



PROJET ANR MURDASP : POUR UNE MOBILITÉ  
DURABLE ET ADAPTÉE À UN CONTEXTE DE  
PANDÉMIE ANR-21-HDF1-0014

**Rapport d'avancement N<sup>o</sup>1**

*Auteurs :*  
Équipe du projet

Septembre, 2021

# Table des matières

1	Aperçu sur l'avancement du projet	2
2	Aperçu sur les éléments techniques	3

# 1 Aperçu sur l'avancement du projet

Au niveau scientifique, nous avons avancé sur la revue de la littérature relative à la propagation de la pandémie et sa prise en compte dans les modèles de transport. A cet effet, un modèle simplifié a été développé nous permettant d'identifier progressivement les mécanismes mettant en relation la mobilité et la pandémie.

Sur un autre plan, la calibration du modèle a été améliorée afin d'aboutir à une représentation plus réaliste. En plus de converger vers des parts modales plus précises (quatre modes de transport sont considérés : marche, vélo, transport public et voiture), la population est subdivisée en trois groupes selon les revenus. Le modèle a été enrichi par l'ajout des flux de transport de marchandises ainsi que les déplacements transfrontaliers. Ces derniers flux sont particulièrement utiles pour la modélisation de la pandémie.

Points	Commentaires
Recrutement	Ousmane Diop, mathématicien de formation et spécialiste des modèles appliqués à la biologie a été recruté pour 12 mois à partir du 01 avril 2021
Investissement	Acquisition des calculateurs qui sont opérationnels depuis le mois de juin 2021
Adaptation du Modèle	Amélioration de la calibration ; segmentation de la population en trois groupes selon les revenus ; prise en compte des flux de transport de marchandises et de la mobilité transfrontalière

TABLE 1 – Synthèse avancement du projet.

La situation budgétaire est résumée sur le tableau 2. Les dépenses budgétaires ont globalement été réalisées. Elles sont, en effet, constituées en grandes parties par l'acquisition d'un équipement de calcul haute performance. Cet équipement est opérationnel depuis juin 2021.

Types	Allocation	Dépenses	Commentaires
Investissement	50 000	46 000	Dernières dépenses prévues pour fin 2021
Fonctionnement	26 000	3 000	Des difficultés de mobilités liées au COVID-19 ont sensiblement limitées les dépenses pour 2021 ; des dépenses sont prévues pour fin 2021 et 2022.

TABLE 2 – Résumé de la situation budgétaire.

## 2 Aperçu sur les éléments techniques

### 2.1 État de l’art

#### 2.1.1 Les modèles compartimentaux du type SIR

L’un des premiers travaux de modélisation d’épidémies remonte de Kermack et McKendrick en 1927[4]. Ces deux pionniers ont mis en place un modèle de type compartimental en subdivisant une population envahie par un agent infectieux (virus) en trois parties. La première, nommée les susceptibles (S) est constituée des individus sains, c’est à dire ceux qui n’a pas encore reçu le virus. La deuxième dit compartiment (I) considère les individus infectés et qui peuvent transmettre le virus à d’autres. En fin la troisième partie dit compartiment (R) représente les individus qui se sont remis de la maladie. Le modèle SIR de Kermack et McKendrick, les individus passent de (S) à (I) avec un taux de contamination  $\beta$  représentant la force de l’infection et de (I) à (R) avec un tau de guérison tous dépendants de la dangerosité du virus. Le modèle SIR, établi sous forme d’équations différentielles ordinaires, permet ainsi d’étudier l’évolution du nombre d’individus dans chaque compartiment au et à mesure de l’épidémie.

Suite aux travaux de Kermack et McKendrick, la modélisation mathématique en épidémiologie a considérablement bien progressé au fil des années et plusieurs adaptations du modèle SIR ont été proposées. D’ailleurs, on peut remarquer que pour certaines épidémies comme la grippe ou les celles à coronavirus telle que la Covid-19, les trois compartiments du modèle SIR ne décrivent pas tout à fait bien la dynamique de l’infection. En effet, les individus infectés peuvent observer un certain délai dit phase d’incubation avant de devenir infectieux. De plus, une fois infectieux, une phase asymptotique peut s’en suivre permettant à l’individu la possibilité de transmettre le virus sans pour autant tomber malade. Ceci conduit à des modèles du type SEIR ou SEIAR où E symbole les individus dits exposés et A représente les individus dans un état asymptotique. En vrai, le nombre de compartiments de ces familles de modèle dépend du comportement de l’épidémie mais également du niveau de détail que l’on souhaite expliquer avec le modèle. Par exemple, dans le cas de la Covid-19, la prise en compte d’un compartiment constitué des cas de contaminations sévères peut permettre d’anticiper sur la capacité d’hospitalisation des hôpitaux.

Dans le modèle de Kermack et McKendrick, une hypothèse phare est que tous les individus ont la même chance d’être infectés. Cette hypothèse sous-entend que le risque d’infection reste pareil pour tous et que les individus de la population ont un comportement semblable (homogénéité). Ainsi, le modèle SIR et ses variantes apparaissent comme étant généralement trop grossiers pour étudier les différentes connexions sociales pouvant être des sources de contamination majeures à savoir les lieux de contact comme les maisons, les transports, les lieux de travail, les loisirs, . . .*etc.*

Afin de contourner ces limitations, d’autres approches telles que les celles dites individus centrés sont proposées. Les modèles individus centrés, appelés également modèles multi-agents, s’appuient sur les utiles informatiques de calcul de haut performance et considèrent chaque individu de la population comme un objet avec un certain nombre d’attributs. Ainsi,

chaque individu, avec ses attributs spécifiques, a une contagiosité et une susceptibilité qui lui sont propres [1]. Par conséquent, le risque d’infection dépendra de plusieurs facteurs tels que l’intensité et la durée de contact entre individus, mais également de façon intéressante aux mesures de protections en vigueur qui participent au contrôle de l’épidémie.

### 2.1.2 Quelques cas d’études avec des approches individus centrés

Dans le cas de la Covid-19, les résultats de Neil Ferguson *et al.*, 2020, [3] ont été l’un des premiers travaux qui tentent d’expliquer la progression de la covid-19 en Europe et de suggérer des stratégies pour la ralentir. Les résultats de leur modèle ont servi dans les politiques de préventions non seulement en Grande Bretagne mais également dans plusieurs pays du monde. Entre autres, ils montrent qu’en l’absence de vaccin, les mesures barrières peuvent potentiellement réduire l’intensité de contact entre les populations et ainsi réduire la transmission du virus. Dans leur étude, les auteurs se sont focalisés sur le cas de la Grande Bretagne et des états unis et deux stratégies de base sont proposées : (a) l’atténuation, qui vise à ralentir, mais pas nécessairement à arrêter, la propagation de l’épidémie en réduisant les pics de demande de soins de santé tout en protégeant les personnes les plus exposées à l’infection, et (b) la suppression, qui vise à inverser la croissance de l’épidémie, en réduisant au fur et à mesure le taux de reproduction de base ( $R$ ) à de faibles niveaux (en dessous de 1) et en maintenant cette situation indéfiniment. Ils constatent que les politiques d’atténuation optimales, combinant l’isolement à domicile des cas suspects, la mise en quarantaine à domicile des personnes vivant dans le même foyer que les cas suspects et l’éloignement social des personnes âgées et des autres personnes les plus exposées au risque de maladie grave, pourraient réduire les pics de demande de soins de santé de  $2/3$  et les décès de moitié. Toutefois, l’épidémie atténuée qui en résulterait serait encore susceptible de provoquer des centaines de milliers de décès et de submerger les systèmes de santé, en particulier les unités de soins intensifs, à plusieurs reprises. De là, il est clair que pour les pays qui en sont capables, la suppression reste donc l’option politique privilégiée. Dans le contexte britannique et américain, la suppression nécessitera au minimum une combinaison de distanciation sociale de la population générale, d’isolement des cas à domicile et de mise en quarantaine des membres de leur famille. Ces mesures peuvent être complétées par la fermeture d’écoles et d’universités, bien qu’il faille reconnaître que de telles fermetures peuvent avoir un effet négatif sur la santé publique.

Dans [6, 7], Müller *et al.*, présentent une approche utilisant des données de mobilités de personnes sur un modèle multi-agents de simulation de transport couplé avec un modèle mécanistique d’infection et propagation d’épidémie. Le modèle multi-agent, basé sur l’outil de simulation MATsim (voir [8]), fournit une approximation du réseau d’interactions des individus de la population. Le modèle mécanistique s’inspire des approches classiques du type SIR avec une subdivision de la population en sept compartiments : Susceptibles, Exposés, Infectieux, Asymptomatiques, Cas confirmés, Cas critiques et Remis. Ce couplage entre le modèle multi-agents et le modèle d’infection est rendu possible grâce à l’outil EpiSim récemment développé dans [6, 7] (voir <https://github.com/matsim-org/matsim-episim-libs> pour plus de détails) qui s’intègre parfaitement dans MATsim.

Dans ce modèle, le risque d'infection d'un individu sachant qu'il est en contact avec un individu infectieux dépend de plusieurs paramètres tels que la force de l'infection qui est intrinsèque au virus, l'intensité de contact, le taux d'excrétion de l'individu infectieux, le taux d'inhalation, la durée du contact ..etc. S'intéressant à la ville de Berlin, ils montrent que les changements de comportements de la population se sont produits principalement avant même l'entrée en vigueur des mesures de distanciation sociale initiées par le gouvernement. Le modèle est ensuite utilisé pour étudier les variations du taux de reproduction de base ( $R$ ) pour les différentes interventions telles que la réduction de la participation aux activités, le port du masque ou le traçage électronique suivi d'une mise en quarantaine à domicile. Ils concluent de façon intéressante que le traçage électronique réduit le  $R$  de 30 à 40% s'il est bien réussi. En plus, lorsque le traçage électronique sature, le taux de reproduction remonte immédiatement et par conséquent des mesures de reconfinement deviennent nécessaires pour le faire chuter à nouveau.

Dans le même sillage que Müller *et al.*, Ouasim et Ciari proposent dans [5] le couplage d'un modèle de simulation multi-agents avec un modèle mécanistique d'infection pour étudier la dynamique de propagation de la pandémie Covid-19 à Montréal. Ils s'intéressent particulièrement à la contribution des différentes activités dans la transmission du virus à l'absence d'une quelconque politique de contrôle de l'épidémie. Comme pour le modèle de Müller, ils considèrent que le risque d'infection dépend presque des mêmes paramètres. Cependant, malgré le choix d'étudier l'épidémie en l'absence de politiques de contrôle, les auteurs ont quand même pris le soin de différencier les intensités de contact pour les lieux très fréquentés tels que les loisirs ou le milieu hospitalier par rapport aux lieux dans lesquels le contact n'est pas aussi important.

Ils montrent que les fréquentations dans les maisons, les lieux de travail et les écoles (les points d'attractions primaires) contribuent de façon très significative de la dynamique de l'infection. Quant aux lieux d'attractions tels que les loisirs, les commerces . . . *etc*, ils contribuent certes à la dynamique d'infection mais ils sont beaucoup moins significatifs que les premiers. Par ailleurs, le risque d'infection dans les espaces de travail dépend des secteurs économiques, par exemple, les personnels de santé sont plus exposés que les autres travailleurs. En plus, le risque d'infection est également sexe dépendant du fait de la répartition des hommes et des femmes dans des secteurs différents. Par exemple, les femmes sont plus exposées dans les services de santé et dans les écoles que les hommes alors que ces derniers le sont plus dans les secteurs de construction, de transport ...etc.

## Présentation du modèle simplifié

On considère un schéma simple constitué d'un réseau de deux voies aller-retour pour le déplacement de personnes entre leur domicile et leur lieu de travail. On suppose que la population empruntant quotidiennement ce réseau est de taille 50 personnes et que 80%, c'est à dire 40 personnes utilisent le transport public répartis dans 18 bus et que les 20% (10 personnes) utilisent quant à elles, leur car (voir Fig. 1).

Afin de permettre une interaction dans les lieux de travail et dans les domiciles, on suppose que les 50 personnes sont réparties dans 10 ménages de 5 personnes chacun et

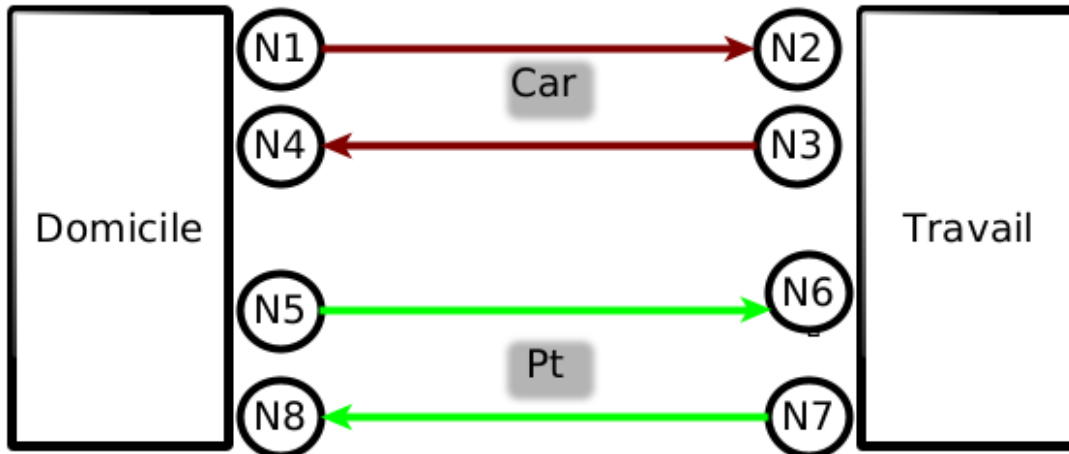


FIGURE 1 – Schéma du réseau

qu’elles travaillent dans 5 entreprises de 10 employés chacune. La répartition des personnes dans les entreprises est faite de sorte que deux personnes partageant le même domicile ne travaillent pas ensemble, permettant ainsi de différencier clairement les infections dans les maisons et celles issues du lieu de travail.

Le modèle d’infection et de propagation d’épidémie utilisé est celui de Müller *et al.*, [7]. Comme évoqué plus haut, le risque d’infection, dans le modèle de Müller, d’un individu sachant qu’il est en contact avec un individu infectieux, dépendant de l’efficacité du virus ( $\Theta$ ), de l’intensité de contact ( $ci$ ), du taux d’excrétion ( $sh$ ), du tau d’inhalation ( $in$ ), de la durée de l’interaction ( $\tau$ ). Dans cet exemple, nous considérons que le paramètre  $\Theta$  est connu et est fixé à  $\Theta = 9^{-10}$ , valeur correspondante à celle choisie dans le modèle de Berlin. Le paramètre  $ci$  modélise l’intensité de contact entre les personnes dans les différentes activités à savoir le domicile, lieu de travail et dans les transports. Les deux paramètres  $sh$  et  $in$  sont en lien avec le port du masque par les personnes. Cependant, ces deux valeurs peuvent être fixées à 1 pour qu’elles n’aient pas d’effet sur le calcul du risque d’infection dans le cas où on souhaite ignorer le port du masque ou tout simplement dans le cas où on souhaite préciser le type de masques portés par les personnes.

Ici, nous nous proposons de tester quelques scénarios en lien avec les mesures de restrictions tels que la distanciation sociale et le port du masque. La prise en compte des règles de distanciations sociales dans certaines activités passe par une simple réduction de l’intensité de contacte  $ci$  dans ces activités. S’agissant du port du masque, nous allons montrer à travers cette exemple que selon le type de masque porté, l’évolution de l’épidémie peut avoir une nette différence par rapport aux scénarios sans restrictions.

## Scénario\_0 : Simple sans restrictions

Considérons un premier scénario sans mesures de restrictions avec les différentes hypothèses suivantes :

- $\theta = 9^{-10}$  (provenant du modèle de Berlin),
- Intensité de contact ( $ci$ ) fixée à 1 pour toutes les activités,
- Taux d'excrétion ( $sh$ ) et taux d'inhalation ( $in$ ) fixées à 1,
- Durée de l'interaction  $\tau$  est donnée par MATSim,
- Nombre d'infectés initialement (patients zéros) :  $I_0 = 1$

En simulant le modèle avec le scénario\_0, l'épidémie ne dure qu'un mois environ et à la fin 40 sur les 50 personnes ont été contaminées (voir Fig. 4, droite). Dans la Fig. 4 (gauche), est représentée la part des trois activités domicile (home), le travail (work) et transport (pt) dans l'infection. On peut remarquer que, le transport public participe plus dans l'infection que les autres activités. Ceci pourrait être expliqué par le fait que les personnes ont tendance à se retrouver simultanément dans les bus aux horaires matinales pour se rendre à leur lieux de travail et en fin d'après midi pour se rendre chez elles, favorisant ainsi la transmission du virus dans les bus.

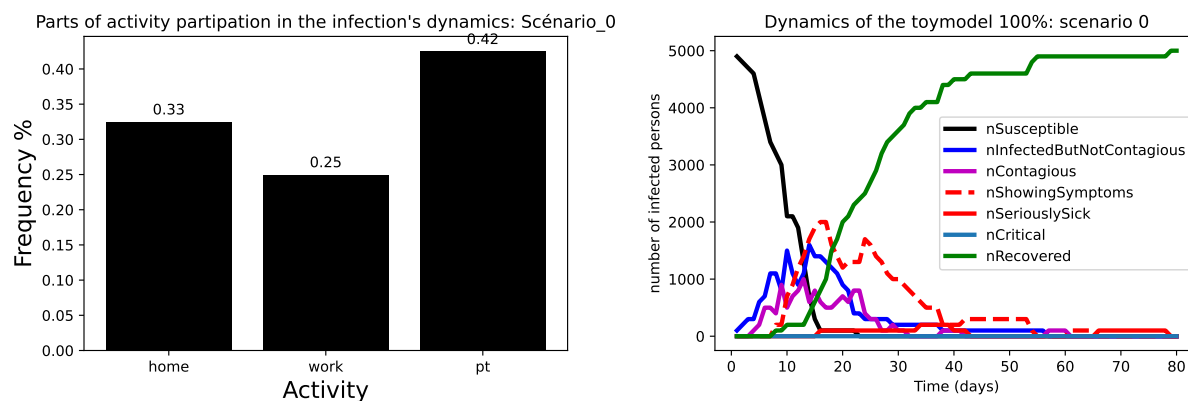


FIGURE 2 – Gauche : part de chaque activité dans la dynamique d'infection ; Droite : évolution du cumul du nombre d'infectés dans le temps.

## Scénario basé sur l'intensité de contact

Il est possible de cibler une activité et de proposer une mesure consistant à réduire l'intensité de contact entre les personnes. Nous proposons deux scénarios à partir du scénario\_0 en considérant les hypothèses suivantes :

$H_1$  := Réduction à moitié de l'intensité de contact de l'activité  $pt$  :  $ci_{pt} = 0.5$ .

$H_2$  := Réduction à moitié de l'intensité de contact de l'activité  $work$  :  $ci_{work} = 0.5$ .

- Scénario\_1 := scénario\_0 +  $H_1$ .
- Scénario\_2 := scénario\_0 +  $H_1$  +  $H_2$ .



La simulation des deux scénarios donne les Fig. 3(7a) et Fig. 3(7b).

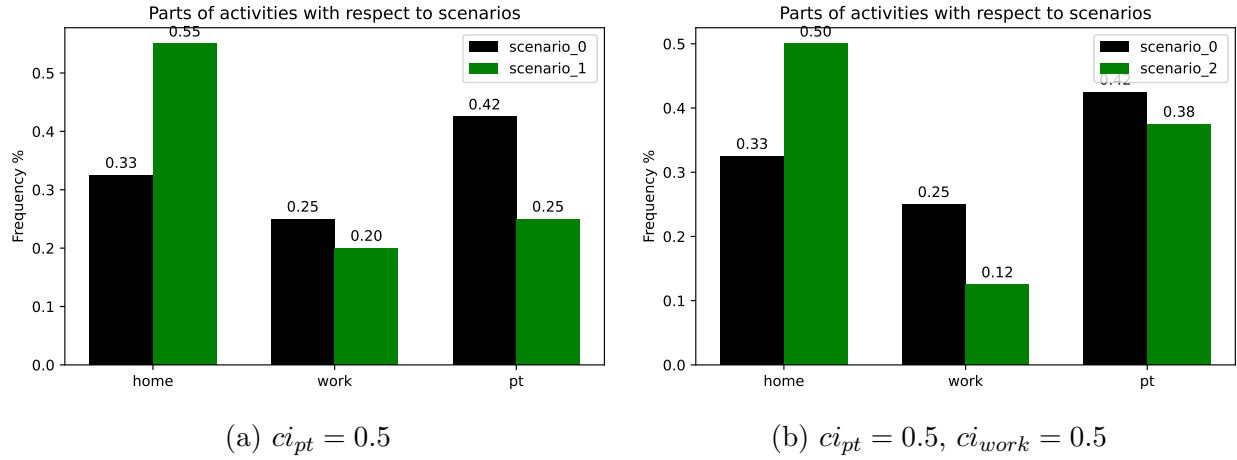


FIGURE 3 – Comparaison des deux scénarios avec le scénario\_0 selon la part de chaque activité dans la dynamique d’infection.

Pour nos deux exemples de scénarios, le nombre de personnes infectés au cours de l’épidémie n’a pas changé par rapport au scénario\_0. Ce constat n’est pas en réalité surprenant d’autant plus que le nombre de personne réduit de notre modèle peut amener l’épidémie à atteindre facilement une immunité collective ; et donc un arrêt de l’épidémie après un certain seuil. Cependant, il est à noter que la part de chaque activité dans l’infection diminue lorsqu’on diminue l’intensité de contact correspondante.

### Scénario basé sur le port du masque

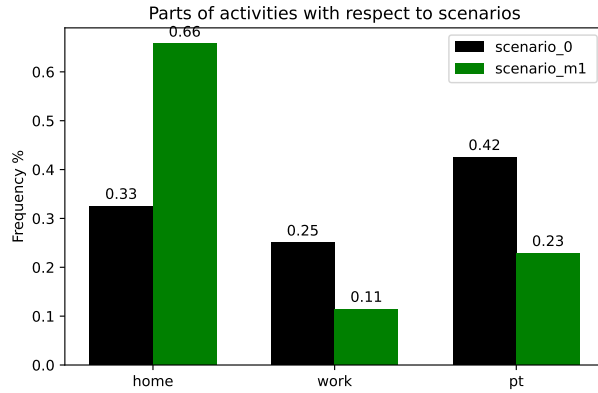
Il est possible d’appliquer des mesures de restrictions liées au port du masque. Dans ce cas, on choisit un type de masques donné et la proportion de personnes qui le porte. Nous illustrons l’utilisation du masque en considérant les trois types de masque proposés **N95**, **chirurgical** et **tissu** qui sont d’efficacités différentes.

Selon [2] (voir [7] également) les masques en tissu réduisent l’excrétion ( $sh$ ) et l’absorption ou l’inhalation ( $in$ ) de 0,6 à 0,5 de leurs valeurs initiales, les masques chirurgicaux de 0,3 à 0,2 et les masques N95 (FFP2) de 0,15 et 0,025. En d’autres termes, on peut retenir que dans l’ordre croissante d’efficacité, nous avons les masques tissus, les masques chirurgical et en fin les masques N95. Afin de tester la cohérence de notre modèle par rapport aux restrictions relatives au port du masque, nous considérons les trois scénarios suivantes :

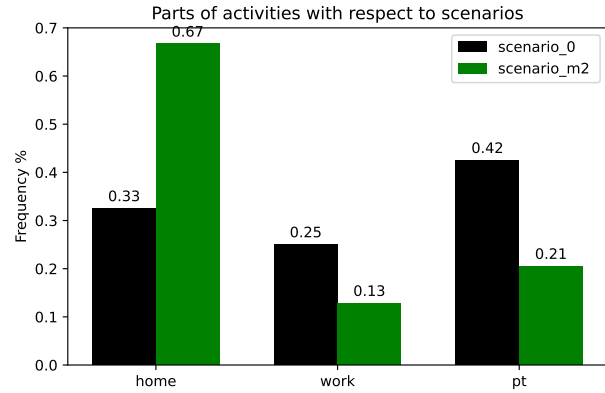
- Scénario\_ $mi$  := scenario\_0 +  $h_i$ 
  - $h_1$  : imposer un masque **N95** à toutes les personnes dans les lieux de travail (work) et dans les transports publics (pt).
  - $h_2$  : imposer un masque **chirurgical** à toutes les personnes dans les lieux de travail (work) et dans les transports publics (pt).

—  $h_3$  : imposer un masque **tissu** à toutes les personnes dans les lieux de travail (work) et dans les transports publics (pt).

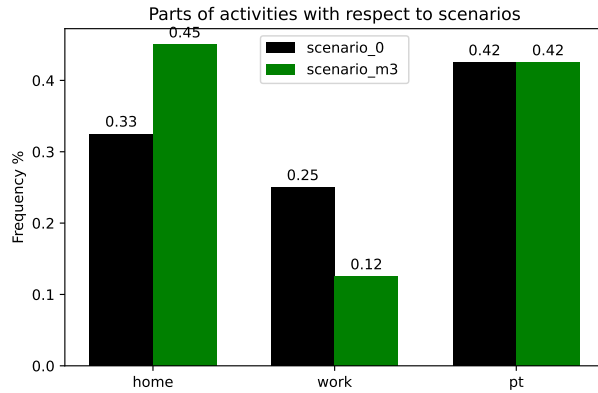
Ces trois scénarios consistent à imposer le port d'un type de masque dans le travail et dans le transport public et laisser libre aux personnes une fois rentrer dans leur domicile.



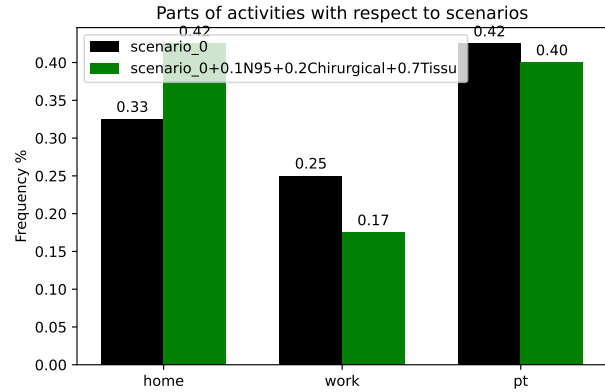
(a) N95 : 35 personnes infectées.



(b) Chirurgical : 39 personnes infectées



(c) Tissu : 40 personnes infectées



(d) N95 10%,chirurgical 20%,Tissu 70% : 40 personnes infectées

FIGURE 4

## 2.2 Différentes versions du modèle

Les versions du modèles se différencient par la prise en compte des déplacements transfrontaliers et la considérations des caractéristiques socio-économiques des populations. Ainsi, à partir du modèle développé par l'équipe à l'Université du Littoral, on distingue plusieurs versions (Table 3).

— A.1 (les grands voyageurs NPdC) : Ce sont les plans journaliers de déplacement des personnes dont un de leurs trajets a été effectué sur au moins une distance de

Modèles	Codes	Designation	Population (10%)
A (homogène)	A.1	Grands voyageurs NPdC	120 393
	A.2	A.1 + trafic transfrontalier + fret	128 976
	A.3	Tout voyageur NPdC	290 451
	A.4	A.3 + trafic transfrontalier + fret	299 034
	A.5	Population des Hauts-de-France	394 397
B (hétérogène)	B.1	A.1 + classification des usagers	120 393
	B.2	A.2 + classification des usagers	128 976
	B.3	A.3 + classification des usagers	290 451
	B.4	A.4 + classification des usagers	299 034
	B.5	A.5 + classification des usagers	394 397
C(innovation technologique)	C.1	A.x + voiture électrique	-
	C.2	A.x + voiture autonome	-
	C.3	A.x + modes partagés	-
P (Situation de pandémie)	P.1	A.3 + trafic transfrontalier	354 872
	P.2	P.1 + classification des usagers	354 872

TABLE 3 – Différentes versions du modèle multimodal de transport du Nord de la France

- 10km. Cette population s'intéresse plus à l'usage du train dans le Nord-Pas-de-Calais (NPdC) par rapport à la voiture dans les déplacements à longue distance.
- A.2 (A.1 + trafic transfrontalier + fret) : Cette version s'intéresse aux déplacements des grands voyageurs associés au trafic transfrontalier entre la France et la Belgique et les flux de marchandises.
  - A.3 (la population tout voyageur NPdC) : Elle concerne le déplacement de tous les voyageurs à long et court trajet dans le NPdC, importante pour l'utilisation des transports publics urbains en plus des questions évoquées par l'étude sur les grands voyageurs.
  - A.4 (A.3 + trafic transfrontalier + fret) : Elle comporte les déplacements de la version A.3 en plus du transport de marchandise ainsi que le flux frontalier avec la Belgique.
  - A.5 (La population des Hauts-de-France) : Elle reprend les déplacements de la version A.4 élargi à la région des Hauts-de-France.
  - Les versions B.1, B.2, B.3, B.4 et B5 sont composées des mêmes populations que respectivement les versions A.1, A.2, A.3, A.4 et A5 précitées. Mais, en plus, elles prennent en considération les revenus des ménages suivant une catégorisation des populations en trois groupes de revenus ( $\leq 11324.39$  pour le groupe 1,  $\geq 11324.39$  et  $\leq 22995.59$  pour le groupe 2 et  $\geq 22995.59$  pour le groupe 3).
  - La version P.1 utilisée actuellement pour le projet MURDASP est basé sur la version A.3 auquel le flux transfrontalier entre la Belgique et le Nord-Pas-de-Calais est intégré. La demande a été par contre réajustée pour tenir compte des personnes dont les déplacements n'ont pas été répertoriés dans l'enquête ménage,

alors que leur contribution dans la dynamique des épidémies est significative. Les modèles A.1, A.2, A.3 et A.4 étudient les déplacements dont l'origine ou la destination est dans le Nord-Pas-de-Calais, qu'ils proviennent du périmètre d'étude ou ailleurs alors que dans le modèle A.5, le réseau et la demande de transport regroupent toute la région Hauts-de-France.

La version A.4 a été simulée à hauteur de 10% pour les besoins du calibrage dont les résultats seront appliqués au modèle pandémique (P.1) pour voir s'il y'a nécessité de les réadapter.

## 2.3 Adaptation du modèle pour une simulation d'épidémies

Afin de rendre le modèle d'épidémie le plus réaliste possible, nous sommes amenés à générer à nouveau la demande. La création de la demande est basée sur deux sources : l'Enquête Régionale Ménage Déplacement (ERMD 2016) et les fichiers détails mobilité professionnelle de l'INSEE (MOBPRO 2016). La première fournit les déplacements en Hauts-de-France et la seconde, quant à elle, permet d'obtenir les déplacements transfrontaliers.

S'agissant des déplacements en Hauts-de-France, les ménages sont générés à 10% selon leur poids afin d'incorporer toutes les informations de l'enquête dans notre modèle. Une fois les ménages générés, les personnes issues de chaque ménage sont créées avec des identifiants uniques. Nous obtenons alors un échantillon de 512 205 personnes représentant environ 12% de la population totale.

Ensuite, pour l'ensemble des personnes créées, leurs trajets peuvent être sélectionnés dans les déplacements observés dans l'enquête. Ces déplacements permettent d'identifier les activités des personnes à partir de leurs motifs de déplacement à l'origine et à la destination. Les activités enregistrées concernent le domicile, le travail, les études, les achats et les loisirs.

Pour l'ensemble de ces activités, on crée ensuite des identifiants de lieux d'attraction qui leur correspondent. Par exemple, pour une personne travaillant dans une commune donnée, on crée un identifiant d'un lieu de travail (une entreprise) avec ses coordonnées dans la commune puis l'attribuer à la personne en question. Ce même procédé est utilisé pour l'ensemble des lieux d'attractions à savoir les commerces, les lieux études, les achats ainsi que les loisirs. Afin de garder une certaine cohérence dans la création des lieux de d'attraction, le nombre de lieux d'attraction par activité est créé en fonction de chaque commune et du nombre de personnes ayant la même activité comme motif de destination dans la commune. Le nombre de lieux d'attraction est ainsi créé de façon à affecter à plusieurs personnes, le même de lieu d'attraction pour permettre suffisamment d'interactions entre les personnes.

Pour les personnes dont les déplacements n'ont pas été répertoriés dans l'enquête ménage, nous avons été amenés à faire l'hypothèse qui consiste à les ranger en deux catégories : les enfants ( $\simeq 65\%$ ) et les personnes âgées (le reste). Ainsi, les enfants auront comme activités leur domicile et l'école, tandis que les personnes âgées n'ont que leur domicile comme activité. Cette hypothèse trouve sa pertinence dans le fait

que les enfants et les personnes âgées, malgré qu'ils se déplacent moins, participent fortement à la dynamique d'infection.

Tout comme pour le modèle initial, le flux de mobilité professionnelle en provenance ou en partance de la Belgique est également pris en compte en utilisant les fichiers détails mobilité professionnelle de l'INSEE (MOBPRO 2016). Dans cette enquête, seul le flux en partance vers la Belgique est donné. Afin de modéliser le flux inverse, c'est à dire les Belges travaillant dans le Nord-pas-de-Calais, nous avons considéré que ce flux correspond à la moitié du flux en partance vers la Belgique. Bien entendu, ce choix peut bien être ajusté mais il faut dire que c'est un bon compromis pour ne pas ignorer le flux inverse. Ainsi, dans la création des lieux de travail, des belges travaillant dans le Nord-pas-de-Calais sont tous logés dans des entreprises situées dans leur commune de déplacement avec bien sûr des Français.

Une fois ces facilities créés, les origines et destinations des déplacements des personnes issues des ménages, ont été logés dans les différentes facilities. Ensuite, les déplacements concernant le périmètre d'étude du Nord de la France sont filtrés. En effet, ils regroupent les personnes dont une origine ou une destination de leur trajet est dans le réseau du Nord-Pas-de-Calais. Ceci permet d'étudier dans un premier temps, la dynamique de propagation du virus en fonction des personnes qui circulent sur le périmètre d'étude avant son extension, dans un second temps, dans la région Hauts-de-France. La demande au format compatible avec l'outil de simulation est alors générée sur la base du résultat de ce filtre.

## 2.4 Calibration

Après la mise en place du modèle composé des données sur l'offre (réseau multimodal, services de transport public) et la demande (plans journaliers des usagers) dans le Nord-Pas-de-Calais, il convient de le calibrer pour le rapprocher au plus de la réalité (5).

Le calibrage est précédé de l'étape d'échantillonnage (scaling) compte tenu de la taille du réseau et dont la simulation de toute la population consomme des temps de calcul et des ressources informatiques trop importantes. En pratique, cette mise à l'échelle se concrétise par une reconfiguration des paramètres de simulation à définir par itération jusqu'à ce que le modèle soit plus réaliste.

La réduction de la population simulée implique une réduction proportionnelle de la capacité du réseau pour éviter des incohérences sur différents paramètres tels que les durées ou distances moyennes des déplacements. Dans MATSim, il s'agit de considérer deux paramètres :

- la capacité de flux (*flowCapacityFactor*) qui indique combien de véhicules peuvent quitter la liaison par unité de temps ;
- la capacité de stockage (*flowStorageFactor*) qui renvoie à combien de véhicules peuvent être stockés en même temps dans le lien

En fonction du facteur d'échelle de réduction de la population ( $f_{scale}$ ), les paramètres de réduction de la capacité du réseau sont donnés les expressions de (1).

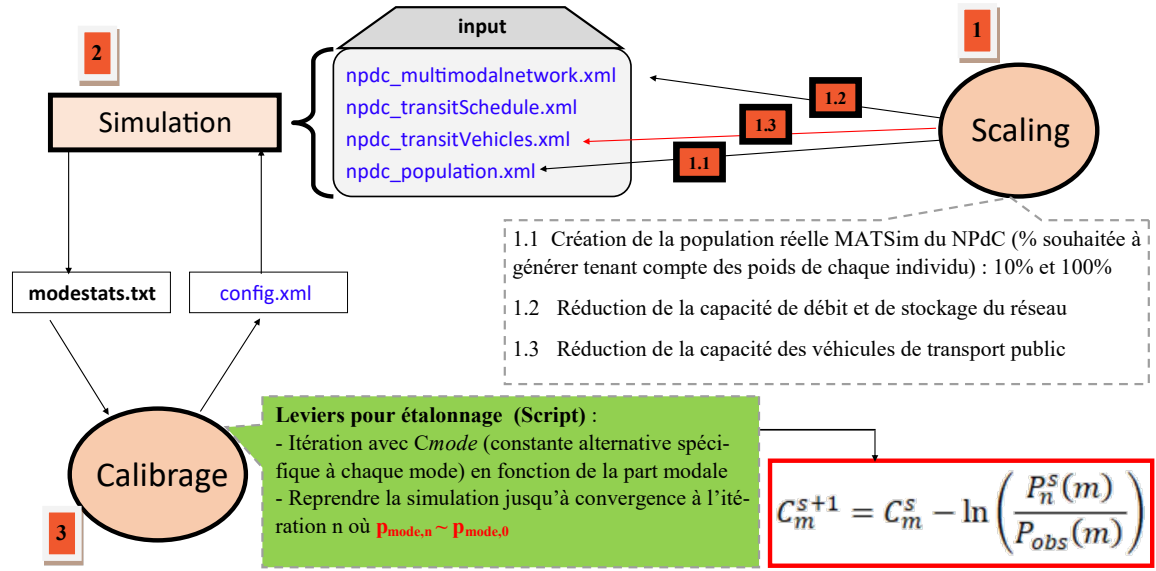


FIGURE 5 – Approche méthodologique de calibration du modèle multimodal de transport

$$\begin{aligned} flowCapacityFactor &= f_{scale} \\ flowStorageFactor &= (f_{scale})^{0.75} \end{aligned} \quad (1)$$

Les deux demandes présentent des populations et des parts modales différentes (Figure 6). La population du modèle initial concerne environ 3 millions d'agents contre 3,5 millions pour la population du modèle pandémique. Les deux modèles sont dominés par l'usage de la voiture (66% pour le premier contre 54% pour le deuxième) et la marche (25% pour le premier contre 38% pour le deuxième). L'importance de la part de la marche à pied pour le modèle pandémique se justifie par l'hypothèse de la participation des enfants dans la dynamique d'infection, non prise en compte dans l'enquête ménage déplacement. L'usage des transports publics (train, métro, tram ou bus) et du vélo est cependant assez similaire dans les deux types de population, de l'ordre respectivement de 6 et 2%. En effet, les hypothèses n'ont pas d'influence sur l'usage de ces modes.

Après calibrage, ces parts modales devront être approximativement retrouvées pour valider le modèle.

La calibration du modèle est effectuée à travers la formule d'utilité qui permet de traduire en score, les heures passées par l'utilisateur dans ses activités ou sur son trajet. La désutilité du déplacement a servi de calibrer le modèle en réajustant les parts modales à travers un coefficient d'attraction d'un mode selon (2) :

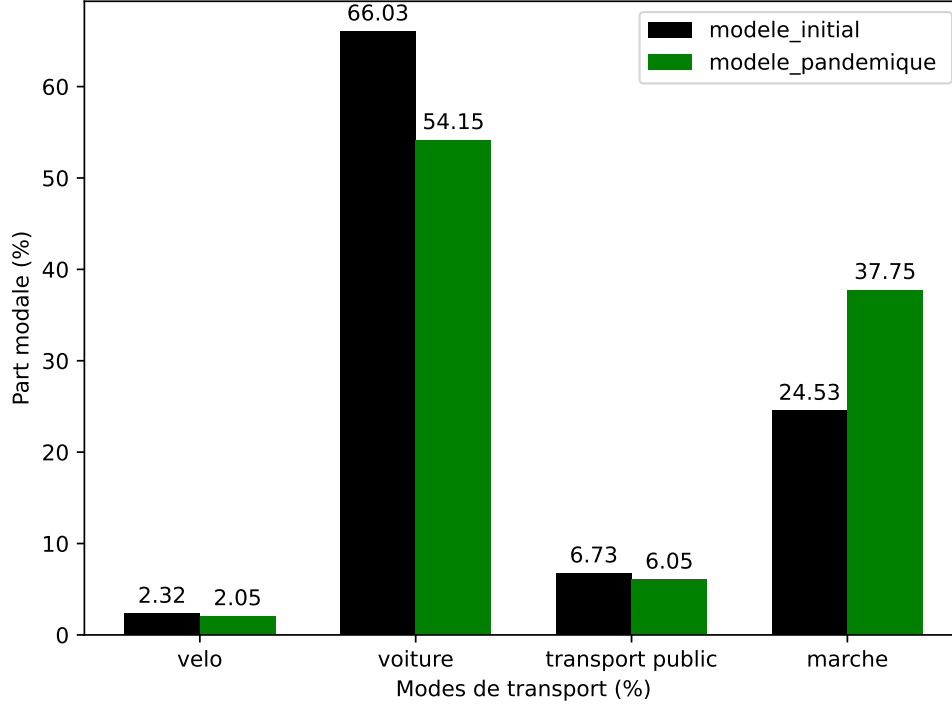


FIGURE 6 – Répartition modale de la population à 10% (modèles initial vs pandémie)

$$S_{trav,q} = C_{mode(q)} + \beta_{trav,mode(q)} \cdot t_{trav,q} + \beta_m \cdot \Delta m_q + (\beta_{d,mode(q)} + \beta_m \cdot \gamma_{d,mode(q)}) \cdot d_{trav,q} + \beta_{transfer} \cdot x_{transfer,q} \quad (2)$$

où,

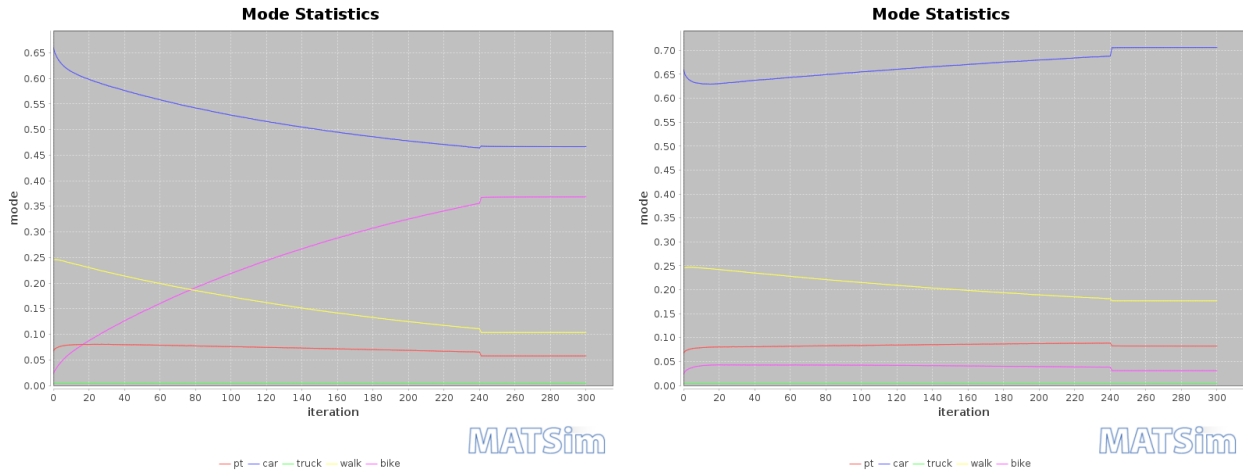
- $C_{mode(q)}$  (constant) est une constante spécifique au mode.
- $\beta_{trav,mode(q)}$  (marginalUtilityOfTraveling\_util\_hr) est l'utilité marginale directe du temps passé à voyager par mode (il prend en compte aussi l'utilité marginale du temps en tant que ressource).
- $t_{trav,q}$  est le temps de trajet entre les lieux d'activité  $q$  et  $q+1$ .
- $\beta_m$  (marginalUtilityOfMoney) est l'utilité marginale monétaire (normalement positive).
- $\Delta m_q$  est la variation du budget monétaire causée par les tarifs ou les péages (normalement négative ou nulle).
- $\beta_{d,mode(q)}$  (marginalUtilityOfDistance\_util\_m) est l'utilité marginale de la distance (normalement négative ou nulle).
- $\gamma_{d,mode(q)}$  (monetaryDistanceRate) est le taux de distance monétaire spécifique au mode (normalement négatif ou nul).
- $d_{trav,q}$  est la distance parcourue entre les lieux d'activité  $q$  et  $q+1$ .

- $\beta_{transfer}$  (waitingPt) sont des pénalités de correspondance pour les transports publics (normalement négatives).
- $x_{transfer,q}$  est une variable 0/1 signalant si une correspondance s'est produite entre l'étape précédente et l'étape actuelle.

La calibration du modèle initial est basée sur la constante spécifique au mode (Figure 5) où  $P_{obs}(m)$  est la part modale du mode  $m$  issue des enquêtes et  $P_n^s(m)$  la part modale du mode  $m$  à l'issue de la simulation  $s$  et de l'itération  $n$ .  $C_n^s(m)$  et  $C_n^{s+1}(m)$  sont respectivement les coefficients spécifiques au mode  $m$  pour les simulations  $s$  et  $s + 1$ .

L'algorithme de calibrage se déroule itérativement de la manière suivante :

- Simulation du scénario avec des valeurs de  $C_{mode} = 0$ ;
- Récupération des parts modales à l'itération  $n$ ;
- Calcul des nouvelles valeurs de  $C_{mode}$ ;
- Création d'un nouveau fichier de configuration;
- Simulation du scénario avec les nouvelles valeurs de  $C_{mode}$ ;



(a) Modèle initial non calibré.

(b) Modèle initial calibré

FIGURE 7 – Répartition des parts modales avant et après calibration du modèle initial

La calibration a permis de rendre le mode vélo moins attractif (Figure 7), de redresser progressivement l'usage de la voiture, tout en gardant des proportions des transports publics assez similaires aux données d'enquête. Dans la suite du projet, les coefficients de calibration seront appliqués au modèle pandémique pour évaluer les parts modales avant et après calibrage. D'autres moyens devront également permettre de réduire davantage l'usage de la voiture et d'augmenter celui de la marche pour qu'elles soient plus réalistes.



## 2.5 Prochaines étapes

- Validation du modèle de simulation MATSim : les paramètres de calibration du modèle initial seront appliqués au modèle pandémique. Si les parts modales issues de l'enquête et des hypothèses effectuées sur les déplacements des enfants et personnes âgées inactives, sont très écartées des valeurs sur le modèle, celui devra être calibré à nouveau.
- Simulation du modèle d'épidémiologie avec différents scénarios : Une fois le modèle MATSim validé, différents scénarios seront simulés pour évaluer la dynamique de propagation du virus en fonction de la mobilité des personnes et des mesures envisagées pour interrompre sa progression.

## Références

- [1] Ramsès Djidjou-Demasse, Christian Selinger, and Mircea T Sofonea. Épidémiologie mathématique et modélisation de la pandémie de covid-19 : enjeux et diversité. *Revue Francophone des Laboratoires*, 2020(526) :63–69, 2020.
- [2] Steffen E Eikenberry, Marina Mancuso, Enahoro Iboi, Tin Phan, Keenan Eikenberry, Yang Kuang, Eric Kostelich, and Abba B Gumel. To mask or not to mask : Modeling the potential for face mask use by the general public to curtail the covid-19 pandemic. *Infectious Disease Modelling*, 5 :293–308, 2020.
- [3] Neil Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, et al. Report 9 : Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand. *Imperial College London*, 10(77482) :491–497, 2020.
- [4] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772) :700–721, 1927.
- [5] Ouassim Manout and Francesco Ciari. The contribution of everyday activities to the spread of covid-19. an agent-based simulation from montreal, canada. 2021.
- [6] Sebastian A Müller, Michael Balmer, Billy Charlton, Ricardo Ewert, Andreas Neumann, Christian Rakow, Tilmann Schlenther, and Kai Nagel. Using mobile phone data for epidemiological simulations of lockdowns : government interventions, behavioral changes, and resulting changes of reinfections. *medRxiv*, 2020.
- [7] Sebastian A Müller, Michael Balmer, William Charlton, Ricardo Ewert, Andreas Neumann, Christian Rakow, Tilmann Schlenther, and Kai Nagel. A realistic agent-based simulation model for covid-19 based on a traffic simulation and mobile phone data. *arXiv preprint arXiv :2011.11453*, 2020.

- [8] Kay W Axhausen, Andreas Horni, and Kai Nagel. *The multi-agent transport simulation MATSim*. Ubiquity Press, 2016.