



BRON ur PSH

## Deux exemples d'estimation de paramètres de modèles EDP à partir de données



## Expansion du moustique tigre : estimation d'un modèle couplé 2D/1D

### Introduction

### Le modèle de couplage 2D/1D

### Application à la propagation du moustique tigre

- Les données

- Les modèles

- Modèle probabiliste

- Estimation des maximums de vraisemblance

- Simulations en  $\Theta^*$

- Comparaison des modèles

## Carpocapse : Estimation de coefficients de diffusion avec des données génétiques

### Le paysage

### Les données génétiques

### Description de la démarche

## Quelques exemples de l'influence des éléments linéaires

- transport de la chenille processionnaire du pin par l'homme (Robinet et al. (2012))
- déplacements des loups le long des lignes de prospection sismiques (McKenzie et al. (2012))
- dispersion du moustique tigre le long du réseau routier (Roche et al. (2015))

H. W. McKenzie, E. H. Merrill, R. J. Spiteri, and M. A. Lewis. How linear features alter predator movement and the functional response. *Interface Focus*, 2012.

C. Robinet, C.-E. Imbert, J. Rousselet, D. Sauvard, J. Garcia, F. Goussard, and A. Roques. Human-mediated long-distance jumps of the pine processionary moth in europe. *Biological Invasions*, 2012.

B. Roche, L. Léger, G. L'Ambert, G. Lacour, R. Foussadier, G. Besnard, H. Barré-Cardi, F. Simard, and D. Fontenille. The spread of aedes albopictus in metropolitan france : Contribution of environmental drivers and human activities and predictions for a near future. *PLoS ONE*, 2015.

## Le modèle

On s'intéresse à un modèle étudié par (Berestycki et al. (2012)) qui propose un couplage 2D/1D :

$$\begin{cases} \partial_t v = d \Delta v + f(v), & t > 0, (x, y) \in \mathbb{R} \times (0, +\infty), \\ \partial_t u = D \partial_{xx} u + \nu v(t, x, 0) - \mu u, & t > 0, x \in \mathbb{R}, \\ -d \partial_y v(t, x, 0) = \mu u(t, x) - \nu v(t, x, 0), & t > 0, x \in \mathbb{R}, \end{cases}$$

avec,

- $v$  (resp.  $u$ ) représente une densité de population surfacique (resp. linéique),
- $d$  (resp.  $D$ ) représente le coefficient de diffusion surfacique (resp. linéique),
- $\mu$  est le taux de transfert du 1D vers le 2D,
- $\nu$  est le taux de transfert du 2D vers le 1D,
- $f$  est le terme de croissance Fisher-KPP de la population dans le 2D.

H. Berestycki, J.-M. Roquejoffre, and L. Rossi. The influence of a line with fast diffusion on fisher-kpp propagation. *Biological Invasions*, 2012.

## Le modèle

On s'intéresse à un modèle étudié par (Berestycki et al. (2012)) qui propose un couplage 2D/1D :

$$\begin{cases} \partial_t v = d \Delta v + f(v), & t > 0, (x, y) \in \mathbb{R} \times (0, +\infty), \\ \partial_t u = D \partial_{xx} u + \nu v(t, x, 0) - \mu u, & t > 0, x \in \mathbb{R}, \\ -d \partial_y v(t, x, 0) = \mu u(t, x) - \nu v(t, x, 0), & t > 0, x \in \mathbb{R}, \end{cases}$$

avec,

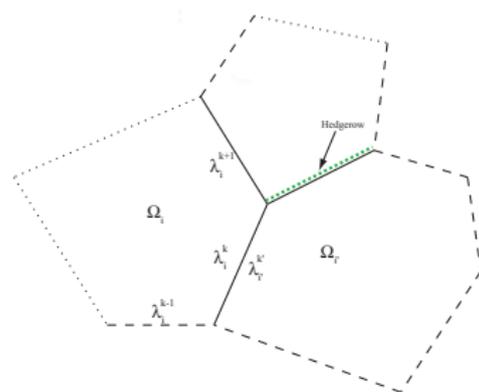
- $v$  (resp.  $u$ ) représente une densité de population surfacique (resp. linéique),
- $d$  (resp.  $D$ ) représente le coefficient de diffusion surfacique (resp. linéique),
- $\mu$  est le taux de transfert du 1D vers le 2D,
- $\nu$  est le taux de transfert du 2D vers le 1D,
- $f$  est le terme de croissance Fisher-KPP de la population dans le 2D.

**Théorème** : Si  $D > 2d$ , alors il y a accélération de la propagation dans tout le demi plan.

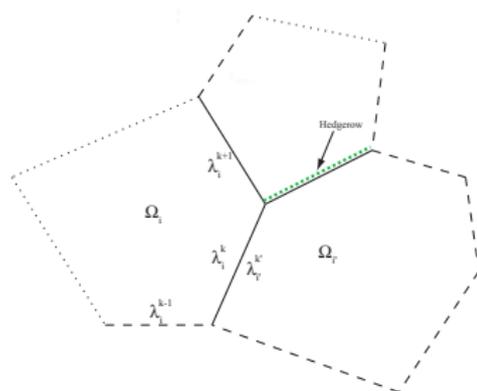
Comment appliquer ce modèle à des paysages plus réalistes ?

H. Berestycki, J.-M. Roquejoffre, and L. Rossi. The influence of a line with fast diffusion on fisher-kpp propagation. *Biological Invasions*, 2012.

## Extension du modèle sur un paysage composé de 3 parcelles



## Extension du modèle sur un paysage composé de 3 parcelles



Dans chaque domaine 2D  $\Omega_i$ ,

$$\partial_t v_i = d \Delta v_i + f_i(v_i),$$

avec les conditions aux bords :

$$d \nabla v_i \cdot \mathbf{n} = \mu_i^k u_i^k(t, x, y) - \nu_i^k v_i(t, x, y),$$

Sur les bordures  $\lambda_i^k$ ,

$$\begin{aligned} \partial_t \tilde{u}_i^k &= D \partial_{zz} \tilde{u}_i^k + \nu_i^k v_i(t, x(z), y(z)) - \mu_i^k \tilde{u}_i^k(t, z) \\ &\quad - \alpha_i^k \tilde{u}_i^k(t, z) + \alpha_i^{k'} \tilde{u}_i^{k'}(t, z) + g_i^k(\tilde{u}_i^k), \quad t > 0, z \in (0, L(\lambda_i^k)), \end{aligned}$$

avec des conditions aux bords adaptées au besoin de modélisation :

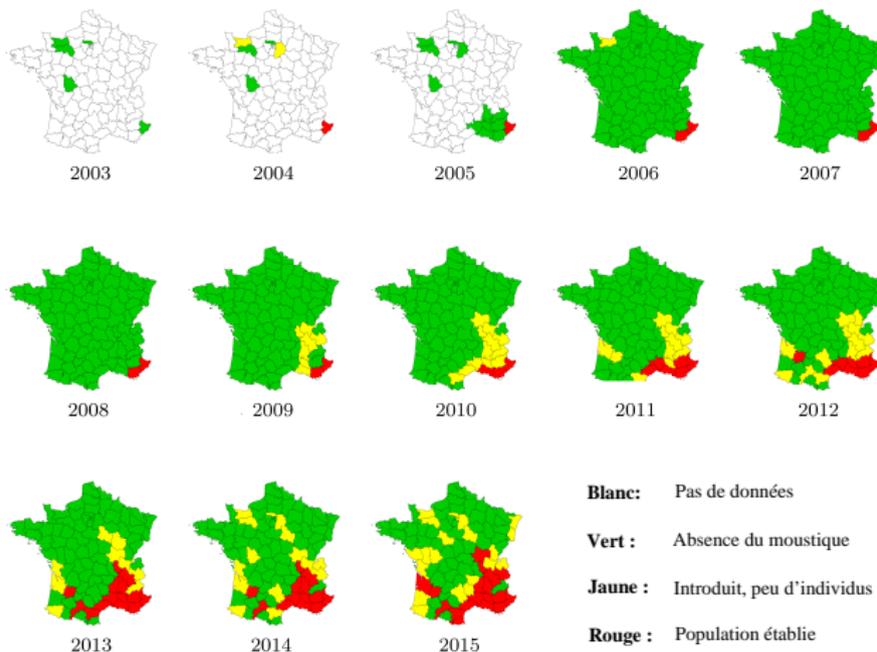
$$\left\{ \begin{array}{l} \tilde{u}_i^k(t, 0) = \tilde{u}_i^{k-1}(t, L(\lambda_i^{k-1})) \\ \partial_z \tilde{u}_i^k(t, 0) = \partial_z \tilde{u}_i^{k-1}(t, L(\lambda_i^{k-1})) \end{array} \right\}, \quad \left\{ \begin{array}{l} \tilde{u}_i^k(t, \cdot) = u_s \\ \partial_z \tilde{u}_i^k(t, \cdot) = 0 \end{array} \right\}.$$

## Notre démarche consiste à :

- Définir un modèle 2D/1D paramétré
- Définir un modèle 2D paramétré
- Définir un modèle probabiliste
- Mettre en oeuvre une procédure d'optimisation
- Comparer les modèles

**Remarque :** Pour cette étude, nous ignorons les effets du climat et du type d'occupation des sols.

# Les données (source : Entente Interdépartementale pour la Démoustication)



On note  $\text{Obs} := \{\mathcal{O}_j^t \in \{0, 1, 2\}, t = 2004, \dots, 2015, j \in J^t\}$



$$\begin{aligned}f_i(v_i) &= r v_i(1 - v_i), g_i^k = 0, \\ \mu_i^k &= \mu, v_i^k = \nu, \\ \alpha_i^k &= \alpha \gg 1,\end{aligned}$$

avec la condition au bord :

$$\begin{cases} u(t, x_s) = u_s, & \text{if } t \in [2003, 2004] \\ \nabla u_x(t, x_s) = 0 & \text{if } t > 2004 \end{cases}$$
$$\Theta = \{r, d, D, \mu, \nu, u_s\}$$

## Les modèles



$$\begin{aligned} f_i(v_i) &= r v_i(1 - v_i), g_i^k = 0, \\ \mu_i^k &= \mu, \nu_i^k = \nu, \\ \alpha_i^k &= \alpha \gg 1, \end{aligned}$$

avec la condition au bord :

$$\begin{cases} u(t, x_s) = u_s, & \text{if } t \in [2003, 2004] \\ \nabla u_x(t, x_s) = 0 & \text{if } t > 2004 \end{cases}$$

$$\Theta = \{r, d, D, \mu, \nu, u_s\}$$



$$\begin{cases} \partial_t v = d \Delta v + r v(1 - v), & x \in \Omega, \\ \nabla v \cdot \mathbf{n} = 0, & x \in \partial\Omega \setminus \Gamma_s, \\ \begin{cases} v = v_s, & \text{if } t \in [2003, 2004] \\ \nabla v \cdot \mathbf{n} = 0 & \text{if } t > 2004 \end{cases} & x \in \Gamma_s, \end{cases}$$

$$\Theta = \{r, d, v_s\}$$

# Modèle probabiliste



**Approche mécanistico-statistique** : Faire le lien entre les données et la densité de population simulée.

**Approche mécanistico-statistique** : Faire le lien entre les données et la densité de population simulée.

On suppose que les variables d'observations  $\mathcal{O}_j^t$  sont indépendantes et suivent la loi binomiale :

$$\mathcal{O}_j^t \sim \mathcal{B}(2, p(V_j^t)), \text{ ie :}$$

$$P(\mathcal{O}_j^t = 0) = (1 - p(V_j^t))^2, P(\mathcal{O}_j^t = 1) = 2(1 - p(V_j^t))(p(V_j^t)), P(\mathcal{O}_j^t = 2) = p(V_j^t)^2,$$

où  $V_j^t$  est la densité moyenne du département  $j$  à l'instant  $t$  et  $p(V_j^t) = \min(V_j^t, 1)$ .

On peut alors écrire la vraisemblance des observations :

$$P(\text{Obs}|\Theta) = \mathcal{L}(\Theta) = \prod_{t=2004}^{2015} \prod_{j \in J^t} \sum_{z=0}^2 \mathbb{1}(\mathcal{O}_j^t = z) \frac{2!}{z!(2-z)!} (p(V_j^t))^z (1 - p(V_j^t))^{2-z}.$$

Nous avons utilisé deux estimateurs de maximum de vraisemblance :

- 1 un algorithme de **recuit-simulé** :

---

## Algorithm 1 recuit-simulé

---

Require:  $X, V, T=2000$

Ensure:  $L$

repeat

$Y \leftarrow \sim N(X, V)$

$DF \leftarrow F(Y) - F(X)$  {avec  $F(X) = -\ln(\mathcal{L}(X))$ }

$ACCEPT \leftarrow [(DF < 0) \text{ ou } \sim \text{Bernoulli}(\exp(-DF/T))]$

if  $ACCEPT$  then

$X \leftarrow Y$

AppendTo[ $L, X$ ]

end if

$T \leftarrow T * 0.999$

until

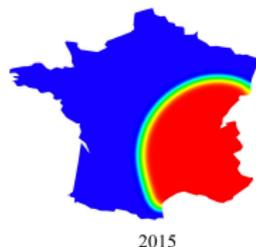
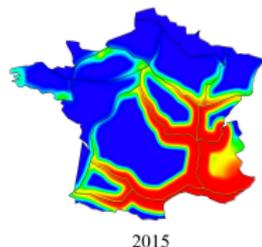
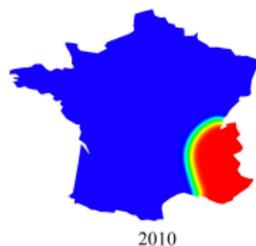
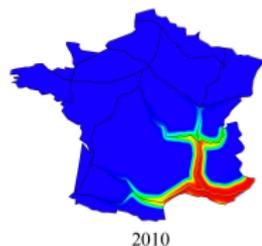
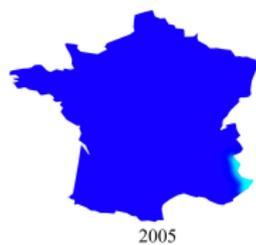
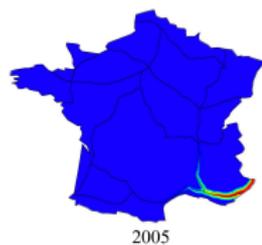
return  $L$

---

- 2 une méthode de **quasi-Newton BFGS** (MATLAB<sup>®</sup> fmincon) :

- évaluation d'un gradient numérique,
- recherche linéaire cubique.

# Simulations en $\Theta^*$



## Comparaison des modèles

- le Bayesian information criterion (BIC) (Schwarz et al. (1978)),  
 $BIC = -2 \ln(\mathcal{L}) + n \ln(N)$ , avec :
  - ▶  $n$ , le nombre de paramètres
  - ▶  $N$ , le nombre d'observations (=959)
- le half-Brier score (BS) (Brier (1950)) et 'Brier skill score'

$$BS_{model}(t) = \frac{1}{2N_t} \sum_{j \in J^t} \sum_{z=0}^2 (P_j^t(z) - \mathbb{1}(O_j^t = z))^2$$
$$BSS(t) = 1 - \frac{BS_{2D1D}(t)}{BS_{2D}(t)}$$

	$-\ln(\mathcal{L}(\Theta^*))$	BIC	BSS(2013)	BSS(2014)	BSS(2015)
2D	3037	6094	-	-	-
2D1D	365	771	0.46	0.44	0.49

G. W. Brier. Verification of forecasts expressed in terms of probability. *Monthly weather review*, 1950.

G. Schwarz et al. Estimating the dimension of a model. *The annals of statistics*, 1978.

# Carpocapse : Estimation de coefficients de diffusion avec des données génétiques

Le travail (Roques et al. (2015)) présente une approche mécanico-statistique pour l'estimation de la diffusion hétérogène en écrivant une vraisemblance basée sur des données génétiques spatialisées.

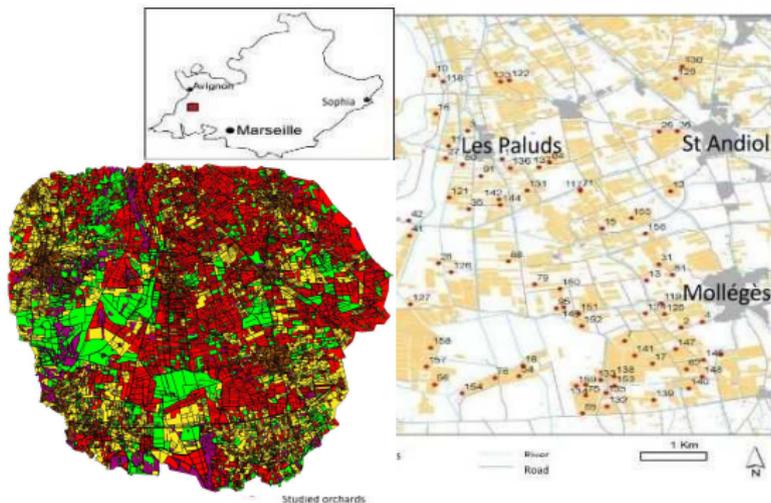
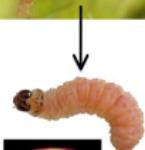
Le stage de Boris Minchenski consiste à adapter cette démarche aux données génétiques du Carpocapse.

L. Roques, E. Walker, P. Franck, S. Soubeyrand, and E. K. Klein. Using genetic data to estimate diffusion rates in heterogeneous landscapes. *Journal of Mathematical Biology*, pages 1–26, 2015. ISSN 1432-1416. doi : 10.1007/s00285-015-0954-4. URL <http://dx.doi.org/10.1007/s00285-015-0954-4>.

## Carpocapse du pommier : *Cydia Pomonella*

Larves → dégâts sur pommes et poires.

- Zone d'étude : 100km<sup>2</sup>
- 5 types d'occupation des sols : culture pérenne/annuelle, bois, urbain, prairie



# les données génétiques

## Échantillonnage :

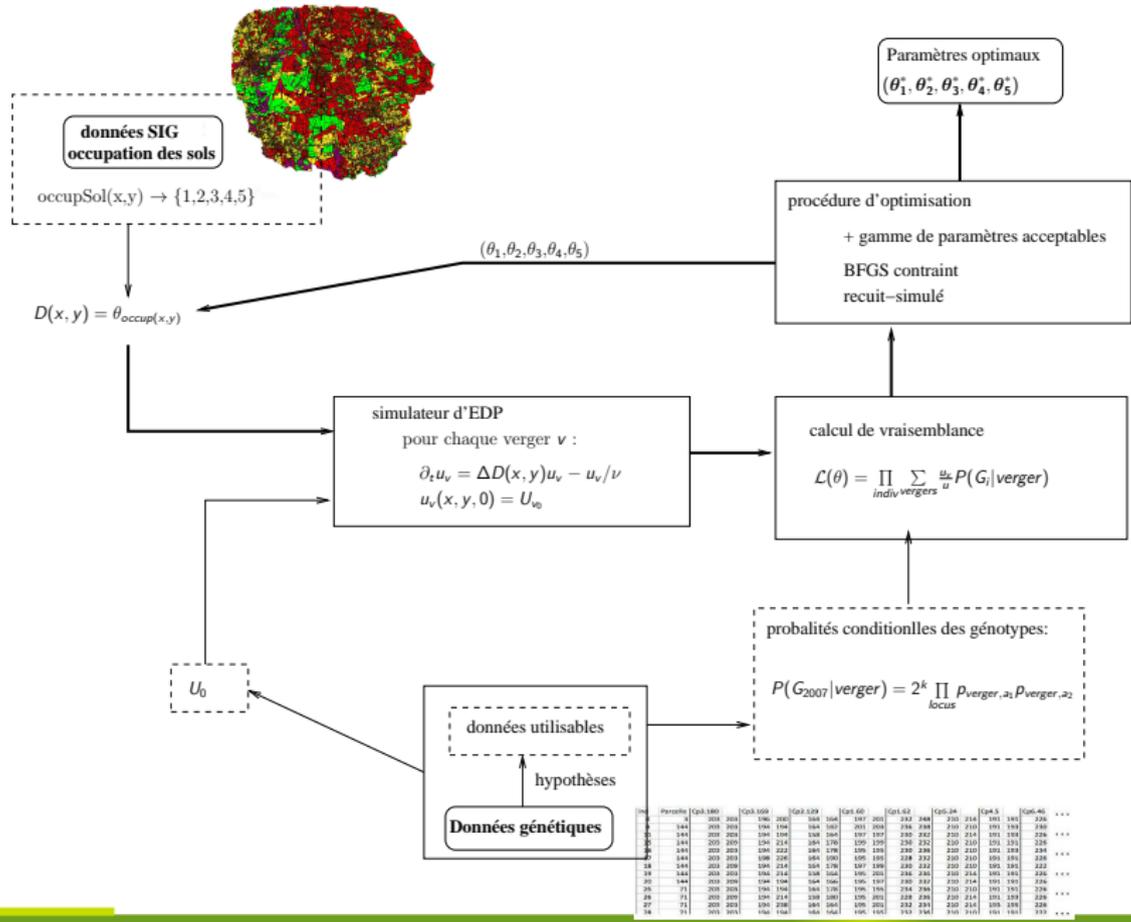
- 1000 individus génotypés en 2006 et 2007
- 51 vergers

## Données génétiques :

- 20 loci
- entre 4 et 51 allèles par locus



Ind	Parcelle	Cp3.180	Cp3.169	Cp2.129	Cp1.60	Cp1.62	Cp5.24	Cp4.5	Cp6.46	...
4	a	203 203	196 200	164 164	197 201	242 248	210 214	191 191	226	
8	144	203 203	194 194	164 182	201 203	236 238	210 210	191 193	230	
11	144	203 203	194 194	158 184	197 197	230 232	210 214	191 193	228	...
15	144	203 209	194 214	164 178	199 199	230 232	210 210	191 191	226	
16	144	203 203	194 222	164 178	195 195	230 236	210 210	191 193	234	...
17	144	203 203	198 226	164 190	195 195	228 232	210 210	191 191	226	
18	144	203 209	194 214	164 178	197 199	230 232	210 210	191 191	222	
19	144	203 203	194 214	158 164	195 201	236 236	210 214	191 191	226	...
20	144	203 209	194 194	164 166	196 197	240 242	210 214	191 191	236	
20	71	203 203	194 194	164 178	195 195	234 236	210 210	191 191	226	
26	71	203 209	194 214	158 180	195 201	228 236	210 214	191 193	226	
27	71	203 203	194 238	164 164	195 201	232 234	210 214	193 195	226	
28	71	203 203	194 164	164 164	195 195	232 236	210 210	191 191	222	...



Individu	Gen1-100	Gen1-101	Gen1-102	Gen1-103	Gen1-104	Gen1-105	Gen1-106	Gen1-107	Gen1-108	Gen1-109	Gen1-110	Gen1-111	Gen1-112	Gen1-113	Gen1-114	Gen1-115	Gen1-116	Gen1-117	Gen1-118	Gen1-119	Gen1-120
1	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
2	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
3	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
4	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
5	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
6	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
7	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
8	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
9	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
10	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
11	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
12	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
13	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
14	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
15	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
16	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
17	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
18	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
19	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
20	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
21	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
22	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
23	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
24	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
25	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
26	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
27	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
28	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
29	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
30	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
31	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120

# QUESTIONS ?

Merci de votre attention