



Horaire de la journée

10:00 - 10:10	Début de l'atelier
10:10 - 11:00	Partie 1: Notion de seuil épidémique
11:00 - 12:00	Partie 2: Introduction au logiciel de régression périodique
12:00 - 13:45	Dîner
13:45 - 15:15	Partie 3: Exercice pratique (individuel ou en sous-groupes)
15:15 - 15:30	Pause
15:30 - 16:00	Partie 3 (suite): Exercice pratique – Interprétation et discussion des résultats
16:00 - 16:30	Partie 4: Autres utilisations et limites
16:30 - 17:05	Discussion et synthèse de la journée

Objectifs de l'atelier

- ▶ Expliquer le concept de seuil épidémique
- ▶ Énumérer et décrire les principales méthodes de calcul de seuil épidémique et leur utilité en vigie et surveillance des maladies infectieuses
- ▶ Utiliser la méthode de régression périodique pour calculer des seuils épidémiques
- ▶ Interpréter les résultats de seuils épidémiques obtenus par la méthode utilisée
- ▶ Discuter des limites de la méthode de régression périodique

Brise glace 1!

- ▶ Avez-vous entendu parler des seuils épidémiques?

Brise glace 2!

- ▶ Décrivez en deux mots ce qui vous vient en tête lorsqu'on parle de seuil épidémique

Brise glace 3

- ▶ Dans quel contexte utilisez-vous les seuils épidémiques?

Partie 1: Notions de seuil épidémique

Notions de seuil épidémique

À partir de combien de cas devons-nous déclarer une écloson ou excès de cas ?



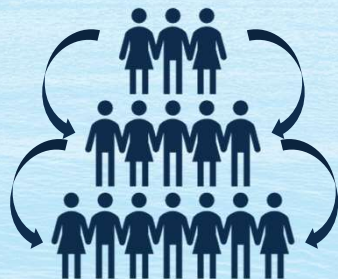
Seuil épidémique

- Clé de voûte lorsqu'une écloson est signalée
- Détermination subjective
- D'autres facteurs peuvent aider à déterminer si l'épidémie sera majeure ou pas

Quizz !

- ▶ Quels sont les facteurs majeurs pouvant servir à déterminer si une éclosion sera majeure ou pas?

Taux de reproduction de base

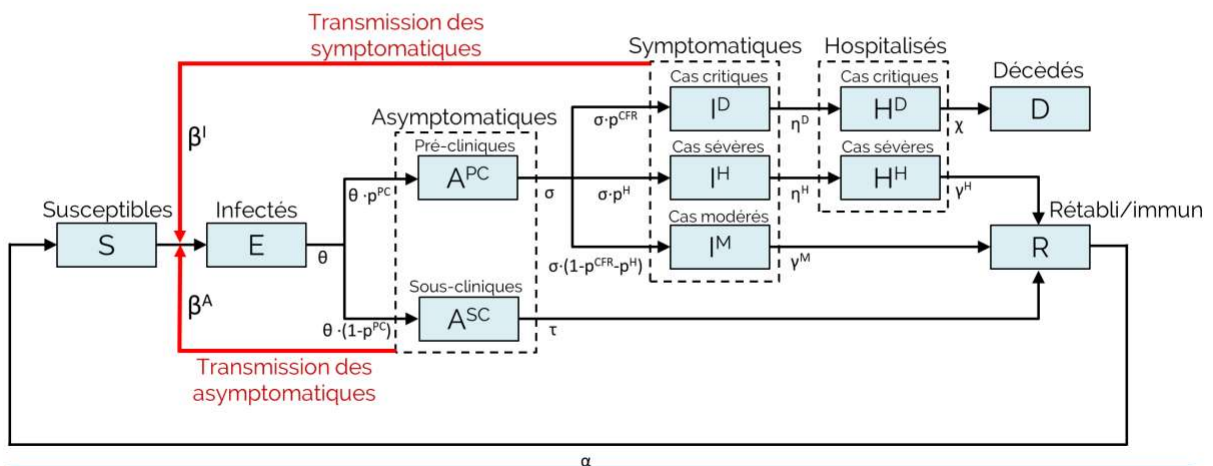


Distribution de la courbe épidémique



Description du modèle

Diagrammes de flux – Modèle dynamique stochastique compartimental

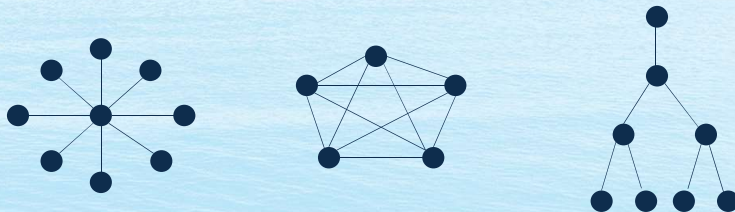


Source: Modélisation de l'évolution de la COVID-19 au Québec.
<https://www.inspq.qc.ca/sites/default/files/covid/inspq-projections-21janvier2021.pdf>

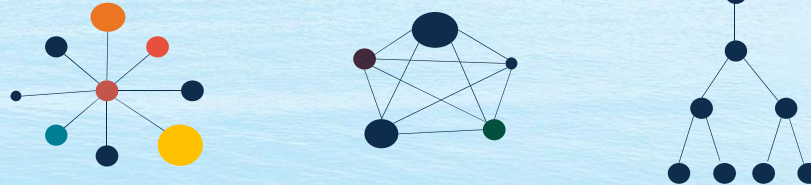
Comportements face à la maladie



Connectivité/structure du réseau de contacts



Hétérogénéité des réseaux



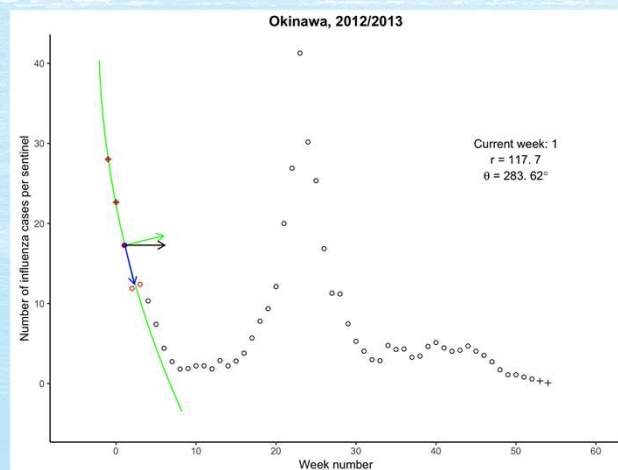
Notions de seuil épidémique...

- ▶ Comment sont calculés les seuils épidémiques ?
- ▶ Les valeurs de seuil sont-elles spécifiques à une région donnée ou est-ce généralisable ?
- ▶ Est-ce possible d'avoir une même valeur seuil pour des agents pathogènes possédant une chaîne de transmission similaire ?

Exemples d'approches méthodologiques

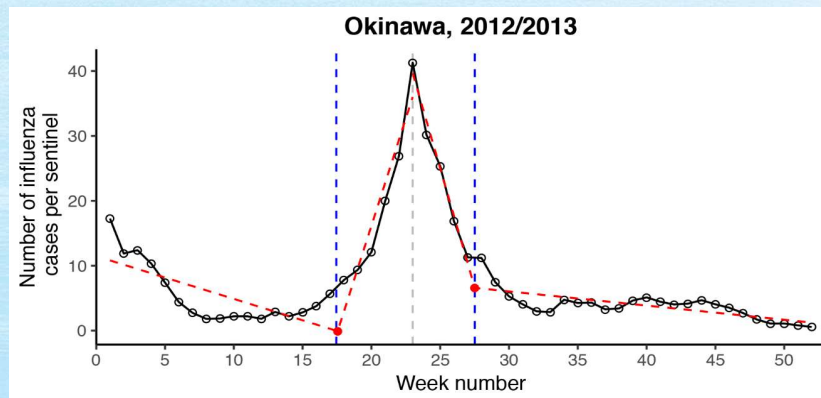
- ▶ Méthode de courbure maximale
- ▶ Méthode de régression segmentée
- ▶ Méthode de régression logistique mobile
- ▶ Modèle de régression périodique

Méthode de courbure maximale



(Cai et al. 2019; <https://github.com/caijun/MCM>)

Méthode de régression segmentée



Méthode de régression logistique mobile

Postulat initial:

- une courbe épidémique symétrique peut être approximée avec un modèle logistique de croissance de population de la forme :

$$y = \frac{k}{1+ae^{-bt}}$$

- ▶ k , a et b : paramètres à estimer (constantes)
- ▶ y : proportion cumulative des spécimens testés pour une maladie X et ayant un résultat de laboratoire positif
- ▶ t : semaine épidémiologique

<https://doi.org/10.1371/journal.pone.0202880>

Méthode de régression logistique mobile...

Étapes d'analyse:

1. Première régression logistique est appliquée sur les données correspondant à la période de pic épidémique et en ajoutant deux semaines de chaque côté de celui-ci;
2. Courbe logistique ajustée (de manière continue au fur et à mesure que les données hebdomadaires de chaque côté du pic sont ajoutés), jusqu'à la fin de la période concernée;
3. Coefficient de détermination (R^2) est calculé par régression linéaire, entre les valeurs observées et estimées.

Le seuil épidémique est ainsi établi à partir de la valeur critique de R^2 .

Modèle de régression périodique

- ▶ **Deux types d'analyses** peuvent être effectués via l'algorithme de régression périodique :
 - ▶ **Rétrospective** afin de localiser et quantifier l'impact des épidémies passées.
 - ▶ **Prospective** utilisée dans le cadre de la surveillance épidémiologique.

Partie 2: Introduction au logiciel de régression périodique

Modèle de régression périodique

▶ <https://periodic.sentiweb.fr/>

The screenshot shows the homepage of the 'Periodic' website. The header is blue with the title 'Periodic - Analysis for periodic epidemics' and navigation links for Home, Datasets, Periodic Regression, and Hidden Markov Model. The main content area is white and features a 'Welcome on Periodic' message. Below this, there are sections for 'Purpose', 'What is new?', and 'Features'. The 'Features' section lists three main capabilities: 'Manage your datasets', 'Periodic Regression Algorithm' (marked as 'ready'), and 'Hidden Markov Model Algorithm' (marked as 'beta').

Periodic - Analysis for periodic epidemics

Home Datasets Periodic Regression Hidden Markov Model

Welcome on Periodic

Purpose
This website aims to provide some epidemic period detection & burden estimation online tools from the works of the French Sentinelles network's research team
This is an enhanced version of "Periodic Regression" website described in Pelat C et al, *BMC MIDW*, 2007

What is new ?
A new website design
New algorithms available : Hidden Markov model (in test)

Features

- Manage your datasets**
View example datasets, add you own data
- Periodic Regression Algorithm** ready
Periodic regression models are used to estimate an expected baseline level for the time series, associated with a prediction interval. They can extract the dates of epidemics and estimate the related morbidity/mortality burden. They can also forecast alert thresholds to perform real time surveillance of diseases.
- Hidden Markov Model Algorithm** beta
HMM is a non parametric algorithm used to determine epidemics periods

Dataset managements Detection Methods

Mise en forme des données

Formats possible :

- Quotidien
- Hebdomadaire
- Mensuel

Concaténer les colonnes pour en former qu'UNE seule en séparant les valeurs par un « ; »
Formule Excel: (=CONCAT(A2;"";B2))

	A	B	C
1	Période (AAAA-SemCDC)	Nbre de cas	Période;Nombre de cas
2	2021-4	12	2021-4;12
3	2021-5	34	2021-5;34
4	2021-6	40	2021-6;40
5	2021-7	23	2021-7;23
6	2021-8	21	2021-8;21
7	2021-9	23	2021-9;23
8	2021-10	NA	NA

NA pour les valeurs manquantes

Modèle de régression périodique

- ▶ Étape 1 : Importer les données en fichier texte OU copier-coller dans l'application et choisir le pas de temps.

Import a personal dataset

Import data Dataset Options

Value Column

Choose the time step of your data set * day month week

Starting date Day: 1 Month: 1 Year: 2019

Dataset name

Dataset units

Submit

Select the column name that holds data values

Choose the time step of your data set

Enter the first date value of your data set.
For month : any date of the first month.
For week : any date of the first week, using ISO8601 week (first day of this kind of week number is monday)

You can optionally set an dataset name (used as the title of the generated plots)

You can optionally set an dataset units (used as label of axis in generated plots)

Modèle de régression périodique

Étape 2: choix de l'analyse (prospective ou rétrospective)

Periodic - Analysis for periodic epidemics

Home Datasets Periodic Regression Hidden Markov Model Give us a feedback

Periodic Regression Select Datasets Analysis method Prune dataset Select model Model results Help page

What kind of analysis do you want to conduct?

Prospective: Surveillance

The baseline model is fitted on a subset of data - the "training period". The model is used to forecast the baseline level and the epidemic threshold for the following year. The baseline level and the epidemic threshold will be output to an excel file. You can download it and use it to perform real-time surveillance by simply pasting your data in the observation column.

Retrospective: Estimation of burden

The model is fitted over all the data to estimate the baseline level and the epidemic threshold for the whole observation period. It allows to extract the dates of the epidemics and to measure the excess during these epidemic periods.

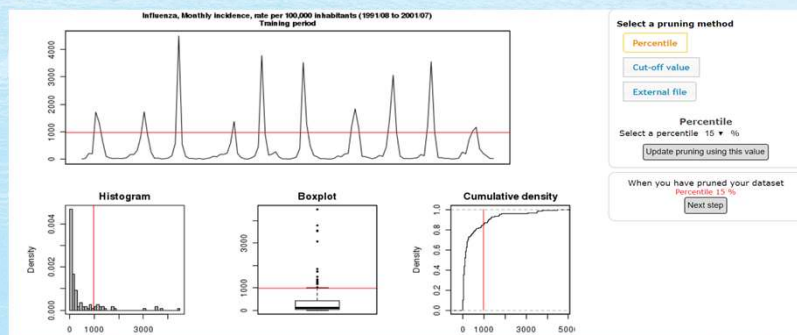
Submit

Two kind of analysis are proposed :

- Prospective analysis: the goal is to timely detect an epidemic event. Information from the past is used in the periodic regression model to forecast the epidemic threshold for the following year.
- Retrospective analysis: the goal is to extract the dates and size of epidemics in historical data. The model is fitted on the whole dataset to estimate the non-epidemic baseline level and the epidemic threshold. A rule is then applied to find the epidemics: e.g. three observations in a row higher than the threshold define an epidemic.

Modèle de régression périodique

Étape 3: visualisation statistique des données et sélection d'une méthode dite de « pruning » pour affiner la décision finale



Modèle de régression périodique

- Étape 4: sélection du modèle qui servira de base en fonction de deux paramètres: tendance et périodicité

Periodicities

1 year	1 year + 6 months	1 year + 3 months	6 months
--------	-------------------	-------------------	----------

Trend

Linear	M1.1	M1.2	M1.3
Quadratic	M2.1	M2.2	M2.3
Cubic	M3.1	M3.2	M3.3

Model selected with the automated selection algorithm

Selected model:
model selection algorithm based on ANOVA comparison and AIC criterion

Choose the prediction interval for the estimated baseline level
 % The upper limit of the prediction interval is used as a threshold to detect epidemics.

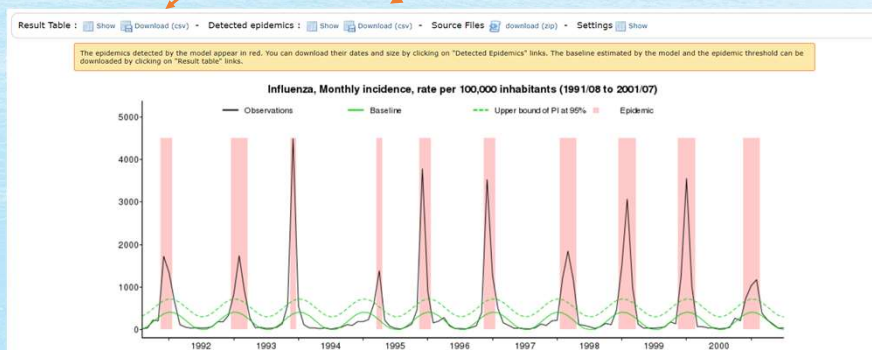
Choose the alert rule
an epidemic is declared when successive observations are upper the estimated threshold.

[Next step](#)

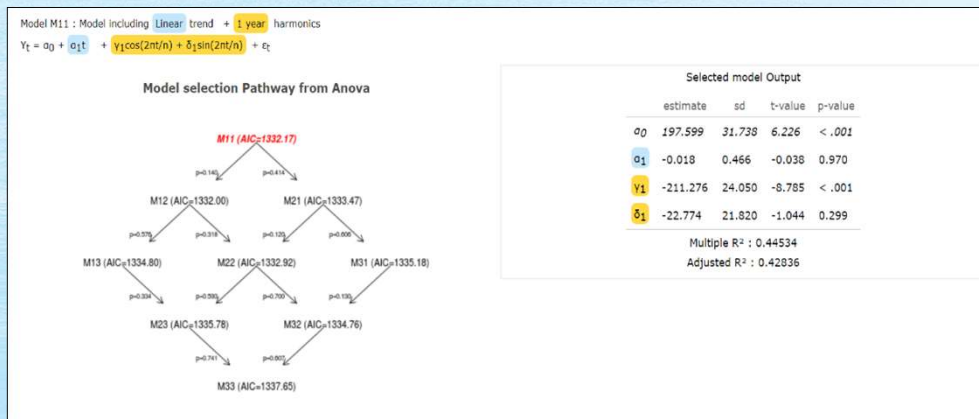
Modèle de régression périodique

- Étape 5: visualisation des résultats (incluant les valeurs seuils)

Télécharger: Ensemble des données ET Épidémies détectées



Modèle de régression périodique



Considérations éthiques

Protection des données

- ▶ Mettez régulièrement à jour vos logiciels;
- ▶ Effectuez des sauvegardes régulières ;
- ▶ Formez les employés aux meilleures pratiques de sécurité;
- ▶ Effacer l'historique et les cookies de votre navigateur web.

Principes éthique



Confidentialité



Transparence



Partage responsable des données



Utilisation responsable des données



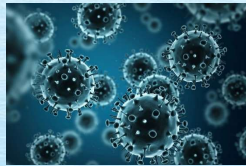
Partie 3: exercice pratique

Explication

- ▶ <https://periodic.sentiweb.fr/>
- ▶ Deux bases de données en parallèle:



Campylobacter

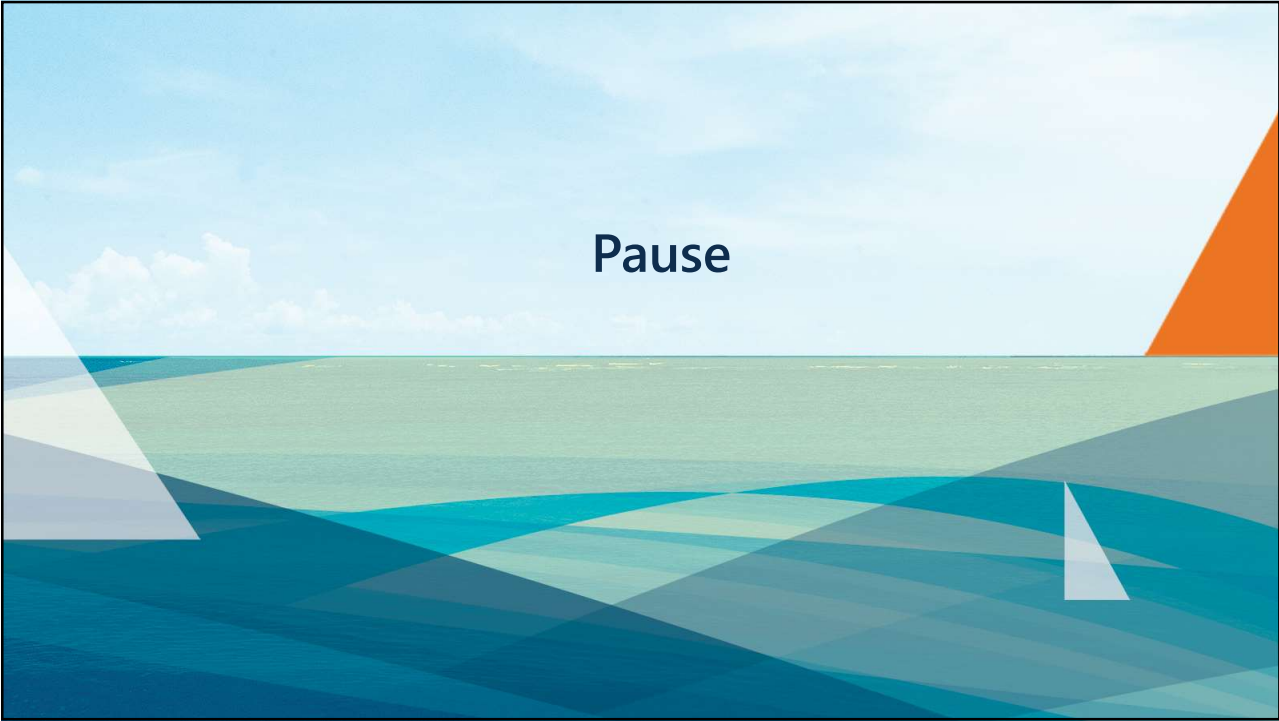


Influenza

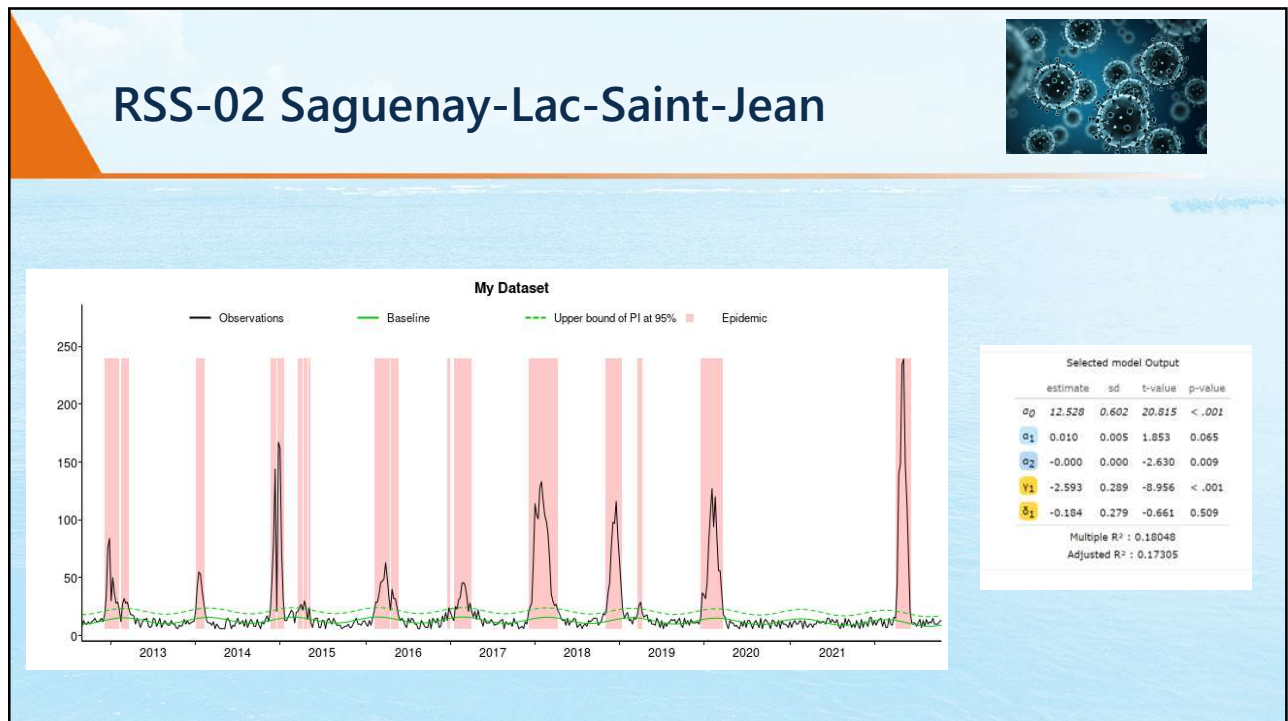
Sources:

<https://www.cdc.gov/campylobacter/index.html>

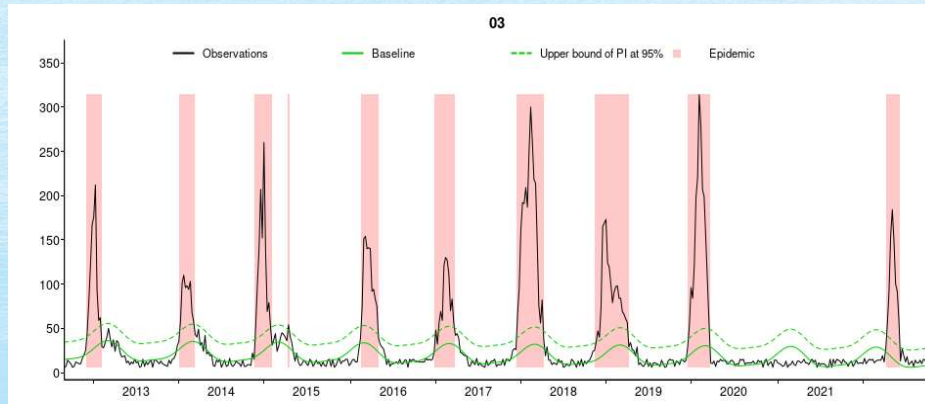
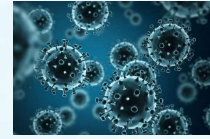
<https://www.nfid.org/infectious-diseases/influenza-flu/>



Résultats: Influenza

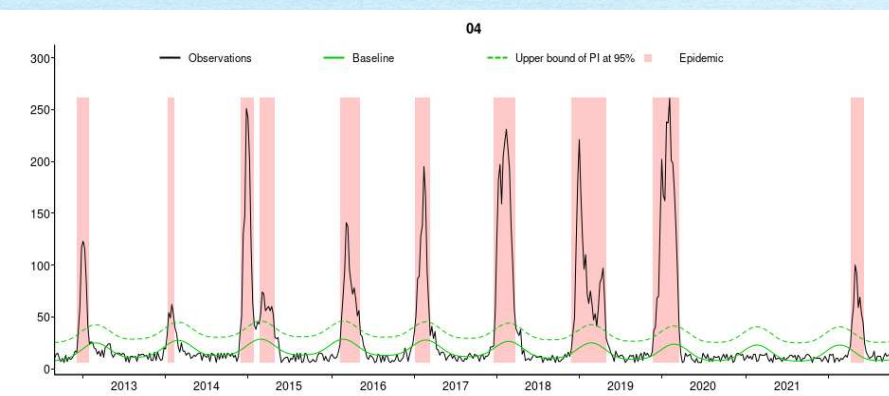
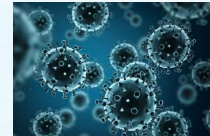


RSS-03 Capitale-Nationale



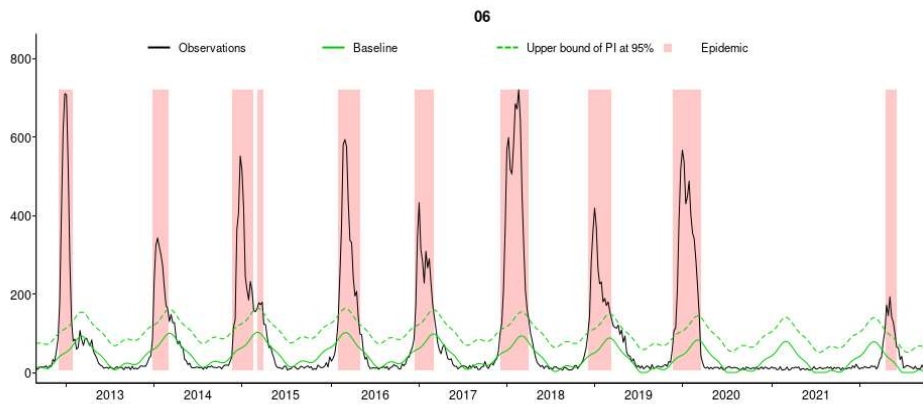
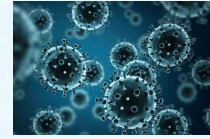
Selected model Output				
	estimate	sd	t-value	p-value
σ_0	22.843	0.947	24.132	< .001
σ_1	-0.015	0.003	-5.083	< .001
γ_1	-10.781	0.718	-15.009	< .001
δ_1	-0.858	0.664	-1.292	0.197
γ_2	2.479	0.684	3.625	< .001
δ_2	1.438	0.681	2.112	0.035
Multiple R ² : 0.37891				
Adjusted R ² : 0.37188				

RSS-04 Mauricie et Centre-du-Québec



Selected model Output				
	estimate	sd	t-value	p-value
σ_0	13.756	1.687	8.155	< .001
σ_1	0.081	0.028	2.911	0.004
σ_2	-0.000	0.000	-2.828	0.005
σ_3	0.000	0.000	2.421	0.016
γ_1	-7.635	0.647	-11.796	< .001
δ_1	-0.463	0.600	-0.772	0.440
γ_2	1.514	0.616	2.457	0.014
δ_2	0.251	0.620	0.405	0.686
Multiple R ² : 0.2724				
Adjusted R ² : 0.26082				

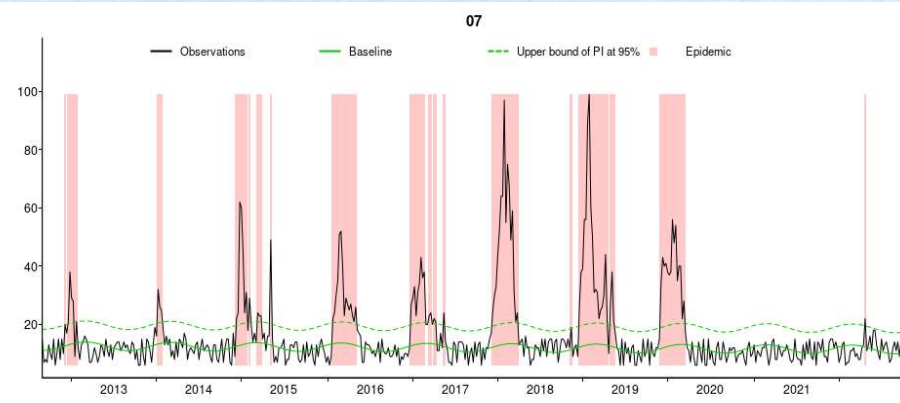
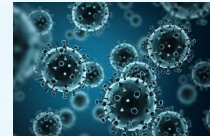
RSS-06 Montréal



Selected model Output				
	estimate	sd	t-value	p-value
σ_0	35.200	6.015	5.852	< .001
σ_1	0.257	0.098	2.614	0.009
σ_2	-0.001	0.000	-2.698	0.007
σ_3	0.000	0.000	2.314	0.021
γ_1	-36.965	2.323	-15.913	< .001
δ_1	-3.367	2.131	-1.580	0.115
γ_2	6.573	2.161	3.042	0.002
δ_2	6.053	2.244	2.697	0.007
δ_3	3.788	2.142	1.768	0.078
δ_4	4.324	2.123	2.037	0.042

Multiple R² : 0.43081
Adjusted R² : 0.41914

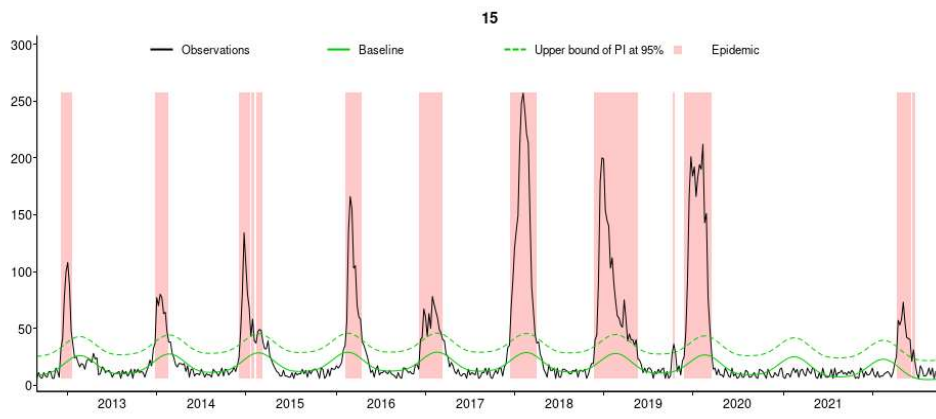
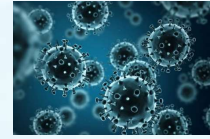
RSS-07 Outaouais



Selected model Output				
	estimate	sd	t-value	p-value
σ_0	12.602	0.340	37.077	< .001
σ_1	-0.002	0.001	-2.110	0.035
γ_1	-1.442	0.261	-5.529	< .001
δ_1	-0.268	0.243	-1.101	0.272

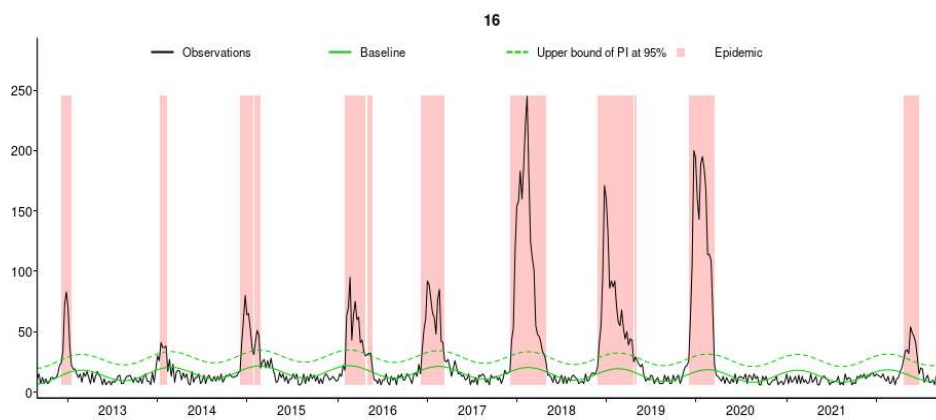
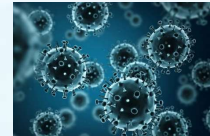
Multiple R² : 0.07907
Adjusted R² : 0.07284

RSS-15 Laurentides



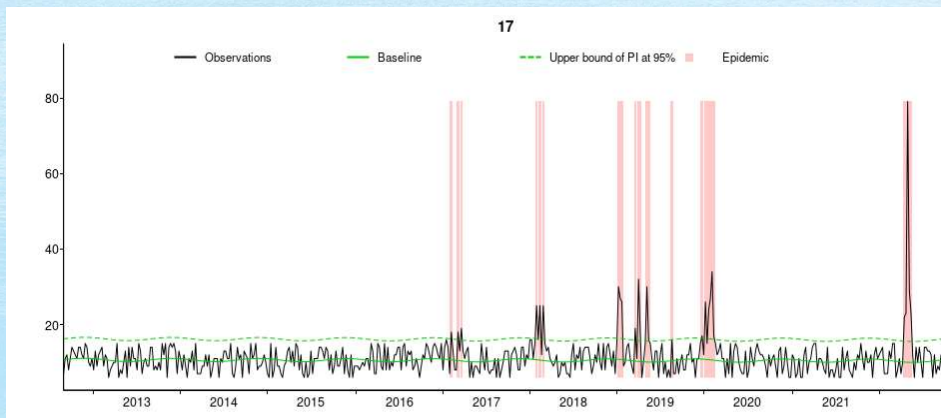
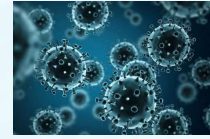
Selected model Output				
	estimate	sd	t-value	p-value
σ_0	15.185	1.200	12.654	< .001
β_1	0.036	0.011	3.358	< .001
σ_2	-0.000	0.000	-4.218	< .001
γ_1	-8.316	0.616	-13.490	< .001
δ_1	0.302	0.575	0.526	0.599
γ_2	1.613	0.581	2.774	0.006
σ_2	0.073	0.593	0.123	0.902
Multiple R ² : 0.31854				
Adjusted R ² : 0.30931				

RSS-16 Montérégie



Selected model Output				
	estimate	sd	t-value	p-value
σ_0	11.380	1.224	9.297	< .001
β_1	0.074	0.020	3.631	< .001
σ_2	-0.000	0.000	-3.395	< .001
β_2	0.000	0.000	2.956	0.003
γ_1	-5.012	0.475	-10.553	< .001
δ_1	0.070	0.441	0.160	0.873
Multiple R ² : 0.22777				
Adjusted R ² : 0.21905				

RSS-17 Nunavik

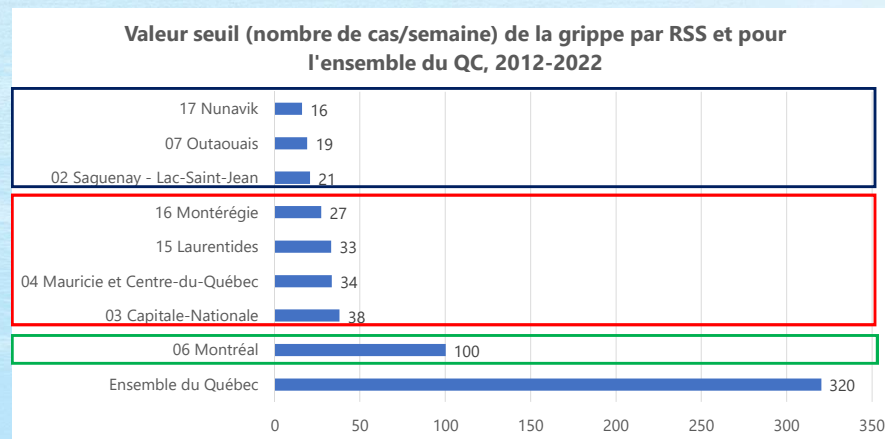
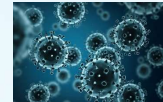


Selected model Output

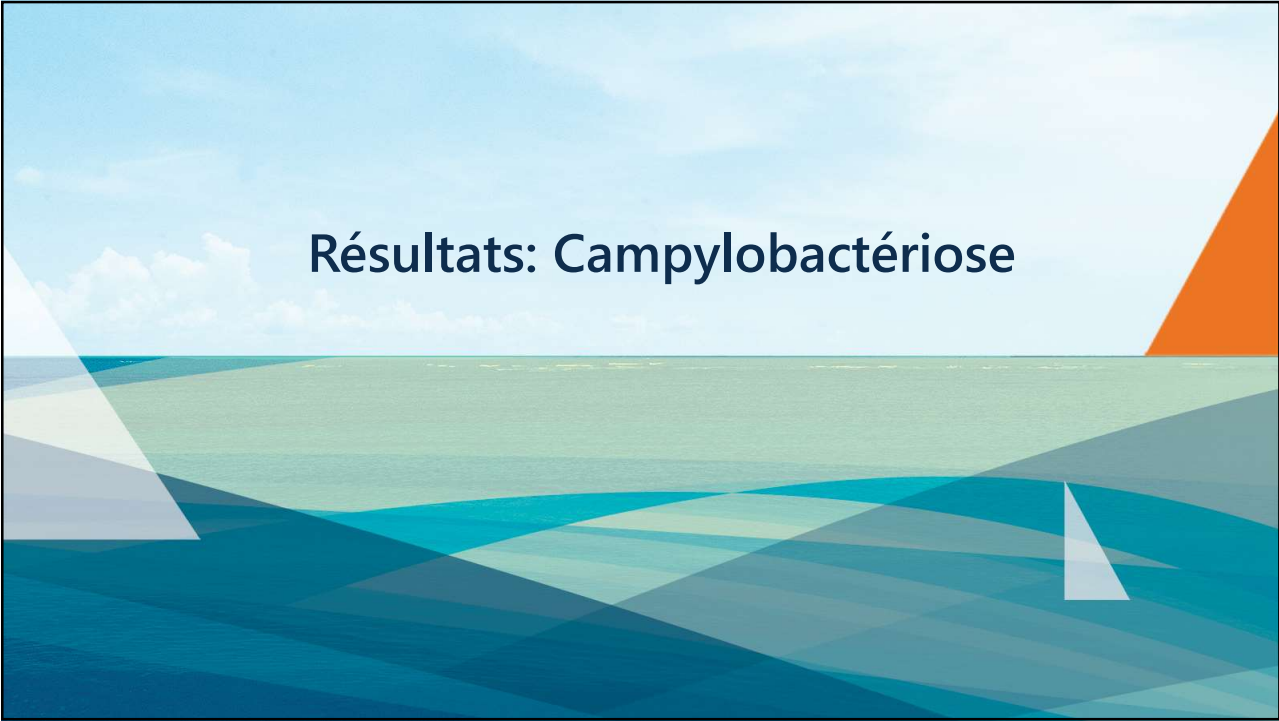
	estimate	sd	t-value	p-value
α_0	10.655	0.249	42.878	< .001
α_1	-0.000	0.001	-0.566	0.571
γ_1	-0.025	0.181	-0.138	0.890
δ_1	0.429	0.179	2.395	0.017

Multiple R² : 0.01235
Adjusted R² : 0.00637

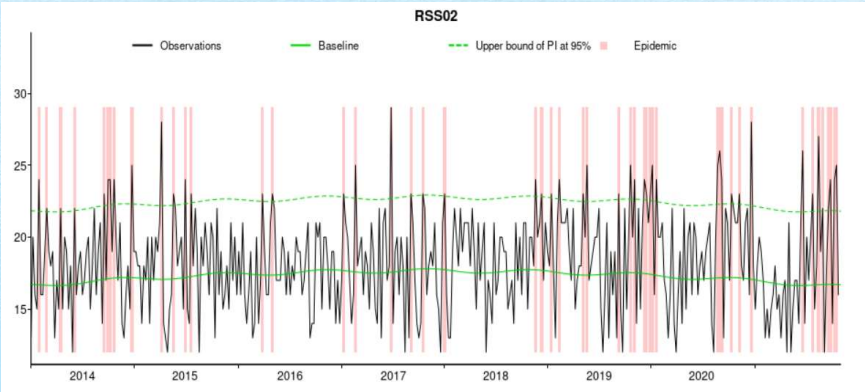
Les valeurs de seuil sont-elles spécifiques à une région donnée ou est-ce généralisable ?



Résultats: Campylobactériose

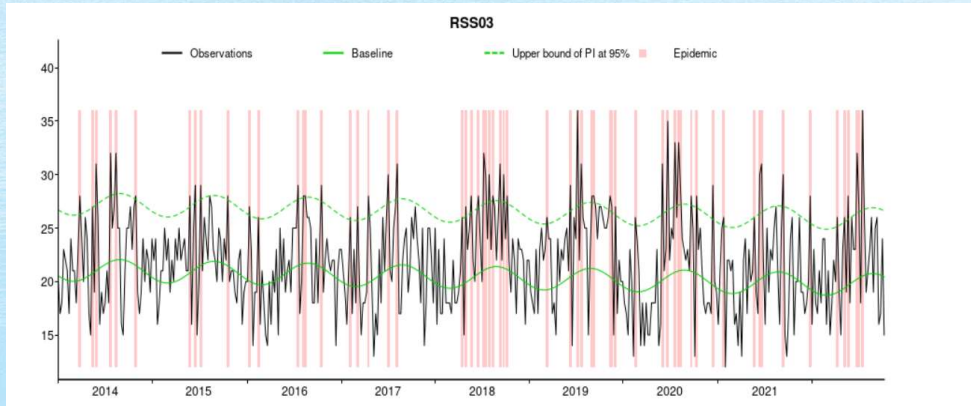


RSS-02 Saguenay-Lac-Saint-Jean



Selected model Output				
	estimate	sd	t-value	p-value
θ_0	16.619	0.434	38.303	< .001
θ_1	0.010	0.005	2.109	0.036
θ_2	-0.000	0.000	-2.200	0.028
γ_1	0.094	0.203	0.462	0.644
δ_1	-0.112	0.204	-0.546	0.585
Multiple R ² : 0.01701				
Adjusted R ² : 0.00524				

RSS-03 Capitale-Nationale

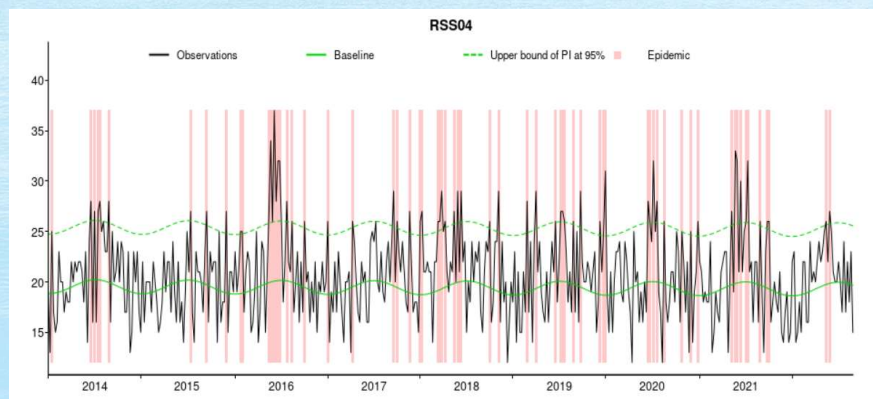


Selected model Output

	estimate	sd	t-value	p-value
θ_0	21.108	0.330	63.991	< .001
θ_1	-0.003	0.001	-2.513	0.012
γ_1	-0.463	0.238	-1.945	0.053
δ_1	-0.948	0.234	-4.044	< .001

Multiple R² : 0.06484
Adjusted R² : 0.05705

RSS-04 Mauricie et Centre-du-Québec

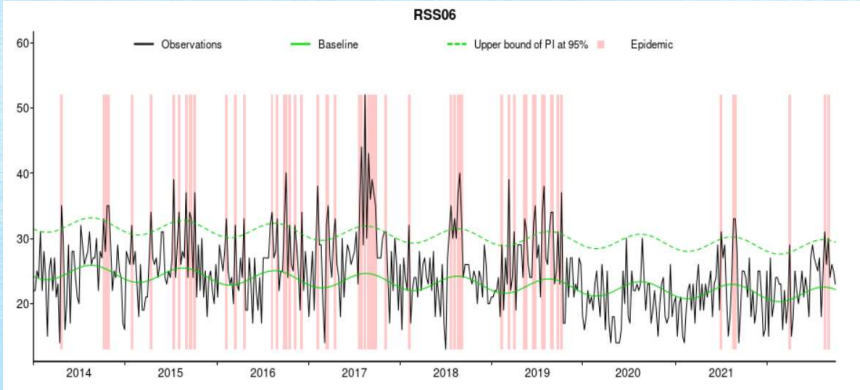


Selected model Output

	estimate	sd	t-value	p-value
θ_0	19.538	0.312	62.531	< .001
θ_1	-0.001	0.001	-0.477	0.634
γ_1	-0.675	0.228	-2.964	0.003
δ_1	-0.124	0.219	-0.567	0.571

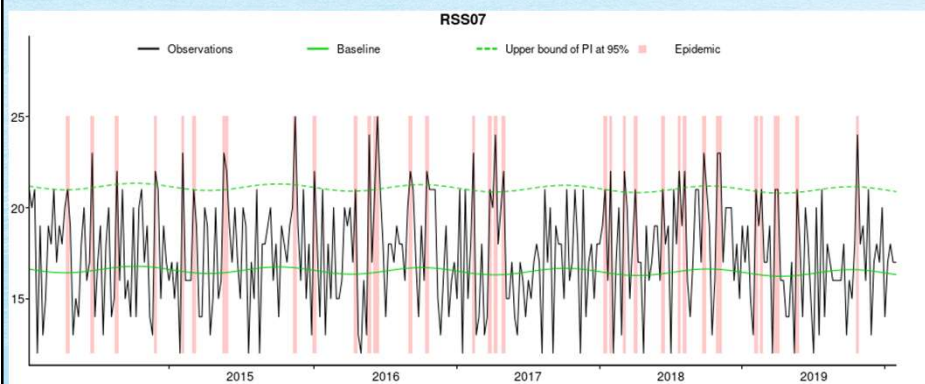
Multiple R² : 0.02475
Adjusted R² : 0.01673

RSS-06 Montréal



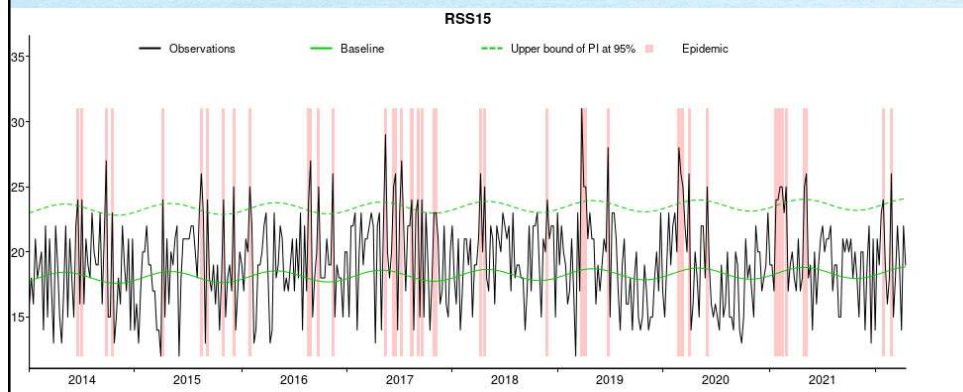
Selected model Output			
	estimate	sd	t-value p-value
α_0	24.941	0.390	63.883 < .001
α_1	-0.008	0.001	-5.658 < .001
γ_1	-0.654	0.266	-2.458 0.014
δ_1	-1.017	0.280	-3.635 < .001
Multiple R ² : 0.11154			
Adjusted R ² : 0.10442			

RSS-07 Outaouais



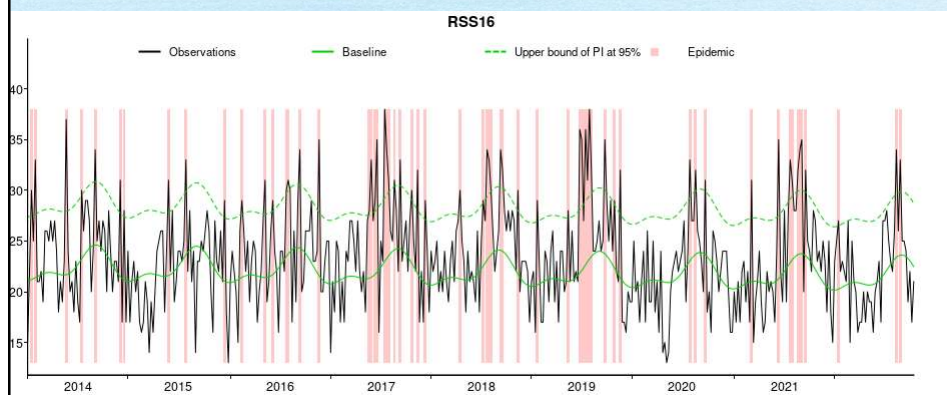
Selected model Output			
	estimate	sd	t-value p-value
α_0	16.621	0.293	56.685 < .001
α_1	-0.001	0.002	-0.457 0.648
γ_1	0.021	0.205	0.101 0.919
δ_1	-0.188	0.212	-0.885 0.377
Multiple R ² : 0.00362			
Adjusted R ² : -0.00829			

RSS-15 Laurentides



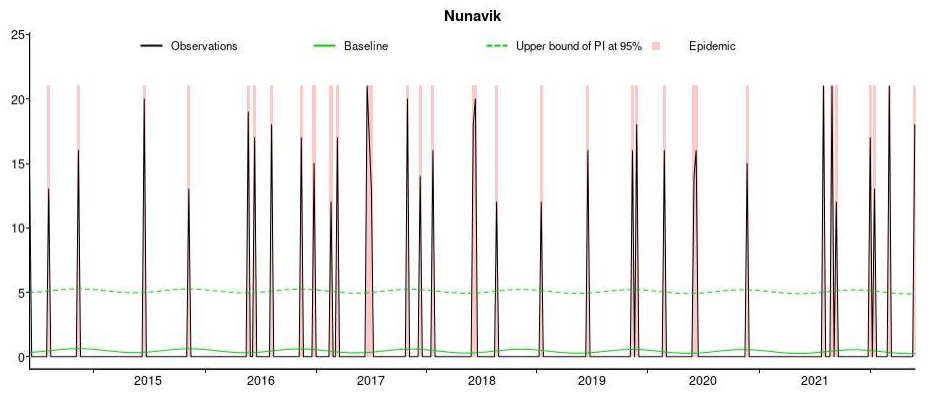
Selected model Output				
	estimate	sd	t-value	p-value
α_0	17.982	0.279	64.477	< .001
α_1	0.001	0.001	0.914	0.361
γ_1	-0.263	0.198	-1.327	0.185
δ_1	0.351	0.202	1.735	0.084
Multiple R ² : 0.01472				
Adjusted R ² : 0.00646				

RSS-16 Montérégie



Selected model Output				
	estimate	sd	t-value	p-value
α_0	22.519	0.331	68.130	< .001
α_1	-0.002	0.001	-1.938	0.053
γ_1	-0.662	0.239	-2.765	0.006
δ_1	-1.261	0.235	-5.360	< .001
γ_2	-0.675	0.231	-2.928	0.004
δ_2	0.401	0.241	1.664	0.097
Multiple R ² : 0.11154				
Adjusted R ² : 0.09959				

RSS-17 Nunavik

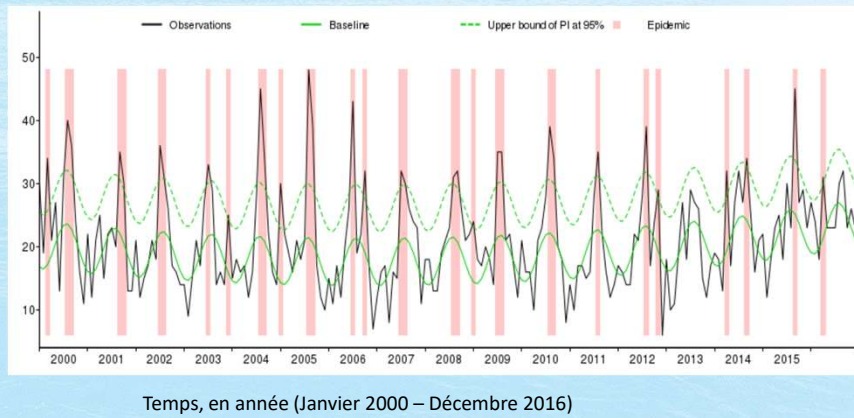


Selected model Output				
	estimate	sd	t-value	p-value
α_0	0.482	0.239	2.016	0.045
α_1	-0.000	0.001	-0.232	0.817
γ_1	-0.145	0.171	-0.851	0.395
δ_1	0.035	0.169	0.208	0.836
Multiple R ² : 0.00214				
Adjusted R ² : -0.00551				

Partie 4: Utilisation et limites

Salmonellose

Exemple: RSS X



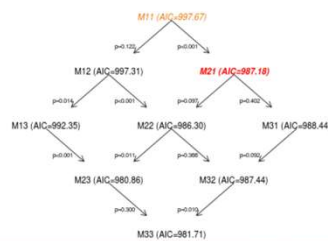
Salmonellose

Exemple: RSS X

Model M21 : Model including Quadratic trend + 1 year harmonics

$$Y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \gamma_1 \cos(2\pi t/n) + \delta_1 \sin(2\pi t/n) + \epsilon_t$$

Model selection Pathway from Anova



Selected model Output

	estimate	sd	t-value	p-value
α_0	20.356	1.094	16.611	< .001
α_1	-0.067	0.024	-2.751	0.007
α_2	0.000	0.000	3.547	< .001
γ_1	-2.307	0.466	-4.948	< .001
δ_1	-2.991	0.510	-5.869	< .001

Multiple R² : 0.31766

Adjusted R² : 0.30102

Periodicities

1 year 1 year - 6 months 1 year - 3 months

Trend

Linear M1.1 M1.2 M1.3

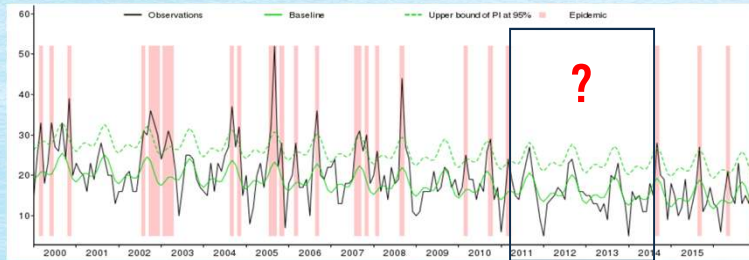
Quadratic M2.1 M2.2 M2.3

Cubic M3.1 M3.2 M3.3

Model selected with the automated selection algorithm

Giardiasis

Exemple: RSS X



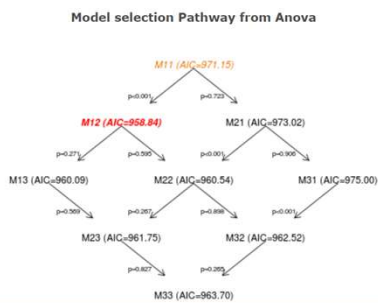
Temps, en année (Janvier 2000 – Décembre 2016)

Giardiasis

Exemple: RSS X

Model M12 : Model including Linear trend + 1 year + 6 months harmonics

$$Y_t = \alpha_0 + \alpha_1 t + \gamma_1 \cos(2\pi t/n) + \delta_1 \sin(2\pi t/n) + \gamma_2 \cos(4\pi t/n) + \delta_2 \sin(4\pi t/n) + \epsilon_t$$



Selected model Output

	estimate	sd	t-value	p-value
α_0	21.692	0.648	33.459	< .001
α_1	-0.037	0.005	-7.156	< .001
γ_1	-0.482	0.408	-1.182	0.239
δ_1	-2.413	0.452	-5.336	< .001
γ_2	-1.707	0.437	-3.905	< .001
δ_2	-0.420	0.417	-1.008	0.315

Multiple R² : 0.33728
Adjusted R² : 0.3172

Periodicities

1 year 1 year + 6 months 1 year + 6 months + 1 year

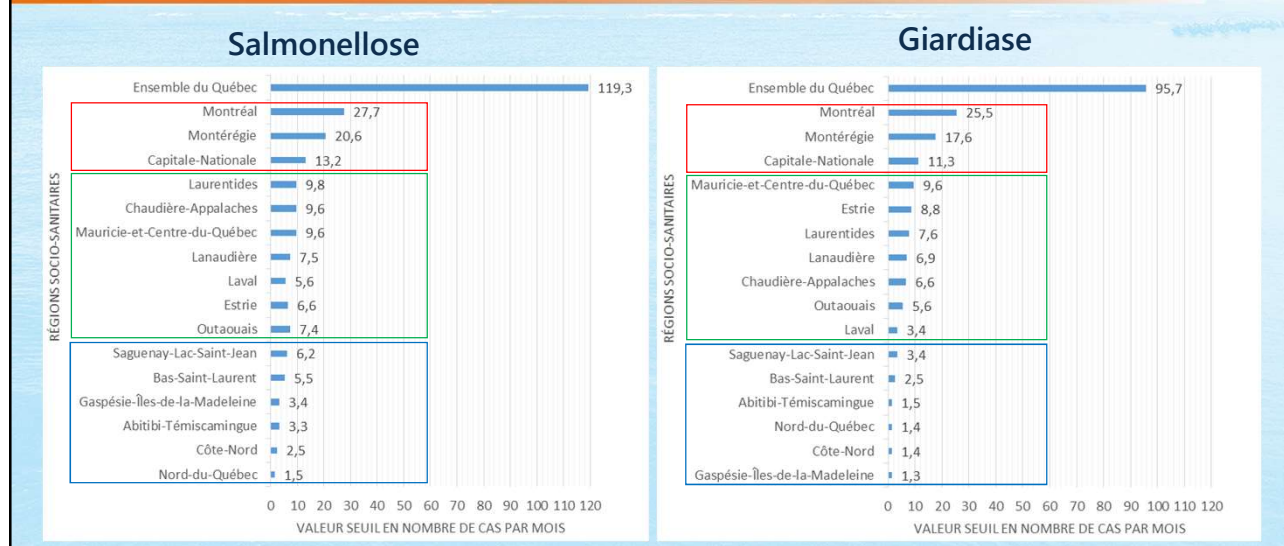
Trend

Linear Quadratic Cubic

M1.1 M1.2 M1.3
 M2.1 M2.2 M2.3
 M3.1 M3.2 M3.3

Model selected with the automated selection algorithm

Est-ce possible d'avoir une même valeur seuil pour des agents pathogènes possédant une chaîne de transmission similaire ?



Limites de la régression périodique

- ▶ Choix du modèle est une supposition;
- ▶ Données épidémiologiques doivent avoir un signal saisonnier et des éclosions;
- ▶ Minimum d'une année de données de suivi épidémiologique → calibration du modèle;
- ▶ Avoir des données avant le pic d'éclosion;
- ▶ Moins performant pour régions avec peu de données.

Synthèse et discussion

Et maintenant...

- ▶ Nommez deux éléments que vous retenez de cet atelier?
- ▶ Prévoyez-vous utilisez cette méthode?



Synthèse

- ▶ Pas de méthode parfaite;
- ▶ Contexte régional à considérer;
- ▶ Diffère selon les agents pathogènes: Grippe vs campylobactériose;
- ▶ Seuils du même ordre de grandeur pour des chaînes de transmission similaires;

Synthèse...

- ▶ Complément aux méthodes traditionnelles;
- ▶ Importance de tester différents paramètres et de déterminer leur influence sur la valeur seuil obtenue;
- ▶ Œil critique: Machine ne remplace pas l'humain!

Merci beaucoup!

Références

- ▶ Cai, J., Zhang, B., Xu, B., Chan, K., Chowell, G., Tian, H., & Xu, B. (2019). A maximum curvature method for estimating epidemic onset of seasonal influenza in Japan. *BMC infectious diseases*, 19(1), 181. doi:10.1186/s12879-019-3777-x.
- ▶ Cheng, X., Chen, T., Yang, Y., Yang, J., Wang, D., Hu, G., & Shu, Y. (2018). Using an innovative method to develop the threshold of seasonal influenza epidemic in China. *PLOS ONE*, 13(8), e0202880. Retrieved from <https://doi.org/10.1371/journal.pone.0202880>.
- ▶ Hartfield, M., & Alizon, S. (2013). Introducing the Outbreak Threshold in Epidemiology. *PLOS Pathogens*, 9(6), e1003277. Retrieved from <https://doi.org/10.1371/journal.ppat.1003277>.
- ▶ Lagorio, C., Migueles, M. V., Braunstein, L. A., López, E., & Macri, P. A. (2009). Effects of epidemic threshold definition on disease spread statistics. *Physica A: Statistical Mechanics and its Applications*, 388(5), 755-763. Retrieved from <https://doi.org/10.1016/j.physa.2008.10.045>.
- ▶ Pelat, C., Boëlle, P.-Y., Cowling, B. J., Carrat, F., Flahault, A., Ansart, S., & Valleron, A.-J. (2007). Online detection and quantification of epidemics. *BMC Medical Informatics and Decision Making*, 7(1), 29. Retrieved from <https://doi.org/10.1186/1472-6947-7-29>.
- ▶ Rakocevic B., Grgurevic A., Trajkovic G., Mugosa B., Sipetic Grujicic S., Medenica S., Bojovic Olivera Lozano Alonso J.E., Vega T. (2109). Influenza surveillance: determining the epidemic threshold for influenza by using the Moving Epidemic Method (MEM), Montenegro, 2010/11 to 2017/18 influenza seasons. *Euro Surveill*. 24(12).