2	5	NA	0	NA	NA	NA	NA	NA	NA	NA	5	-2	5	3	4	4
10720	2	6	0	0	4	2	4	2	3	5	34	NA	NA	NA	NA	NA	NA
10720	2	7	NA	0	NA	NA	NA	NA	NA	NA	NA	4	2	4	2	3	5
10720	2	8	0	0	4	2	5	2	2	5	34	NA	NA	NA	NA	NA	NA
10720	2	9	0	0	5	3	2	1	2	6	34	4	2	5	2	2	5
10720	2	10	0	0	6	3	2	1	2	5	34	5	3	2	1	2	6
10720	2	NA	NA	0	NA	NA	NA	NA	NA	NA	NA	6	3	2	1	2	5
10720	3	1	NA	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
10720	3	2	0	0	3	2	5	2	4	3	34	NA	NA	NA	NA	NA	NA
10720	3	3	0	0	3	-2	5	2	3	2	34	3	2	5	2	4	3

# Postulats du modèle

- Egalité des intervalles temporels entre deux mesures (beep aléatoire donc pas d'égalité parfaite)
- Stationarité des séries temporelles
  - 77% des séries selon le Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test
  - BIC de modèles sans tendance est meilleur que celui de modèles avec tendance linéaire
- Nombre de décalages
  - Décalage d'ordre 3 est le meilleur modèle... mais les coefficients de décalage 1 sont très proches dans tous les modèles
  - “the impulse response functions also did not reveal any substantive effects of interest, which could have warranted a more complex analysis, we proceeded with the order-1 results.”

# Estimation des coefficients

Package `lme4` de R

Cf. infra pour l'écriture des modèles