

Aurez-vous la grippe cet hiver ?

Atelier MATHs en JEANS

de l'École internationale de Manosque^(*)

Grégory Depaepe, Mathilde Mus & Zoé Hance^(**)

Notre sujet consiste à étudier la propagation du virus de la grippe afin d'en créer un modèle mathématique. L'objectif scientifique est de mieux comprendre la propagation de la grippe afin de mieux contrôler la prochaine épidémie.

1. L'équipe

Nous sommes un groupe de six élèves de 15-16 ans. Ce projet est né d'une discussion en classe autour des affiches d'un stage « Hippocampe » sur les mathématiques et la médecine (à l'IREM d'Aix-Marseille). Nos professeurs, M. Marchal et Mme Bonacucina, avaient contacté une chercheuse, Mme Chapuisat, qui travaille à l'Université de Marseille, sur la modélisation en médecine. Ils ont décidé de choisir comme projet « MATHs.en.JEANS » pour cette année 2012/2013, un sujet lié à la propagation d'une maladie. Elle est venue à notre école en Septembre et nous a parlé de son travail et comment la modélisation pourrait aider dans le domaine médical. Elle nous a informés de l'existence du site Web « Sentinelles ». Nous avons donc travaillé avec le médecin scolaire M. Scola et nous étions sous la direction de deux professeurs de mathématiques de l'école.

2. Notre sujet

Très vite, le choix de la grippe est devenu évident car :

- Elle revient chaque année, donc nous pouvons essayer de savoir comment la prédire.
- C'est une maladie connue que de nombreuses personnes attrapent.
- C'est une maladie dangereuse, au moins pour certaines catégories de personnes.
- Nous avons un accès régulier (en ligne) au nombre de personnes malades.
- Sa transmission est bien connue (bien que pas si facile à modéliser !).

De plus, nous avons décidé de limiter notre étude à la région PACA et à un « pas » hebdomadaire parce que les données auxquelles nous avons accès ne sont mises à jour qu'une fois par semaine.

(*) Lycée public ouvert en 2007 pour accueillir les enfants du personnel du projet ITER de Cadarache

(**) Ces trois élèves ont présenté le projet au concours national Cgénial, où ils ont obtenu un premier prix qui leur a permis de participer au concours international CASTIC (China Adolescents Science and Technology Innovation Contest) cet été à Nankin, où ils ont obtenu un second prix. Trois autres élèves ont également participé à ce projet MATH.en.JEANS, Nadège Chojnacki, Céliane Lacour et Robin Marty.

3. Recherche de paramètres influents

Pendant le premier mois, nous avons défini les paramètres susceptibles d'influer sur l'épidémie de grippe. Chacun de nous a fait une liste et nous avons vite remarqué que certains paramètres étaient plus importants que d'autres. Nous les avons classés en quatre catégories.

La première catégorie est la plus importante pour nous, la seule que nous avons prise en compte à ce moment-là. C'est le nombre d'habitants dans la région, le nombre de personnes malades et le nombre de personnes en bonne santé, ainsi que le nombre de personnes immunisées (qui ont été malades ou vaccinées). Nous devons en plus tenir compte de la façon dont le modèle lui-même change en fonction de l'évolution de la propagation.

Dans la seconde catégorie, nous avons les habitudes sociales, c'est-à-dire les contacts avec les gens (bien sûr, si vous êtes à la retraite dans une petite ferme, vous avez moins de chance d'attraper la grippe), et les habitudes de travail, les habitudes de voyage (les migrants, les touristes) et bien sûr la répartition géographique de la population (densité).

Dans la troisième catégorie, nous plaçons les conditions météorologiques et le fait de considérer des sous-ensembles plus fins de population (comme les personnes âgées ou les bébés).

Enfin, dans les facteurs moins influents que nous avons examinés, on peut citer les campagnes de prévention ou les gens qui souffrent déjà d'une autre maladie.

Il est apparu rapidement que certains paramètres, comme la météo, bien qu'ils soient importants, seraient difficiles à quantifier numériquement pour les prendre en compte dans notre futur modèle.

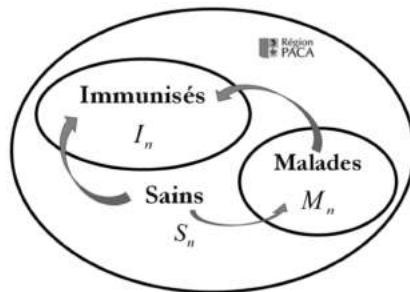
4. Premier modèle

Une fois que nous avons identifié nos variables, l'étape la plus importante a été de les relier par des équations. Donc nous avons fait un diagramme pour représenter la propagation de la grippe.

Les personnes en bonne santé (S_n) peuvent devenir malades et ensuite être immunisées, car elles ont développé les anticorps nécessaires. Les personnes en bonne santé peuvent être immunisées de manière directe par la vaccination.

Nous supposons que la population totale est constante dans l'échelle de temps que nous considérons (de Septembre à Avril).

La durée de la maladie (chez un patient donné) est un paramètre important et doit être réglable.

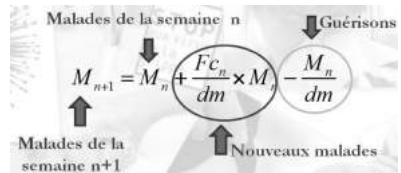


4.1. Notations

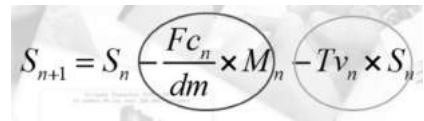
M_n = nombre de personnes malades
 I_n = nombre de personnes immunisées
 S_n = nombre de personnes saines
 $T_n = M_n + I_n + S_n$ = nombre total de personnes = cste = T_0
 Fc_n = facteur de contamination
 Tv_n = taux de vaccination hebdomadaire
 dm = durée de vie de la maladie (chez un malade donné)

4.2. Équations

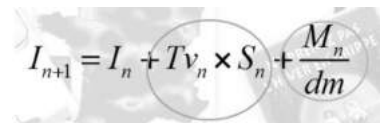
Le nombre de personnes malades de la semaine $(n + 1)$ est égal au nombre de personnes malades de la semaine précédente, auquel on ajoute le nombre de nouveaux malades et dont on retranche le nombre de ceux qui sont guéris :



Nous faisons d'une manière similaire pour le nombre de personnes en bonne santé à la semaine $n + 1$: les personnes en bonne santé de la semaine précédente, déduction faite des nouvelles personnes malades et de celles qui se sont fait vacciner.



Idem pour I_{n+1} : nombre de personnes immunisées la semaine d'avant, plus le nombre de nouveaux vaccinés, plus le nombre de personnes guéries.

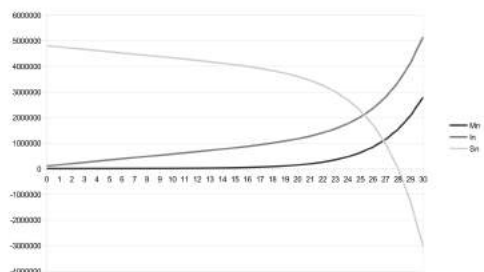


De cette façon, il est facile de vérifier que la somme (qui représente la population totale) est constante (ce qui est ajouté à une suite est soustrait à une autre).

4.3. Premiers essais

Nous avons choisi d'utiliser un tableur pour étudier les variations de ces variables semaine par semaine.

Quand nous avons essayé de faire fonctionner notre modèle sur ces équations, des choses étranges se sont produites. Tout d'abord, notre nombre de personnes en bonne santé est devenu négatif (ce qui est



assez bizarre : il ne peut pas y avoir -20 humains sur Terre !). Nous aurions pu forcer la variable à être positive, mais cela n'aurait pas été une solution satisfaisante.

Ensuite, notre nombre de personnes malades augmentait beaucoup trop et trop vite ; en quelques semaines, la population entière était immunisée (après avoir été malade). C'était aussi impossible, car tout le monde n'attrape pas la grippe.

Nous avons donc décidé de rendre le taux de contamination variable (il était constant dans notre premier modèle).

5. Alors, qu'est-ce qu'on fait ?

Nous avons contacté Mme Chapuisat. Elle nous a dit d'écrire avec attention nos « hypothèses biologiques » liées à nos équations (ce qu'elles signifient dans la réalité).



Par exemple :

- La population est constante, alors il y a autant de morts que de naissances.
- La grippe « vit » pendant environ deux semaines (dans le corps des malades).
- Une personne qui attrape la grippe est ensuite immunisée pour l'hiver entier.

Et puis de les faire vérifier par un médecin.

Donc nous avons contacté le médecin de notre école qui nous a donné des renseignements de base au sujet de la grippe. Qu'est-ce que la grippe ?

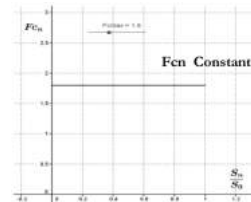
- C'est une infection causée par un virus. Elle se propage par la toux et le contact de la main. Le virus pénètre dans la cellule et la fait souffrir.
- Nous pouvons avoir différents symptômes : maux de tête, toux, fièvre.
- Les médicaments comme le paracétamol ou la vitamine C font disparaître les symptômes, mais ne tuent pas le virus.
- Le virus est mutant, c'est-à-dire qu'il change chaque année.
- Les personnes les plus vulnérables sont les personnes très âgées et les très jeunes enfants.



5.1. Facteur de contamination

En parallèle, nous avons poursuivi la réflexion sur l'élément central du modèle : le processus de contamination.

Notre facteur de contamination est le nombre de personnes en bonne santé qu'un malade contamine pendant le temps qu'il a la grippe. Au début il était constant, mais nous n'avons vite réalisé que le nombre



de personnes contaminées dépend à la fois du nombre de personnes malades et de personnes saines. Pourquoi ?

Bien sûr, cela dépend du nombre de personnes malades. Plus ils sont nombreux, plus ils contaminent de gens. Mais cela dépend aussi du nombre de personnes en bonne santé. Considérons la population comme une urne avec des gens en bonne santé à l'intérieur ; en termes de probabilité, plus les gens en bonne santé sont nombreux, plus une personne malade peut en contaminer au hasard dans cette urne. S'il y a 10 000 personnes malades mais une seule personne en bonne santé, les malades ne contamineront pas beaucoup de gens.

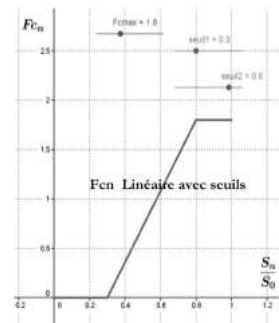
Le médecin nous a dit que le taux moyen de contamination est d'environ 2 pour notre grippe.

Nous avons essayé différentes formules pour représenter les variations de ce facteur en fonction du nombre de sains S_n .

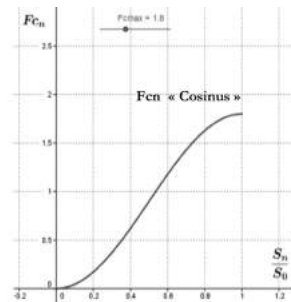
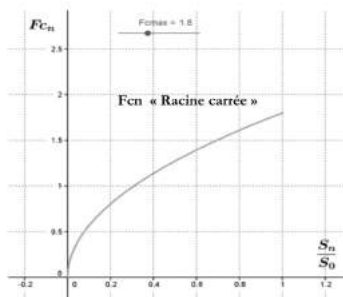
Notre premier choix a été de le laisser constant, à 1,8 par exemple.

Ensuite, nous avons essayé de réfléchir : comment le phénomène de contamination opère-t-il dans la réalité ? Pour prendre en compte le phénomène de double dépendance décrit juste au-dessus, nous voulions un facteur de contamination ayant une valeur maximale (FC_{max}) quand S_n est grand (au début de la maladie), puis qui diminue à mesure que S_n devient plus petit. Notre fonction FC_n est donc croissante en fonction de S_n et décroissante en fonction du temps écoulé n .

Nous avons choisi une fonction linéaire à seuils (ou fonction linéaire par morceaux). Vous devez lire de droite à gauche (le nombre maximum de personnes en bonne santé vaut S_0 au départ et décroît vers 0). Nous avons divisé S_n par S_0 pour obtenir un nombre entre 0 et 1, ce qui est plus commode à traiter. Tant qu'il existe un grand nombre de personnes en bonne santé, le taux reste à son maximum. Ensuite, il diminue jusqu'à ce qu'il atteigne 0 (avant que la population saine n'atteigne elle 0 car tout le monde n'attrape pas la grippe).



Nous avons également envisagé deux autres fonctions ayant le même comportement : racine carrée et cosinus.



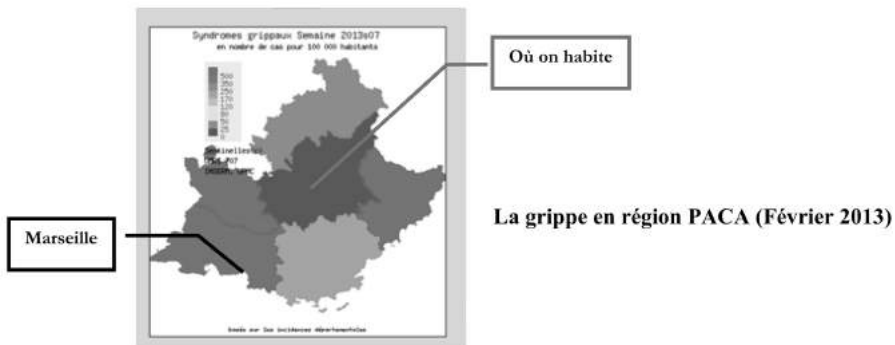
Ce taux ne dépend que de S_n avec ces formules, mais nous multiplions ce taux par M_n afin d'avoir notre double dépendance.

5.2. Comparaison avec la réalité

Pour mesurer l'efficacité de notre modèle (et voir qui des fonctions de contamination précédente « marche » le mieux), nous l'avons comparé avec le nombre réel de cas de grippe dans la région, que l'on trouve en ligne sur le site Sentinelles. Ce site recense le nombre de personnes malades par région, chaque semaine. Nous voulions que notre courbe soit similaire à celle-ci, pour être le plus proche possible de la réalité. Pour juger de la qualité de notre modèle, nous mesurons la distance entre notre courbe et celle de Sentinelles.

Les données sont disponibles sous la forme soit de cartes, soit de tableaux que nous pouvons importer dans notre modèle pour effectuer nos comparaisons. Le recensement est assuré par des médecins et des hôpitaux locaux.

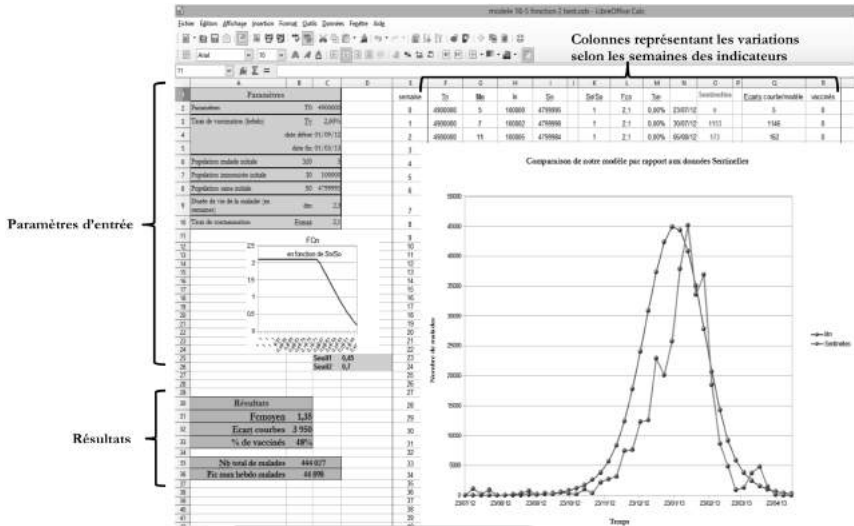
Remarque : Nous étudions ici le virus classique de la grippe et non le célèbre H1N1 redouté en Extrême-Orient.



6. Qu'est-ce que ça donne maintenant ?

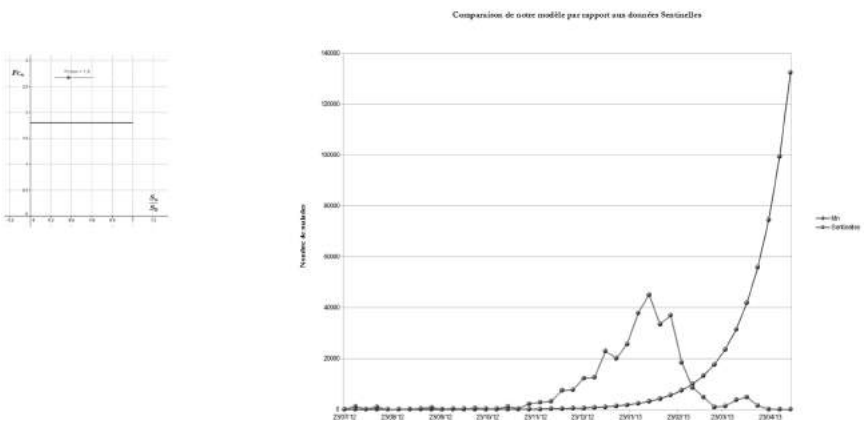
Nous comparons donc notre courbe avec celle de Sentinelles. Nous avons les paramètres d'entrée (population totale, le taux de vaccination, ... dans la zone grisée supérieure) que nous pouvons modifier. Puis nous avons les colonnes montrant l'évolution de nos variables (le nombre de personnes immunisées ou malades par exemple), semaine après semaine.

Nos résultats (indicateurs permettant de mesurer la manière dont le modèle correspond à la réalité) sont affichés dans la zone grisée inférieure. Il y a le facteur moyen de contamination (celui qui doit être aussi proche que possible de 2), l'écart entre les deux courbes, le taux global de vaccination, le nombre total de personnes malades au cours de la période et le plus haut sommet hebdomadaire atteint par ce nombre.

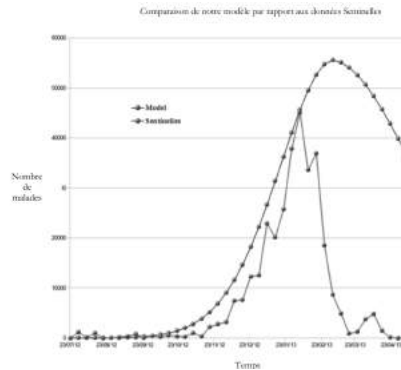
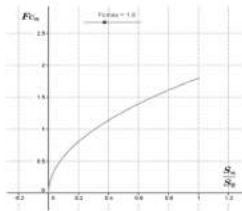


En fait c'est beaucoup mieux que nous le pensions ! Nous allons maintenant comparer les quatre courbes (avec nos différentes fonctions de contamination) avec le meilleur ajustement obtenu :

Première courbe : Celle avec une fonction de contamination constante. Nous voyons que ce n'est pas vraiment bon ; dans ce cas la suite M_n est géométrique d'où l'augmentation exponentielle du nombre de personnes malades.

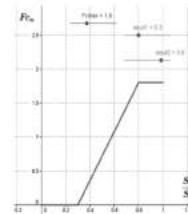


Deuxième courbe : fonction de contamination type racine carrée. C'est mieux puisque ça permet de briser la propagation et de faire apparaître un pic, mais la diminution du facteur de contamination arrive trop tard.

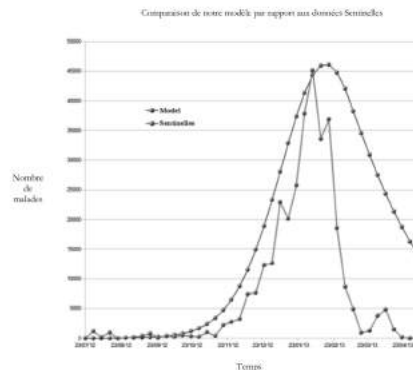
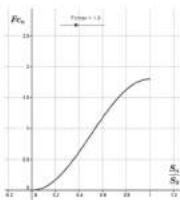


Troisième courbe : Voici (voir page précédente) notre meilleur résultat, avec la fonction linéaire à seuils.

La différence moyenne hebdomadaire entre le modèle et la réalité est de 3 950 personnes, ce qui est assez bon (sur environ 10 800).



Quatrième courbe : Celle en forme de cosinus. C'est assez bon en termes de forme de la courbe obtenue, mais il n'est pas possible de faire un pic qui soit assez raide



Notre fonction avec des seuils donne les meilleurs résultats, avec des paramètres d'entrée très réalistes :

- La durée de la maladie (le temps de contagiosité) est de 2,3 semaines.
- Le taux de vaccination global est de 48% (nous avons été capables de calculer le taux réel, grâce à un arbre de probabilités, dans notre région où de nombreuses personnes âgées vivent et la valeur est de 43%).
- Le facteur moyen de contamination est 1.35 donc ce n'est pas si loin de 2.

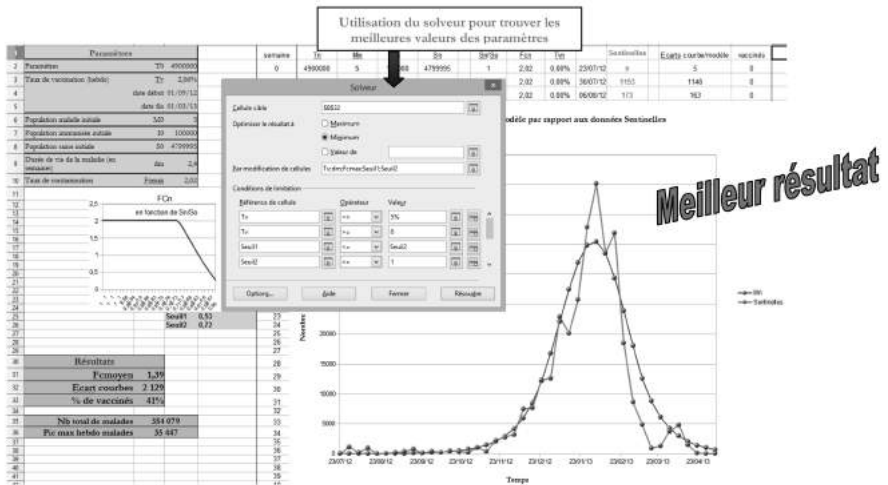
7. Amélioration des réglages

7.1. Utilisation du solveur

Jusqu'à présent, nous réglions les valeurs de nos paramètres à la main afin de mieux comprendre leurs rôles respectifs dans la qualité des résultats.

Pour obtenir des réglages plus fins, nous utilisons maintenant la fonction « solver » du logiciel. Cette fonction consiste à balayer toutes les combinaisons de paramètres (en ayant la possibilité d'ajouter des contraintes) afin d'optimiser (minimiser pour nous) une « cellule cible » (pour nous l'écart entre les courbes).

De cette façon, nous avons pu réduire l'écart entre les courbes de 50%



7.2. Analyse de la sensibilité des paramètres

Maintenant que notre modèle est optimisé, nous devons étudier la sensibilité de nos différents paramètres. Pour mieux visualiser les choses, nous avons fait un tableau représentant l'élasticité des différents paramètres.

L'élasticité d'un paramètre est le rapport entre la variation relative de l'écart entre les courbes et la variation relative de ce paramètre. Bien sûr, plus la valeur est élevée, plus la sensibilité au paramètre est importante et plus il est intéressant de jouer sur ce paramètre pour influencer sur le comportement de la maladie.

Pas de surprise, le taux de contamination et le taux de vaccination sont les plus sensibles (la durée de la maladie aussi, mais elle est inhérente à la maladie et nous ne pouvons pas la changer), mais maintenant nous pouvons quantifier cette sensibilité.

	Elasticité des différents paramètres						
	variation relative du paramètre		variation relative du résultat		élasticité		
	valeur initiale	valeur finale	variation	valeur initiale	valeur finale	variation	
taux de vaccination	2,06%	2,08%	1,00%	2174	2133	-1,89%	-1,89
	2,06%	2,04%	-1,00%	2174	2305	6,03%	6,03
	2,06%	2,16%	5,00%	2174	2297	5,66%	1,13
taux de contamination	2,02	2,04	1,00%	2174	2439	12,19%	12,19
	2,02	2	-1,00%	2174	2288	5,24%	-5,24
	2,02	2,12	5,00%	2174	6226	186,38%	37,28
population malade initiale	5	5,05	1,00%	2174	2173	-0,05%	-0,05
	5	4,95	-1,00%	2174	2180	0,28%	-0,28
	5	5,25	5,00%	2174	2221	2,16%	0,43
durée de vie maladie	2,4	2,42	1,00%	2174	2247	3,36%	3,36
	2,4	2,38	-1,00%	2174	2179	0,23%	-0,23
	2,4	2,52	5,00%	2174	2776	27,69%	5,54

8. Comment mieux contrôler la prochaine épidémie ?

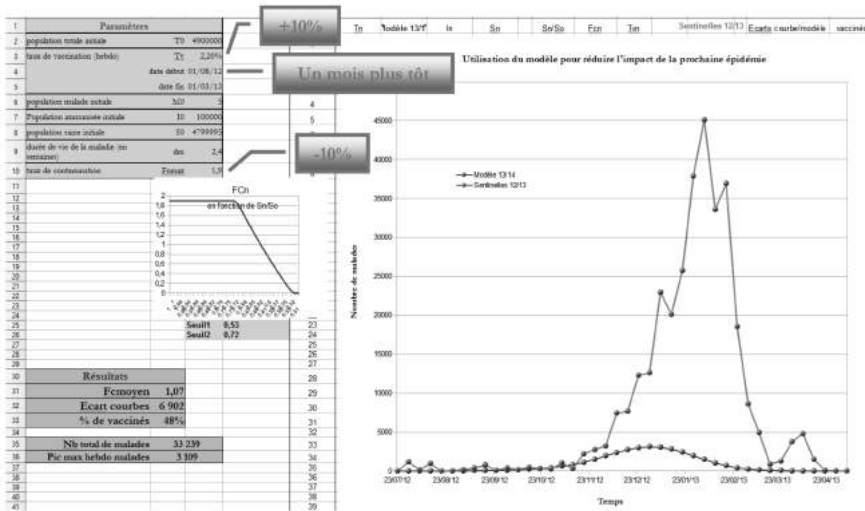
Maintenant que notre modèle est bien ajusté et que nous connaissons les paramètres les plus pertinents à modifier, nous pouvons l'utiliser pour voir comment prévenir la prochaine épidémie de grippe.

Nous n'avons été capables de faire baisser le nombre total de malades de 90%, en changeant seulement :

- le taux de vaccination hebdomadaire de 10%,
- diminuant notre FC_{max} de 10%,
- démarrant la campagne de vaccination un mois plus tôt.

Ce qui est tout à fait réalisable en effet !

La première en communiquant sur l'importance de la vaccination et le second en soulignant l'importance pour les gens de prendre davantage de mesures d'hygiène et/ou en leur permettant de rester à la maison quelques jours afin de réduire la propagation de la maladie.



9. Qu'est-ce qu'il resterait à faire ?

La principale chose pour améliorer notre modèle numérique serait de prendre en compte la météo et la répartition géographique de la population. Cela représente un bond énorme puisque cela signifie que nous devons maintenant faire un réel programme (et pas seulement un tableur) pour diviser notre région en zones beaucoup plus petites et modéliser la propagation à l'intérieur et entre ces zones.

Le problème vient beaucoup moins de l'écriture du programme (par exemple en utilisant Python ou Scilab) que de l'existence (ou notre capacité à les trouver) de données fiables avec la précision nécessaire (pour le nombre de personnes malades ou la météo). Nous pourrions alors reprendre notre réglage de la fonction de contamination, qui est l'élément central du modèle. Nous devrions aussi réfléchir sur la signification de nos seuils en termes d'épidémiologie de la maladie ou de psychologie de la population.