



HAL
open science

**SAMUFLUX : une démarche outillée de diagnostic et
d'amélioration à base de doubles numériques :
application aux centres d'appels d'urgence de trois
SAMU**

Eva Petitdemange

► **To cite this version:**

Eva Petitdemange. SAMUFLUX : une démarche outillée de diagnostic et d'amélioration à base de doubles numériques : application aux centres d'appels d'urgence de trois SAMU. Autre [cs.OH]. Ecole des Mines d'Albi-Carmaux, 2020. Français. NNT : 2020EMAC0012 . tel-03267854

HAL Id: tel-03267854

<https://theses.hal.science/tel-03267854>

Submitted on 22 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Fédérale



Toulouse Midi-Pyrénées

THÈSE

en vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

délivré par

IMT – École Nationale Supérieure des Mines d'Albi-Carmaux

présentée et soutenue par
Eva PETITDEMANGE

le 06 novembre 2020

**SAMUFLUX : une démarche outillée de diagnostic et
d'amélioration à base de doubles numériques :**
application aux centres d'appels d'urgence de trois SAMU

École doctorale et discipline ou spécialité :

EDSYS : Génie industriel et Informatique

Unité de recherche :

Centre Génie Industriel, IMT Mines Albi

Directeurs de thèse :

Matthieu LAURAS, Professeur, IMT Mines Albi

Franck FONTANILI, Maître-assistant, IMT Mines Albi

Autres membres du jury :

Thibaud MONTEIRO, Professeur, INSA Lyon, (*Rapporteur*)

Christophe MERLO, Professeur, ESTIA, Bidart, (*Rapporteur*)

Maria DI MASCOLO, Professeur, Université Grenoble Alpes, (*Examinatrice*)

Grégory ZACHAREWICZ, Professeur, IMT Mines Alès, (*Examineur*)

Xavier LORCA, Professeur, IMT Mines Albi, (*Président*)

Elyes LAMINE, Maître de Conférences, IMT Mines Albi et INU Champollion, (*Encadrant*)

Dr. Yann PENVERNE, CHU Nantes, Nantes, (*Invité*)

Remerciements

« Si on t'a donné les clés du camion, c'est que t'es capable de le conduire. » - Q. Schoen

Avant de vous laisser vous plonger dans la lecture de cette thèse, il me semble important de remercier un certain nombre de partenaires et de personnes sans qui cette traversée n'aurait été ni possible, ni aussi sympathique.

Tout d'abord, je tiens à remercier la région Occitanie pour le financement de cette thèse ainsi que nos établissements partenaires : les SAMU de Haute-Garonne, de l'Aveyron et du Tarn. Sans leur soutien je n'aurais jamais eu l'occasion de monter à bord de cette barque que j'ai menée durant trois ans de thèse.

Un grand merci à Christophe et Thibaud d'avoir accepté d'être rapporteurs, pour leur lecture attentive, leurs remarques précieuses ainsi que les échanges qui en ont découlés lors de la soutenance. Merci également à Maria et Gregory d'avoir endossé le rôle d'examineur/examinatrice : j'ai apprécié l'intérêt que vous avez porté à mes travaux et je vous remercie également pour vos questions qui me permettront d'aller plus loin. J'ai eu la chance d'avoir la présence d'un docteur en médecine dans mon jury, merci Yann de m'avoir apporté ton expertise métier de médecin urgentiste. Enfin merci Xavier d'avoir présidé mon jury, d'avoir été présent pour ce jour spécial de la soutenance mais aussi pour tous les autres jours durant ces trois belles années.

À Matthieu, je ne saurais assez te remercier pour ta présence, ta bienveillance et ton humanité, mais si ce n'était que ça ! Ce ne serait pas rendre hommage à ton esprit scientifique affûté toujours prêt à trouver les bons mots pour formuler la problématique à traiter. J'ai tant appris en travaillant avec toi durant ces trois années et j'ai aussi apprécié te découvrir dans le cadre du théâtre et j'espère qu'on pourra rapidement remonter sur les planches ensemble ! Franck, sans toi, cette aventure n'aurait jamais été possible, merci d'être venu me chercher de l'autre côté de l'Atlantique. Merci d'avoir été pour moi un modèle de rigueur, de précision, et de travail acharné. Ce sont de belles qualités que je ne cesserai de cultiver, merci pour ton enthousiasme et ta présence, j'ai vraiment apprécié partager ce projet, ce temps, ces années et cette passion avec toi, il nous reste encore beaucoup à faire ensemble ! Je connais ton amour du bricolage, et je suis heureuse de pouvoir participer encore un peu à la construction de l'axe IOS, j'ai confiance, cette structure va être magnifique !

Merci Elyes pour la mine d'or d'informations que tu sais toujours amener afin d'ouvrir de nouveaux horizons. Enfin, je tiens à remercier Uche avec qui la rédaction d'article devient un plaisir ! Un grand merci à toi pour ces bons moments parsemés de rire !

Si la thèse est semblable à une traversée en haute mer, alors il y a des jours de beaux temps où la mer est calme mais aussi des tempêtes qu'il faut traverser. Pour nous aider, il y a des lumières qui brillent la nuit, qui permettent de nous orienter, de nous rassurer et qui nous accompagnent durant la traversée. Durant mon voyage, j'ai eu la chance de croiser de nombreuses lumières, les plus brillantes sont peut-être les plus proches, celle du laboratoire, qui au delà d'être un environnement de travail absolument fantastique est surtout un lieu où j'ai la chance d'avoir, certes des collègues, mais surtout des amis et ça, c'est très précieux. C'est un endroit où je me sens chez moi, où je saurai toujours trouver une porte ouverte si j'ai une question, une considération scientifique, une idée ou encore juste envie de prendre un café. Je pourrais tous vous nommer mais cela serait bien long, alors à vous tous, les étoiles du CGI, ne cessez jamais de briller et continuez à guider les doctorants dans leur traversée !

Si la thèse est semblable à l'ascension d'une montagne alors vous n'êtes pas sans savoir que les derniers mètres sont souvent les plus durs, les plus fatigants. C'est sans compter les encouragements de nos compagnons de voyage. Les amis qui nous accompagnent, qui cheminent avec nous. J'ai eu la chance de faire de très belles rencontres sur le discord PhD Student. Paradoxalement, cette pandémie ne m'a pas isolée ni coupée du monde mais elle m'a donné des amis pour la route. Il y a ceux qui sont déjà en haut, qui nous attendent, témoins vivants du fait que le sommet existe, ils sont source d'espoir, on puise en eux la force d'aller de l'avant. Il y a ceux qui marchent à nos côtés avec qui on avance, un pas après l'autre ils brisent la monotonie et la solitude du voyage, ils illuminent la route de leurs rires et de leur bonne humeur. Un grand merci de m'avoir poussée, soutenue, épaulée durant cette difficile étape de rédaction. Maintenant que j'y suis, je peux vous assurer que la vue est belle d'ici, et que les efforts fournis valent la peine. À vous, qui commencez votre ascension, je serai là pour vous encourager, vous tendre la main, pour vous aider à franchir les derniers pas avant le sommet, comme on l'a fait pour moi.

Merci à ma famille, pour leur soutien et leur support indéfectible, pour m'avoir fourni toutes les bases nécessaires à ce que j'ai accompli aujourd'hui. Merci Johan et Amélie, vous que je considère aujourd'hui comme ma famille, merci pour votre présence, votre soutien tout au long de ce parcours, votre amitié m'est extrêmement précieuse.

Durant les longues traversées, il y a aussi des lumières terrestres qui s'éteignent pour rejoindre celles du ciel, Mamie, Mémère, j'espère que vous êtes fières, je l'ai fait ! Je pense aussi à Isa, sache que je continue à voir ton sourire dans les rayons de soleil.

Le doctorat est certes une belle aventure scientifique, mais c'est avant tout une aventure humaine et un voyage personnel qui m'a apporté beaucoup et m'a fondamentalement changée. Je ne suis plus la même qu'il y a trois ans, j'ai traversé des océans, effectué l'ascension de sommets, et c'est le résultat de ces pérégrinations que vous vous apprêtez à lire.

Table des matières

<i>Remerciements</i>	iii
<i>Table des matières</i>	v

Introduction générale	3
Contexte	3
Problématique générale	6
Méthodologie de recherche	9
Organisation du manuscrit	11
1 Services d'Aide Médicale Urgente : Enjeux et défis	15
1.1 La diversité des services de soins d'urgence	15
1.1.1 Des modèles formels et informels	15
1.1.2 La chaîne des urgences : Amont et Aval	16
1.1.3 L'accès aux services d'urgences : numéro unique ou numéros spécialisés ?	17
1.1.4 Quels sont les différents modèles de décrochés des appels	18
1.1.5 Des stratégies pré-hospitalières différentes : <i>Stay and Play</i> ou <i>Scoop and Run</i> ?	19
1.2 Le modèle français du SAMU : Service d'Aide Médicale Urgente	23
1.2.1 Le SAMU français : Naissance et Histoire	23
1.2.2 Les différents flux	25
1.2.3 Le cycle de vie d'un appel entrant	26
1.2.4 Les ressources	27
1.2.5 Processus de traitement d'un appel	29
1.2.6 Les systèmes d'information des SAMU	31
1.2.7 Évaluation des SAMU	33
1.3 Revue de littérature sur la gestion des SAMU	37
1.3.1 Les problèmes stratégiques	39
1.3.2 Les problèmes tactiques	42
1.3.3 Les problèmes opérationnels	45

1.3.4 Synthèse : Un déséquilibre des études entre la partie amont et aval.	50
2 Positionnement et questions de recherche	55
2.1 Un besoin métier à traiter	55
2.2 Les SAMU : un objet d'étude très hétérogène à l'échelle nationale . .	58
2.2.1 Les données brutes agrégées sur l'année 2018	59
2.2.2 Définir des groupes de SAMU sur la base des données nationales	61
2.3 Les verrous scientifiques	67
2.4 Formulation de la question de recherche	70
3 Proposition d'un cadre de diagnostic	73
3.1 Introduction du besoin	73
3.1.1 La notion de diagnostic	73
3.1.2 Contexte et justification de la question de recherche : Vers une méthode outillée de diagnostic organisationnel des SAMU	75
3.2 Présentation de la méthodologie	77
3.2.1 Étape 1 : Collecter et préparer les données	78
3.2.2 Étape 2 : Interpréter et intégrer	81
3.2.3 Étape 3 : Diagnostiquer	86
3.2.4 Étape 4 : Améliorer	89
3.3 Synthèse	90
4 Influence des types d'organisation sur les performances d'un SAMU	93
4.1 Étape 1 : Collecter et préparer les données	94
4.1.1 Les approches pour la découverte et la cartographie des processus	94
4.1.2 Sous-étape 1.a : Collecter et nettoyer les données	96
4.1.3 Sous-étape 1.b : Découvrir le processus métier	101
4.2 Étape 2 : Interpréter et intégrer	104
4.2.1 Sous-étape 2.a : Fusionner les modèles	106
4.2.2 Sous-étape 2.b : Extraire les connaissances	109
4.3 Synthèse	119
5 Expérimentations SAMU	121
5.1 Étape 3 : Diagnostiquer	122
5.1.1 Sous-étape 3.a : Paramétrer le double numérique ; Hypothèses et Modèle conceptuel	122
5.1.2 Sous-étape 3.b : Évaluation de l'AS-IS : Un Modèle conceptuel pour guider la construction du modèle	128
5.1.3 Sous-étape 3.c : Validation du modèle	135
5.2 Étape 4 : Améliorer; Scénarios de Simulation, Analyses et Résultats .	139
5.2.1 Comment hiérarchiser les types d'appels, quel impact sur les performances ?	139
5.2.2 Des propositions pour le traitement des appels : Le Triage et La Prépriorisation	139

5.2.3	1ère Expérimentation : Quels sont les effets de la prépriorisation et du triage ?	144
5.2.4	Une piste d'amélioration : modifier la priorisation des appels sortants	156
5.2.5	Gérer une crise : comment le centre d'appels réagit-il confronté à une augmentation soudaine et significative des appels ?	163
5.3	Synthèse	177
6	Conclusion et perspectives	179
6.1	Bilan du projet de recherche	179
6.2	Perspectives	181
6.2.1	Perspectives applicatives	181
6.2.2	Des perspectives académiques	182
<hr/>		
A	Recommandations du rapport Marcus	185
B	Profil des appels entrants à Rodez	189
C	Modèle de simulation agrandi	191
<hr/>		
	<i>Table des figures</i>	193
	<i>Liste des tableaux</i>	197
	<i>Bibliographie</i>	199

*Hâtez-vous lentement ; et, sans perdre courage,
Vingt fois sur le métier remettez votre ouvrage :
Polissez-le sans cesse et repolissez-le ;
Ajoutez quelques fois, et souvent effacez.*

Nicolas Boileau (L'Art Poétique)

Introduction générale

Existe-t-il pour l'homme un bien plus précieux que la santé ?

Socrate

Contexte

Les services d'urgences médicales (Emergency Medical Systems - EMS) et particulièrement leurs centres d'appels peuvent être considérés comme la porte d'entrée du système de soin pré-hospitalier pour toute demande urgente ou non programmée. Leur bon fonctionnement est donc crucial pour une prise en charge rapide du patient. La gestion de l'urgence est la mission principale de ces organisations et le moindre impair peut avoir des conséquences très graves sur la vie des patients.

Durant les dernières années, en France, on note des difficultés au niveau des SAMU à décrocher tous les appels d'urgences suffisamment rapidement. Soucieux de toujours s'améliorer et d'apporter à la population un service de qualité, plusieurs études ont été entreprises pour mieux comprendre le fonctionnement du SAMU. Elles révèlent un certain nombre de difficultés bien réelles. La santé est un sujet sensible et qui fait rapidement réagir le grand public. En réponse à cela, les gouvernements successifs de ces dernières années ont annoncé des réformes du système de soins d'urgence, notamment avec une unification de la formation des Assistants de Régulation Médicale (ARM) qui, jusqu'à présent, n'était pas unique. Cela s'inscrit dans une volonté d'uniformiser les pratiques des SAMU à travers le cadrage qualité national opéré par le Ministère de la santé.

La grippe H1N1 de 2009 avait déjà, à l'époque, mis à l'épreuve les capacités des SAMU à répondre à l'urgence et à réagir rapidement. À l'issue de cette crise, des mesures ont été prises afin d'améliorer le fonctionnement global de nos systèmes d'urgence. Un état des lieux du fonctionnement des SAMU a été fait et consigné dans

le rapport Mardegan¹. Ce rapport souligne la grande hétérogénéité des SAMU sur les plans techniques, structurels, et sur celui de la gestion des ressources humaines. Les conclusions de ce rapport préconisent de renforcer les liens entre les SAMU au niveau national et soulignent la nécessité de les moderniser. C'est notamment à cette occasion que le projet SI-SAMU a été lancé afin d'uniformiser les systèmes d'information. Cet outil commun a pour ambition d'uniformiser les systèmes d'informations des SAMU en proposant un outil centralisé qui laisse augurer des évolutions organisationnelles à l'échelle d'un territoire de santé. Ce projet est encore en cours fin 2020, et cette nouvelle solution technique est petit à petit déployée dans une dizaine de SAMU pilotes avant de l'être dans tous les SAMU de France.

Toutefois, la récente crise du coronavirus de 2020 (COVID-19) a de nouveau mis à l'épreuve nos systèmes de santé à l'échelle de la planète. Les premières semaines de cette pandémie virale ont engendré une forte augmentation du nombre d'appels d'urgences. En ce qui concerne la France, dans une interview² pour le journal « La Croix » en date du 6 mars 2020, un médecin régulateur du SAMU de Lille annonçait que le nombre d'appels avait plus que triplé avec la pandémie. Il relate aussi qu'il est arrivé que plus de 40 appels soient mis en attente. À Toulouse, une cellule de crise a été mise en place dès le 3 mars³ pour endiguer l'augmentation de 40% de leur activité. Sur la journée du 9 mars ce sont 627 dossiers de régulations qui ont été ouverts uniquement pour des appels à propos du coronavirus. On compte en temps normal environ 900 dossiers durant la semaine et 1100 durant le week-end. Les centres d'appels ont dû réagir rapidement pour augmenter leurs effectifs, adapter leurs manières de fonctionner afin d'être capable de répondre au besoin d'information de la population, mais aussi de gérer ce qu'on peut appeler l'urgence quotidienne. Pour cela, des étudiants en médecine ainsi que des médecins retraités sont venus agrandir les rangs des centres d'appels⁴. Un tel événement met en évidence qu'il est encore nécessaire de prendre du recul sur le fonctionnement de nos services d'urgences ainsi que sur la capacité d'adaptation des équipes à répondre aux demandes sanitaires en quelques heures afin de les améliorer et de se préparer au mieux à ce genre de situations.

Si l'on entend souvent que la santé n'a pas de prix, il semble essentiel de nuancer ce propos, car elle a néanmoins un coût. Dans son dernier rapport du 20 décembre 2019 (OCDE, 2019), l'Organisation de Coopération et de Développement Économique (OCDE) constate une augmentation globale des dépenses dans le domaine de la santé. Ces dernières dépasseront la croissance du PIB au cours des 15 prochaines années dans presque tous les pays de l'OCDE. En effet, les dépenses de santé par habitant augmenteront à un taux annuel moyen de 2.7 % et pourront atteindre jusque 10.2 % du PIB d'ici 2030 par rapport à 8.8 % en 2018. Cette donnée pourrait encore augmenter du fait de la pandémie de COVID-19 qui pourrait conduire les gouvernements à moins couper dans le budget de la santé. Ces augmentations dans les

1. Rapport de la Mission DGOS relative à la modernisation des SAMU - Pierre Mardegan, juillet 2010, www.sante.gouv.fr/IMG/pdf/mission_dgos-rapport_modernisation_des_samu-07-2010.pdf

2. La Croix, no. 41651, France, vendredi 6 mars 2020, p6, <https://tinyurl.com/ya5n64qm>

3. La Dépêche, 10/03/2020, <https://tinyurl.com/y86z3v8m>

4. Journal 20 Minutes (Nantes), 17/03/2020, <https://tinyurl.com/yalhtu8d>

dépenses de soins peuvent être expliquées par un vieillissement de la population, ainsi qu'une amélioration des moyens de dépistage et une spécialisation des traitements, ce qui les rend plus coûteux. Les récents événements qui ont impacté le domaine de la santé montrent que les systèmes de santé n'étaient pas préparés et ne disposaient pas de suffisamment de matériels ni de ressources pour gérer au mieux une telle crise. Ces dépenses, bien que nécessaires, pourraient certainement être rationalisées, notamment en optimisant le fonctionnement des organisations de santé. C'est pourquoi il est important que les professionnels de santé et les chercheurs en organisation travaillent de concert afin d'optimiser ces dépenses tout en cherchant à améliorer la qualité du service pour le patient.

Les centres d'appels des SAMU sont la pierre angulaire du système de soin pré-hospitalier. Leur bon fonctionnement est un maillon essentiel de la chaîne des urgences. Chaque seconde compte pour que la prise en charge du patient se fasse dans les meilleures conditions. C'est pourquoi il est impératif d'avoir une très bonne compréhension des enjeux et des défis qui attendent les centres d'appels dans les années à venir.

Toutefois, avant de se projeter, il est intéressant de regarder la genèse des centres d'appels. En France, le SAMU ne reçoit les appels du public que depuis 40 ans. Le premier centre d'appels a en effet été créé en 1979 par le Professeur Louis Lareng à Toulouse. Depuis cette date, la téléphonie, l'informatique et les moyens de télécommunication de manière générale ont beaucoup progressé. Le SAMU est une institution relativement jeune, qui utilise des technologies qui évoluent très vite. Si bien que les centres d'appels d'urgences n'arrivent pas toujours à suivre bien que ces dernières années d'importantes évolutions peuvent être notées sur l'utilisation des systèmes d'informations dans certains centres.

À cela s'ajoute une augmentation du volume d'appels pouvant aller jusque 15% par an pour certains départements (France et al., 2015). En 2018, les SAMU ont reçu 29 millions d'appels. Néanmoins, on ne comptabilise que 25 millions d'appels décrochés (Moutard et al., 2019). Cela signifie que 12,9% des appels n'ont pas été décrochés en 2018, ce sur quoi s'est appuyée l'étude du magazine « Le Point » mentionnée précédemment. En réalité, il est nécessaire de nuancer ce chiffre, car dans ces appels non décrochés, une partie est en fait des appels abandonnés qui ne nécessitent pas de régulation urgente. Il faudrait être capable de différencier les appels abandonnés (qui n'ont pas vocation à être traités) des appels perdus (qui auraient dû être décrochés). Cela permettrait d'avoir un taux d'appels non décrochés qui soit plus proche de la réalité. Toutefois, ce sont tout de même plusieurs millions d'appels qui ne sont pas décrochés tous les ans. Or, les objectifs préconisés par SAMU et Urgences de France donnent une cible à atteindre de 99%⁵ des appels entrants décrochés. Le rapport MARCUS (Moutard et al., 2019) propose d'ailleurs de mettre en perspective ces objectifs et proposent de se baser sur des indicateurs différents tel que la qualité de service à 15 et à 30 secondes pour se rapprocher des critères d'évaluations internationaux des centres d'appels d'urgence. Il est donc nécessaire d'avoir une vision claire de l'activité des SAMU afin de les aider à atteindre les objectifs de performance fixés.

5. Référentiel Métier des SAMU - SudF

En général, plus un pays est développé et plus sa politique, son infrastructure de santé et ses services d'urgence sont organisés et avancés. (Sikka et al., 2005) ont proposé une approche pour catégoriser la variété des services médicaux d'urgence dans le monde. Ils ont identifié deux types de services médicaux d'urgence : formels et informels. Un système formel « *inclut un numéro d'accès, un système de prise d'appels, un système de répartition, une réponse, des soins et un transport* ». Un système informel se caractérise par l'absence d'intervention organisée et planifiée en cas d'urgence. Dans le présent document, nous nous concentrons sur le système formel, qui comporte un numéro d'accès et un système de prise en charge d'appels. Le système de prise en charge des appels diffère beaucoup d'un pays à l'autre. Essentiellement, il existe des systèmes médicaux et non médicaux. D'un côté nous avons l'exemple du système fondé sur le médecin qui est notamment le modèle français où la philosophie (souvent appelée « *Stay and Play* » (Smith et al., 2009)) consiste à déplacer le médecin (et même une unité mobile de soins hospitaliers) vers le patient avant d'envisager son transport vers un hôpital. La spécificité de ce modèle est d'avoir un médecin à chaque étape des soins pré-hospitaliers, de la régulation de l'appel en passant par l'intervention des moyens sur place. De plus, les numéros d'accès sont multiples selon les besoins : en France, 112 qui est le numéro européen des urgences, 15 pour une urgence médicale, 18 pour une urgence incendie et 17 pour la police. Chacun de ces numéros est spécifique à un type d'intervention. Même si quelques cas de mutualisation existent, chaque corps de métier est indépendant et dispose de son propre centre d'appels, avec ses préposés aux appels et ses répartiteurs. De l'autre côté, nous avons le modèle nord-américain qui consiste à ramener le patient le plus rapidement possible vers un hôpital en fonction de la nature de son besoin. Ce modèle est désigné généralement sous l'appellation anglophone « *Scoop and Run* » (Smith et al., 2009). Aux États-Unis, lorsque l'on appelle le 911, l'appel arrive dans un centre d'appels de la sécurité publique. L'appel est analysé par l'opérateur, puis acheminé à l'intervenant le plus approprié. L'objectif est de porter une assistance à l'appelant le plus rapidement possible. Dans le cas d'une intervention d'ordre médical, il s'agit de transporter le patient à l'hôpital aussi vite que possible. Dans ce modèle et contrairement au modèle français, il n'y a ni médecin lors de la régulation téléphonique de l'appel, ni médecin envoyé sur place.

Problématique générale

Les SAMU évoluent dans un contexte politique où l'augmentation de la demande de soins les oblige à décrocher toujours plus d'appels. Cette augmentation doit être gérée en garantissant un niveau de qualité optimal auprès des patients. Bien souvent, l'augmentation du nombre d'appels n'est pas suffisante pour justifier une augmentation proportionnelle du nombre de ressources pour la prise en charge des appels. Dans un contexte financier restreint, si l'on ne peut pas étendre les moyens par manque de dotation, il est nécessaire de s'adapter sur le plan organisationnel et structurel. Toutefois, avant de mettre en place de nouvelles organisations, il est préférable de les avoir étudiées, analysées et évaluées. On peut difficilement se permettre d'établir une démarche d'apprentissage par l'erreur et d'expérimenter de nouvelles solutions directement sur le terrain lorsque des vies humaines sont dans

la balance. C'est pourquoi, avant d'appliquer des changements dans l'organisation actuelle, il faut être certain qu'ils apporteront une amélioration et non pas une dégradation, ce qui est particulièrement crucial pour des organisations telles que les SAMU.

Pour cela, nous proposons de créer un modèle numérique de l'organisation existante et des flux d'un centre d'appels du SAMU afin de disposer d'un environnement virtuel de simulation que l'on peut qualifier de double numérique. Ce double numérique permettra d'observer et d'analyser en détail les forces et faiblesses de l'organisation existante, puis de mettre en place virtuellement des solutions d'amélioration avant de les évaluer objectivement. Si les solutions testées avec le double numérique mettent en évidence une amélioration possible de l'organisation, leur implémentation dans le monde réel sera facilitée et présentera un risque d'échec réduit. Pour cela, il importe de disposer d'un double numérique dont le comportement dynamique soit aussi proche possible que celui du monde réel qu'il simule.

L'approche de diagnostic organisationnel qui sera développée dans cette thèse de doctorat s'apparente au travail du médecin. En effet, si l'on considère que les difficultés du SAMU à décrocher tous les appels sont des symptômes, tout comme un patient qui aurait des valeurs anormales dans son hémogramme, alors le SAMU est malade, et nécessite d'être ausculté! Bien entendu la maladie dont souffre le SAMU est une maladie organisationnelle. Toutefois il est important de prendre en compte la juste allocation des effectifs au regard de l'activité attendue dans l'étude de notre organisation. Pour être capable de diagnostiquer des problèmes chez un patient, le médecin doit connaître le fonctionnement du corps humain et les interactions entre les différents systèmes qui le composent. C'est en connaissant et en analysant ces mécanismes complexes qu'il peut réaliser un diagnostic lui permettant d'identifier la source d'une maladie ou la cause de ces symptômes. Dans le domaine de l'ingénierie organisationnelle, pour effectuer un diagnostic organisationnel du SAMU, une démarche comparable peut être suivie : il faut d'abord comprendre comment les différentes activités d'un centre d'appels fonctionnent et s'articulent, afin de connaître en quelque sorte l'anatomie du SAMU. De ce fait, la première question à laquelle nous souhaitons répondre est la suivante :

Quels sont les éléments constitutifs d'un centre d'appels urgence permettant de formaliser et caractériser son fonctionnement et son organisation ?

Dans notre étude, une attention particulière sera portée au modèle français de gestion des appels.

Une fois que nous aurons pu établir comment est structuré un centre d'appels, nous serons à même de comprendre son fonctionnement interne et donc d'être en mesure de faire un diagnostic organisationnel conduisant à l'identification de potentiels dysfonctionnements. Cela revient par exemple à faire des « prises de sang » du SAMU pour effectuer un bilan et vérifier que tout va bien, grâce à des données extraites du système d'information qui permettent de tracer le déroulement d'un appel. Une fois extraites, ces données peuvent être analysées et des indicateurs

de performance peuvent être calculés afin de vérifier si leurs valeurs sont dans la plage du « bon » fonctionnement pour le niveau de performance visé. En ingénierie organisationnelle, on va mesurer le degré de dysfonctionnement du SAMU grâce à des indicateurs de performance organisationnelle, comme par exemple la Qualité de Service (QoS) ou le temps d'attente d'un appel avant d'être décroché. Si les valeurs de ces indicateurs de performance sont en dehors des valeurs visées, il faut rechercher les causes pour pouvoir ensuite agir sur elles plutôt que sur les seuls symptômes. Comme en médecine, cette recherche des causes n'est pas toujours aisée, ce qui amène parfois à mettre en place des traitements symptomatiques plutôt qu'étiologiques. Cela permet de calmer la douleur ponctuellement, mais le mal subsiste en profondeur... Grâce au double numérique que nous proposons d'utiliser pour ce travail, nous allons mener des expérimentations virtuelles qui faciliteront la recherche de causes de dysfonctionnement racines en identifiant, plus facilement que dans le monde réel, plus vite et sans risque, les facteurs d'influence ou de sensibilité à la performance. Agir sur ces facteurs après les avoir identifiés permettra de mesurer les effets sur la performance du double numérique du centre d'appels du SAMU. La nécessité de disposer d'un double numérique au comportement le plus proche possible du monde réel qu'il cherche à reproduire est donc indispensable.

Les centres d'appels d'urgence sont soumis à des exigences spécifiques élevées en matière de performance.

Une fois le diagnostic organisationnel établi, il est possible de passer à la recherche d'un traitement adapté, toujours à l'aide du double numérique. Il pourra notamment être nécessaire de modifier l'organisation des SAMU pour les améliorer. En effet, dans le rapport de la Mission de Modernisation de l'Accessibilité et de la Réception des Communications d'Urgence pour la Sécurité, la Santé et les Secours (MARCUS)⁶ de décembre 2019, né d'une étude lancée par l'Inspection Générale de l'Administration et de l'IGAS pour faire un état des lieux du fonctionnement des services de secours, de soins et de sécurité, une des préconisations majeures est :

« Harmoniser les axes organisationnels principaux des centres d'appels d'urgence »

Cette recommandation provient du fait qu'il existe une grande hétérogénéité des organisations dans les différents centres d'appels d'urgence. Cette hétérogénéité n'est pas forcément objectivée. Ainsi, il apparaît nécessaire de retravailler l'organisation des centres d'appels d'urgence dans le but d'améliorer leur efficacité (leur capacité à faire ce pour quoi le système est prévu) et leur efficience (leur capacité à réaliser leurs objectifs en exploitant au mieux les moyens mis à leur disposition). Il s'agit alors de trouver d'autres moyens d'action comme modifier la gestion des files d'attente d'appels, ou encore spécialiser les ressources. Une seconde question de recherche que l'on se pose est donc la suivante :

Sur quels leviers organisationnels peut-on agir pour améliorer les performances des centres d'appels ?

6. http://www.urgences-113.fr/marcus/Rapport_MARCUS3.pdf

Nous avons vu rapidement dans la présentation du contexte que l'on peut trouver une diversité d'organisations dans les centres d'appels d'urgence. De plus, comme nous le verrons dans la section 1.3, les études existantes dans la littérature à propos des SAMU sont bien souvent mono-centriques, c'est-à-dire, ne concernent qu'un seul SAMU. L'hétérogénéité de ces centres et les spécificités du domaine des services d'urgence permettent d'expliquer les niveaux de performance variables. En plus des difficultés pouvant apparaître en situation normale, on note des difficultés à anticiper le comportement de l'organisation en situation de crise ponctuelle et prévisible (épisode de grippe saisonnière par exemple), de crise ponctuelle et imprévisible (attentat ou catastrophe industrielle par exemple) ou de crise durable (épidémie de type COVID-19 par exemple). Une contribution de ces travaux de recherche consiste à relever ce défi complexe :

Comment prendre en compte les spécificités des SAMU afin de leur proposer des recommandations adaptées à leur fonctionnement, que ce soit en régime nominal mais aussi en situation de crise ?

Méthodologie de recherche

Pour répondre à ces questions, nous avons mis en place une méthodologie de recherche en adaptant le cadre de recherche en système d'information de Hevner et al. (Hevner et al., 2004) et les étapes méthodologiques proposées par (Caillaud et al., 2015). Pour Hevner, la pertinence du travail de recherche mené dépend de la capacité à répondre et à satisfaire le besoin métier. Dans notre cas, ce travail est pertinent car le but de cette recherche est d'apporter des éléments de réponse aux problèmes organisationnels avérés des SAMU pour améliorer la performance des centres d'appels. La Figure 1 est adaptée de la méthodologie d'Hevner appliquée à notre cas. Dans cette méthodologie, nous nous appuyons sur la connaissance de l'environnement pour extraire les besoins métiers du domaine. Dans notre cas, l'environnement se compose du système d'aide médicale urgente français, de l'étude des données de téléphonies des centres d'appels et plus spécifiquement sur les données de trois SAMU de la région Occitanie. De plus, cet environnement se trouve imbriqué dans un système de santé qui évolue et qui doit s'adapter aux situations imprévues (comme la pandémie de COVID-19 de 2020). Notre démarche, essentiellement basée sur des expériences menées sur le double numérique, fait appel à un ensemble d'outils parmi lesquels on trouve notamment des tableaux de bord, des techniques d'exploration de processus (*Process Mining*) et de fouille de données (*Data Mining*), d'ingénierie dirigée par les modèles, notamment des modélisations des processus de prise de décisions dans les centres d'appels, des modélisations de leurs organisations et différentes méthodologies de diagnostics organisationnels. C'est donc sur de la connaissance générée par ces outils que nous allons réaliser nos expérimentations ainsi que leurs évaluations pour répondre au besoin métier. Cette démarche est itérative, car l'évaluation d'un scénario peut conduire à rechercher une nouvelle amélioration qu'il faudra alors de nouveau tester puis évaluer.

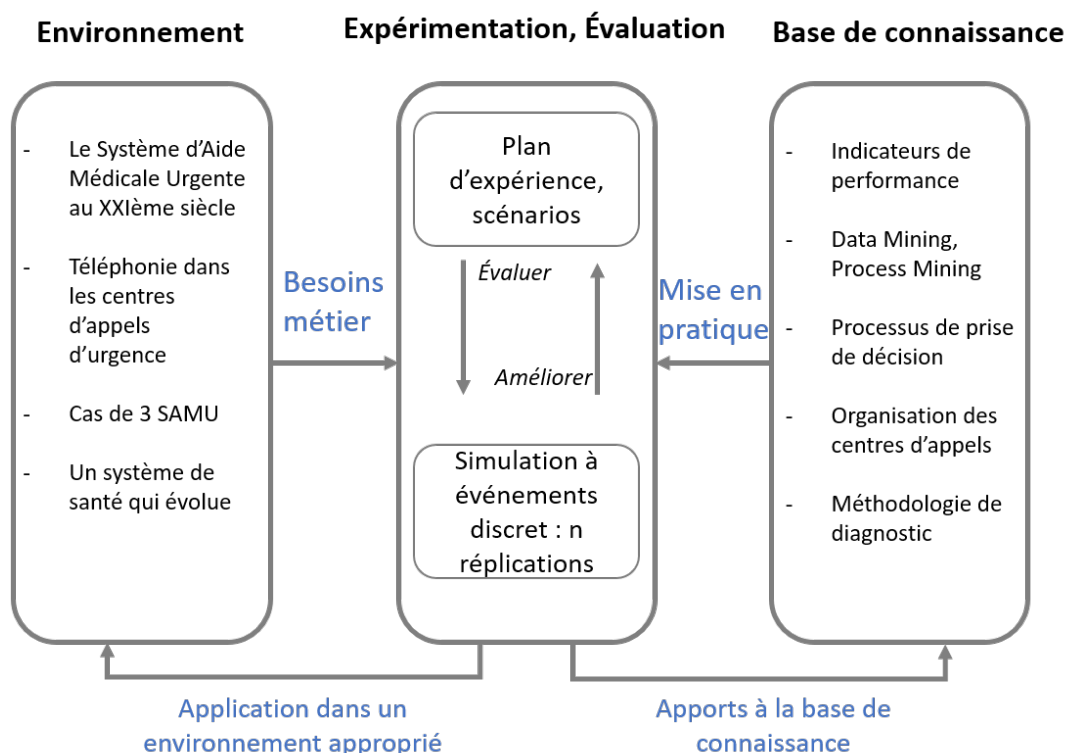


Figure 1 – Méthodologie de recherche adaptée de Hevner et al. (2004)

Notre recherche s’appuie donc sur une démarche de recherche en science de la conception (Design science research methodology - DSR) en plusieurs étapes :

- Une étude du terrain est faite afin d’avoir une bonne compréhension du contexte, des enjeux et défis à relever pour les experts métiers. Cela nous permet d’identifier les particularités et les caractéristiques de ce domaine par rapport à d’autres secteurs d’activités. Une analyse quantitative des performances actuelles est conduite afin de pouvoir évaluer les enjeux concrets et objectifs de notre travail de recherche. Cette phase nous permet notamment de définir les problématiques industrielles.
- Une étude bibliographique vient compléter l’étude terrain pour identifier les problématiques scientifiques et métiers traitées dans le domaine étudié ainsi que les outils utilisés pour les résoudre. Cette étude permet ainsi d’identifier les manques et de justifier les outils et les méthodes employés.
- La combinaison des deux études précédentes nous permet de formuler le problème scientifique à la lumière des manques de la littérature et des verrous scientifiques afin d’établir les problématiques que l’on souhaite adresser et de présenter les hypothèses faites.
- La description des modèles est détaillée et argumentée et les méthodes de modélisation utilisées sont justifiées. Cette étape permet aussi la vérification de la répliquabilité des résultats.
- Enfin en s’appuyant sur l’analyse des expérimentations sur les modèles produits lors de l’étape précédente, une liste de préconisations pour les experts métiers est fournie.

Afin de mieux comprendre le fonctionnement des centres d'appels et pour les aider à améliorer leur gestion des appels d'urgence, nous avons pu bénéficier de la collaboration de trois SAMU de la région Occitanie : La Haute-Garonne, L'Aveyron et le Tarn. Trois départements avec des caractéristiques différentes en termes de démographie et d'offre de soins. En effet, la Haute-Garonne présente une grande concentration d'habitants au niveau de Toulouse. L'Aveyron est un territoire plus rural et de moyenne montagne avec certaines zones difficiles d'accès. Le Tarn est un département semi-rural avec la présence de zones plus urbanisées telle Albi ou Castres. La rencontre avec les acteurs de ces différents SAMU nous a permis de mieux comprendre les problématiques rencontrées par les centres d'appels d'urgences. Prendre en compte les spécificités des territoires dans l'organisation des centres d'appels pourrait permettre d'améliorer la prise en charge des appels en proposant des solutions qui soient adaptées à leur fonctionnement et leurs particularités.

Organisation du manuscrit

Ce manuscrit s'organise en quatre parties présentées dans la figure 2. La première partie correspond à cette introduction. La deuxième partie se décompose en deux chapitres.

Le chapitre 1 s'intéresse à la description de *l'Espace des Problèmes* relatif à l'étude. Dans sa première moitié, il fait le point sur les différents modes de fonctionnements de centre d'appels d'urgence au niveau international. Puis il décrit plus précisément le fonctionnement des centres d'appels français ainsi que les différences qui peuvent exister au niveau national. Il traite des ressources utilisées pour le bon fonctionnement d'un centre d'appels ainsi que du « cycle de vie » des appels d'urgence. La deuxième moitié traite de *l'Espace des Solutions* proposé dans la littérature scientifique à propos des centres d'appels : Comment ce sujet a-t-il été abordé et traité par la communauté scientifique ? Quels sont les outils utilisés pour traiter le sujet et pourquoi ont-ils été choisis ?

À la lumière des faits exposés dans le chapitre 1, le chapitre 2 croise les défis issus du métier et les faiblesses de la littérature scientifique afin de faire émerger le problème scientifique à traiter dans ce manuscrit. Les hypothèses de recherche que nous souhaitons valider sont présentées ainsi que les contributions scientifiques envisagées pour cette thèse de doctorat.

La deuxième partie est composée des chapitres 3, 4 et 5. Elle développe nos propositions pour apporter des éléments de réponse concrets et rigoureux aux problèmes scientifiques identifiés dans le chapitre 2. Chacun de ces chapitres contiendra une revue de littérature en lien avec les outils utilisés pour les réaliser.

Dans le chapitre 3, nous présentons notre première contribution qui consiste à définir un cadre méthodologique outillé en quatre étapes pour le diagnostic et l'amélioration des centres d'appels d'urgence. Pour cela, nous détaillons d'abord les manques méthodologiques pour l'étude des centres d'appels dans la littérature pour justifier l'intérêt d'une telle proposition. Ensuite, nous décrivons les étapes des deux parties de notre méthodologie en spécifiant et justifiant les outils sélectionnés.

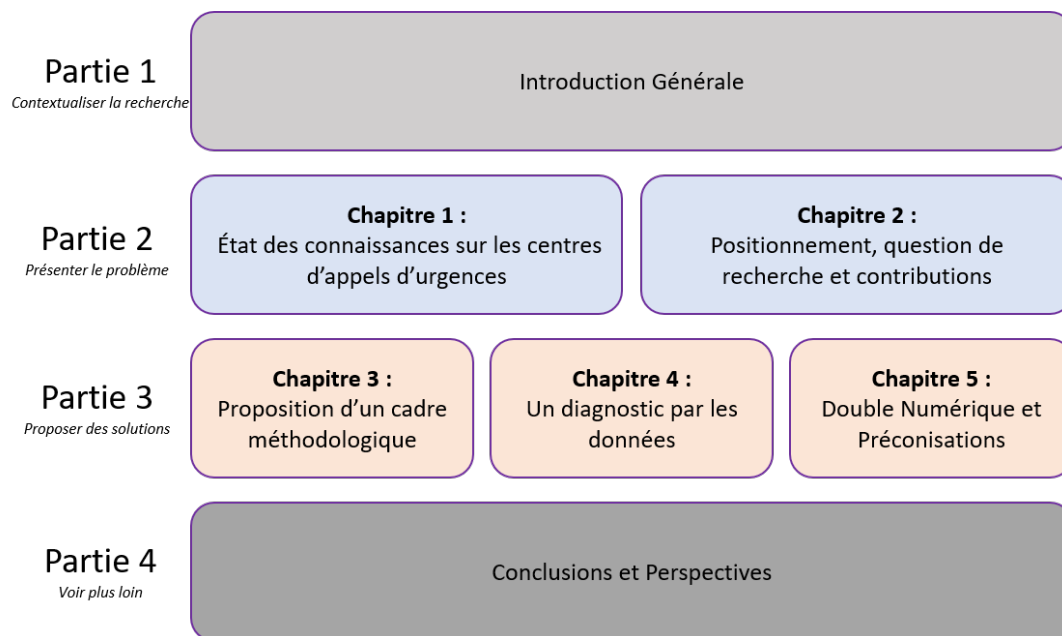


Figure 2 – Plan du manuscrit

Le chapitre 4 développe la première partie de la méthodologie proposée dans la partie 3. Il correspond à la première moitié du cadre méthodologique qui étudie le système réel. Une seconde contribution consiste à proposer ici une démarche de diagnostic organisationnel s'appuyant sur l'analyse des données réelles des centres d'appels. Ces données proviennent de deux sources : d'une part des données agrégées à l'échelle nationale pour tous les départements de France métropolitaine et d'autre part des données collectées auprès des trois centres d'appels d'urgence en Occitanie. Dans une première partie, nous étudions les données nationales afin de proposer une classification des SAMU suivant une typologie définie par des méthodes de regroupement (*clustering*) de données. Dans la seconde partie, les données de téléphonies sont utilisées pour effectuer un diagnostic organisationnel plus pointu pour un SAMU. Cette étude nous permet d'extraire les caractéristiques spécifiques à ce SAMU, et nous sert comme données d'entrée pour le chapitre suivant.

Dans le chapitre 5, nous développons le double numérique du SAMU à base de simulation. Ce modèle nous permet d'avoir un environnement d'expérimentation virtuel pour tester de nouvelles organisations. Ainsi, nous proposons ici une troisième contribution avec l'étude de l'impact de la gestion des files d'attente sur les performances de décrochés des centres d'appels. Cela nous permet également d'étudier le comportement présumé d'un centre d'appels en cas d'augmentation soudaine des appels. Nous présentons nos travaux de simulation comprenant la construction du modèle, sa validation ainsi que les différents scénarios testés. L'impact de ces scénarios est évalué sur des indicateurs de performances pertinents vis-à-vis du référentiel métier du SAMU. Nous testons à la fois le délai à travers la qualité de service, mais aussi l'efficacité en évaluant le taux d'occupation des ressources. De nouvelles organisations sont finalement proposées pour l'amélioration de l'efficacité

et de l'efficacité du décroché des appels d'urgence, notamment avec la proposition de la mise en place d'un triage au décroché de l'appel.

Enfin la dernière partie constituée du seul chapitre 6 présente les conclusions et les nombreuses perspectives de recherche de cette thèse de doctorat.

1

Services d'Aide Médicale Urgente : Enjeux et défis

Ce chapitre a pour vocation de présenter l'espace des problèmes ainsi que la richesse des sujets traités et à traiter dans les services d'aide médicale urgente (SAMU). Dans un premier temps, nous présentons la diversité des systèmes de santé à l'échelle internationale. Nous proposons plusieurs critères sur lesquels nous appuyer pour caractériser ces systèmes. Dans un second temps, nous décrivons le fonctionnement détaillé de notre cas d'étude : le modèle français. Enfin, nous nous intéressons aux problèmes spécifiquement rencontrés par les SAMU et la manière dont la communauté scientifique s'y est intéressée et les a adressés.

1.1 La diversité des services de soins d'urgence

Les services de soin d'urgence diffèrent d'un pays à un autre et parfois même au sein d'un même pays. Dans cette partie, nous nous intéressons à ces différences et à la manière de les caractériser.

1.1.1 Des modèles formels et informels

En fonction du niveau de développement d'un pays, la maturité du système de santé ne sera pas la même. En effet, cela va dépendre des infrastructures à disposition, du niveau de vie et d'éducation de la population, de l'accessibilité au système de soin, de l'équipement disponible, du personnel formé, etc. C'est pourquoi (Sikka et al., 2005) proposent de distinguer deux types d'EMS (Emergency Medical System) : les modèles informels et formels.

Les modèles informels se caractérisent par une absence de système pré-hospitalier. Il n'y a pas d'organisation établie pour la réponse à l'urgence ou alors ce qui existe est très rudimentaire. On constate plusieurs niveaux dans les modèles informels : un premier avec une absence totale d'organisation pré-hospitalière où le patient

doit s'appuyer sur sa famille, ses amis, ou toute autre personne en dehors de ses proches pour le transporter vers un établissement de santé en cas de besoin. Certains systèmes évoluent lorsque des acteurs dont le rôle premier n'est pas le soin d'urgence commencent à porter secours. Ce peut être le cas lorsque les pompiers ou la police commencent à transporter des patients vers l'hôpital alors que ce n'est pas leur cœur de métier (Mock et al., 2002). Un premier pas dans l'amélioration des systèmes informels consiste à commencer à former des citoyens à l'usage des premiers secours.

Les modèles formels se caractérisent par une organisation coordonnée entre plusieurs ressources capables de répondre et assurer le transport des patients depuis le terrain jusqu'à l'établissement de santé adapté à son besoin lorsque cela est nécessaire. La plupart comprennent un numéro d'accès, un système de prise d'appels, un système de répartition, de réponse, de soin et de transport.

Dans la suite de ce manuscrit, nous réduisons notre périmètre à l'étude des systèmes formels qui disposent de moyens organisés et coordonnés pour la gestion des appels d'urgence.

1.1.2 La chaîne des urgences : Amont et Aval

L'enchaînement des actions faisant intervenir plusieurs ressources afin de porter secours au patient s'appelle **la chaîne des urgences** (Yersin, 2006). Celle-ci peut être décomposée en plusieurs étapes. Sikka et al. proposent une décomposition en quatre étapes (Sikka et al., 2005) comme suit :

1. Notifier les autorités compétentes lorsqu'il y a une urgence.
2. Réceptionner cette information puis envoyer des moyens adaptés pour porter secours à l'appelant.
3. Appliquer les soins de première nécessité sur place.
4. Transporter le patient dans une structure adaptée disposant de l'équipement et des équipes nécessaires pour traiter le patient.

Toutefois, nous avons préféré retenir pour nos travaux la décomposition en 8 étapes issues du référentiel métier du système français qui nous semble plus précise (France et al., 2015) (cf. figure 1.1). Dans cette figure, nous présentons notre vision de la chaîne des urgences qui vient préciser celle proposée par Sikka et al. et qui s'appuie sur le référentiel métier du système français.

1. **Appeler** : Composer un numéro d'urgence qui notifie les autorités compétentes lorsqu'il y a urgence.
2. **Qualifier et Orienter** : Réceptionner cette information, lui affecter une priorité et transférer l'appel au praticien adéquat.
3. **Réguler** : Diagnostiquer et décider des moyens adaptés pour porter secours.
4. **Engager des moyens** : Mobiliser les moyens adaptés et transmettre les informations aux équipes de secours.
5. **Rejoindre et Stabiliser** : Se rendre auprès du patient et stabiliser son état.

6. **Orienter le patient** : Identifier l'établissement de santé le plus adéquat pour envoyer le patient en prenant en compte les places disponibles ainsi que la volonté du patient.
7. **Transporter le patient** : Transporter le patient jusqu'à l'établissement de santé choisi.
8. **Hospitaliser** : Procéder à l'admission du patient dans le service idoine.

Dans cette chaîne des urgences, on identifie deux macro-étapes. Une première qui concerne la réception, le tri et l'orientation de l'appel, jusqu'à la décision d'envoyer des moyens. Toutes ces activités ont lieu au niveau du centre d'appels. En France, ce dernier est appelé Centre de Réception et de Régulation des Appels (CRRA). Nous nommerons cette partie **l'amont de la chaîne des urgences**. La deuxième macro-étape se compose de tout ce qui concerne l'intervention sur place, le transport jusqu'au patient, sa stabilisation ainsi que son transfert jusqu'à ce qu'il soit admis dans un service de santé si cela est nécessaire. Nous nommerons cette partie **l'aval de la chaîne des urgences**.

La gestion du soin urgent peut être décomposée en deux macro-étapes : l'amont et l'aval. Elles différencient les activités au niveau du centre d'appels et les actions des équipes de secours sur place. Au sein de ces deux grandes catégories, on trouve plusieurs décompositions possibles avec un niveau de détail pouvant être différent.

Dans la suite de cette étude, nous nous sommes appuyés sur la décomposition du référentiel métier français et nous nous sommes plus spécifiquement intéressés aux différences que l'on peut retrouver au sein de l'Union Européenne (UE).

1.1.3 L'accès aux services d'urgences : numéro unique ou numéros spécialisés ?

La première étape de la chaîne des urgences est donc *Composer un numéro d'urgence qui notifie les autorités compétentes*. Cette première étape présente déjà des différences notables d'un pays à un autre. L'*European Emergency Number Association* (EENA) est une organisation non gouvernementale dont la mission est de contribuer à

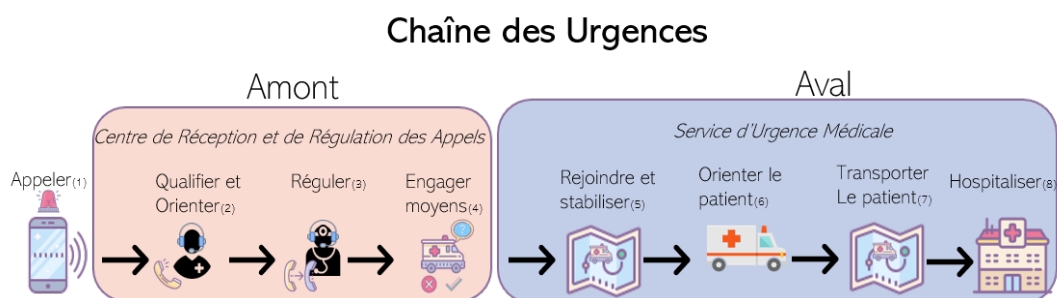


Figure 1.1 – La chaîne des urgences

l'amélioration de la gestion de l'urgence pour les citoyens. Une de ses activités consiste notamment à réaliser une veille technologique sur les différentes modalités d'appels dans plusieurs pays. Dans son dernier rapport ¹ de décembre 2019, elle mentionne différentes pratiques au sein de l'UE concernant l'accès aux soins d'urgence. En ce qui concerne l'accès aux centres d'appels d'urgences, trois émergent plus particulièrement :

1. **Un numéro unique** pour toutes les urgences, qu'elles soient médicales, civiles ou criminelles. 11 pays de l'UE adoptent ce mode de fonctionnement : L'Irlande, le Portugal, les Pays-Bas, le Danemark, la Suède, la Finlande, l'Estonie, la Roumanie, la Bulgarie, Chypre et Malte.
2. **Un numéro distinct** pour chaque type d'urgence. Ce mode de fonctionnement concerne 12 autres pays de l'UE : la France, l'Espagne, l'Italie, L'Autriche, la République tchèque, la Pologne, la Lituanie, la Lettonie, la Slovaquie, la Hongrie, la Croatie et la Grèce.
3. **Un numéro commun** simplifié pour les urgences médicales et les pompiers, la sécurité civile ayant gardé son propre numéro. C'est le cas pour 4 pays, la Belgique, le Luxembourg, l'Allemagne et la Slovénie.

Parmi ces trois moyens de contacter les secours, on constate que 41% des pays de l'UE disposent d'un numéro unique, 44% ont un numéro distinct dépendant du type d'urgence et 15% utilisent un modèle mixte avec un numéro commun pour les urgences médicales et les pompiers et un numéro particulier pour contacter la police. Il n'y a donc pas de système qui prédomine à l'échelle des pays européens, et encore moins d'harmonisation.

1.1.4 Quels sont les différents modèles de décrochés des appels

Une fois l'appel composé en utilisant le bon numéro, il est acheminé vers le centre d'appels adéquat pour être décroché. Plusieurs questions se posent : vers quel centre d'appels doit-il être envoyé s'il y en a plusieurs ? Quelles sont les ressources présentes sur la plateforme téléphonique ? Est-ce que ce sont des ressources généralistes, capables de décrocher tous les types d'urgence ? Ou est-ce que sont des ressources spécialistes d'un type d'urgence en particulier ? L'*European Emergency Number Association* (EENA) ² recense six modèles différents de prise en charge des appels. L'étude *EUROCALL* menée dans 21 centres d'appels de la sécurité civile de 18 pays différents montre l'importance de l'accès au centre d'appels (Nikolaou et al., 2017). En effet, le fait d'avoir une organisation mono ou bi-niveau aura un impact en terme d'accès.

1. **Le modèle décentralisé** présenté dans la Figure 1.2. Lorsqu'un appelant est dans le besoin, il choisit le numéro qu'il doit composer pour être mis en

1. Public Safety Answering Points – 2019 Global edition - <https://eena.org/knowledge-hub/documents/>

2. <https://eena.org/knowledge-hub/documents/>

relation avec le centre d'appels qui est le plus à même de répondre à son besoin : police, pompier ou service médicaux. C'est le cas de la France, de l'Autriche, de l'Allemagne, de l'Italie et de la Norvège.

2. **Le premier modèle centralisé (version V1)** est présenté dans la Figure 1.3. Lorsqu'un appel arrive, il est d'abord pris en charge par un opérateur de niveau 1 qui demande à l'appelant avec quel centre d'appels d'urgence (*gathering*) il souhaite être mis en relation, puis il effectue le transfert (*dispatching*). La phase de récupération des données et d'envoi de moyens est gérée par les spécialistes. C'est le cas notamment de l'Irlande.
3. **Le deuxième modèle centralisé** présenté dans la Figure 1.4. Lorsqu'un appel arrive, il est aussi décroché par un opérateur de niveau 1, mais celui-ci commence déjà à récupérer de l'information et c'est lui qui décide à quel(s) effecteur(s) (pompiers, policiers ou urgentistes) il doit transférer l'appel. C'est le cas de la Roumanie.
4. **Le modèle centralisé et unifié** présenté dans la Figure 1.5. Ce modèle est le même que le précédent sauf que tous les acteurs des centres d'appels sont dans un lieu commun. C'est le cas de certaines régions d'Espagne, de la Belgique.
5. **Le modèle généraliste** présenté dans la Figure 1.6. Dans ce modèle il n'y a qu'un centre d'appels géré par des civils sans spécialisations. Ce sont eux qui gèrent le décroché, la récupération des informations et l'envoi de moyens. La Finlande fonctionne avec ce modèle.
6. **Le modèle inter-connecté** présenté dans la Figure 1.7. Ce modèle est une variante collaborative où tous les centres d'appels sont connectés entre régions. S'il n'y a pas de ressources disponibles pour décrocher dans le centre d'appels le plus proche, l'appel est redirigé vers un autre centre d'appels. C'est le cas de la Bulgarie, de la République Tchèque et de la Suède.

1.1.5 Des stratégies pré-hospitalières différentes : *Stay and Play* ou *Scoop and Run* ?

Concernant la prise en charge préhospitalière du patient dans une situation d'urgence, on distingue deux catégories distinctes : Les soins vitaux de base (Basic Life Support,

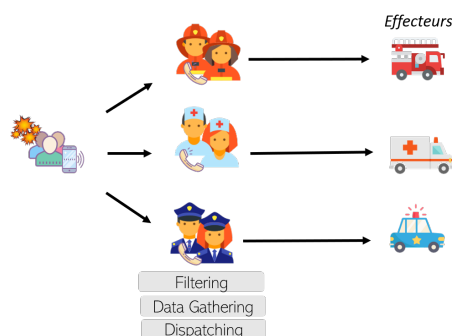


Figure 1.2 – Le modèle de décroché décentralisé

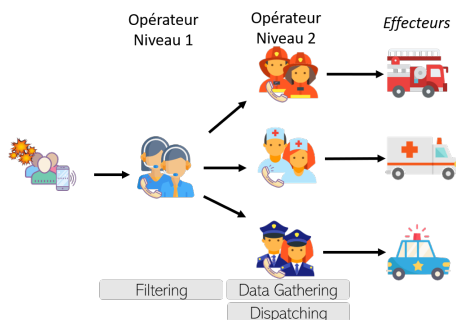


Figure 1.3 – Le modèle de décroché centralisé : première version

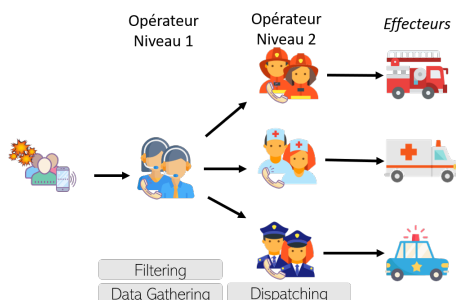


Figure 1.4 – Le modèle de décroché centralisé : deuxième version

BLS), et les soins vitaux avancés (Advanced Life Support, ALS). Le maintien des fonctions vitales de base (BLS) suppose la prise en charge des patients blessés par des méthodes non invasives. Cela suppose le dégagement des voies respiratoires, l'apport d'oxygène, la réanimation cardio-pulmonaire, l'arrêt de l'hémorragie externe, l'immobilisation des fractures et de la colonne vertébrale. L'objectif principal du BLS est de maintenir les fonctions cardiaques et respiratoires pendant le transport du patient vers le centre de traumatologie, sans causer de dommages supplémentaires (Beuran et al., 2012). Les soins vitaux avancés (ALS) impliquent l'utilisation de procédures invasives pour la prise en charge initiale des patients traumatisés : procédures plus sophistiquées pour la gestion des voies aériennes, la surveillance et la défibrillation cardiaques, le cathétérisme intraveineux et l'administration de médicaments (Beuran et al., 2012). Ce faisant, ces différences de pratique engendrent des philosophies de traitement du soin d'urgence qui divergent. On trouve deux stratégies aux extrêmes du spectre du soin préhospitalier qui sont :

- Le *Rejoindre et stabiliser* connue sous le nom "*Stay And Play*" (SAP) ou "*Stay And Stabilize*" qui implique le soin et la stabilisation de l'état du patient sur place avant de le transporter vers un établissement de santé. Ce mode de fonctionnement implique plus souvent l'utilisation d'ALS.
- Le *Charger et courir* connue sous la désignation "*Scoop And Run*" (SAR) ou "*Load And Go*" où l'objectif est de transporter le patient le plus rapidement possible dans un établissement de santé pour lui prodiguer les soins nécessaires. Cela nécessite de gérer les blessures qui mettent immédiatement la vie en danger dans l'ambulance pendant le transport.

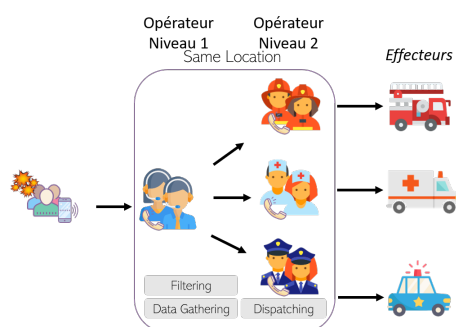


Figure 1.5 – Le modèle de décroché centralisé et unifié

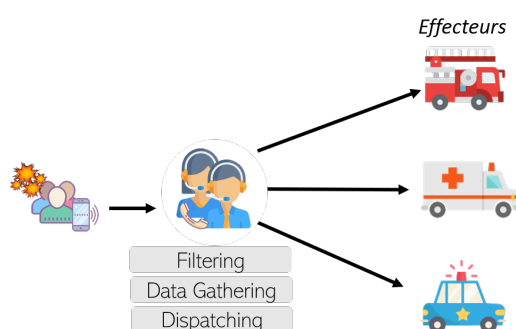


Figure 1.6 – Le modèle de décroché généralisé

Si on prend l'exemple du modèle français, on se trouve dans un cas de SAP. En effet, les véhicules du Service Mobile d'Urgence et de Réanimation (SMUR) sont gérés par une équipe médicale avec à son bord un médecin urgentiste. De plus, le modèle français va plus loin, car le patient est suivi par un médecin à chaque étape de sa prise en charge lorsque la médicalisation sur place est nécessaire. De la prise d'appels à son arrivée aux urgences en passant par la stabilisation du patient sur le lieu de l'urgence, il y a un médecin pour prendre en charge le patient.

Le modèle SAP a la particularité d'inclure un médecin à chaque étape de l'urgence préhospitalière. (Bieler et al., 2017) dans leur étude, comparent les modèles allemands et autrichiens au modèle anglo-saxon des États-Unis et du Royaume Uni. En Allemagne et Autriche, un médecin urgentiste est envoyé sur le terrain dans 95% des cas nécessitant une intervention. L'étude de Javaudin et al. (Javaudin et al., 2020) permet de documenter le système français et ses avantages. A contrario, de l'autre côté de la Manche et de l'Atlantique, ce sont des équipes para-médicales qui s'occupent de stabiliser le patient sur place et durant le transport vers un établissement de santé. L'objectif de cette étude est de mesurer les bénéfices potentiels de chacun des modèles. Les résultats laissent le débat entier sur l'intérêt de la présence d'un médecin, car il n'y a pas de différences significatives entre les deux modèles sur le temps de prise en charge ou sur le taux de mortalité dans les 24h ou durant hospitalisation. La conclusion est qu'il existe toujours un besoin d'analyse pour objectiver la procédure de prise en charge du patient la plus appropriée en termes de soins d'urgence. Malgré la grande quantité d'études existantes sur le sujet, ((Hannay et al., 2014) (Levy et al., 2007) (McCoy et al., 2013) (Smith et al., 2009)) nous ne

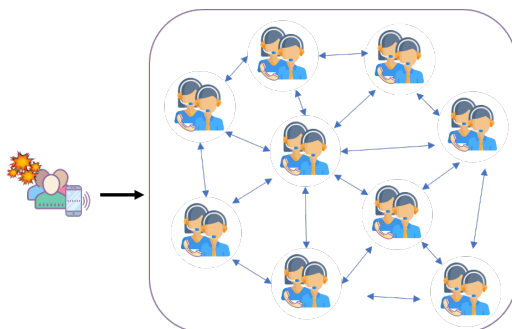


Figure 1.7 – Le modèle de décroché interconnecté

sommes toujours pas capables de statuer sur le modèle le plus bénéfique du point de vue de la survie des patients. Une étude sur plusieurs centres d'urgence européens montre néanmoins que les deux modèles sont utilisés en Europe et qu'il n'y a pas d'uniformité entre les pays. Le contenu de l'étude ne permet pas de conclure sur une meilleure performance d'un système par rapport à l'autre (Cnossen et al., 2019).

Dans cette partie, nous avons fait plusieurs constats. Les systèmes de santé pré-hospitaliers sont multiples et hétérogènes sur plusieurs niveaux :

- On trouve des modèles informels pour lesquels la gestion du soin d'urgence a un niveau de maturité assez faible. On trouve aussi des modèles formels qui sont plus structurés, organisés, et qui disposent de plus de moyens pour gérer le soin d'urgence.
- Au sein des modèles formels, on définit la chaîne des urgences qui se découpe en deux parties : amont et aval correspondant respectivement aux activités du centre d'appels et des services d'urgence médicale mobiles.
- Concernant l'accès à cette partie amont, les modalités varient selon les pays, certains adoptent un numéro unique, d'autre un numéro distinct, ou encore un modèle mixte avec un numéro commun pour les services médicaux et les pompiers, la police étant indépendant.
- On trouve aussi plusieurs modes de décrochés qui peuvent être plus ou moins centralisés avec des niveaux d'opérateurs téléphoniques différents.
- Enfin, les types de réponses pour la gestion du patient en aval avec le *Stay and Play* et le *Scoop and Run*.

Dans la suite de notre étude, nous nous focaliserons plus particulièrement sur le modèle français qui est un modèle formel. Le mode d'accès est avec un numéro distinct pour chaque type de besoin, et le mode de décroché peut varier au sein du territoire mais avec une prédominance pour le décroché décentralisé. Dans ce manuscrit, nous nous intéresserons donc aux modèles « *Stay and Play* » dont fait parti le modèle français.

1.2 Le modèle français du SAMU : Service d'Aide Médicale Urgente

Puisque cette étude a permis d'observer trois SAMU de la région Occitanie (Toulouse, Albi et Rodez), nous nous sommes intéressés plus spécifiquement au fonctionnement du modèle français que nous présentons dans cette section.

1.2.1 Le SAMU français : Naissance et Histoire

Le SAMU tel que nous le connaissons apparaît pour la première fois en 1955. Son succès se généralise et il se multiplie à travers la France dès 1964. La première forme du SAMU et de son Service Mobile d'Urgence et de Réanimation (SMUR) est un véhicule d'urgence dans lequel se trouvent un infirmier et un médecin. En 1965, les services d'urgence et de réanimation sont créés et dissociés des SMUR. C'est en 1968 que la première salle de régulation est créée à Toulouse par le Pr Louis Lareng : c'est la naissance du SAMU tel qu'on le connaît aujourd'hui. Il faudra attendre 1978 pour que le 15, numéro unique et national d'urgence, soit créé. La loi n°86-11 du 6 janvier 1986 officialise et rend obligatoire le SAMU dans chaque département. Elle définit aussi les différentes missions du SAMU comme le montre la Figure 1.8 :

- Réception et tri des appels via un numéro unique (15) auquel se rajoute le 112 dans certains départements.
- Assurer une écoute constante et apporter un conseil médical.
- Assurer l'envoi et la coordination des moyens de secours les plus adaptés dans les délais les plus rapides.
- Assurer la coordination des SMUR et de leurs Unités Mobiles Hospitalières (UMH).
- Vérifier la disponibilité des lits d'hospitalisation publics ou privés afin d'orienter le patient vers la structure la mieux adaptée à sa pathologie.
- Participer à l'élaboration et à l'organisation des différents plans de secours (plan blanc, plan rouge).
- Assurer la formation des personnels médicaux et paramédicaux aux soins d'urgence via les Centres d'Enseignement des Soins d'Urgences (CESU).

En 2009, la création de la Permanence Des Soins Ambulatoires (PDSA) par la loi HPST³ (Hôpital, Patient, Santé, Territoire) crée un nouveau flux d'appels dans les centres 15 : les appels de médecine ambulatoire. La PDSA est accessible durant les horaires de fermeture des médecins généralistes, soit le soir de 20 h à 8 h et le samedi de 12 h jusqu'à 8 h le lundi suivant.

Cet ajout ainsi que la démocratisation des SAMU ont engendré une augmentation structurelle du nombre d'appels dans les centres 15 pouvant aller de 5 à 15% par an (Penverne et al., 2009). Le livre blanc du SAMU-Urgences de France (SUdF) souligne⁴ que les Centres de Réception et de Régulation des Appels (CRRA) du

3. Loi n° 2009-879 du 21 juillet 2009 portant réforme de l'hôpital et relative aux patients, à la santé et aux territoires

4. (FRA 2015) SAMU Urgences de France Paris FRA. Livre blanc : organisation de la médecine d'urgence en France : un défi pour l'avenir. Français. Tech. rep. 2015, 48p

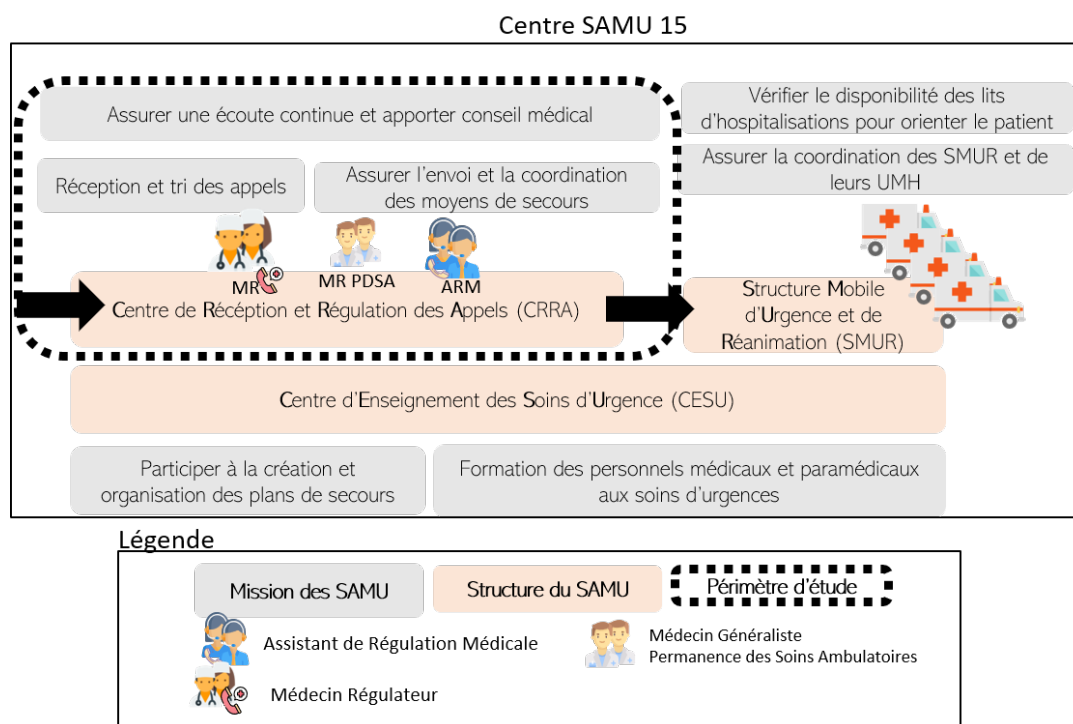


Figure 1.8 – Missions et structure du SAMU

SAMU font face à une augmentation importante du nombre d'appels entrants allant jusqu'à 10% par an. Il faut par ailleurs prévoir une nouvelle augmentation de la volumétrie des appels dans les prochaines années dues à l'intégration de nouveaux canaux de communication tels que les lunettes connectées, les appels en visio, les e-mails et texto, etc. Ces canaux de communication commencent à être testés dans certains centres d'appels et les discussions pour les intégrer aux SAMU sont déjà en cours depuis la sortie du pacte de refondation des urgences de 2019. Il faut aussi prendre en compte les appels E-call qui augmenteront en flèche d'ici 2025 (Oorni et al., 2017). En effet, depuis janvier 2020, toutes les voitures neuves doivent être équipées d'un système de communication directe avec les plateformes d'urgence. Ces changements vont obliger les SAMU à s'interroger sur la pertinence de leur organisation pour garantir un niveau de service à la hauteur de leurs missions et des enjeux associés.

Chaque SAMU se compose de trois parties :

- Une plateforme de gestion des appels, le CRRA (Centre de Réception et Régulation des Appels), qui est l'objet de ce travail de recherche.
- Un Centre d'Enseignement des Soins d'Urgence (CESU).
- Un ou plusieurs SMUR (Structure Mobile d'Urgence et de Réanimation) qui sont répartis dans le département pour assurer la meilleure couverture possible du territoire. Chaque SMUR dispose d'une ou de plusieurs équipes médicales d'intervention, pouvant être spécialisée (pédiatrie par exemple). Les équipes médicales sont généralement composées d'un médecin urgentiste, d'une infirmière (anesthésiste ou pas) et d'un ambulancier.

Pour décrire le modèle des SAMU français, nous nous appuyons sur trois sources principales :

- *Le référentiel et guide d'évaluation des SAMU Centres 15* publiés par *SAMU et Urgences de France* (SUDF) et la *Société française de Médecine d'Urgence* (SFMU) de mars 2015. (France et al., 2015)
- Le rapport MARCUS publié en décembre 2019 par Yann Penverne (médecin urgentiste) et le Lieutenant-colonel Romain Moutard (pompier sécurité civile) sur *la modernisation de l'accessibilité et de la réception des communications d'urgence pour la sécurité, la santé et les secours*. (Moutard et al., 2019)
- Les observations terrain effectuées dans les trois SAMU partenaires du projet.

Le rapport MARCUS est le résultat d'un programme interministeriel lancée par l'Inspection Générale de l'Administration et L'Inspection Générale des Affaires Sociales (IGAS) pour faire l'état des lieux du fonctionnement des services de secours, de soins et de sécurité. Ce programme commandé par le ministère de la santé et le ministère de l'intérieur a pour objectif d'émettre des préconisations concrètes d'améliorations des centres d'appels et de fournir une feuille de route afin d'améliorer la qualité du service et l'accessibilité à la population des services de secours. Les trois objectifs principaux recherchés par cette étude sont :

- Décrocher tous les appels adressés aux différents services d'urgence.
- Les décrocher plus rapidement.
- Garantir l'orientation de chaque appel vers le service pertinent.

De cette étude (rapport MARCUS), résulte un certain nombre de constats et de recommandations sur lesquels nous nous appuyerons afin de valider la pertinence de notre démarche d'un point de vue métier. Un premier constat effectué est que l'activité opérationnelle de réponse aux appels d'urgence n'est pas standardisée. Il apparaît donc nécessaire de pouvoir décrire le traitement des appels d'urgence en utilisant une sémantique et un vocabulaire commun et homogène.

Dans la suite de ce chapitre nous présentons les différents éléments qui constituent et caractérisent les SAMU. Ces éléments permettent de comprendre leur fonctionnement et leur complexité tout en introduisant le vocabulaire utilisé dans le domaine. De plus, nous proposons d'enrichir ces descriptions avec les observations terrain effectuées dans nos SAMU partenaires. Ces observations permettent de venir mettre en avant les différences qui existent entre les SAMU et corroborent les challenges relevés par le rapport MARCUS.

1.2.2 Les différents flux

Plusieurs types et flux d'appels transitent dans les centres d'appels d'urgences. Ils sont représentés dans la figure 1.9 Les centres d'appels du SAMU sont mixtes. Cela signifie qu'ils gèrent à la fois des appels entrants, et des appels sortants.

- Dans **les appels entrants** on considère, les appels de la population (aussi appelé primo appels), mais aussi les appels des professionnels de l'urgence (police, pompiers, ambulances) et les appels des centres d'appels d'autres régions.

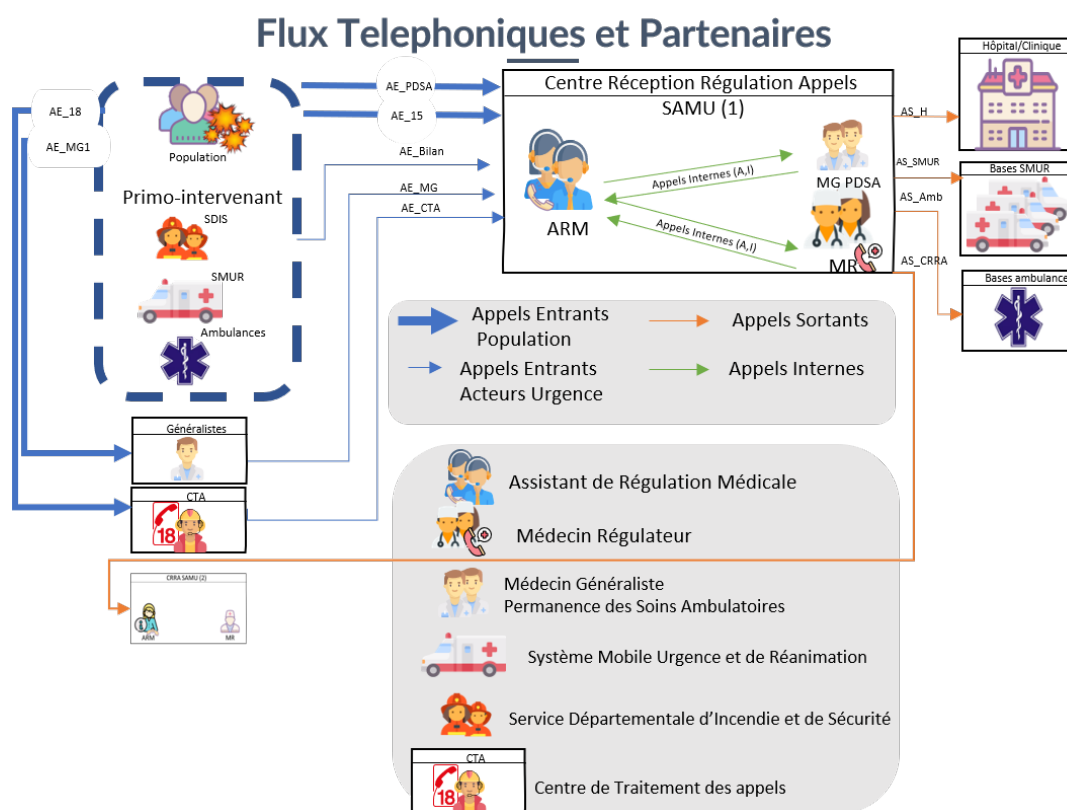


Figure 1.9 – Flux et partenaires du SAMU

Il y a aussi les appels bilans qui sont effectués par les primo-intervenants lorsqu'ils arrivent auprès du patient.

- Dans **les appels sortants**, on considère les communications vers les cliniques, hôpitaux et autres établissements de santé, les appels aux équipes de primo-intervenants pour les envoyer sur place, les appels vers les centres d'appels voisins.
- On considère aussi **des appels internes** entre les opérateurs du centre d'appels. Ceci peut arriver par exemple lorsqu'un Assistant de Régulation Médicale (ARM) appelle un Médecin Régulateur (MR) pour lui transmettre un appel particulier.

1.2.3 Le cycle de vie d'un appel entrant

Les SAMU ont défini plusieurs états pour un appel d'urgence afin de décrire au mieux son cycle de vie. La figure 1.10 résume tous les états de l'appel.

- **Un appel initié** est un appel qui est composé vers le SAMU.
- **Un appel non entrant** est un appel qui est rejeté par l'autocommutateur de manière automatisée.
- **Un appel entrant** est un appel initié qui arrive à joindre l'autocommutateur.
- **Un appel abandonné** est un appel entrant qui est interrompu par l'appelant (erreur, faux numéros, etc.).

- **Un appel traité automatiquement** est un appel entrant dirigé vers un serveur vocal interactif avec traitement automatique de l'appel.
- **Un appel perdu** est un appel présenté qui n'est ni décroché, ni traité automatiquement, ni rejeté et qui aurait dû faire l'objet d'un traitement.
- **Un appel présenté** est un appel entrant en mesure d'être décroché par un agent du CRRA.
- **Un appel décroché** est un appel présenté pris en charge par un agent du CRRA. une fois que l'appel est décroché, la communication est établie.

Le processus de gestion de l'appel nous permet de définir plusieurs durées caractéristiques du cycle de vie de l'appel. Ces différentes durées, une fois identifiées, serviront dans le calcul d'indicateurs de performance dont nous parlerons ultérieurement. Ces durées sont notamment représentées dans la figure 1.9.

- **Durée de l'appel** : Cette durée englobe toute la durée de vie de l'appel depuis son arrivée dans le centre d'appels jusqu'au moment où l'appel est raccroché à la fin de son traitement.
- **Temps d'attente** : c'est la durée d'attente de l'appel avant d'être décroché par un opérateur. On commence à mesurer ce temps à partir du moment où l'appel a passé l'autocommutateur et le disque d'accueil du SAMU. Ainsi, ce temps d'attente peut être décomposé en deux temps : **Le temps de réponse** et **le temps de sonnerie**. **Le temps de réponse** correspond à l'accueil vocal du SAMU. C'est ce que l'on appelle le « disque du SAMU » : un message automatique de bienvenue est passé à tous les appelants. **Le temps de sonnerie** correspond au moment où le téléphone de l'appelant sonne en attendant d'être décroché par un opérateur.
- **Durée de communication** : Cette durée est mesurée dès que l'appel est pris en charge par un opérateur. Elle s'étend jusqu'à ce que l'appel soit fini et raccroché. Il se peut que l'appel soit entrecoupé de temps d'attente notamment avant d'être transféré vers un médecin.
- **L'acheminement** correspond à la durée entre le moment où le patient compose l'appel et où celui-ci est visible dans le centre d'appel. Ce temps dépend majoritairement de l'opérateur téléphonique.

Ces termes présentés dans le référentiel métier de SUdF permettent d'uniformiser le vocabulaire utilisé pour décrire le traitement des appels du SAMU. C'est en effet un point très important permettant de faciliter l'échange. Nous avons pu le vérifier lors des observations menées sur le terrain. De plus, le rapport MARCUS souligne l'importance de mettre en place une sémantique homogène pour caractériser le traitement de l'appel d'urgence entre les métiers. Ceci constitue une des premières étapes pour se diriger vers une uniformisation des pratiques et permettre davantage de collaboration entre les SAMU.

1.2.4 Les ressources

Dans un CRRA, on trouve différents agents qui participent à la réception et au traitement de l'appel. Ce sont les acteurs internes du centre d'appels. On distingue :

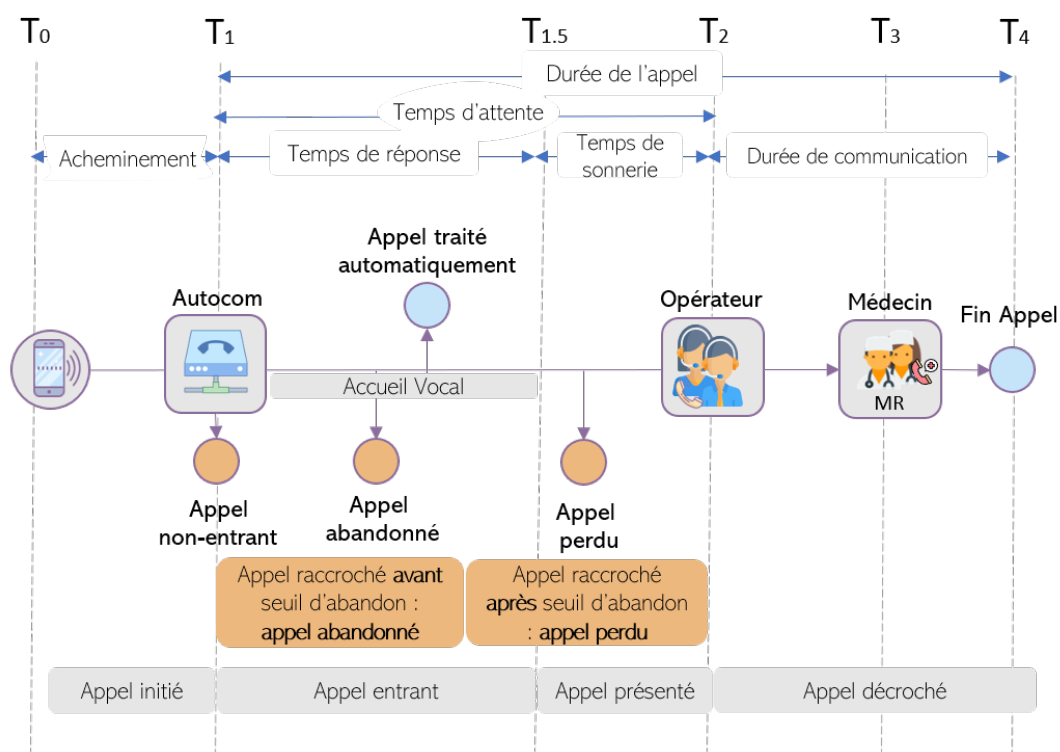


Figure 1.10 – Cycle de vie d'un appel : les étapes.

- Les **Assistants de Régulation Médicale (ARM)** qui décrochent et qualifient l'appel.
- Les **Médecins Régulateurs (MR)** qui traitent les appels nécessitant une régulation et décident de l'envoi de moyens adaptés.
- Les **Médecins Généralistes** de la PDSA qui traitent les appels de médecine générale.
- Les **Gestionnaires des Moyens (GM)** (facultatif) qui s'occupent de l'envoi effectif des moyens. À titre d'exemple, ils sont en charge d'appeler les compagnies d'ambulances, de contacter les pompiers pour les envoyer, de notifier les équipes SMUR de leur départ, etc.
- Les **Superviseurs** (facultatif) qui garantissent un management opérationnel conforme aux recommandations de l'Agence Nationale d'Appui à la Performance (ANAP). Leur rôle est d'assurer un usage des ressources cohérent vis-à-vis de la charge de travail.

Le management des ressources varie d'un département à un autre. Par exemple, parmi les SAMU observés, il existe des différences notables. Toulouse dispose d'un GM et d'un superviseur, ce qui n'est pas le cas pour Rodez et Albi. Rodez fonctionne avec un seul ARM et un seul MR pour gérer l'Aide Médicale Urgente (AMU). Durant les horaires de la PDSA, un ARM et un MR sont ajoutés pour gérer ces appels. Toulouse et Albi ont un effectif constant d'ARM sur 24 heures pour décrocher les appels, qu'ils soient AMU ou PDSA.

Puisque le CRRA fait partie du réseau des urgences, il interagit aussi avec des partenaires externes. Une partie des intervenants externes va générer des appels

entrants. C'est le cas pour les centres d'appels de la police ou de ceux de la gestion des incendies (pompiers). De même, les SAMU des départements voisins peuvent être amenés à appeler le centre d'appels pour solliciter l'aide de leur voisin. Les médecins généralistes, SOS Médecins, les infirmier(e)s libéraux et les maisons médicales de garde sont aussi des partenaires qui peuvent appeler le SAMU. Dans les appels entrants, on trouvera aussi tous les appels de suivi des cas en cours, les appels bilans passés par les primo-intervenants.

Ensuite, on retrouve les partenaires en aval de la chaîne des urgences. Tous les intervenants de terrain en font partie : les équipes du SMUR, les compagnies d'ambulances privées et les équipes d'interventions des pompiers. Les établissements de santé et les cliniques privées font aussi partie de ce réseau et peuvent être les destinataires d'appels sortants. Ils sont contactés afin d'obtenir des informations sur les places disponibles dans leurs services. Ces informations permettent au MR d'orienter efficacement le patient vers le meilleur établissement.

1.2.5 Processus de traitement d'un appel

Lorsqu'un appel arrive au CRRA, le processus de traitement d'un appel est lancé et suivra les différentes étapes correspondant à la chaîne des urgences décrite plus haut :

1. *Qualifier et Orienter* : Lorsqu'un ARM reçoit un appel, il effectue une première sélection. Dans le cas où l'appel n'est pas une demande médicale, l'ARM peut donner des renseignements administratifs. Sinon, ses tâches peuvent se décomposer en plusieurs sous-tâches comme précisées dans le référentiel métier de SuDF ((France et al., 2015)(chap. 3.6.1 p25) : «*Un ARM (1) reçoit l'appel initial (2) localise et identifie le patient (3) circonscrit la demande (4) en détermine le degré d'urgence et (5) engage si besoin des moyens avant (6) d'informer le MR et (7) d'éventuellement lui transférer l'appel.*». L'ARM engage des moyens lui-même seulement dans des cas bien précis qui sont identifiés dans des procédures métiers. C'est le cas notamment pour les Arrêts Cardio-Respiratoires (ACR).
2. *Réguler* : La régulation concerne le MR qui va pouvoir analyser la situation et diagnostiquer le patient par téléphone. Il effectue un interrogatoire médical pour ensuite décider des mesures à prendre : conseil médical ou envoi de moyens.
3. *Engager des moyens* : En fonction de la gravité de l'appel et de son diagnostic, le MR va décider des moyens à mettre en œuvre pour répondre à la situation du patient. Le type de moyen dépend de la priorité de l'appel, les échelles de priorités seront présentées un peu plus loin. Une fois que le MR a pris la décision d'envoyer une équipe SMUR, il identifie la base qui est géographiquement la plus proche et qui dispose d'une équipe disponible pour leur transmettre la notification d'intervention avec les informations nécessaires. Une fois cette étape finie, le processus bascule dans la partie aval (ou externe) de la gestion de l'appel.

4. *Rejoindre et stabiliser* : Une fois la notification reçue, l'équipe SMUR prépare l'intervention avec le véhicule adapté à la situation. Les Unités Mobiles Hospitalières (UMH) sont des ambulances disposant du matériel et des médicaments nécessaires pour transporter un patient alité. Les Véhicules Légers (VL) contiennent du matériel médical et des traitements pour la prise en charge et la stabilisation du patient, néanmoins ils ne permettent pas de transporter le patient. Enfin, l'hélicoptère sanitaire permet d'accéder très rapidement à des lieux reculés ou à des régions difficiles d'accès ou impraticables par voie routière. Une fois sur place, l'équipe de soignants stabilise le patient puis contacte le CRRA pour transmettre un bilan au MR afin qu'il puisse orienter le patient.
5. *Orienter* : Une fois le patient stabilisé et le bilan transmis, le MR dispose de plus d'information pour décider du devenir du patient. S'il décide d'envoyer le patient dans un établissement de santé, il doit s'assurer de la place disponible pour l'accueil du patient. Le choix de l'établissement de santé dépendra des places disponibles, de la spécialité nécessaire, mais aussi du choix du patient.
6. *Transporter* : Étape de transfert du patient du lieu de l'intervention vers l'établissement de santé.
7. *Hospitaliser* : L'équipe SMUR reste avec le patient jusqu'à son admission. Puis elle retourne vers sa base et son statut est de nouveau disponible, et peut au besoin être renvoyée en intervention.

Prioriser les appels

En fonction de leur priorité, les appels suivront tout ou partie des étapes présentées ci-dessus. Toutefois, les questions qui peuvent se poser maintenant sont notamment : dans quel ordre traiter les appels ? Quels sont les appels prioritaires par rapport aux autres ? Selon quels critères ? Est-ce que les appels entrants sont plus importants que les appels sortants ? Est-ce que tous les appels entrants sont plus importants que les appels sortants ? etc.

Le référentiel métier du SAMU (France et al., 2015) apporte des éléments de réponse à certaines de ces questions. Effectivement, il y a une hiérarchie dans les appels entrants. Des niveaux de priorité sont définis. La « priorisation » de l'appel se décline selon 3 niveaux de priorité, de P0 à P2.

- P0 : Pour les appels présentant d'emblée un caractère d'extrême urgence et autorisant un départ réflexe « flash » (SMUR, VSAB médicalisé) sans régulation médicale préalable (selon un protocole préétabli), suivi d'une régulation prioritaire ;
- P1 : Pour les appels nécessitant une régulation médicale prioritaire, immédiate, par un urgentiste ;
- P2 : Pour les appels relevant de la PDSA ou de transferts de patients vers une structure médicale, pouvant être temporisés.

La Haute Autorité de Santé (HAS) propose une classification différente (Santé, 2011). Bien qu'elle reconnaisse la pertinence de l'échelle de priorité proposée par SAMU

et Urgences de France, la HAS propose une décomposition en 4 niveaux. Les deux premiers concernant l'Aide Médicale Urgente (AMU) et les deux suivants concernant les médecins généralistes.

- Le niveau R1 correspond à une urgence absolue et vitale. On y envoie une équipe SMUR.
- Le niveau R2 correspond à une urgence sans détresse vitale. Il exige une intervention médicale prompte (temps d'intervention inférieur à 30 minutes).
- Le niveau R3 correspond à la Permanence des Soins Ambulatoires (PDSA). Ce niveau requiert un avis médical dans les 12 à 24 h.
- Le niveau R4 correspond à un conseil médical. Il peut donner lieu à la programmation d'un rendez-vous avec le médecin traitant.

Les ARM sont en mesure d'identifier les appels relevant de l'AMU (urgences de niveaux R1 et R2) de ceux relevant de la PDSA (degrés R3 et R4) et de les transférer vers le bon médecin qu'il soit urgentiste ou généraliste. Ils peuvent satisfaire de plus aux requêtes administratives distinctes du conseil médical.

Dans la suite, nous utiliserons le référentiel de SUdF auquel nous rajouterons un niveau P3 correspondant aux appels qui ne donnent pas lieu à un suivi par un médecin.

1.2.6 Les systèmes d'information des SAMU

La gestion informatique des SAMU se décompose en deux parties supportées par deux logiciels métiers différents. La partie de gestion des appels téléphoniques qui s'occupe de la gestion des files d'attente, du paramétrage du disque d'accueil, du routage des appels vers les ARM est supportée par le Logiciel de Téléphonie (LT). La gestion des dossiers de régulation est faite grâce au Logiciel de Régulation Médicale (LRM).

Les Logiciels de Téléphonie

Tous les SAMU de France ne sont pas équipés du même système de téléphonie. Il en existe de plus ou moins performants décomposés en quatre types d'après l'étude menée pour la mise en place du projet SI-SAMU⁵ :

- Des SAMU équipés de téléphonie très peu évoluée, avec des téléphones fixes et des serveurs non sauvegardés. Cela implique que la récupération des données de téléphonies est impossible, car il n'y a pas d'endroit où les données sont stockées.
- Des SAMU équipés de téléphonie standard, avec des téléphones fixes, mais une sauvegarde des serveurs. C'est notamment le cas pour le SAMU de Rodez où le décroché des appels se fait sur le téléphone fixe. Les ARM n'ont pas de visibilité sur la file d'attente des appels si ce n'est un voyant lumineux lorsqu'il y a un ou plusieurs appels en attente d'être décroché.

5. <https://esante.gouv.fr/projets-nationaux/si-samu>

- Des SAMU équipés de téléphonie sans fil avec sauvegarde et serveurs avec sauvegarde locale assurant la redondance des informations. C'est le cas par exemple pour le SAMU d'Albi.
- Des SAMU équipés de téléphonie avancée, interfacée avec un ordinateur pour permettre de décrocher et de gérer l'appel depuis le poste informatique. Le SAMU de Toulouse est dans ce cas, il dispose d'un logiciel de téléphonie interfacé avec le poste informatique. De ce fait, le décroché peut se faire directement depuis l'ordinateur. De plus, ils ont une visibilité sur les appels présents dans la file d'attente.

La gestion des dossiers patients

Lorsqu'un patient appelle le SAMU, il est affilié automatiquement à un dossier Patient. Il en existe plusieurs types (D, DR, DRM) en fonction de la situation du patient et de la réponse qui lui est apportée. Dès qu'un appel arrive au SAMU, le LRM crée un Dossier D en lien avec cet appel. Ce dossier sera qualifié de Dossier de Régulation (DR) lorsqu'il contient les informations renseignées par l'ARM lors de la qualification de l'appel. Le DR concerne les demandes d'information à caractère médical, médico-social ou sanitaire. Un dossier est qualifié de Dossier de Régulation Médicale (DRM) lorsqu'il a fait l'objet d'un acte de régulation médicale par un MR.

À la suite de nos observations sur le terrain, nous avons pu constater que la communication des données entre ces deux logiciels n'est pas évidente. Le lien entre un appel dans le LT et un dossier dans le LRM n'est pas automatique. Il n'existe pas forcément d'identifiants uniques permettant de les relier. Toutefois, nous pouvons espérer qu'avec le projet SI-SAMU d'uniformisation des systèmes d'information du SAMU, cela ne soit plus un problème. Intéressons nous un instant à ce fameux projet SI-SAMU.

Le projet SI-SAMU : Programme de modernisation des systèmes d'information et télécommunication des SAMU

Le projet SI-SAMU est un projet mené par l'agence du numérique en santé (appelée agence des systèmes d'information partagés de santé (ASIP) avant 2019) qui propose de concevoir un nouveau système d'information commun pour tous les SAMU. Ce projet ambitieux cherche à répondre à plusieurs objectifs :

- Contribuer à des soins et un service de qualité homogènes sur l'ensemble du territoire.
- Assurer la régulation médicale et le pilotage à tous les échelons, local, régional et national des appels d'urgences.
- Être en capacité de gérer une crise sanitaire d'envergure nationale.
- Améliorer l'interaction avec les systèmes d'information des partenaires comme avec les logiciels de santé.
- Garantir la confidentialité des données de santé.
- Offrir de nouvelles opportunités pour la recherche clinique.
- Concilier l'efficacité de la dépense publique et la performance de l'activité du SAMU.

SI-SAMU veut regrouper le LRM, le LT, les flux multicanaux (comme les ondes hertziennes), et le système de géolocalisation sous un seul logiciel unique et commun à tous les centres d'appels du territoire. Cette uniformisation a pour ambition une mise en réseau des SAMU, une interopérabilité accrue entre les acteurs du réseau des urgences, un meilleur pilotage de l'activité nominale, mais aussi de la gestion de crise. Cette solution métier est très attendue comme le montre ce témoignage du Dr Marc Giroud, cofondateur de SAMU France en 1975 :

« Si la pandémie de grippe A H1N1 a souligné l'immense intérêt des centres de régulation médicale en 2009, elle a aussi mis en évidence l'insuffisance et le cloisonnement de leurs solutions de télécommunication et de leurs systèmes d'information, dont le rôle est pourtant critique pour prendre en charge des patients en situation de détresse vitale. Le travail conduit sur le SI-SAMU par l'ASIP Santé depuis 2014 est sans précédent. Non seulement il s'est fait à un niveau d'expertise et de concertation inédit, mais il a suscité l'adhésion de la communauté métier élargie. Une adhésion croissante des urgentistes hospitaliers, dont certains avaient initialement manifesté des réserves et qui sont aujourd'hui impatients d'accéder aux progrès attendus, en termes de qualité et de sécurité ».

Ce projet, commencé en 2014, est aujourd'hui en 2020 dans la phase des tests pilotes sur certains sites. D'après le calendrier initial, le déploiement du SI-SAMU devrait se généraliser sur tout le territoire entre 2020 et 2024.

1.2.7 Évaluation des SAMU

Nous avons auparavant souligné les difficultés du SAMU à décrocher tous les appels. Un nombre non négligeable d'appels abandonnés ou perdus a été relevé durant l'année 2018. À noter qu'un seuil d'abandon en deçà duquel l'appel est considéré comme une erreur de l'appelant est fixé à 15 secondes pour le SAMU. Cela signifie que les *appels abandonnés* avant 15 secondes ne sont pas considérés comme des appels qui auraient nécessité une réponse, là où les appels décrochés après 15 secondes sont des *appels perdus* qui auraient dû recevoir une réponse (France et al., 2015).

Plusieurs indicateurs de performances (KPI) sont utilisés pour évaluer les centres d'appels. On peut distinguer plusieurs catégories d'indicateurs suivant le triptyque classique coût, qualité et délais (Olsen, 1971). Parmi eux on retrouve les suivants :

Les indicateurs de délais

1. **La Durée Moyenne de Communication** (DMC) est le rapport entre le temps de communication et le nombre d'appels, elle correspond au temps moyen de communication passé sur chaque appel par agent. Cela permet d'évaluer leur productivité. (cf. équation 1.1)

$$DMC = \frac{\text{temps de communication total}}{\text{nb d'appels traités}} \quad (1.1)$$

2. **Le temps de réponse** est défini comme la durée entre l'arrivée de l'appel sur la plateforme téléphonique et la prise en charge du patient par des effecteurs sur place (Aboueljinane, 2014). Il peut être calculé de plusieurs manières. On peut considérer le temps de réponse moyen, mais aussi le **Taux de couverture dans un délai T** (Ingolfsson et al., 2003) où T est une valeur prédéfinie qui est soit fixée par les auteurs (cf. équation 1.3), ou fixée par la loi en vigueur dans le pays. On retrouve aussi des **Taux de non-couverture au-delà d'un délai T** qui se définissent comme le pourcentage d'appels pris en charge au-delà du seuil T (cf. équation 1.4). La valeur de T varie d'un pays à un autre suivant les recommandations gouvernementales ou régionales. En Italie, 100% des appels doivent être gérés en moins de 8 min dans les villes et leurs périphéries. En Amérique du Nord, 90% des appels doivent être couverts en moins de 9 minutes (Ingolfsson et al., 2003).

$$\text{Temps de réponse} = \text{Instant de prise en charge du patient} - \text{Arrivée de l'appel} \quad (1.2)$$

$$\text{Taux de couverture dans un délai } T = \frac{\text{Nbre d'appels traités en moins de } T \text{ sec}}{\text{Nbre appel total}} \quad (1.3)$$

$$\text{Taux de non couverture dans un délai } T = \frac{\text{Nbre d'appels traités en plus de } T \text{ sec}}{\text{Nbre appel total}} \quad (1.4)$$

3. **Le temps d'accès au traitement** est défini comme le délai entre l'arrivée de l'appel dans le centre et l'arrivée du patient dans un établissement de santé. Il peut être mesuré de différentes manières : Su et al. utilisent la durée entre la demande d'envoi d'une équipe d'effecteur et l'admission du patient à l'hôpital (Su et al., 2003). Lee et al. le définissent d'une manière similaire au temps de couverture avec un seuil T. Ils mesurent alors la part des appels dont le temps d'accès au traitement est inférieur à ce temps T (Lee et al., 2012).
4. **Le temps de service** correspond à la durée ininterrompue pendant laquelle une ressource est engagée à réaliser un service. Pour la partie aval, cela correspond au temps passé en service par l'équipe d'effecteurs entre le moment où elle est notifiée de l'intervention et le moment où elle est de nouveau disponible. Pour la partie amont dans le centre d'appels, le temps de service correspond au temps de communication, c'est le temps où les ressources sont en ligne ou en train d'effectuer une activité en lien avec l'appel. Cet indicateur est important car il permet d'assurer que la charge de travail est bien répartie. Goldberg, montre l'importance sociale de cet indicateur : une équité dans la répartition de la charge de travail permet de diminuer le sentiment de pénibilité et d'injustice ressenti par le personnel (Goldberg, 2004). Ce qui amène directement à définir un taux d'utilisation des ressources.

5. **Le temps de déclenchement** correspond au délai entre l'arrivée de l'appel au CRRA et la notification de la mission aux effecteurs. Cet indicateur permet de mesurer l'efficacité du maillon amont de la chaîne des urgences, la capacité du centre d'appels à traiter et à décider d'envoyer des moyens (Aboueljinane, 2014). Pour le maillon aval, c'est l'indicateur suivant qui sera utilisé.
6. **Le temps de transport** au lieu d'intervention correspond au délai entre la notification de la mission aux effecteurs et l'arrivée de l'équipe sur place. Ce délai est sujet à des perturbations extérieures comme les conditions de circulation, ou la difficulté d'accès et de localisation de la zone d'intervention (Su et al., 2003).
7. **Le temps d'attente** se définit à la fois pour la partie amont et la partie aval. Pour la partie amont, c'est le temps avant qu'un appel ne soit décroché. Pour la partie aval, il correspond au temps passé à attendre avant qu'une équipe d'effecteurs soit affectée à un appel. De même, **la taille de file d'attente** correspond au nombre d'appels en attente de l'affectation d'une équipe d'effecteurs. Silva et al. utilisent cet indicateur pour évaluer la coordination entre les activités de l'amont et l'aval. De plus, ils utilisent la valeur maximale de la file d'attente pour évaluer la résilience du système lors d'une augmentation des appels (Silva et al., 2010). La résilience peut être définie comme la capacité d'un système à revenir à son état initial après avoir subi une perturbation (Leflar et al., 2013).

Les indicateurs de qualité

1. **Le taux d'accueil** est défini comme le nombre d'appels répondus sur le nombre d'appels entrants. Il exprime la capacité du centre d'appels à être capable de décrocher tous les appels. D'après SUdF en 2015, il n'y avait pas de recommandation professionnelle concernant l'objectif du taux d'accueil. Toutefois le rapport MARCUS de 2019 précise que suite au comité stratégique du 7 novembre 2019, la cible de 99,95% de taux d'accueil est fixée. Cet indicateur de taux d'accueil permet de mesurer l'accès de la population aux centres d'appels d'urgences. Il est généralement calculé selon la formule 1.5.

$$Taux\ Accueil = \frac{Nb\ d'appels\ décrochés}{nb\ d'appels\ entrants - nb\ d'appels\ abandonnés} \quad (1.5)$$

2. **La qualité de service** aussi appelé temps de décroché est le pourcentage d'appels décrochés en moins de x secondes. x peut prendre plusieurs valeurs selon le pays. L'état français fixe que 99% des appels doivent être décrochés en moins de 60 secondes (Perverne et al., 2017). La NENA⁶ pour le 911 aux USA définis que 95% des appels doivent être décroché en moins de 20 secondes. (cf. équation 1.6)

$$QS_x = \frac{Nb\ d'appels\ décrochés\ avant\ x\ sec}{nb\ d'appels\ entrants - nb\ d'appels\ abandonnés} \quad (1.6)$$

6. National Emergency Number Agencies

3. **Le nombre d'appels servis par base/équipe** permet de compléter la mesure du temps de service ainsi que le taux d'utilisation pour assurer une certaine équité entre les équipes. Cela permet que ce ne soit pas toujours la même équipe qui parte en intervention, cela permet aussi de lisser le taux d'utilisation entre les équipes. Cet indicateur est transposable aux personnels du CRRA avec le nombre d'appels décrochés/traités par ressource (Penverne et al., 2009). Les recommandations de SUDF sont qu'un ARM devrait traiter une moyenne de 7,2 appels par heure. De la même manière, la charge cible pour un MR est de 6 dossiers de régulations médicales par heure⁷.
4. **Le pourcentage de non-réponse** correspond à l'échec de prise en charge d'un appel dans un temps limite fixé. Il est complété par le taux d'abandon, qui correspond aux appels qui ne sont pas décrochés. Il faudrait aussi rajouter les appels qui n'arrivent pas dans le système, car les files d'attente sont saturées avant le décroché. (Inakawa et al., 2010)
5. **Le taux de survie** est un indicateur à part entière qui est plus difficile à mesurer que les différents délais. Il correspond au pourcentage de patients qui survivent après avoir fait appel au service de secours. C'est un indicateur qui reste très difficile à quantifier dans la mesure où beaucoup de facteurs extérieurs peuvent intervenir dans le décès potentiel d'un patient. Évaluer les impacts des modifications organisationnelles sur le taux de survie est donc assez complexe, ce qui explique le manque de représentation de cet indicateur dans la littérature. (Sacco et al., 2005)
6. **La collaboration** est une famille d'indicateurs définis par (Belaidi et al., 2007) qui évalue un SAMU au sein du réseau de soin des urgences vis-à-vis des autres acteurs du réseau. Les concepts d'évaluation de la collaboration sont définis, mais ne sont pas forcément mis en application.

Les indicateurs de coût

1. **Le taux d'utilisation** ou encore **taux de charge** concerne tant les ressources du centre d'appel que les effecteurs. Il correspond à la fraction entre le temps de service et le temps total de travail (cf. équation 1.7). Les taux d'utilisation des équipes comme des ressources du CRRA sont généralement bas pour maximiser les chances d'être disponibles lors d'une urgence. Dans le système français, le taux de charge cible pour les médecins devrait être fixé autour de 25% d'après les recommandations de la MeaH (Hospitalier (MeaH)., 2008). Concernant les ARM, le taux de charge cible est se situe entre 20 et 40 %⁸.

$$\text{Taux d'utilisation} = \frac{\text{temps de service}}{\text{temps total de travail}} \quad (1.7)$$

7. Samu-Urgences de France. Définitions & standardisation des données : Activité des SAMU Centre 15. 2009

8. Samu-Urgences de France. Définitions & standardisation des données : Activité des SAMU Centre 15. 2009 : http://www.samu-de-france.fr/documents/actus/155/559/definitions_apl_drm_2009.pdf; http://www.samu-de-france.fr/documents/actus/2155/2559/thesaurus_apl_drm_2009.xls

2. **La distance parcourue** permet d'évaluer le coût kilométrique des interventions, cela permet d'évaluer l'efficacité du déploiement des équipes. Elle correspond à la distance totale parcourue par une équipe SMUR depuis son départ de la base jusqu'à son retour après intervention. (Koch et al., 2003)
3. **La rentabilité** correspond à l'évaluation du coût de mise en place des modifications organisationnelles proposées. C'est un indicateur plutôt complexe dans le contexte d'un service d'urgence. Le but de cet indicateur n'est pas tant de mesurer la rentabilité économique des services d'urgences que d'être en mesure d'évaluer le coût et les bénéfices des solutions proposées pour l'améliorer. Bien souvent les améliorations proposées ne sont pas ou peu évaluées du point de vue économique. Aboueljinane et al. l'ont déjà relevé à travers leur revue de littérature sur le sujet (Aboueljinane et al., 2013). La littérature est assez succincte sur cette question sûrement à cause de la difficulté de chiffrer les améliorations proposées, bien que des tentatives de classifications de coûts standards aient été proposées par (Savas, 1969) ou encore par (Lerner et al., 2007).

La Figure 1.11 présente une synthèse des indicateurs sur la chaîne des urgences.

1.3 Revue de littérature sur la gestion des SAMU

Maintenant que nous avons fixé les éléments de caractérisation du système d'étude, intéressons-nous aux principaux travaux de recherche qui ont été conduits sur le sujet. Les SAMU ont suscité l'intérêt de la sphère académique de par leurs problématiques qui ouvrent un champ d'investigation à la fois scientifiquement riche, mais aussi très humain (Aringhieri et al., 2017). La complexité des interactions humaines lors d'une situation d'urgence, croisée avec la nécessité d'agir rapidement et le tout au sein d'un réseau multi-acteurs (Belaidi, 2009) font qu'il y a de nombreux défis à relever. De nombreux travaux scientifiques ont étudié les centres d'appels d'urgences en prenant en compte leurs différences et leurs particularités. Parmi les problématiques adressées, on trouve notamment la définition du nombre de ressources (Addis et al., 2012) pour décrocher les appels, mais aussi la définition d'indicateurs de performance adéquats pour évaluer l'activité (Passmore et al., 2013; Penverne et al., 2017), le dimensionnement des flottes de véhicules pour répondre à la demande de soin (Aboueljinane et al., 2014). Il faut ensuite traiter la localisation et la répartition de ces véhicules pour avoir un maillage optimal du territoire (Bélangier et al., 2016). Ce sont des exemples de challenges auxquels font face les SAMU et que nous allons étudier plus en profondeur dans la suite.

Toutefois, tous les problèmes ne concernent pas l'intégralité de la chaîne des urgences. En effet, on observe des problèmes bien distincts entre la partie amont liée à la gestion des appels et la partie aval liée à la gestion des effecteurs et de la prise en charge du patient. La figure 1.12 cherche à représenter les types de problèmes traités en fonction de l'étape de la chaîne des urgences. Une manière de les classer et de découper en

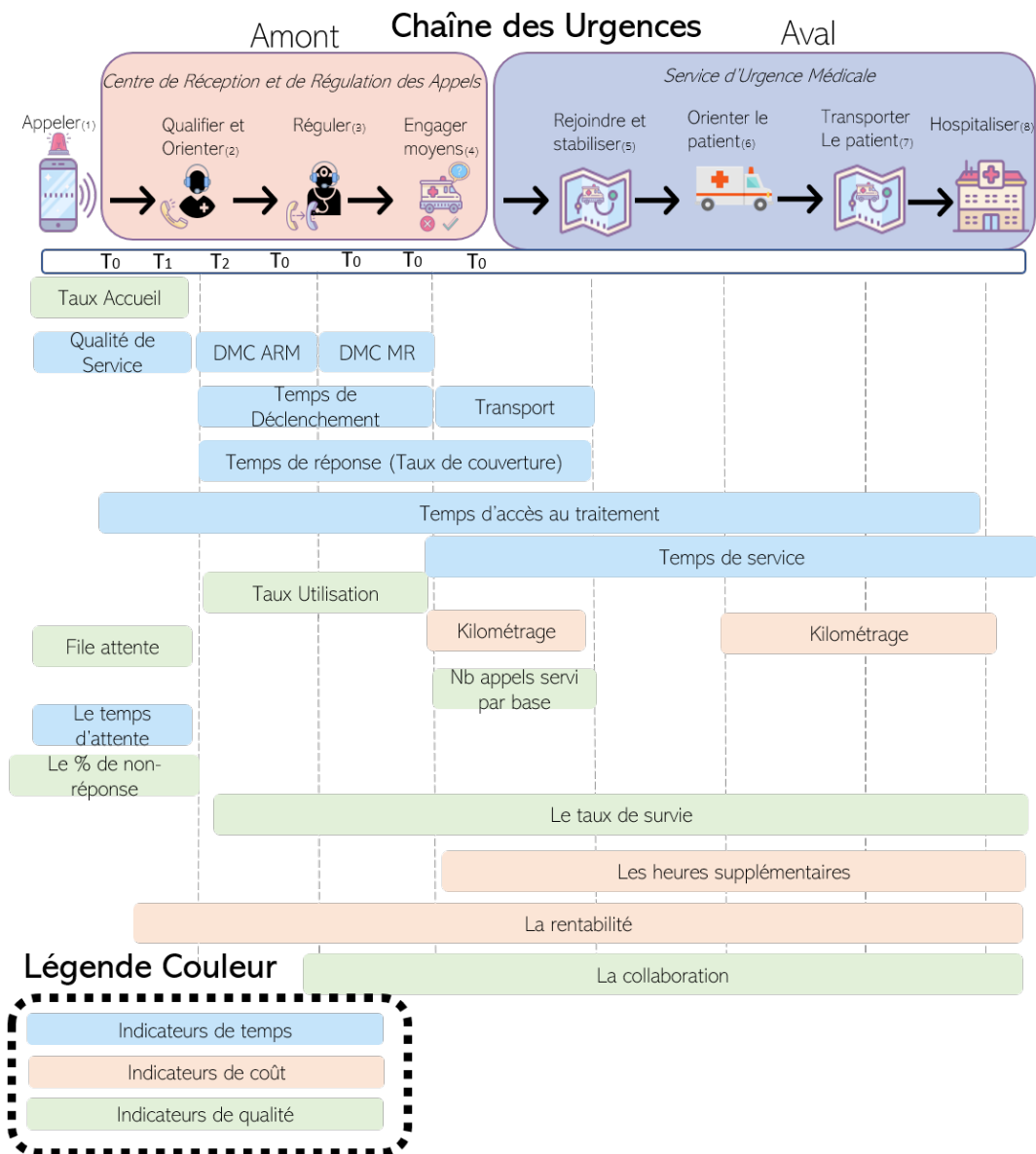


Figure 1.11 – Indicateurs de performance

trois niveaux de décisions différents : stratégique, tactique et opérationnel (Gendreau et al., 1997). Le niveau stratégique correspond à un horizon de temps assez long avec des décisions structurelles qui nécessitent un investissement conséquent. On retrouve les problèmes de localisation des bases SMUR, de dimensionnement des flottes de véhicules et de leurs équipes. On retrouve aussi les problèmes prospectifs comme la gestion des nouveaux supports de communication. Le niveau tactique s'intéresse notamment à la prévision des appels pour pouvoir définir les plannings des ressources, les affectations des équipes aux différentes bases. Le niveau opérationnel concerne le pilotage de l'activité quotidienne. Dans cette catégorie on va trouver les problèmes de répartition (*dispatching*), de routage (*routing*) et de relocalisation des équipes d'intervention pour la partie aval. Pour la partie amont on peut y ranger les règles de décroché et de priorisation des appels.

1.3.1 Les problèmes stratégiques

Les problèmes de dimensionnement

Les services médicaux d'urgences ont besoin de ressources hautement qualifiées que ce soit au niveau amont du centre d'appels ou au niveau aval des intervenants terrains afin d'apporter le meilleur soin possible au patient. C'est pourquoi les problèmes de dimensionnement des équipes de soin (Passmore et al., 2013) sont primordiaux dans la conception et la gestion des systèmes de soins urgents. En effet, il faut trouver un juste équilibre entre un nombre suffisant de ressources de manière à garder un temps d'attente bas et une éventuelle sous-utilisation de ces ressources hautement qualifiées avec un coût salarial important. (Addis et al., 2012) Dans les problèmes de dimensionnement, on trouve aussi ceux qui concernent les flottes de véhicules (Savas, 1969 ; Trudeau et al., 1989). Ces problèmes cherchent à répondre à la question « Combien d'équipes et de véhicules faut-il pour être capable de répondre à la demande d'intervention ? ». Ces problèmes sont complexes, car ils sont soumis à une grande incertitude. En effet, le dimensionnement des équipes dépend directement de la prévision du nombre d'interventions potentielles (Setzler et al., 2009). Plusieurs chercheurs se sont penchés sur ce problème pour identifier le nombre de bases nécessaires pour couvrir une zone géographique le plus efficacement possible. Nous pouvons citer notamment : (Aringhieri et al., 2017) (Iskander, 1989), (Su et al., 2003), (Swoveland et al., 1973), (Trudeau et al., 1989). Savas, 1969 a utilisé la simulation afin de tester l'effet de l'ajout ou de la suppression d'équipes et de bases potentielles sur la qualité et le coût du service d'urgence. L'objectif était de réduire le temps de réponse moyen des équipes de l'hôpital Kings County à New York. Il montre que l'ajout de nouvelles équipes dans la base initiale (soit l'hôpital) permet de réduire le temps de réponse moyen de 18 secondes. Au-delà de dix équipes, il n'y a plus de différence sur le temps moyen. De plus, il montre qu'en jouant sur les emplacements des équipes il est possible d'améliorer le temps de réponse jusqu'à 19% du temps initial. Cette étude est une des premières à utiliser la simulation pour un service d'urgence. Une autre étude (Gunes et al., 2005) a évalué l'impact qu'aurait l'extension de planning d'un hélicoptère de 5 à 7 jours pour un état du nord des États-Unis. Leurs résultats montrent que cela induirait

une diminution du taux d'utilisation de l'hélicoptère de 43% avec un impact très faible et non significatif sur le temps de réponse. (Aboueljinane, 2014), (Inakawa et al., 2010) et (Ingolfsson et al., 2003) ont développé des modèles de simulation pour mesurer l'effet de la modification du nombre d'équipes et de bases pour les services d'urgence de la région du Val-de-Marne en France, de Seto au Japon et d'Edmonton au Canada respectivement. Les résultats ne sont pas les mêmes d'une étude à l'autre. À Edmonton, il serait plus intéressant de rajouter deux nouvelles équipes plutôt qu'une nouvelle base, cela permettrait d'améliorer le pourcentage des appels répondus en moins de 9 minutes de 87% à 90%. Au Japon et dans le Val-de-Marne, les études montrent que l'ajout de nouvelles bases et la répartition des équipes dans celles-ci permettent d'obtenir de meilleurs résultats que l'ajout de nouvelles équipes.

Les problèmes de localisation

Les problèmes de localisation ont pour but de décider de l'emplacement des bases SMUR d'où partent les véhicules d'intervention (Savas, 1969) (Aboueljinane et al., 2013). Comme on a pu le constater dans le paragraphe précédent, ils vont bien souvent de pair avec les problèmes de dimensionnement. L'objectif est d'avoir un maillage optimal pour maximiser la couverture du territoire. Cela signifie que l'on doit être capable de joindre chaque point du territoire dans un délai défini au préalable. La localisation va être liée aux emplacements des bases. Les problèmes de dimensionnement répondent à la question « Combien d'infrastructures dois-je déployer ? ». La localisation, quant à elle, répond à la question « Où dois-je les déployer ? ». Il est nécessaire de trouver un compromis pour maximiser la couverture tout en minimisant le nombre de bases. Ces problèmes de localisation et de dimensionnement vont donc avoir un impact sur plusieurs types d'indicateurs :

- Des indicateurs de temps : Si la localisation est bien faite et que la couverture du territoire est bonne, le temps de prise en charge des interventions sera réduit, car il sera plus rapide de se rendre sur place.
- Des indicateurs de coût : Si la localisation est optimale, elle permet d'avoir une bonne couverture du territoire en minimisant le nombre d'infrastructures. Cela permet de réduire les coûts associés au nombre de kilomètres parcourus.

(Başar et al., 2012) proposent de classer les problèmes de localisation suivant trois critères principaux : le type de problème, la modélisation, et les méthodologies de résolution utilisées. Pour le type de problème, ils les ont répartis en quatre sous-groupes comprenant le type d'urgence, la structure du modèle (s'il est déterministe, stochastique ou robuste), la variation du modèle dans le temps (s'il est statique ou dynamique) et enfin le nombre d'objectifs à atteindre. Concernant la modélisation, elle est caractérisée par la définition de la fonction objectif, les paramètres utilisés ainsi que le type de programmation mathématique utilisée (entière, dynamique, non linéaire, etc.) Quant aux méthodologies elles se répartissent dans plusieurs catégories : les méthodes exactes, les heuristiques, les méta heuristiques et la simulation.

La gestion des nouveaux supports

Dans les années à venir de nouveaux supports de communication vont être utilisés par les centres d'appels d'urgence que ce soit des supports vidéo, SMS, e-mail, ou encore les appels E-call issus des voitures. L'avancée des technologies de communication pourrait avoir un impact sur le fonctionnement des SAMU. Gupta et al. listent les différents challenges et problèmes en lien avec la diversification des canaux de communication notamment dans le cadre de la gestion de crise : « Comment faire évoluer l'architecture du 911 pour intégrer les technologies sans fil ? La géolocalisation ? », « Comment gérer la continuité de la prise en charge d'une série de SMS ? », « Comment éviter la transmission de virus par SMS et MMS dans de telles situations ». « Comment intégrer la gestion de support image et vidéo ? », « Est-ce que l'utilisation de capteurs dans les maisons notamment de personnes âgées est envisageable ? ». Toutes ces questions se posent à l'heure où la technologie de l'information et de la récupération des données se développe extrêmement rapidement. La grande quantité de données que va générer la multiplication des canaux de communication induit de nouveaux besoins pour la visualisation, l'analyse et l'utilisation des données. Il y a aussi un besoin d'interopérabilité entre les différents acteurs pour la transmission des données pertinentes. Un autre challenge concerne la localisation de l'appelant pour localiser le lieu de l'accident, mais aussi localiser le centre qui doit gérer l'appel. Notamment pour le cas des téléphones portables qui sont plus difficiles à localiser lorsqu'ils sont proches de frontières ou de grands axes. Un autre problème est la gestion des appels redondants dans le cas de crises majeures. Comment les traiter, les rediriger vers d'autres acteurs de la sécurité civile comme la croix rouge par exemple ? Une preuve de concept a été faite sur le nouveau modèle envisagé du NG-9-1-1 où le transfert d'appels d'un centre d'appels vers un autre a été testé avec succès (Gupta et al., 2010).

Markakis et al. posent différentes questions comme « Comment intégrer les médias sociaux à la gestion des appels d'urgence ? » ou encore « Comment préparer les services d'urgences la prochaine génération de réseau (Next Generation Networks, NGN) ». Cet article propose une plateforme pour intégrer les nouveaux réseaux de communication au service des urgences. Cela permet alors de rassembler les communications téléphoniques, radiophoniques, smartphones, tablettes, E-Call, mais aussi les différents médias sociaux existants. (Markakis et al., 2017). (Weinlich et al., 2018) développent et testent un système de géolocalisation par GPS et Wi-Fi afin de contacter rapidement les urgences. À des fins de test, il a été déployé dans 11 pays différents. Les résultats de leur travail montrent qu'un tel système améliore significativement l'arrivée des urgences (jusqu'à 2 heures peuvent être gagnés) surtout pour des patients désorientés ou avec une difficulté à s'exprimer.

Oorni et Goulart rappellent qu'à partir de 2020, tous les véhicules commercialisés devront être équipés du module E-call permettant de joindre directement les centres d'appels d'urgence par pression du bouton ou automatiquement lors du déclenchement des airbags. Avec cette évolution, les auteurs se demandent si les plateformes de gestion des appels d'urgences sont prêtes et équipées pour recevoir les E-Call. À travers cet article les auteurs présentent le concept des appels E-call et la manière

dont il est implémenté en comparant deux systèmes. Cet article étudie plus le côté architecture système des centres d'appels d'urgence (Oorni et al., 2017).

Les problèmes de dimensionnement et de localisation sont très liés et bien souvent traités ensemble dans la littérature scientifique. En effet, il est évident qu'il faut connaître le nombre de ressources disponibles avant de pouvoir les répartir. La difficulté de ces problèmes réside dans le fait de trouver une réponse optimale pour satisfaire des critères de temps tout en ayant des contraintes de moyens.

L'apparition de nouveaux supports de communication soulève des challenges au niveau de l'architecture du réseau, mais aussi au niveau de la gestion et du traitement de ces sources d'information. Qui devra s'en occuper ? Et comment les intégrer aux outils déjà existants sont des questions auxquelles des réponses restent à apporter.

1.3.2 Les problèmes tactiques

Les problèmes de prévisions des appels

Pour résoudre les problèmes mentionnés précédemment il est nécessaire d'avoir une vision assez précise du nombre d'appels, du nombre d'interventions qu'il peut y avoir par heure, jour, semaine, etc. Ainsi, un autre pan de la recherche s'intéresse à la prédiction du volume d'appels et aux lois probabilistes suivies par l'arrivée des appels. Pour déterminer au mieux le nombre d'appels dans le futur, il est aussi intéressant de connaître les paramètres extérieurs et environnementaux qui peuvent influencer l'arrivée des appels. Les vacances scolaires peuvent générer une augmentation des appels d'urgence dans certaines régions. Ceci est dû au tourisme et à l'augmentation de la population dans une zone donnée sur une courte période. Les épidémies saisonnières comme la grippe auront aussi un impact sur l'arrivée des appels. La démographie d'une région aura aussi une influence sur le nombre d'appels : si un département est très urbanisé avec une population plutôt dense, le nombre d'appels d'urgence augmentera en conséquence.

Ainsi, les problèmes de prévision des appels touchent aussi à l'épidémiologie, la géographie et la sociologie. Pour être capable de définir un modèle de prévision des appels fiable, il faut être en mesure de prendre en compte tous ces paramètres. C'est pourquoi une pratique courante est de travailler avec des données historiques (Sariyer, 2018 ; Viglino et al., 2017) ce qui permet d'avoir l'assurance que le comportement de l'arrivée des appels est pertinent (puisqu'il a déjà eu lieu). Cette méthode est pertinente lorsqu'il s'agit de développer de nouvelle manière de fonctionner. Pour une gestion opérationnelle, il reste limité. C'est pourquoi des modèles de prévision des appels basés sur les données historiques ont été développés. Une bonne prévision des appels est très importante, car c'est sur celle-ci que les responsables d'équipes vont se baser afin de constituer les plannings des ressources notamment. (L. H. Brown et al., 2007)

On trouve plusieurs types de modèles de prévision. Les premiers modèles rencontrés sont ceux qui se basent sur des statistiques descriptives, (Hall, 1971) ils sont néanmoins très vite limités par la quantité et la qualité des données disponibles. Certains d'entre eux sont particulièrement intéressants, car ils prennent en compte des données socio-démographiques pour effectuer leurs prévisions (Aldrich et al., 1971). McConnel et al. ont également développé un modèle de régression prenant en compte l'âge de la population (McConnel et al., 1998).

Un autre type de modèle de prévision se base sur des séries temporelles, ce qui permet de corriger certaines faiblesses des techniques classiques de régression. Le modèle de (Baker et al., 1986) a la particularité de prendre en compte les appels urgents ainsi que les appels non urgents dans ses prévisions en utilisant un modèle de lissage exponentiel de type Holt Winters. (Channouf et al., 2007) quant à eux testent plusieurs modèles de série temporelle pour être capable de générer des appels heure par heure pour les systèmes de santé de Calgary en Alberta (Canada). Une évolution importante dans les prédictions des appels vient avec l'étude de (Setzler et al., 2009) qui pour la première fois se rendent compte que l'arrivée des appels dépend du jour de la semaine et de l'heure de la journée. Ils développent alors un réseau de neurones afin de prévoir l'arrivée des appels. Les performances de leur réseau sont comparées aux données historiques afin d'évaluer la précision de cette méthode. Une autre étude propose de prédire les appels dans le temps et dans l'espace afin d'avoir une prévision géographique d'où l'appel pourrait survenir (Vile, 2013). Pour cela l'auteur utilise une technique non paramétrique pour l'analyse des séries temporelles connue sous le nom d'analyse à simple spectre (*single spectrum analysis*), il teste ce modèle sur un service d'ambulance en Écosse et leur méthode permet d'obtenir des prévisions sur le long terme de meilleure qualité ainsi que des prévisions sur le court terme de qualité comparable aux méthodes déjà existantes. D'autres études comme celle de (Grekousis et al., 2019) utilisent une approche par réseaux de neurones pour prévoir la localisation géographique d'un appel. Ils proposent de considérer les appels d'urgence comme des ensembles de points interconnectés dans l'espace qui créent des trajectoires dans le temps. Ils combinent ensuite leur réseau de neurones à un algorithme évolutif pour prédire la localisation future de la demande en soin.

Les problèmes de planning

Les problèmes de planning se décomposent globalement en deux parties, la création du planning et l'affectation des équipes sur ce planning. Dans la littérature on retrouve ces deux problèmes respectivement sous le nom de *scheduling* et *rostering*. (Ernst et al., 2004a)(Ernst et al., 2004b) proposent une revue de littérature autour de ces problèmes en analysant les applications, les méthodes et les modèles utilisés pour résoudre ces problèmes. Dans les systèmes de santé, les problèmes de planning du personnel infirmier et médical ont été très largement étudiés (voir la revue de littérature de (Cheang et al., 2003)).

On trouve cependant un peu moins de travaux concernant les problèmes de planning directement en lien avec les EMS. Bradbeer et al. proposent d'utiliser un algorithme évolutif pour fournir un planning et des affectations acceptables en prenant en compte le nombre et le placement des effecteurs comme données et non comme paramètres

(Bradbeer et al., 2000). On trouve des études qui cherchent à résoudre à la fois le problème de dimensionnement, de localisation, et de planification comme dans les travaux effectués par (Erdoğan et al., 2010). Dans leurs travaux, ils appliquent dans un premier temps une recherche de voisinage pour résoudre le problème de localisation, puis ils utilisent les résultats dans des modèles de programmation en nombres entiers pour résoudre le problème de planification. Rajagopalan et al. (Rajagopalan et al., 2011) présentent également une approche en deux étapes pour l'affectation des plannings aux équipages. Au cours de la première étape, ils résolvent un modèle dynamique de couverture en utilisant un algorithme tabou, tandis qu'un modèle de programmation en nombres entiers est présenté au cours de la deuxième étape pour résoudre le problème d'affectation des plannings. Enfin, Li et Kozan (Li et al., 2009) proposent des modèles en deux étapes qui utilisent aussi des techniques de programmation non linéaires en nombres entiers. Dans la première étape, on détermine les plages horaires travaillées et le nombre d'équipes nécessaires sur chacune d'elles. Les résultats de la première étape sont utilisés comme entrée pour la deuxième étape, qui détermine un planning équilibré pour les effecteurs.

Néanmoins, il y a assez peu de travaux sur l'effet des plannings sur la partie amont, on note quelques études comme celle de Montassier et al. (Montassier et al., 2015) qui ont travaillé sur l'impact de la planification du travail sur les performances des centres d'appels d'urgence en France. Ils ont notamment montré que la durée passée en poste pouvait avoir un impact négatif sur le temps de traitement des appels.

Le planning des opérateurs du centre d'appels est à faire en adéquation avec le volume d'activité du centre d'appels. Il doit aussi respecter les contraintes légales qui peuvent exister dans le milieu : comme le nombre d'heures consécutives maximales autorisées ou encore des temps de pause de 10 minutes toutes les deux heures. (Moutard et al., 2019). (Sariyer, 2018) a par ailleurs travaillé sur les données historiques d'un centre d'appels d'une grande ville en Turquie. Il développe une méthode basée sur le problème du vendeur de journaux pour définir des plannings optimisés pour le centre d'appels. Cette méthode leur permet de constituer des plannings prenant en compte la variation de la demande à l'échelle de la semaine.

En France, le nombre d'ARM nécessaire est défini par le volume des appels entrants et selon un taux d'occupation maximal cible des ARM. Pour rappel, la Mission d'Expertise et d'Audit Hospitaliers (MeaH) (Hospitalier (MeaH)., 2008) préconise un taux d'occupation des ARM entre 20 et 40% de leur temps de travail. Penverne et al. conseillent un taux d'occupation des ARM de 23% (Penverne et al., 2017). La recommandation pour fixer le nombre d'ARM est qu'ils devraient décrocher en moyenne 7,2 appels par heure. En ce qui concerne les MR, leur nombre est fixé en fonction du nombre de dossiers de régulations médicales. En effet, un MR ne doit pas dépasser 6 dossiers de régulations médicales par heure (France et al., 2015).

La prévision des appels est un problème fondamental de l'étude des centres d'appels d'urgence, car ce sont ces prévisions qui serviront de base pour résoudre les problèmes de dimensionnement des ressources, de localisation des équipes d'effecteurs, de planning des équipes, etc. Une difficulté de ce domaine d'étude est qu'il faut être en mesure de prévoir les appels aussi

bien dans le temps que dans l'espace. Ainsi, il est tout aussi important de prévoir quand l'appel va arriver, que la zone géographique dont il sera issu. Aringhieri et al. soulèvent le manque de solutions technique à ce sujet dans la littérature actuelle et notent que des propositions originales sur le sujet pourraient être un véritable atout pour le pilotage en temps réel des systèmes d'urgence (Aringhieri et al., 2017).

1.3.3 Les problèmes opérationnels

Les processus de gestion des appels

Un centre d'appels peut être considéré comme un enchaînement de files d'attente et de serveurs. Les SAMU reçoivent plusieurs types d'appels, par des canaux différents. Ce qui donne lieu à un traitement de l'appel qui peut varier en fonction des caractéristiques de ce dernier. Ces spécificités se doivent d'être modélisées en matière de processus. Cela permet alors de prendre du recul sur la gestion des appels au sein de l'organisation et d'envisager des améliorations. La gestion des appels nous fait nous questionner sur le rôle des ressources et notamment celui des ARM qui décrochent l'appel. Doivent-ils être purement généralistes et être capables de traiter tous les appels ? Ou encore faut-il mieux qu'ils soient spécialisés afin de traiter plus rapidement certains types d'appels ? Enfin, est-il possible d'envisager un compromis entre spécialisation et généralisation pour faire en sorte que les appels soient traités le plus rapidement possible ?

La partie amont concernant l'organisation et le management des processus des centres d'appels présente quelques travaux intéressants. Aboueljinane travaille ainsi sur le SAMU du Val-de-Marne et analyse l'impact de l'augmentation du nombre de ressources pour le décroché des appels. (Aboueljinane, 2014) Elle montre que l'ajout de ressources n'a pas forcément un impact significatif sur les indicateurs de performance. Son analyse montre aussi qu'un gain de quelques secondes entre le décroché de l'appel et l'envoi de moyens peut avoir un effet positif sur le temps de prise en charge total des appels et sur le taux de couverture des interventions en moins de 20 minutes. Toutefois, dans ces mêmes travaux, on ne trouve pas de proposition pour atteindre cette diminution du temps de gestion des appels.

Lewis et al. s'intéressent aux problèmes de management des ressources dans un centre d'appels d'urgence (Lewis et al., 2007). Comme le coût des agents correspond à 70% des coûts d'un centre d'appels, c'est un sujet d'intérêt majeur pour les centres d'appels. Plusieurs facteurs rendent ce terrain propice à la recherche tels que la complexité mathématique et opérationnelle, les files d'attente multiples, le caractère aléatoire de l'arrivée des appels ainsi que le caractère aléatoire de la durée de ceux-ci, mais aussi le côté multicompetences des agents et la diversité grandissante des types d'appels. Les auteurs utilisent la simulation à base d'agents pour modéliser le centre d'appels et gérer les files d'attente. Le modèle bien que défini dans cet article n'est pas testé et les résultats ne sont pas présentés.

Van Buuren et al. testent l'impact de la variation du nombre d'opérateurs sur le temps d'attente avant décrochés (M. v. Buuren et al., 2015). Les auteurs ont divisé

les appels en trois catégories, forte, moyenne et basse priorité. La simulation tourne sur une journée qui reçoit 2000 appels. L'étude montre qu'un centre d'appel avec 3 ARM peut décrocher jusqu'à 1600 requêtes par jour, 2 ARM peuvent prendre jusqu'à 650 requêtes et 1 ARM peut prendre 100 requêtes. Ils notent que leur étude pourrait être améliorée en prenant en compte le lien entre temps de prise en charge de l'appel et charge de travail. En réalité, ils ont constaté que le temps de prise en charge d'un appel est plus faible lorsque la charge de travail est élevée et que cela pourrait avoir une influence sur les performances des centres d'appels. De plus, il suggère qu'il serait intéressant de comparer leur étude avec des centres d'appels où les ressources font à la fois le décroché de l'appel et l'envoi de moyen afin de comparer les performances.

Van Buuren et al. comparent trois modèles de centre d'appels avec des classes d'ARM différents (M. v. Buuren et al., 2017). Ils différencient les ressources qui prennent les appels de celles qui s'occupent d'envoyer les moyens. Ils testent trois scénarios (1) les ressources ont des tâches séparées (2) les ressources sont généralistes (3) certaines ressources sont spécialisées, d'autres sont généralistes. Ils font l'hypothèse que le temps de prise en charge est indépendant de la charge de travail. Ce qui pourrait être amélioré dans un nouveau modèle. Les auteurs, pour des recherches futures, proposent de considérer les centres d'appels d'urgence comme des centres d'appels multiressources avec certaines assignées à des appels de gravité basse et d'autres pour des appels d'urgence vitale. De plus, ils proposent de tester ce modèle sur d'autres centres d'appels qui ne sont pas dédiés à l'urgence pour voir si les résultats peuvent être généralisés.

Montassier et al. proposent enfin une étude quantitative sur les horaires de travail des ARM en France sur le SAMU de Nantes (Montassier et al., 2015). Leurs données s'étendent sur une année entière sur quatre créneaux de travail dans la journée (matin, journée, soir et nuit profonde). Ils s'intéressent à l'impact de la durée des créneaux de travail sur la durée des appels ainsi qu'à la rotation des équipes sur ces horaires. Ils observent que la durée d'appel tend à augmenter (jusqu'à 11,9% d'augmentation) durant les créneaux de nuits et les créneaux de plus de 8h. Ces créneaux de nuit tendent à augmenter la moyenne de temps de communication de 14 secondes. Le fait d'avoir des créneaux de 8h tend à augmenter le temps de communication de 6,4%. Considérant cette augmentation, les auteurs suggèrent que des améliorations pourraient être faites sur les durées de créneaux de travail afin de limiter cette augmentation du temps de traitement.

Les outils d'aides à la décision pour les problématiques cliniques

De nombreux articles dans la littérature sont issus de revues médicales et traitent l'amélioration de la phase de diagnostic, notamment en utilisant la linguistique.

Penn et al. montrent que le choix des mots au décroché de l'appel peut faire gagner jusqu'à 4 secondes sur le temps moyen de traitement de l'appel. Les premiers résultats obtenus les incitent à continuer les recherches sur l'analyse des conversations dans les centres d'appels d'urgence (Penn et al., 2017).

Riou et al. montrent qu'avec l'étude des bandes enregistrées des prises d'appels il est possible d'étudier l'impact des mots choisis sur le temps de diagnostic et la durée

de l'appel. Un autre sujet abordé concernant le diagnostic est « Comment éviter le sur-triage et le sous-triage en utilisant un diagnostic basé sur des critères ? ». Encore aujourd'hui le diagnostic effectué dépend pour beaucoup du médecin qui l'effectue. Toutefois des recherches sont menées pour extraire les critères spécifiques à certaines pathologies afin de créer des procédures de diagnostic (Riou et al., 2017) .

Dami et al. soulignent qu'il n'existe pas de métrique universelle pour évaluer la pertinence de l'envoi de moyen ce qui rend difficile l'évaluation des systèmes d'urgences préhospitalières. Ils mettent en place un système d'évaluation de l'envoi de moyens et évaluent la pertinence de ces choix sur une période d'un mois ayant donné lieu à 29000 envois primaires. Selon leurs critères les taux de sur triage (le fait de trier des appels comme urgent alors qu'ils ne le sont pas) et de sous triage (le fait de trier des appels comme non-urgent alors qu'ils le sont) sont respectivement de 78 et 4,6% (Dami et al., 2015).

Snooks et al. cherchent à identifier les critères permettant d'identifier les chutes chez les personnes âgées. Ils mènent une étude qualitative sur des données passées. Le but est de définir des codes d'envoi de moyen pour identifier rapidement les chutes et envoyer des moyens adaptés (Snooks et al., 2011).

Blomberg et al. proposent une nouvelle méthode utilisant du machine learning afin d'identifier les cas d'arrêts cardiaques en se basant sur les bandes enregistrées. Leurs résultats montrent que l'algorithme permet d'obtenir d'avoir une sensibilité supérieur à celle d'un ARM, toutefois les ARM obtiennent de meilleurs résultats en matière de spécificité concernant l'identification des arrêts cardiaques. Ce genre d'outil d'aide à la décision basé sur de l'Intelligence Artificielle pourrait jouer, très rapidement, un rôle très important dans l'amélioration des performances des centres d'appels (Blomberg et al., 2019).

Le support au diagnostic est un sujet qui fait l'objet de nombreux travaux cliniques. La plupart des publications identifiées ont été faites dans des journaux médicaux. Le constat est que si les centres d'appels ont des fonctionnements différents, les protocoles de diagnostic sont également différents. Bohm et al. dans leur étude, montre qu'il demeure toutefois une absence de niveau de preuve de l'exactitude des systèmes de triage (Bohm et al., 2018).

Les problèmes de routing et de relocalisation

Une fois qu'on a identifié le nombre de bases SMUR, qu'on a décidé où les placer, et que l'on a décidé du nombre d'équipes dans chacune des bases. Il reste encore des problèmes à résoudre pour optimiser la gestion (ou le pilotage) d'une intervention donnée. En effet, en fonction du lieu de l'intervention, il est question de déterminer la base la plus appropriée pour faire partir l'ambulance. Quels chemins doit emprunter cette ambulance ? Une fois que son intervention est terminée, est-ce qu'elle rentre directement à son entrepôt ? Est-ce qu'elle doit rentrer au même entrepôt ? Et si une nouvelle intervention intervient avant qu'elle ne rentre, est-ce que je modifie l'itinéraire de cette ambulance ? Est-ce que j'en envoie une nouvelle ? C'est sur ce genre de questions que des scientifiques tels que (Aboueljinane, 2014 ; M. v. Buuren

et al., 2018; Koch et al., 2003) travaillent. L'objectif de toutes ces questions est de réduire le temps de prise en charge des interventions ainsi que le temps pour se rendre sur le lieu de l'intervention. Toutefois un autre objectif est de réduire les coûts, car si on peut optimiser les trajets, on réduit le kilométrage parcouru et par conséquent les frais d'essence, d'entretien, etc. (Aboueljinane et al., 2014) (M. v. Buuren et al., 2015).

Lorsqu'on parle de problème de *dispatching* cela fait référence à l'affectation d'une équipe à la gestion d'un appel. Cette notion concerne essentiellement la partie aval lorsque la décision d'envoyer des effecteurs est prise. L'enjeu ici est de choisir l'équipe qui pourra arriver le plus rapidement sur place ou qui parcourra le moins de kilomètres. Plusieurs règles de déploiement ont alors été testées. Dans leur projet, (Koch et al., 2003) étudient les règles de déploiement de la Croix-Rouge autrichienne. Ils testent une première règle où toutes les équipes mobiles peuvent intervenir sur l'ensemble du territoire en utilisant une heuristique proche optimale pour l'affectation équipe-intervention. Ils utilisent la règle de l'équipe disponible la plus proche. Une deuxième règle de déploiement teste une réponse plus régionalisée où les équipes sont affectées à des zones géographiques pré-définies et ne peuvent intervenir que si l'appel se trouve dans cette zone. Les résultats montrent que la première règle plus globale donne de meilleurs résultats en termes d'attente des patients, et de kilométrage total parcouru.

Une sous catégorie du problème de localisation correspond à la relocalisation des véhicules d'urgences. Cela consiste à redéployer les véhicules non utilisés lorsque les autres équipes sont en intervention. La particularité de ce problème est sa dimension dynamique, car il faut connaître l'emplacement des équipes de secours en temps réel. De ce fait, on trouve deux grandes manières d'aborder ces problèmes : les problèmes mixtes à plusieurs périodes ou les problèmes véritablement dynamiques. Le coeur de ce sujet est de se concentrer sur la reconfiguration de la flotte de véhicules disponibles pour assurer un service optimal en tout point du territoire. Bien que ce soit un problème dynamique en temps réel, les chercheurs l'ont approché dans un premier temps en proposant des reconfigurations multi-périodiques. Cela permet de fournir aux décideurs l'emplacement des véhicules pour chaque période à l'aide d'un tableau de conformité. L'inconvénient de cette méthode est que les véhicules doivent beaucoup se déplacer afin de maintenir un niveau de conformité acceptable, ce qui peut rajouter des kilomètres parcourus sans pour autant améliorer la qualité de service pour les appels (Bandara et al., 2012). De nombreux travaux sont disponibles à ce sujet, on peut citer notamment ceux de Gendreau et al. et leur modèle multi-période pour la relocalisation des ambulances de Montréal (Gendreau et al., 2001). Leur méthode consiste à résoudre un programme d'optimisation mixte en nombres entiers en temps réel. Leurs résultats montrent que si les appels d'urgence sont très rapprochés, le temps de calcul pour le redéploiement peut être trop important ce qui compromet l'efficacité de la relocalisation. De plus, il faut que le redéploiement des équipes soit possible, car il peut y avoir des cas où la mutualisation des ressources pour les transports pré et intra hospitalier impose aux équipes de retourner à leur base après une intervention. Ce modèle a été amélioré par d'autres auteurs pour prendre en compte plus de paramètres comme l'incertitude temporelle liée au temps de trajets des ambulances (Schmid et al., 2010). Belanger et al proposent une revue

de littérature pour évaluer les différentes méthodes de redéploiement (Bélanger et al., 2016). Le constat aujourd'hui est qu'il existe beaucoup de méthodes de redéploiement, mais que nous ne sommes pas en mesure d'évaluer laquelle serait la meilleure en pratique. De plus, ce sont des problèmes qui nécessitent de bonnes capacités de calcul, les méthodes doivent donc continuer d'évoluer pour être en mesure de traiter ces problèmes dans des temps raisonnables.

Choix de l'établissement d'accueil

Le choix de l'établissement d'accueil est un problème crucial, car le système de soin d'urgence va contribuer à l'engorgement des urgences physiques au sein du système hospitalier. Ainsi il faut être en mesure de prendre en compte cet effet pour guider le patient vers l'établissement de santé qui lui apportera le meilleur soin adapté à son besoin réel et celui qui pénalisera le moins le système de santé dans son ensemble. Une règle assez commune dans la littérature scientifique internationale consiste à envoyer le patient dans l'hôpital le plus proche du lieu de l'intervention (M. v. Buuren et al., 2012 ; Fitzsimmons, 1971). Néanmoins, la proximité n'est pas le seul critère qu'il faille prendre en compte. Dans les faits les patients ne sont pas toujours transportés dans le service d'urgence le plus proche : il faut prendre en compte le besoin de soin réel du patient avant la proximité de l'établissement le plus proche (Ingolfsson et al., 2003). D'autres critères comme la disponibilité des lits, les spécialités et installations présente dans l'établissement de santé ou encore l'engorgement du service d'urgence de l'établissement cible sont parfois considérés. De plus, la prise en compte du trafic en temps réel pourrait avoir une incidence sur le choix de l'établissement d'accueil. C'est pour cela que d'autres auteurs ont fait le choix de se baser sur des répartitions empiriques pour leurs études (Ingolfsson et al., 2003). Les situations de crise de grande ampleur modifient les paramètres pour le choix de l'établissement, que ce soit lors d'attentats, de catastrophes naturelles à grande échelle (Wang et al., 2007), ou encore de pandémie comme nous avons pu le voir récemment avec des transferts inter-régionaux. Il est alors intéressant de travailler avec des listes d'établissements potentiels. Une autre étude intéressante est celle de (Miller et al., 2020) qui analyse l'effet de la fermeture des hôpitaux de proximité sur le temps de réponse, le temps d'accès au traitement, et le temps de service des systèmes d'urgences. Leur étude se base sur des données nationales récoltées entre 2010 et 2016 aux Etats-Unis issus du *National EMS Information System, Area Health Resource, and Center for Medicare & Medicaid Provider of Service files*. Leurs résultats montrent que la fermeture des hôpitaux locaux augmente le temps moyen de transport de 2,6 minutes et le temps de réponse moyen augmente de 7,2 minutes. Cela montre à quel point le choix de l'établissement d'accueil peut avoir une incidence sur le temps de prise en charge du patient.

La gestion opérationnelle de l'appel au niveau du centre d'appels est un sujet qui est assez peu étudié dans la littérature. Les principaux défis à relever concernent la gestion des ressources humaines du centre d'appels. L'influence des horaires de travail ainsi que de la segmentation (spécialisation) ou la polyvalence des ressources sont des sujets qui doivent encore être investigués.

L'étape de diagnostic de l'appel pourrait être aussi améliorée en travaillant sur des procédures de diagnostic téléphonique plus ciblées ou encore en développant des outils d'aide à la décision à destination du personnel médical.

Concernant la partie aval, ce sont les problèmes de déploiement et de relocalisation en temps réel qui présentent un intérêt. Les difficultés résident dans la prise en compte des facteurs extérieurs tels que le trafic ou la météo ainsi que la gestion dynamique en temps réel des véhicules.

Enfin, le choix de l'établissement d'accueil pour le patient nécessite aussi de prendre en compte les critères extérieurs qui peuvent rallonger le trajet entre le patient et l'établissement de santé. Une fois encore, il serait intéressant de développer des outils d'aide à la décision basés sur différents critères permettant de prendre en compte l'état et l'engorgement des établissements de santé ainsi que le choix du patient.

1.3.4 Synthèse : Un déséquilibre des études entre la partie amont et aval.

Le dénominateur commun de toutes ces études est d'améliorer les performances du système de soin préhospitalier. Cela consiste en une double approche d'une part qualitative avec une meilleure analyse du besoin de soin et des décisions adaptées d'autre part quantitative par la gestion des délais. Les délais tentent d'être réduits en faisant en sorte que les solutions proposées réduisent le temps entre le moment où un appel d'urgence est lancé et la prise en charge du patient dans un établissement de santé. On constate un nombre significatif d'études qui travaillent sur la réduction des temps de trajets dans les interventions.

Tableau 1.1 – Synthèse des défis à relever

	Type de problème	Challenge
Stratégique	Dimensionnement	Définir le nombre de ressources nécessaires
	Localisation	Placer les équipes de manière à optimiser le maillage
	Gestion des nouveaux supports	Comment intégrer les nouvelles technologies ?
Tactique	Prévisions des appels	Prévoir l'arrivée temporelle et spatiale des appels
	Planning	Planifier les équipes de manière à répondre à la demande
Opérationnel	Gestion des appels	Proposer des processus de gestion pour minimiser l'attente et maximiser le temps d'occupation
	Aide à la décision (clinique)	Développer des outils pour aider le diagnostic des praticiens
	Déploiement et relocalisation	Gérer les équipes d'effecteurs en temps réel pour diminuer les temps de prise en charge
	Choix établissement accueil	Choisir l'établissement d'accueil en prenant en compte les contraintes extérieures et le choix du patient

Un premier constat est que la littérature est riche d'une grande quantité de travaux sur la partie amont pour la résolution de problèmes d'affectation, de tournées de véhicules, de relocalisation. La résolution de ces problèmes permet de réduire le temps d'intervention des équipes, ainsi que le nombre de kilomètres parcourus. Ces problèmes sont principalement étudiés depuis la fin des années 1980 (Fujiwara et al., 1987). En comparaison, les problèmes concernant la partie aval est plus récente avec une volumétrie plus modeste. Pour illustrer ce propos, nous avons effectué deux requêtes sur la plateforme de bases de données bibliographiques et bibliométriques Web of Science⁹. Nous avons fait un tri sur les thématiques de Gestion, d'Informatique

9. <https://www.webofknowledge.com>

et de Génie industriel. Nous avons analysé le nombre de publications par années sur les dernières décennies. Les figures 1.13 et 1.14 montrent les résultats de cette requête. On constate que les sujets concernant la partie aval sont plus nombreux (coefficient de 1 pour 10) et qu'il y a une augmentation des études avec le temps. Du côté des centres d'appels, les articles sont moins nombreux et plus récents avec des publications surtout après les années 2000. Bien que ceci ne constitue pas une revue de littérature systématique au sens strict du terme, cela illustre bien l'écart qui existe entre les deux parties de la chaîne de l'urgence.

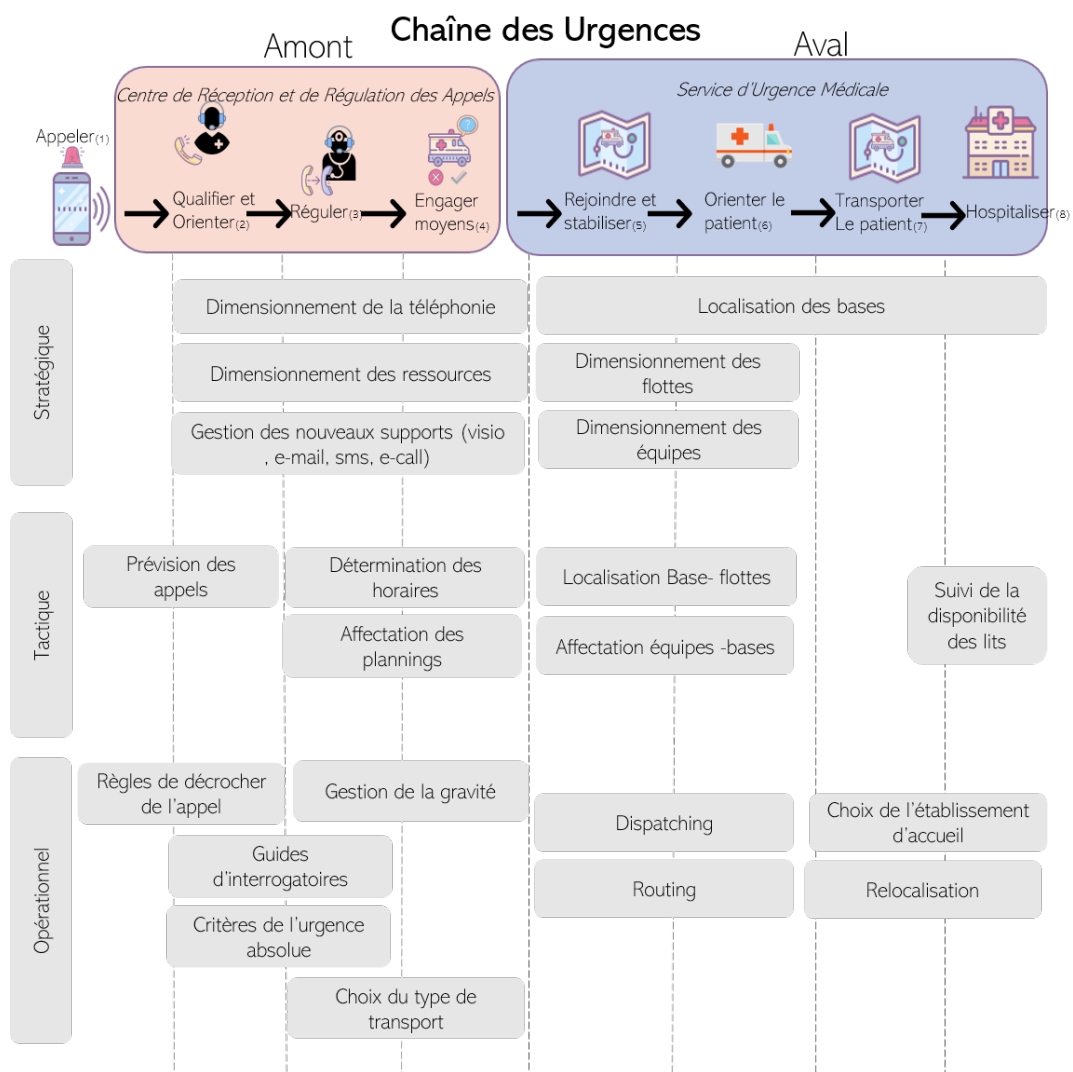


Figure 1.12 – Positionnement des types de sujets en fonction de l'activité de la chaîne des urgences

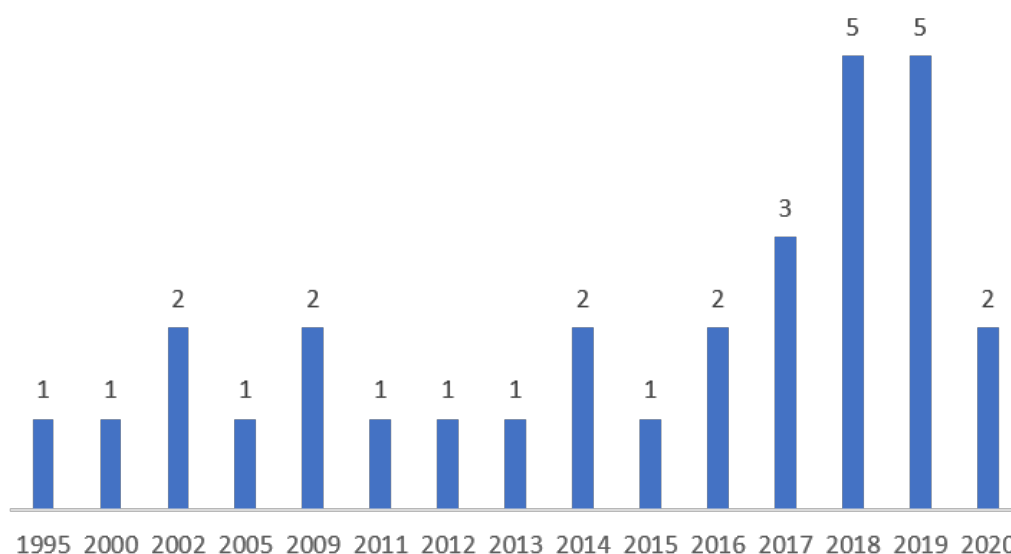


Figure 1.13 – Résultats de recherche du 02/06/2020 sur Web of Science avec la requête :(("Emergency Medical Service*" OR EMS) AND((Call OR Contact OR Communication) AND Center))

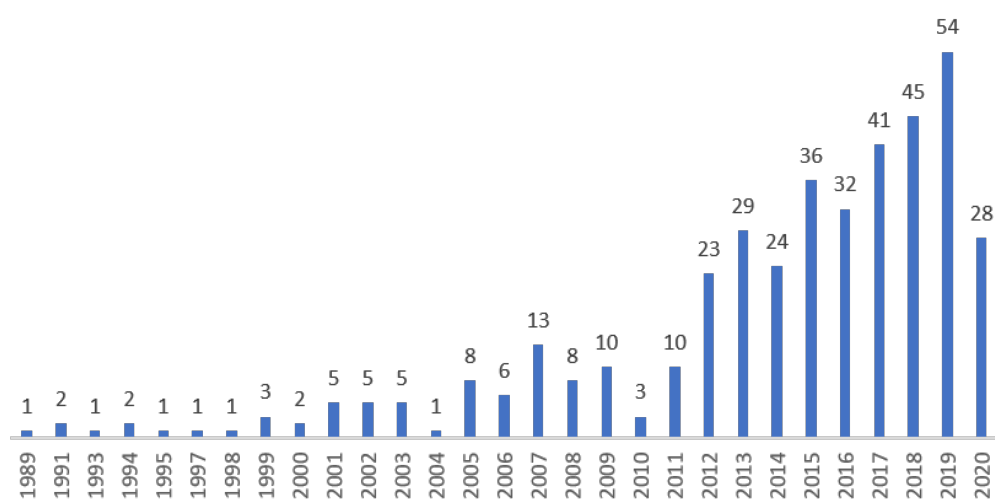


Figure 1.14 – Résultats de recherche du 02/06/2020 sur Web of Science avec la requête : "Emergency Medical Service*" OR EMS) AND (Ambulance* OR transport*)

2

Positionnement et questions de recherche

Ce chapitre a pour objectif de formuler et justifier la problématique de recherche adressée dans ce manuscrit. À la lumière de l'étude bibliographique et de la description des activités métiers effectuées précédemment, nous précisons ici le positionnement de notre sujet ainsi que les questions de recherches abordées.

2.1 Un besoin métier à traiter

L'objectif du rapport de la Mission de Modernisation de l'Accessibilité et de la Réception des Communications d'Urgence pour la Sécurité, la Santé et les Secours (MARCUS) est de faire une analyse qualitative et quantitative des centres d'appels d'urgences français (Moutard et al., 2019). À la lumière de ces analyses et observations, il s'agit de proposer des scénarios de réforme de l'organisation de la prise en charge des communications d'urgences. 35 recommandations ont été faites pour améliorer l'état des services d'urgences français (ces recommandations sont disponibles en annexe A). Nous proposons d'utiliser ce rapport comme base pour analyser les recommandations que nous traitons ou sur lesquelles nous estimons avoir une incidence avec nos travaux.

Les recommandations de ce rapport sont classées selon les sept catégories suivantes :

1. **L'accessibilité des centres d'appels** : le constat est que la volumétrie des appels d'urgence est encore difficile à objectiver aujourd'hui. Il est nécessaire d'harmoniser le vocabulaire ainsi que les méthodes de récupération des données. Une harmonisation à ce niveau permettra d'avoir une vision et une évaluation précise de l'accessibilité des centres d'appels.
2. **L'amélioration de l'activité opérationnelle** : Face à l'hétérogénéité des SAMU, il est nécessaire de revoir la manière dont la prise en charge des appels se déroule. Il faut définir un cadre de travail qui permette un meilleur suivi, une meilleure évaluation de l'activité afin d'être en mesure de proposer des solutions adaptées à la prévision de l'activité opérationnelle.

3. **La formation des opérateurs des centres d'appels** : Aujourd'hui, la formation des acteurs des centres d'appels d'urgence est insuffisante et hétérogène. Il est nécessaire de repenser la formation et de définir des objectifs pédagogiques clairs afin que le personnel des centres d'appels puissent bénéficier d'une formation de qualité, leur permettant de répondre aux besoins métiers.
4. **La mise en place d'une démarche qualité** : Il y a un besoin de structurer l'organisation du travail afin d'améliorer la qualité du service rendu aux patients mais aussi pour améliorer la qualité de vie au travail des agents.
5. **L'interopérabilité des services d'urgence** : Afin d'améliorer la gestion des appels d'urgence, il est crucial de mettre en place des outils et des systèmes d'information qui puissent faciliter et fluidifier la communication entre les acteurs de la chaîne des urgences.
6. **L'éducation de la population à l'utilisation des numéros d'urgence** : Il y a un besoin d'harmoniser et de repenser la communication autour des numéros d'appels d'urgence, qui en France sont nombreux, afin d'ancrer durablement le bon usage des numéros d'urgence auprès de la population.
7. **Envisager de nouveaux modèles organisationnels** : Afin d'améliorer le traitement des communications d'urgence, il est nécessaire de revoir la manière dont les appels sont traités. Il faut aussi revoir la manière dont les territoires collaborent pour gérer les communications d'urgence.

Dans chaque catégorie on retrouve plusieurs points d'intérêts et recommandations qui devront être traités dans les années à venir. Nos travaux visent principalement à contribuer aux points 2 et 7 de cette liste. Les catégories concernant l'accessibilité, l'amélioration de l'activité opérationnelle, l'interopérabilité des services d'urgence ainsi que la réflexion sur les nouveaux modèles organisationnels nous ont aussi interpellés. Parmi ces recommandations certaines ont particulièrement attiré notre attention :

- Diffuser et généraliser l'emploi d'une sémantique homogène du traitement de l'appel d'urgence entre les métiers, afin d'harmoniser la description opérationnelle.
- Fixer un seuil d'abandon des appels à 10 secondes.
- Définir des indicateurs clefs de performance de l'activité opérationnelle.
- Historiser les données en perspective de la réalisation de modélisations d'activité.
- Harmoniser les axes organisationnels principaux des centres d'appels d'urgence.
- Adapter le dimensionnement des centres d'appels d'urgence pour optimiser le traitement des appels et répondre aux objectifs de performance établis.
- Renforcer les stratégies d'articulation et d'entraide entre les centres d'appels d'urgence pour gérer les pics d'appels.
- Faire évoluer les modalités de distribution des flux téléphoniques grâce à la virtualisation pour permettre l'interconnexion des centres d'appels d'urgence.
- Structurer les processus et les outils de pilotage opérationnel des centres d'appels d'urgence afin de suivre l'activité en temps réel aux niveaux local et national.

- Encourager le développement et l'utilisation des nouvelles technologies afin d'améliorer le traitement des appels d'urgence.
- Associer un flux urgent et un flux non urgent répondant à des demandes spécifiques.
- Mettre en place un premier niveau de décroché assurant les fonctions de filtrage et/ou d'orientation des appels d'urgence.
- Confirmer la nécessité de prioriser les appels d'urgence lors d'un triage rapide de premier niveau (appelé « N1 »), selon un référentiel co-construit par les différents services métiers.
- Placer les fonctions de filtrage et/ou d'orientation de l'appel d'urgence (N1) à un niveau supra-départemental couvrant un bassin populationnel de 3 à 6 millions d'habitants et recevant entre 10 à 18 000 appels/jour (hors territoires ultramarins).

Considérant ces différentes recommandations adressées à l'échelle nationale, il y a un besoin métier assez net qui se dessine :

- Il est nécessaire d'objectiver et de cadrer la manière de fonctionner des centres d'appels.
- Il est nécessaire d'évaluer la performance des centres d'appels qui se basent sur des données collectées sur le terrain et ceci dans le but d'optimiser l'activité.
- On note enfin une volonté forte des SAMU de se moderniser et d'utiliser tout le potentiel que les nouveaux outils numériques offrent pour améliorer la gestion des centres d'appels.

Il y a en France, une diversité avérée des fonctionnements des centres d'appels. Les outils de gestion du système informatique varient, la gestion des horaires des agents dépend de chaque centre d'appels, etc. Il est difficile de trouver des recommandations officielles pour le management et l'organisation interne des centres d'appels. On trouve aussi en France une grande diversité démographique des territoires. Nos observations régionales menées sur les trois SAMU des départements de la Haute-Garonne, du Tarn et de l'Aveyron en sont une parfaite illustration. En effet, l'Aveyron est un département majoritairement rural avec une faible densité de population. Son terrain montagneux rend certaines zones difficiles d'accès. Le Tarn est une zone semi-rurale avec une densité de population concentrée autour de la région d'Albi et de Castres. La Haute-Garonne est un territoire très étendu du nord au sud avec la majorité de sa population concentrée dans la zone toulousaine qui est très urbaine. Les offres de soins entre les différents territoires sont également très différentes en termes de volumétrie des appels reçus et de l'hétérogénéité des organisations. Ainsi, il y a sans doute un lien entre les caractéristiques d'un département, son offre de soins et la manière dont son SAMU est géré.

Ce que nous pouvons retenir de nos observations sur trois SAMU régionaux est que ce sont des systèmes organisationnels très hétérogènes qui présentent, tous, des marges de progrès en matière de performance opérationnelle. Ce

constat a été également confirmé par les gestionnaires des SAMU que nous avons pu interroger pour ces travaux de recherche.

Nous l'avons vu, le volume d'activité des centres d'appels d'urgence en régime standard augmente et va continuer à augmenter (Penverne et al., 2009). Lors du confinement engendré par la pandémie de Covid-19, les téléconsultations se sont largement généralisées, et il ne serait pas surprenant que les pratiques évoluent dans ce sens à l'avenir.

La demande de soins et le nombre d'appels devant être traités par les centres d'appels d'urgences augmentent tous les ans et vont continuer durant les prochaines années. Les SAMU ont donc besoin de s'organiser pour absorber dans de bonnes conditions cette augmentation.

La crise sanitaire de 2009 à propos de la grippe H1N1 a révélé l'importance des centres d'appels dans la gestion d'une crise. Cela a conduit à la création du projet SI-SAMU afin d'unifier les outils utilisés et de standardiser les processus de gestion des appels au sein du territoire. Aujourd'hui, ce besoin est encore plus fort alors que nous traversons la pandémie de COVID-19. Certains départements déclarent avoir vu une augmentation du nombre d'appels de 100%¹. Des cellules de crise ont été ouvertes pour gérer cet afflux exceptionnel d'appels. Du personnel médical s'est spontanément présenté afin d'aider pour le décroché des appels. Dans ce contexte, le besoin d'améliorer la performance du fonctionnement du SAMU est encore plus présent. Pour ce faire, il est nécessaire de savoir comment les centres d'appels d'urgences peuvent réagir à la crise. Quelles sont les actions qui peuvent être mises en place rapidement afin de traiter une augmentation aussi soudaine de la charge de travail ? Ces informations vont permettre de se préparer à ce genre d'événements, en créant, par exemple, des procédures de gestion de crise.

Les périodes de crises sanitaires mettent en avant les difficultés du système de santé ainsi que leurs fragilités. Il y a donc un besoin criant de disposer d'outils de diagnostic organisationnel et de préconisations organisationnelles afin d'améliorer la performance future des SAMU en période de crise ainsi qu'en période normale.

2.2 Les SAMU : un objet d'étude très hétérogène à l'échelle nationale

Afin d'illustrer et d'objectiver l'hétérogénéité des SAMU, nous avons collecté des données sur les SAMU à l'échelle nationale. Au travers de cette étude, nous faisons l'hypothèse que les SAMU peuvent être regroupés suivant des critères structurels.

1. https://www.francetvinfo.fr/sante/maladie/coronavirus/samu-le-nombre-d-appels-explose_3864987.html

Pour ce travail, nous n'avons pas inclus les SAMU des DOM-TOM français par manque de données les concernant. Ainsi notre étude ne sera valable que pour les départements de métropole ainsi que la Corse. Dans cette section, nous explorons les données de ces 93 SAMU afin de proposer une classification.

2.2.1 Les données brutes agrégées sur l'année 2018

Pour faire cette analyse, nous nous appuyons sur les données nationales pour chaque département français. La figure 2.1 présente les données utilisées classées en trois catégories différentes. Les données démographiques des départements sont issues du site de l'INSEE, les données sur les informations structurelles et les performances des SAMU proviennent de la SAE.² Il est important de noter que la SAE est une base de donnée déclarative réalisée à partir de la remontée d'informations fournie par les SAMU issu de systèmes d'informations non harmonisés et non centralisés. Nous détaillons ci-dessous la liste des données et leur signification. Ces critères peuvent être répartis suivant plusieurs catégories : des critères démographiques qui traitent des caractéristiques géographiques du département, des critères structurels sur la composition du SAMU et comment il est équipé, cela correspond aux moyens de l'urgence, et des critères sur les caractéristiques de la demande de soin urgent.

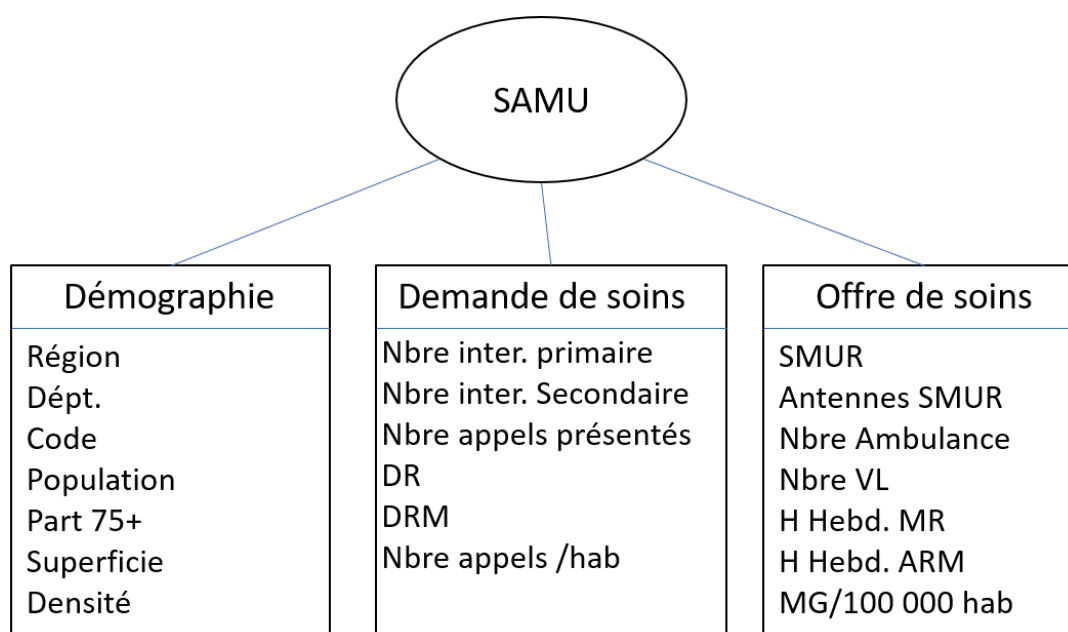


Figure 2.1 – Présentation des données nationales

Les critères démographiques

Les critères démographiques permettent de renseigner sur les grandeurs du périmètre géographique d'interventions ainsi que sur les caractéristiques de la population.

2. <https://www.sae-diffusion.sante.gouv.fr/sae-diffusion/recherche.htm>

- *Région* : Ce champ correspond à la région administrative à laquelle le département est relié, cet item ne nous sert pas pour la classification à proprement parlé, mais nous permet de situer le département à l'échelle régionale.
- *Dept.* : Ce champ correspond au nom du département.
- *Code* : Ce champ correspond au code de numérotation du département.
- *Population* : Ce champ renseigne le nombre d'habitants dans le département en 2018.
- *75ans et +* : Ce champ renseigne le pourcentage d'habitants de plus de 75 ans sur le territoire.
- *Superficie* : Superficie du territoire exprimé en km^2 .
- *Densité* : Densité de la population du territoire en habitants / km^2 .

Les critères structurels sur la composition du SAMU : L'offre de soins

L'offre de soin permet de caractériser les moyens disponibles au niveau du département pour répondre aux situations d'urgences. Nous qualifions l'offre de soin par les ressources existantes et non pas les ressources engagées dans les interventions

- SMUR général correspond au nombre de bases SMUR sur le département.
- Antenne SMUR correspond à des bases SMUR ouvertes de manière ponctuelle (selon l'affluence touristique saisonnière par exemple).
- Ambulance : le nombre d'UMR à disposition du SAMU.
- Véhicule léger : le nombre de VL à disposition du SAMU.
- Heures moyennes hebdomadaires des MR.
- Heures moyennes hebdomadaires des ARM.
- Moyens aériens : le nombre de moyens aériens de type hélicoptère.
- Moyens maritimes : Pour les SAMU côtiers qui disposent d'un accès maritime.
- Le nombre de Médecins Généralistes (MG).
- Le nombre de MG pour 100 000 habitants.

Les caractéristiques de la réponse de soin d'urgence : La demande de soins

Cette catégorie permet de renseigner les caractéristiques du besoin de la population en terme de services d'urgence.

- Sorties terrestres primaires.
- Sorties terrestres secondaires.
- Sorties terrestres dans le cadre de transports infirmiers inter- hospitalier.
- Appels présentés.
- Dossiers de Régulation.
- Dossiers de Régulation Médicale.
- Le nombre d'appels par habitant.

Le territoire Français est varié, on y trouve des zones très urbanisées ainsi que des territoires plus ruraux. Le relief est tout aussi varié avec des zones montagneuses, des littoraux, des zones qui combinent les deux, des plaines et des vallées. Notre territoire est riche et d'une grande diversité. Cette diversité pourrait avoir un impact sur la

manière dont on consomme et dispense le soin d'urgence. Ainsi, nous souhaitons vérifier s'il existe des types de SAMU qui dépendent de cette diversité. La figure 2.2 présente un extrait de cette base de données ainsi constituée et des données présentées précédemment que nous avons utilisées.

Une première manière de découvrir ces données est d'utiliser une approche de type statistique descriptive afin de voir ce que l'on peut obtenir pour chacune des catégories. Le tableau 2.1 présente les statistiques obtenues sur ces critères tels que la moyenne, la déviation standard, les valeurs minimums et maximums ainsi que les différents quartiles.

2.2.2 Définir des groupes de SAMU sur la base des données nationales

Application du clustering hiérarchique agglomérant

Afin d'aller plus loin dans la découverte de nos données que l'étude des statistiques descriptives, nous nous sommes penchés sur l'apprentissage automatique, plus connu sous le nom de *Machine Learning* (Mitchell, 1999). Le Machine Learning peut se définir comme un champ d'études de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'« apprendre » à partir de données. Avec cette technique on distingue deux approches principales : l'apprentissage supervisé et l'apprentissage non supervisé (Kaufman et al., 2009). Dans chacune des catégories, on retrouve plusieurs outils, plusieurs algorithmes permettant de traiter des données. L'apprentissage supervisé est intéressant lorsqu'on connaît déjà les catégories formées par nos données. L'apprentissage non supervisé est utilisé pour des données qui ne sont pas classées suivant des catégories. De ce fait, on ne dispose que des individus formés par les données et ces derniers ne sont pas regroupés suivant leur similitude. Les algorithmes d'apprentissage non supervisés ont pour objectif de former des groupes homogènes d'individus selon leurs attributs. Une fonction de distance sera utilisée pour calculer la distance entre des paires d'exemples. Une fois ces groupes formés, c'est à l'humain de donner du sens au résultat obtenu et d'identifier les motifs récurrents dans les groupes, ce qui les rassemble et ce qui les différencie.

Dans notre cas, l'apprentissage non supervisé semble tout à fait adapté puisque nous avons plusieurs individus (les SAMU) avec des caractéristiques, et nous cherchons à les regrouper suivant des groupes homogènes pour voir quels sont les critères qui permettent de les rassembler. On parle alors de méthode de regroupement (*clustering*) (Kaufman et al., 2009). Plusieurs algorithmes sont courants au sein du clustering. Lorsqu'on connaît le nombre de groupes, on peut utiliser l'algorithme du *k-mean*. Or, ici nous ne savons pas a priori combien de groupes de SAMU nous allons obtenir, ainsi nous avons plutôt opté pour *une approche hiérarchique* qui a l'avantage de laisser le choix sur le nombre de groupes à former.

Nous avons utilisé un algorithme de clustering hiérarchique agglomérant (*hierarchical agglomerative clustering*) (Müllner, 2011). Les systèmes de clustering agglomératif

Région	SAMU	Code	SMUR général	Antenne SMUR	Population	Part des 75ans et plus	Superficie	Densité	Ambulance	Véhicule légers	Moyens aériens	Sorties terrestres primaires	Sorties terrestres secondaires	Appels présentés	DR	DRM	MR Heure hebdo moy	ARM Heure hebdo moy	MG	Appels/hab	MG pour 100000 hab
Auvergne-Rhône-Alpes	Ain	1	4	0	631 877	8,0	5 762	109,66	4	4	0	3 230	407	254 017	141 331	136 986	404	622	388	0,40	61,09
Hauts-de-France	Aisne	2	6	1	538 659	9,0	7 369	73,10	6	8	1	4 908	976	330 637	98 734	90 219	248	360	351	0,61	65,16
Auvergne-Rhône-Alpes	Allier	3	3	0	341 613	13,0	7 340	46,54	5	6	0	3 940	902	234 242	125 574	97 346	272	660	243	0,69	71,13
Provence-Côte d'Azur	Alpes-de-Haute-Provence	4	3	0	161 799	12,5	6 925	23,36	1	5	0	1 140	198	77 626	40 730	31 821	283	188	159	0,48	98,27
Provence-Alpes-Côte d'Azur	Hautes-Alpes	5	2	0	140 916	11,1	5 549	25,39	0	3	1	1 038	344	107 210	82 172	34 746	276	252	157	0,76	111,41
Provence-Alpes-Côte d'Azur	Alpes-Maritimes	6	5	2	1 082 440	12,2	4 299	251,79	6	7	1	7 452	3 600	392 776	246 269	201 352	603	1 008	1 144	0,36	105,69
Auvergne-Rhône-Alpes	Ardèche	7	3	0	324 209	11,3	5 529	58,64	0	1	0	1 814	105	127 428	92 450	76 241	848	304	248	0,39	76,49
Grand Est	Ardennes	8	4	1	277 752	9,8	5 229	53,12	5	5	0	4 023	611	111 250	71 482	52 272	168	266	229	0,40	82,45
Occitanie	Ariège	9	2	0	152 499	12,5	4 890	31,19	2	3	0	1 350	561	134 950	35 112	25 911	195	216	142	0,88	93,12
Grand Est	Aube	10	2	1	309 056	10,1	6 004	51,48	2	4	0	2 449	468	141 208	97 490	59 303	216	371	206	0,46	66,65
Occitanie	Aude	11	2	2	366 957	11,9	6 139	59,77	2	5	2	3 791	670	118 624	92 932	71 714	225	552	322	0,32	87,75
Occitanie	Aveyron	12	5	0	279 169	13,7	8 735	31,96	3	5	2	2 243	249	103 148	54 065	50 429	268	302	221	0,37	79,16

Figure 2.2 – Extrait de la base de données

Tableau 2.1 – Statistiques descriptives sur les départements pour chacun des critères

	count	mean	std	min	25%	50%	75%	max
Code	96	47,33	28,07	1,00	23,75	47,50	71,25	95,00
SMURgeneral	96	3,84	1,70	1,00	2,75	4,00	5,00	8,00
Population	96	669800,22	511265,55	76309,00	302609,50	537809,00	843113,00	2605238,00
Partdes75ansetplus	96	10,42	2,12	5,12	8,95	10,52	11,99	14,99
Superficie	96	5673,85	1942,30	105,00	5147,00	5947,00	6775,00	10725,00
Densite	96	564,51	2458,75	14,77	50,15	82,66	161,14	21014,17
Ambulance	96	4,97	3,80	0,00	2,00	4,00	6,25	22,00
Vehiculelegers	96	6,98	3,81	1,00	4,75	6,00	9,00	25,00
Moyensaeriens	96	0,74	0,91	0,00	0,00	1,00	1,00	5,00
Moyensmaritimes	96	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Sortiesterrestresprimaires	96	5757,64	4585,81	723,00	2749,75	4167,50	7712,50	27080,00
Sortiesterrestressecondaires	96	1450,45	1243,04	55,00	621,00	978,50	1939,50	6996,00
Appelspresentes	96	296664,42	219427,66	0,00	140041,75	250273,50	402146,75	1141250,00
DR	96	159537,61	111394,50	0,00	91209,00	128381,50	203449,75	540796,00
DRM	96	124066,74	78411,49	0,00	70196,25	110923,50	170224,25	343051,00
MRHeurehebdomoy	96	378,81	240,88	0,00	219,00	294,50	475,00	1347,00
ARMHeurehebdomoy	96	625,22	377,16	0,00	360,00	529,00	848,75	1680,00
Plateformecommune	96	0,19	0,39	0,00	0,00	0,00	0,00	1,00
MG	96	536,25	441,04	58,00	215,00	400,50	700,75	2315,00
Appels/hab	96	0,50	0,45	0,00	0,38	0,42	0,50	4,22
MGpour100000hab	96	79,75	13,35	53,75	68,58	80,50	87,10	111,41

partent de la partition de l'ensemble de données en nœuds simples. Puis ces données sont fusionnées avec les données les plus proches les unes des autres afin de former un nouveau nœud. Le processus est répété jusqu'à ce qu'il n'y ait qu'un seul nœud qui comprend l'ensemble des observations.

La représentation en sortie de l'algorithme est un dendrogramme par étape dont voici la définition (Müllner, 2011) :

Définition. Soit un ensemble fini S_0 de cardinalité $N = \|S_0\|$, un dendrogramme par étape est une liste de $N - 1$ triplet (a_i, b_i, δ_i) ($i = 0, \dots, N - 2$) tel que $\delta_i \in [0, \infty)$ et $a_i, b_i \in S_i$, où $S_i + 1$ est récursivement défini comme $S_i \setminus a_i, b_i, \cup n_i$ et $n_i \notin S \setminus a_i, b_i$, un repère pour un nouveau nœud.

Détaillons cette définition, l'ensemble S_0 correspond à l'ensemble des points étudiés (dans notre cas, les 96 départements français métropolitains). À chaque étape, n_i correspondant au nouveau nœud formé par le regroupement des nœuds a_i et b_i espacé de la distance δ_i . Cela revient à dire que deux départements sont proches, car la distance entre leurs vecteurs de composantes est proche. L'algorithme contient $N - 1$ étapes, car lors de la dernière étape on obtient un nœud qui regroupe les N nœuds initiaux (les 96 départements).

Pour expliquer le fonctionnement du clustering hiérarchique, nous proposons un exemple abstrait présenté dans la figure 2.3. Afin d'être capable de représenter la distance graphiquement entre les individus, cet exemple n'utilise que deux critères. Nous avons 8 individus définis par ces deux critères, ce qui nous permet de faire une représentation graphique comme sur la partie gauche de la figure. La première étape de l'algorithme considère chaque individu comme un groupe. La deuxième étape va chercher le plus proche voisin de chaque point, de manière à regrouper les points deux à deux en minimisant leur distance euclidienne. On obtient alors trois doublets (4, 5), (1, 2), (6, 8) et deux monoïdes (3) et (7). La troisième étape va essayer de rassembler ces groupes deux à deux, toujours en cherchant à minimiser leur distance. On obtient alors 3 groupes qui rassemblent les doublets précédents avec le monoïde le plus proche [7, (6, 8), (3, (4, 5), (1, 2))]. L'étape suivante rassemble ces trois groupes de manière à obtenir deux groupes formés du doublet (1, 2) d'un côté et [(7, (6, 8)), (3, (4, 5))] de l'autre. Enfin la dernière étape forme un seul groupe contenant tous les individus. Si on présente les résultats sous la forme d'un dendrogramme comme sur la partie droite de la figure 2.3 on obtient un arbre avec quatre niveaux. Chaque niveau présente un nombre de groupes différents qui contient plus ou moins d'individus. C'est donc au décideur de choisir le nombre de groupes le plus pertinent dans son cas d'étude en fonction des critères étudiés.

Le dendrogramme de représentation des SAMU met en évidence des clusters sur plusieurs niveaux. Ainsi c'est à la personne qui lit le graphique de décider où couper dans l'arbre afin de former ces groupes. Les couleurs correspondent à la valeur de la composante. Toutes les composantes ont été normalisées en réduisant la valeur par la moyenne et en la divisant par la déviation standard afin de s'assurer, pour chaque composante, une moyenne de 0 et une variance de 1.

Le dendrogramme ainsi obtenu est présenté dans la figure 2.5. Une première observation est le fait que Paris n'appartient à aucun groupe, il forme un singleton. Sa

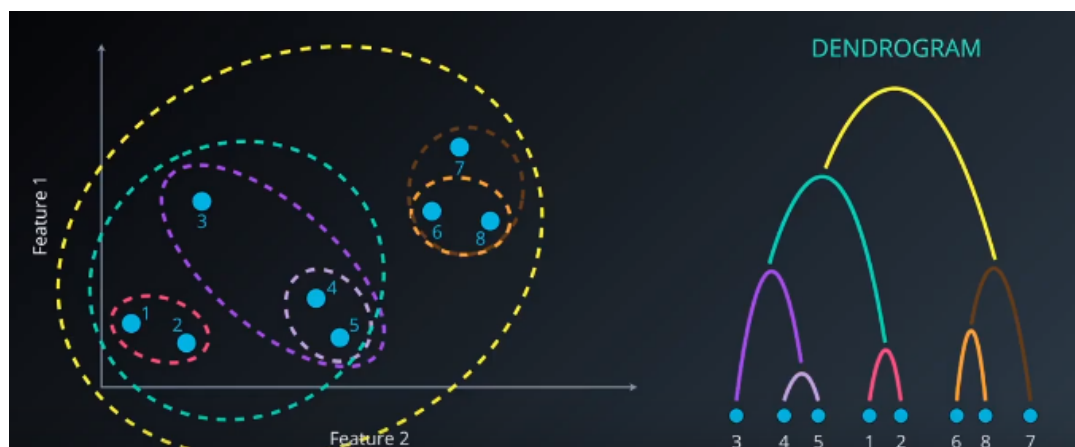


Figure 2.3 – Exemple de clustering hiérarchique sur deux critères avec une distance simple.

Clustering proposé selon la démographie, l'offre et la demande de soin des départements français

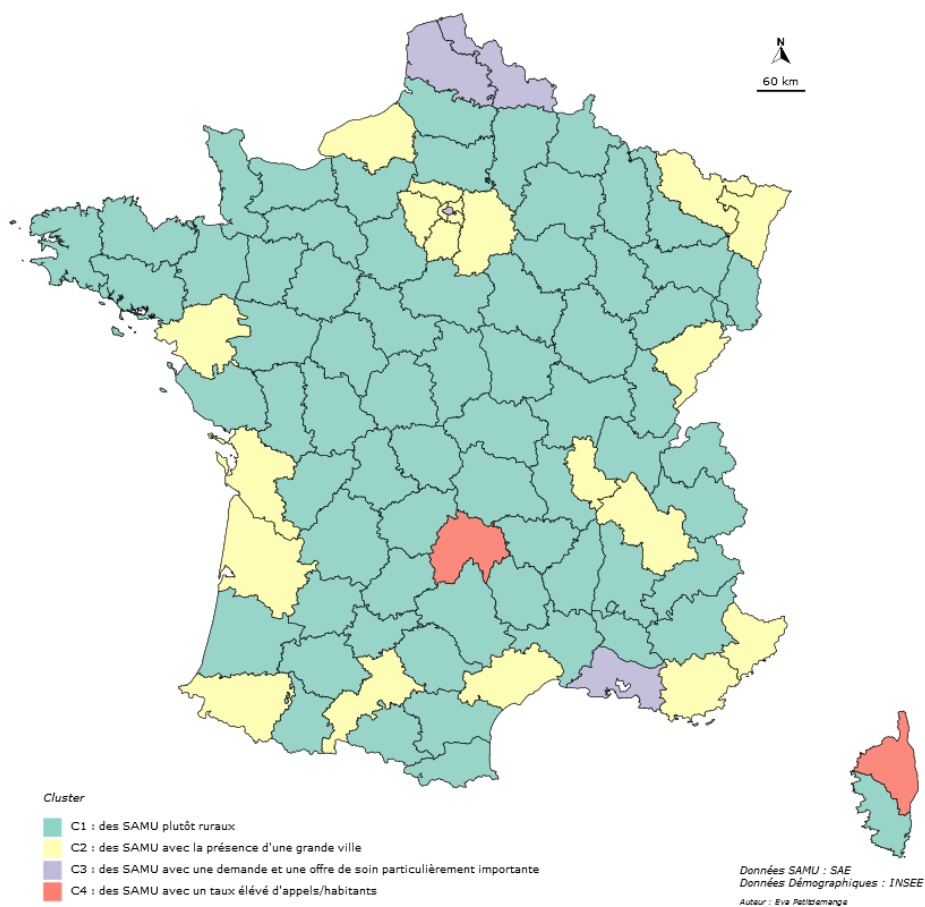


Figure 2.4 – Clustering proposé des départements français

caractéristique principale, qui le différencie des autres, est sa forte densité qui dénote beaucoup avec les autres départements (21014 habitants par km², pour une densité médiane de 89 habitants par km² à l'échelle nationale). Un autre cas à la marge est la Corse-du-Sud dont le nombre d'appels par habitant est bien plus élevé que sur le reste du territoire. On est à 4,22 appels par habitant, alors que la moyenne nationale est de 0,53 appel par habitant.

La Corse-du-Sud correspond aussi à un cas particulier dû à son nombre d'appels par habitant qui sort de l'ordinaire. Une piste d'explication serait la saisonnalité touristique. La Corse voit sa population augmenter fortement durant la période estivale. Ce qui doit avoir un impact non négligeable sur le nombre d'appels d'urgence. En 2015 l'INSEE³ mesure une augmentation de la population en Corse pouvant aller jusqu'à 4 fois sa population habituelle. Il serait alors intéressant de rajouter cette variation touristique de la population dans les données. Malheureusement nous n'avons pas trouvé de source nous permettant de récupérer ces données pour tous les départements.

Ensuite on distingue un groupe contenant les départements du Nord, des Bouches-du-Rhône et du Pas-de-Calais. Le tableau 2.2 regroupe les caractéristiques de ces départements par rapport à la moyenne nationale. Ce groupe se caractérise par une quantité plus élevée d'ambulances (de 2 à 4 fois supérieur à la moyenne nationale), de véhicules légers (VL) et de ressources humaines. On a donc une offre de soins qui est plus importante que pour les autres départements. Toutefois, cela est à mettre en regard avec la demande de soins. On constate aussi un volume d'appels présentés, de dossiers de régulation médicale (DRM) et de sorties terrestres primaires plus élevés, ce qui explique en partie les moyens plus importants. On peut se demander d'où vient cette différence. Pour les Bouches-du-Rhône comme pour la Corse précédemment, cela pourrait être dû à la saisonnalité touristique et à l'augmentation de population qu'il peut y avoir durant les congés estivaux. Mais cela n'est qu'une hypothèse qu'il faudrait vérifier avec des données plus précises à une échelle moins macroscopique que celle de l'année. Pour les départements du Nord et du Pas-de-Calais, on constate un nombre d'appels par habitant légèrement plus important pour le Pas-de-Calais, 0,60 appel par habitant pour une moyenne nationale à 0,50. Il serait intéressant de regarder l'évolution du nombre d'interventions primaires par rapport au nombre de ressources disponibles. Cela pourrait être juste dû au fait que si un SAMU dispose de plus de matériel, de plus d'équipes de soins, il est plus facile d'envoyer des moyens auprès du patient, alors que les SAMU disposant de moins de moyens envoient moins facilement des moyens internes et font plus appel aux ambulances privées ainsi qu'aux pompiers. Pour tester cette hypothèse, nous avons tracé les graphiques pour les nombres de sorties primaires et secondaires en fonction du nombre d'ambulances et de VL. Les résultats sont présentés dans la figure 2.6, la taille des cercles pour chaque département correspond au nombre d'appels présentés, plus le cercle est gros et plus le nombre d'appels reçus est important. On constate que les trois départements susnommés dénotent effectivement avec les autres départements avec un nombre particulièrement important de sorties et de moyens. Cela vient corroborer notre hypothèse que plus on a de moyens, plus on a tendance à les utiliser facilement, alors

3. <https://www.insee.fr/fr/statistiques/1894595>

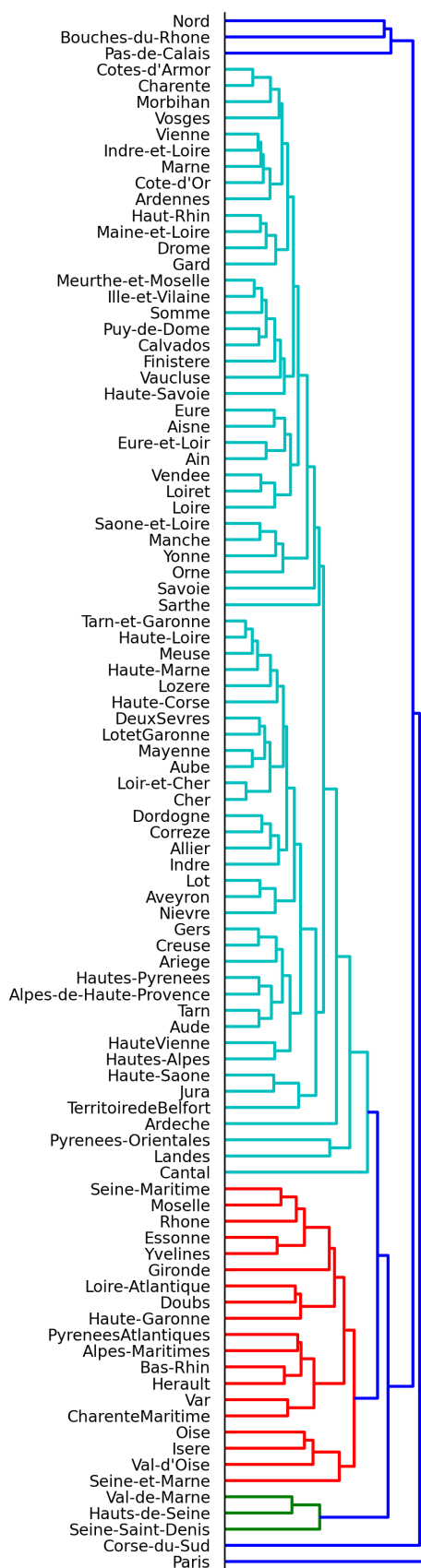


Figure 2.5 – Dendrogramme des départements français

que dans le cas inverse, les SAMU tendent à gérer a priori une pénurie potentielle et donc être plus prudent sur les engagements de moyens. Néanmoins, sans informations contextuelles supplémentaires et détaillées à propos de l'activité du soin d'urgences dans ces départements, il est difficile de conclure sur la cause de ces différences.

Le groupe suivant contient la Seine-Saint-Denis, les Hauts-de-Seine et le Val-de-Marne, qui sont tout les trois des départements d'Île-de-France et se caractérisent par leur forte densité et leur population plutôt jeune.

Enfin, on distingue au cinquième niveau deux groupes. Le premier comporte 18 départements : Seine-et-Marne, Val-d'Oise, Isère, Oise, Charente-Maritime, Var, Hérault, Bas-Rhin, Alpes-Maritimes, Pyrénées Atlantique, Haute-Garonne, Doubs, Loire-Atlantique, Gironde, Yvelines, Essonne, Rhône, Moselle et Seine-Maritime. Ces départements soit se situent en Île de France, soit on constate la présence d'une grande ville dans le département (Nantes, La Rochelle, Bordeaux, Toulouse, Montpellier, Nice, Toulon, Grenoble, Lyon, Besançon, Strasbourg, Metz, Le Havre). Ce groupe est représenté en jaune comme le cluster C2 sur la carte de la figure 2.4. Tous les autres départements sont représentés en vert par le cluster C1. Ils correspondent à des départements qui pourraient être qualifiés de plus ruraux.

En conclusion, on constate en France trois grands types de SAMU : les SAMU plutôt ruraux, les SAMU avec la présence d'une grande ville et les SAMU d'Île-de-France. On note aussi quelques cas particuliers comme Le Nord, Le Pas-de-Calais, les Bouches-du-Rhône et la Corse-du-Sud. Ces différences s'expliquent notamment, mais pas seulement, par des volumétries d'appels par habitants très disparates impliquant une hétérogénéité des pratiques.

2.3 Les verrous scientifiques

L'état de l'art réalisé dans la section 1 démontre que l'emphase a été principalement mise sur la partie aval de l'organisation alors même que des problèmes critiques ont été démontrés sur la partie amont. Il y a donc un manque évident dans la littérature de solutions concrètes permettant d'adresser la problématique métier discutée précédemment.

Des verrous scientifiques doivent donc être levés pour être en mesure d'évaluer l'impact des types d'organisation sur les performances des centres d'appels. Il n'est pas encore aisé aujourd'hui de statuer de manière objective sur la prédominance d'un modèle par rapport à un autre. Il est donc nécessaire d'apporter des éclairages sur l'effet des organisations sur les indicateurs de performance des centres d'appels.

Afin d'être en mesure d'évaluer ces organisations sur la base de faits objectivés, il est nécessaire d'aborder prioritairement la question de l'accessibilité des données ; ce qui, aujourd'hui dans le domaine de la santé, n'est pas évident. En effet, les systèmes d'information en santé sont souvent limités et pas toujours à jour. Cela rend l'accès à la donnée complexe. Avec la modernisation récente des systèmes de santé et l'accès

Nombre de sorties terrestres primaires et secondaires en fonction du nombre d'ambulances et de véhicules légers

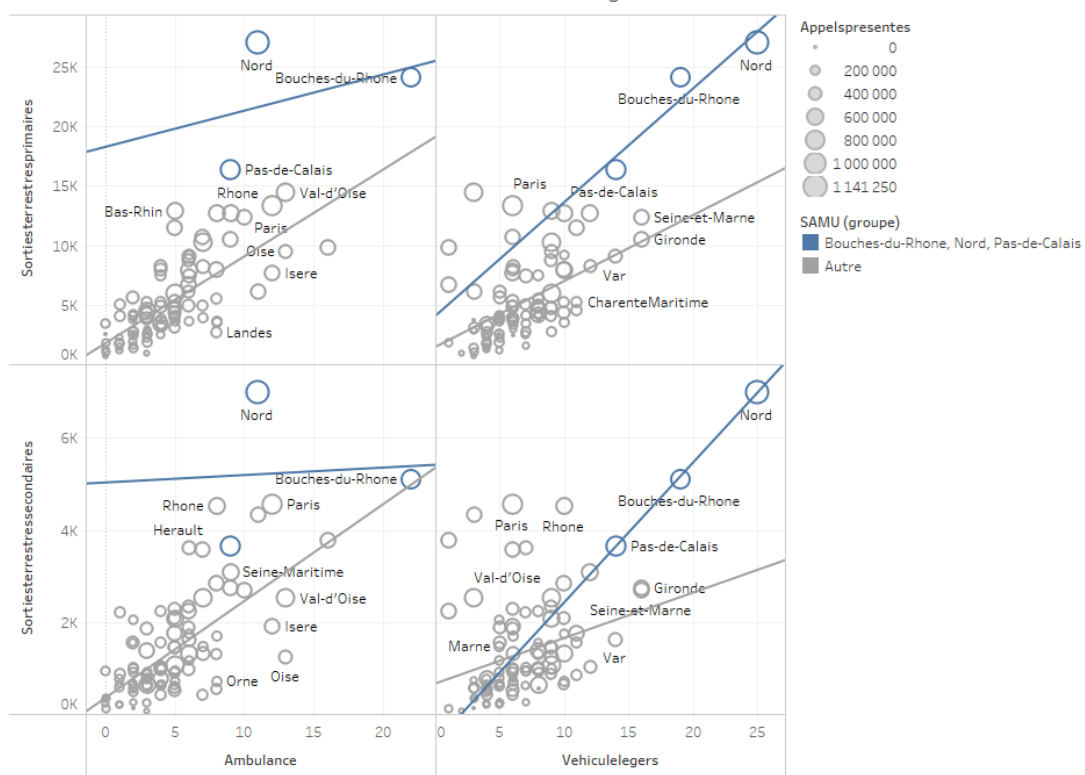


Figure 2.6 – Nombre d'interventions en fonction du nombre de moyens

Tableau 2.2 – Tableau récapitulatif des données pour les Bouches-du-Rhône, le Nord et le Pas-de-Calais comparé à la moyenne nationale

	Bouche du Rhône	Nord	Pas-de-Calais	<i>Moyenne nationale</i>
SMUR général	6	1	7	4
Antenne SMUR	7	-	-	1
Population	2 000 000	2 600 000	1 470 000	669 800
Part des 75ans	10 %	7 %	8 %	10 %
Superficie	5 087	5 743	6 671	5 674
Densité	396	454	221	565
Ambulance	22	11	9	5
Véhicule léger	19	25	14	7
Sorties terrestres primaires	24 124	27 080	16 397	5 758
Appels présentes	786 191	1 141 250	907 616	296 664
DR	540 796	449 947	386 661	159 538
DRM	343 051	342 475	209 709	124 067
MR heures hebdo moy	504	861	1 347	379
ARM heures hebdo moy	1 176	1 653	1 623	625
MG	2 160	2 315	1 191	536
Appels/hab	0,39	0,44	0,62	0,50
MGpour100000hab	107	89	81	80

aux données qui l'accompagne, de nouvelles opportunités apparaissent. L'analyse des données doit permettre de mieux comprendre le fonctionnement des organisations de santé d'une part, et les résultats de ces analyses doivent permettre de créer de la connaissance, de modéliser les systèmes de santé pour les challenger, les éprouver et enfin les améliorer d'autre part.

Le fait d'être capable de reproduire le comportement des centres d'appels à l'aide de doubles numériques par exemple devrait permettre de faciliter de telles démarches de progrès permanent. Ce genre d'environnement virtuel est en effet un atout majeur dans le développement de nouvelles manières de fonctionner car il ne présente aucun risque pour les usagers, au contraire d'une expérimentation terrain.

La littérature met par ailleurs en évidence que le niveau de détail dans la modélisation des centres d'appels d'urgence se limite le plus souvent à la prise en compte des appels entrants uniquement. Les appels sortants ou les appels internes ne sont que rarement pris en compte, ce qui peut influencer sur la précision des modèles par rapport au monde réel.

L'intégration de nouveaux outils d'analyse, d'optimisation et de pilotage des organisations des systèmes de gestion des appels d'urgence apparaît donc comme une évidence qu'il convient d'accompagner. Le présent travail de recherche ambitionne d'apporter sa contribution à cet objectif.

2.4 Formulation de la question de recherche

Aussi, sur la base des éléments précédents, il est possible de formuler la question de recherche générale de nos travaux de thèse comme suit :

Comment structurer et outiller une démarche de progrès permanent des centres d'appels d'urgence ?

Cette question de recherche peut se décomposer en 3 sous-questions correspondants aux 3 contributions qui seront développées dans la suite de ce manuscrit :

Quelle méthodologie et quels outils peut-on proposer pour diagnostiquer et améliorer le fonctionnement d'un centre d'appels d'urgence ?

Un premier constat lors de l'étude des modèles de centres d'appels et de gestion de l'urgence à travers le monde est qu'ils sont fortement hétérogènes. Les études qui visent à objectiver leur fonctionnement en vue d'améliorer leurs performances sont bien souvent applicables à un seul type de centre d'appels. En effet, les données utilisées pour la modélisation proviennent bien souvent d'un seul et unique site. De ce fait, les analyses et propositions sont difficilement généralisables. Toutefois, même si les différences sont nombreuses, il est possible d'identifier des caractéristiques qui légitiment la tentative d'un classement des modèles organisationnels de gestion de l'urgence.

En considérant la diversité de fonctionnement des SAMU, ainsi que leur besoin d'avoir une analyse fine et objective de leur activité, nous proposons donc dans la suite un cadre méthodologique outillé pour l'analyse, le diagnostic et l'amélioration des centres d'appels. L'intérêt de ce cadre est qu'il suggère de partir des données terrains afin de créer un double numérique le plus fidèle possible à la réalité. Ce cadre comprend une démarche complète allant de la récupération des données, leur nettoyage, jusqu'au test de nouvelles organisations en passant par la création et la validation du modèle. Un bénéfice majeur, et nous y reviendrons plus tard en détail, est qu'il permet de prendre en compte tous les appels pris en charge par les centres d'appels (et pas seulement ceux entrants) et donc de refléter au mieux les contraintes du monde réel.

Quelle influence ont les organisations en place dans les centres d'appels d'urgence vis-à-vis de leur niveau de performance ?

Une contribution scientifique du présent travail de recherche concerne l'analyse des données de téléphonie des centres d'appels. En effet, la compréhension des registres de données de téléphonie ainsi que l'identification des différentes activités est fondamentale pour les gestionnaires des centres d'appels. Le traitement de ces données est un atout essentiel qui permet de comprendre objectivement quelles sont les étapes qui conduisent à de potentiels dysfonctionnements. L'utilisation d'outils

issus du *Data-Mining* et du *Process Mining* permet ainsi de fournir une image objective des processus de l'organisation. Cela permet également de faire une analyse objective des performances du centre d'appels sur la période étudiée. L'avantage par rapport à une étude de données agrégées telle qu'existant classiquement dans la littérature est qu'il est possible d'extraire les indicateurs de performances sur des créneaux de la longueur que l'on souhaite afin d'augmenter le niveau de précision de l'étude. Cela permet notamment d'étudier les relations existantes entre le nombre d'appels, la qualité de service et le taux d'occupation des agents sur des fenêtres de temps plus étroites que celles de la journée entière. L'étude des données de téléphonie rend possible l'extraction du profil des appels, de la durée de session de ceux-ci ainsi que des répartitions des types d'appels au sein d'une journée. Toutes ces informations sont aussi utilisées pour paramétrer le double numérique du SAMU et contribuent à le rendre le plus fidèle possible.

Quels schémas organisationnels préconiser pour des centres d'appels d'urgence en situation nominale et en situation de crise ?

Une fois le double numérique paramétré nous proposons de l'utiliser pour répondre à la question « comment améliorer les performances des SAMU en régime nominal ainsi qu'en réponse à une augmentation significative des appels lors d'une crise sanitaire ? » Des préconisations originales de gestion des files d'attente constituent une contribution essentielle de ce travail de recherche. Nous proposons également de tester l'impact de l'ajout d'un système de triage des appels comme on peut le trouver dans les accueils des urgences physiques des hôpitaux.

3

Proposition d'un cadre de diagnostic

3.1 Introduction du besoin

Ce chapitre présente notre proposition de cadre méthodologique (cf. figure 3.1) pour le diagnostic organisationnel et l'évaluation des centres d'appels d'urgence.

Cette méthodologie a fait l'objet d'un article de journal publié en 2019 (Petitdemange et al., 2019) dans la revue *IEEE Transactions on Engineering Management* et ce chapitre emprunte une partie de ses résultats à cette publication. Ainsi, une première partie du chapitre est dédiée à la justification du besoin d'une telle méthodologie et une seconde à la justification des outils employés ainsi qu'à la présentation détaillée des étapes de la méthodologie proposée.

3.1.1 La notion de diagnostic

Le concept de processus de diagnostic ne se trouve pas uniquement en médecine où il est défini comme le « temps de l'acte médical permettant d'identifier la nature et la cause de l'affection dont un patient est atteint. »¹. En effet, il peut être transposé à d'autres domaines. En gestion par exemple, cela correspond à l'évaluation des points forts et des points faibles d'une entreprise. En industrie automobile, c'est le contrôle au moyen de dispositifs embarqués ou non permettant de déterminer l'état du véhicule lors de révisions ou lors de pannes ou d'incidents pour en déterminer les causes. (Française, 2020). On retrouve aussi la notion de diagnostic en automatique qui peut se définir comme « l'ensemble de mesures, de contrôles faits pour déterminer ou vérifier les caractéristiques techniques d'un système à des fins de maintenance ou d'amélioration. » On trouve en automatique toujours l'usage de la boucle commande-contrôle qui à l'aide de boucle de rétroaction permet de donner de l'information sur le système afin d'identifier rapidement des problèmes (Le Gallais, 1997).

1. Larousse - 2020

Le diagnostic médical

Lorsqu'un patient ressent une gêne, qu'il ne se sent pas bien, il va identifier ses symptômes, les choses qui le dérangent, qui ne sont pas normales. Il va partager ses symptômes avec son médecin qui va l'ausculter, afin d'avoir des informations supplémentaires (prise de tension, prise de sang, température, etc.). Sur la base des symptômes fournis par le patient (manifestation d'une maladie perçue par le malade) et des signes (manifestation d'une maladie constatable objectivement) mesurés par le médecin, un diagnostic va être fait. De ce fait, le diagnostic est déjà en cours, car il peut être considéré comme un processus, un enchaînement d'activités dans le but d'identifier le mal dont souffre le patient afin de le traiter (Jutel, 2009). Jutel dans son étude sociologique sur le diagnostic distingue deux définitions : Le diagnostic peut à la fois être un processus et une catégorisation. L'auteur considère que c'est une catégorisation car il correspond à l'agrégation de symptômes et de signes qui deviennent clairs, objectifs et justifiés dès lors qu'on peut les relier à une étiquette : le diagnostic. Le diagnostic comme processus correspond à la méthode d'évaluation et de traitement des maux du patient. C'est cette définition de processus que nous utiliserons pour la suite. En médecine, le processus de diagnostic a été défini par (National Academies of Sciences et al., 2015) comme un enchaînement complexe d'activités de collaboration, centrées sur le patient, qui impliquent la collecte d'information et le raisonnement clinique, dans le but de déterminer le problème de santé d'un patient.

Le diagnostic d'un patient est un processus en plusieurs étapes : Le patient constate qu'il a un problème de santé (1), il prend alors contact avec le système de santé (2). S'en suit un cycle itératif de récolte de l'information (3), d'analyse et d'interprétation de l'information (4), et de formulation du diagnostic en tant que tel (5). Si la quantité d'information recueillie n'est pas suffisante pour identifier le problème de santé du patient, le cycle reprend jusqu'à réussir à valider un diagnostic (ici, cela comprend la définition de catégorisation de la maladie). L'étape suivante est la communication et l'explication du problème de santé au patient (6). Enfin, un traitement s'appuyant sur le diagnostic est planifié et mis en place (7). Ce traitement fait l'objet d'un suivi via le cycle de récupération d'information afin de s'assurer de son effet (8). (National Academies of Sciences et al., 2015).

Le diagnostic organisationnel

Si on considère le SAMU comme un patient et que l'on file la métaphore, en se référant aux éléments discutés dans les chapitres précédents, on peut considérer que le SAMU présente des symptômes de dysfonctionnement, il est malade. Cette hypothèse se renforce si on considère les articles de presse généraliste relatifs aux SAMU qui se multiplient ces dernières années tels que : « Il faut prendre des mesures pour éviter que le drame du SAMU de Strasbourg ne se reproduise »², « Appels aux urgences : les SAMU accusés de ne pas toujours décrocher »³, « Classement

2. Le Monde, 15 mai 2018

3. Le Monde, 24 août 2018

des SAMU : quand le 15 ne répond pas »⁴. On constate dès lors un problème qu'il semble nécessaire d'analyser, de comprendre, et autant que faire ce peut, de résoudre

3.1.2 Contexte et justification de la question de recherche : Vers une méthode outillée de diagnostic organisationnel des SAMU

En un mot, les centres d'appels sont confrontés à de nombreux défis liés à leur performance, et plus particulièrement à leur réactivité. C'est extrêmement préoccupant, car une mauvaise performance dans ce secteur peut avoir un impact énorme sur la vie des gens. Pour améliorer leur performance, les centres d'appels doivent remettre en question le service de prise d'appels, c'est-à-dire la façon dont ils traitent et répondent aux appels. L'hétérogénéité de ces organisations et les spécificités du domaine des services d'urgence permettent d'expliquer les niveaux de performance variables dans des conditions de fonctionnement « normales » ainsi que les difficultés à anticiper le comportement de l'organisation en situation « extrême ». La contribution de ces travaux de recherche consiste en partie à relever ce défi complexe

Concernant les modèles formels, nous avons décidé de discuter ces différences selon les dimensions présentées au chapitre 1.1 : **la stratégie pré-hospitalière, l'accès au centre d'appels** et le modèle de décroché que l'on décline en quatre dimensions : **le type de plate-forme, la spécialité du centre d'appels, la gestion des files d'attente et les compétences en matière de ressources.**

La stratégie pré-hospitalière

La stratégie pré-hospitalière définit la façon dont le Centre de Réception et de Régulation des Appels (CRRA) répond aux besoins de secours. Deux approches principales ont été identifiées : *Scoop And Run* et *Stay And Play*. En France, le modèle est principalement du SAP. Malgré la grande quantité d'études existant sur le sujet, (Hannay et al., 2014 ; Levy et al., 2007 ; McCoy et al., 2013 ; Smith et al., 2009) nous ne sommes toujours pas capable de savoir quel modèle est le mieux pour la survie du patient. Ainsi, lors de l'étude d'un centre d'appels d'urgence, il est intéressant d'identifier le type de stratégie pré-hospitalière dans laquelle nous nous trouvons.

L'accès

L'accès au centre d'appels définit la manière dont le centre d'appels peut être contacté. On considère plusieurs moyens d'accès aux centres d'appels d'urgences : un accès universel si le numéro est unique, et spécialisé s'il y a plusieurs numéros pour chaque type d'urgences différentes (Polices, Pompiers, Urgences Médicales). Les différents types de modèles que l'on retrouve majoritairement ont été discutés dans le chapitre 1.1. La différence majeure réside dans le fait que le patient ait besoin de choisir le

4. Le Point, 22 août 2018

bon numéro en fonction du type de secours requis ou si cette décision revient aux professionnels de l'urgence. De plus, dans le cas où le patient a besoin d'effectuer lui-même la détermination de la gravité de la situation, on observe en Europe 10% d'erreur dans les organisations qui opposent les services d'urgence et les services non urgents. Il y a ici un enjeu de simplification et de lisibilité sur l'utilisation des services d'urgences pour les citoyens.

La plate-forme

Le type de plate-forme peut être partagé ou spécifique. Elle est partagée si tous les appels sont traités au même endroit indépendamment qu'ils soient destinés aux urgences médicales, civiles, ou de sécurité incendie. Elles sont spécifiques si chaque spécialité dispose de son propre site pour la répartition et l'envoi des interventions. Par exemple, le SAMU de Carcassonne regroupe à la fois la plate-forme d'appels des pompiers et celle du SAMU, cette plate-forme est donc partagée. En revanche, les SAMU d'Albi, Rodez, Nantes ou Toulouse disposent de leur propre plateforme téléphonique indépendante de celle des pompiers. Au sein, d'un même pays les pratiques peuvent différer. Toutefois, en France les plate-formes sont majoritairement spécifiques, la configuration de plateforme commune ne concerne que 17 départements sur 101. Il n'y a actuellement pas d'études qui permettent de connaître l'impact de ces modèles sur les performances des centres d'appels. Il pourrait être intéressant d'étudier l'effet qu'une plate-forme commune peut avoir sur le temps de prise en charge des appels.

Le gestion des files d'attente

La gestion des files d'attente fait référence aux règles de gestion des appels. Fondamentalement, il peut être avec ou sans triage. Sans triage implique que les appels sont gérés selon le principe du premier arrivé, premier servi pendant tout le processus de l'appel, alors qu'avec le triage signifie que les appels sont classés par ordre de priorité après avoir été pris en charge en fonction de leur gravité. Ainsi, les appels les plus graves sont traités en priorité.

La spécialité

La spécialité renvoie aux types d'appels couverts par le centre d'appels : qu'il s'agisse d'appels urgents ou d'appels de médecine générale. Nous ajoutons cette catégorie, car cela apporte une nuance pour la priorisation des appels. Car dans le cas où le même centre d'appels gère à la fois de la médecine d'urgence et de la médecine générale, nous sommes face à deux types d'appels qui n'ont pas les mêmes besoins. In fine, ils n'ont pas non plus la même chaîne de valeur. En effet, les appels de médecine générale sont moins critiques en matière de temps en comparaison à ceux de médecine d'urgence. Ainsi, cela aura un effet sur la gestion des priorités dans le centre d'appels.

La gestion des ressources

Les compétences en matière de ressources font référence au fait que les ressources au sein d'un CRRA peuvent être spécialisées (pour une tâche spécifique, un flux d'appels, etc.) ou polyvalentes. (M. v. Buuren et al., 2017)

Étant donné qu'il y a deux possibilités pour chacune des dimensions organisationnelles que nous venons d'évoquer, nous avons une combinatoire de 64 (2^6). Cette combinatoire conduit mécaniquement à une grande hétérogénéité de systèmes SAMU, et à une grande complexité. Ainsi, il y a un enjeu majeur à disposer d'une méthode outillée complète permettant de poser un diagnostic étiologique^a organisationnel pertinent et complet.

a. ("Le diagnostic étiologique, consiste à identifier la cause de l'affection" - Larousse 2020)

Afin de confirmer la pertinence de notre objectif, nous avons effectué une analyse documentaire systématique à l'aide de ces mots clés sur le Web of Science :

(Methodology OR Method OR Method OR Framework OR Tool)
AND (« Emergency Call Center » OR « Emergency Contact Center »
OR « Emergency Dispatch Center » OR « Public Safety Answering
Point ») AND (Evaluation OR Diagnosis * OR Assess *) AND
(Responsiveness OR Performance OR Efficiency).

Cette requête a généré 1240 résultats. Parmi ces résultats, Gardett et al. (Gardett et al., 2016) passent en revue la littérature systématique sur la « recherche sur les dépêches d'urgence » depuis les années 1970 jusqu'en 2013. Dans leur revue, ils soulèvent un manque d'études comparatives sur le terrain, un cadre d'évaluation universel et une description précise du protocole utilisé dans les études individuelles. Nous avons donc concentré nos investigations de recherche à des articles publiés après 2013. Après un tri sur les titres et les résumés, nous avons sélectionné 10 articles qui traitent de l'organisation ou de la méthodologie des centres d'appels d'urgence. Toutefois, aucun d'entre eux ne présente de solution détaillée permettant d'adresser notre problématique. C'est pourquoi nous pensons que la proposition d'un cadre outillé est pertinente et vient compléter utilement la littérature tant pour les chercheurs qui veulent étudier les centres d'appels que pour les praticiens du domaine afin d'évaluer leur structure organisationnelle.

3.2 Présentation de la méthodologie

Notre proposition est donc une méthodologie exploitant différents outils. Elle peut être utilisée pour diagnostiquer et améliorer les centres d'appels tout en tenant compte de leurs spécificités. Cela permet d'effectuer une analyse plus précise en vue d'améliorer notamment la réactivité des centres d'appels. À partir du système réel, l'objectif est de pouvoir construire un modèle de connaissance. Ce modèle permet

de comprendre et de formaliser le système réel afin de créer son double numérique : la réplique numérique du centre d'appels. Le système réel et le double numérique couplé aux six dimensions déjà présentées dans la section 3.1.2 constituent le cadre proposé comme indiqué dans la Figure 3.1. Tout le principe est de partir du système réel, d'en extraire les données utiles afin de les transformer en connaissance. Tout ce travail dans le système réel nous permettra, dans un second temps, de construire et paramétrer le double numérique du système étudié. Ce double numérique va agir comme une zone expérimentale virtuelle, sur laquelle nous pourrions nous appuyer pour améliorer l'organisation du système réel.

La méthodologie est divisée en quatre phases que nous allons développer dans la suite :

1. Collecter et préparer les données, ce qui permet de comprendre le processus opérationnel du système réel et d'en extraire des données significatives.
2. Interpréter et intégrer les informations afin de transformer les données en connaissances.
3. Diagnostiquer, ce qui permet d'obtenir tous les paramètres d'entrée du modèle de simulation ainsi que d'analyser la performance actuelle du centre d'appels.
4. Améliorer, ce qui permet de tester de nouvelles configurations organisationnelles et d'évaluer l'impact sur la performance du système.

3.2.1 Étape 1 : Collecter et préparer les données

1.a Collecter et nettoyer les données

La complexité d'obtenir le journal des événements des données réside dans la variété de la structure des données, qui peut être différente d'une organisation à l'autre. Afin de faciliter ce processus il est préférable que les données soient préparées selon un format standardisé. Van Dongen et van der Aalst (Dongen, van B.F. et al., 2005) ont développé un métamodèle pour le traitement par fouille de données (*Data Mining*). Chaque événement du journal doit avoir un nom, un horodatage et un identifiant (ID) de trace. Dans notre cas spécifique, nous ajoutons la ressource par qui l'activité a été effectuée et l'origine de la trace. La structure de nos données (que l'on appellera *event log*) comprendra ainsi les éléments suivants :

- **Événement** : le nom de l'événement qui se produit ;
- **Horodatage** : l'heure de l'événement - année, mois, jour, heure, minute, seconde ;
- **ID** : une code unique pour identifier tous les événements liés à un appel ;
- **Ressource** : le numéro d'identification de la ressource exécutant l'événement ;
- **Origine** : le numéro composé par l'appelant pour accéder au centre d'appels

La figure 3.2 reprend les grands principes de cette première étape. En entrée, nous avons besoin des données de téléphonie brutes, sur lesquelles nous allons appliquer des algorithmes d'associations (Hand et al., 2001) pour rassembler les événements correspondant au même appel. Prenons un exemple, un appel à lui seul va générer

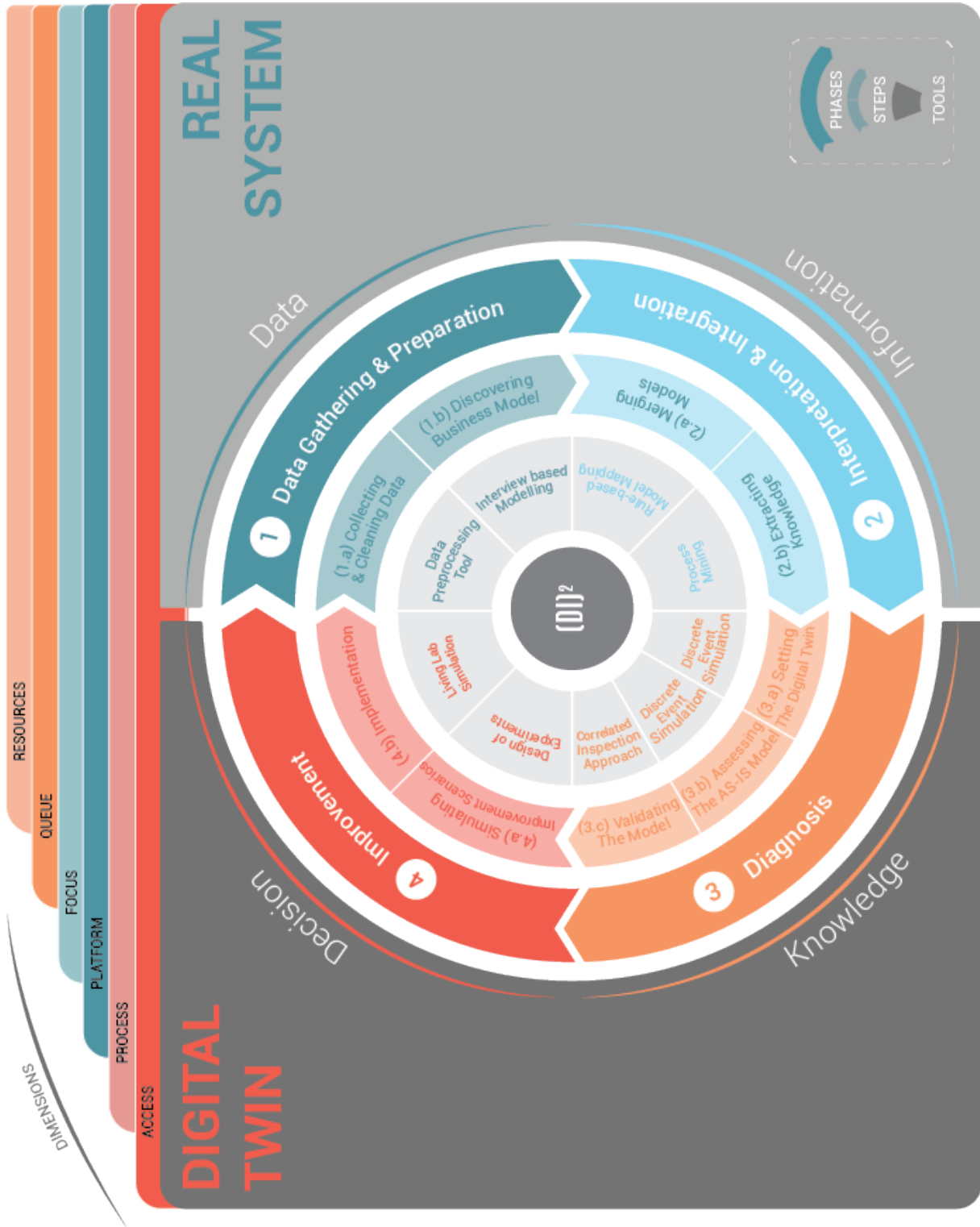


Figure 3.1 – La méthodologie outillée proposée (Petitdemange et al., 2019)

plusieurs événements. Afin de pouvoir retracer le cycle de vie de cet appel, il faut que chaque événement en lien avec ce dernier ait un identifiant commun qui permette de le relier à cet appel précisément. Cette suite d'événements horodatés en rapport avec un et un seul appel se nomme une trace. Cela va nous permettre de faire un tri sur les traces incomplètes en utilisant de l'analyse de séquence (Van der Aalst et al., 2004) et de les mettre de côté. Les traces considérées comme incomplètes sont celles qui ne vont pas avoir d'événement de début, ou qui n'auront pas d'événement de fin. Ainsi nous obtenons des données qui sont nettoyées et complètes pour la suite.

Le but de cette première étape est d'obtenir des données qui puissent être utilisées en entrée des prochaines étapes où l'on va appliquer du *Process Mining*.

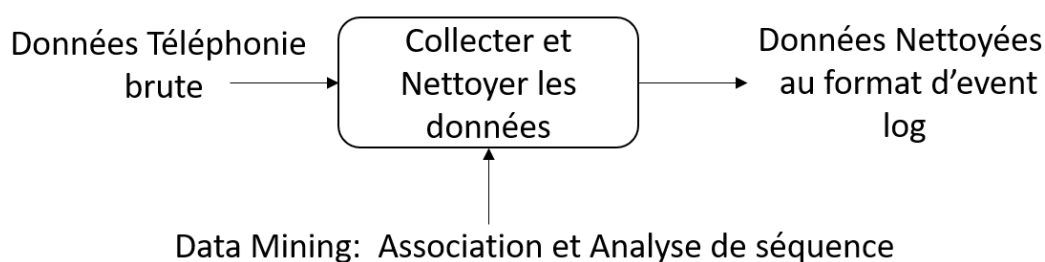


Figure 3.2 – Collecter et nettoyer : entrées, sorties et outils

1.b Découvrir le processus métier

L'étape suivante consiste à comprendre le processus opérationnel et à obtenir plus d'information sur les opérations du centre d'appels. Comme nous l'avons déjà mentionné, les centres d'appels peuvent être organisés de différentes façons d'une région à une autre (Penverne et al., 2019). L'une des difficultés est de saisir et de comprendre ces différences. Dans ce cas, nous avons besoin d'information sur les ressources :

- Combien de types de ressources sont utilisés ?
- Quelles sont les différences entre elles ?
- Quelles sont les activités réalisées par chaque type ?

Nous avons également besoin d'information sur les flux entrants et la gravité des appels. Ensuite, nous devons comprendre les données afin de pouvoir relier le numéro d'identification de l'agent à sa fonction, et le numéro d'identification du flux entrant à sa catégorie. Cela implique d'avoir un glossaire des numéros d'identification ainsi qu'un accès au journal du centre d'appels.

Une partie de ces informations pourrait être récupérée à travers les journaux (logs) de données. Toutefois, il est nécessaire de les compléter avec de l'information provenant directement du système réel afin d'avoir une meilleure compréhension globale du système. Cette connaissance peut s'obtenir en analysant les processus, donc lors de démarches d'observation. Pour collecter ces données, on retrouve deux types de méthodes : la recherche quantitative et la recherche qualitative.

La recherche qualitative est une méthode d'exploration qui permet de décrire, de comprendre et d'expliquer un phénomène particulier en répondant aux questions « quoi ? », « qui ? » et « comment ? », en opposition à la recherche quantitative qui

répond à des questions du type « combien ? » et « à quelle fréquence ? » (Jones, 1999). La recherche qualitative se caractérise par l'utilisation de méthodes particulières comme les interviews semi-structurées afin de récupérer des données ou d'observer des événements.

Elle est surtout utilisée pour étudier les processus, pour permettre de comprendre comment des actions sont effectuées, par quelles étapes on doit passer pour réaliser l'action (Ma, 2000). Gill et Stewart nous montrent qu'il s'agit d'une méthode de collecte de donnée très utilisée dans le domaine de la santé (Gill et al., 2008). C'est pourquoi nous avons effectué des interviews avec des experts métiers. Lors de ces interviews (voir Figure 3.3), il s'agit de récupérer des informations sur le processus métier, sur la manière dont s'enchaînent les différentes activités et ainsi obtenir un premier processus formalisé des activités. On doit aussi utiliser ces moments de rencontre privilégiés afin de donner un sens métier aux données de téléphonies. Par exemple, en récupérant la liste des identifiants des agents on pourra par la suite identifier quel agent est un assistant de régulation médicale (ARM) ou un médecin régulateur (MR).

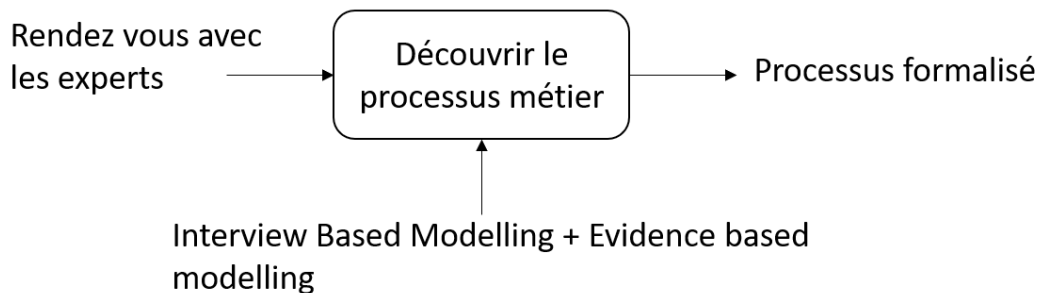


Figure 3.3 – Découvrir le processus : entrées, sorties et outils

3.2.2 Étape 2 : Interpréter et intégrer

2.a Fusionner les modèles

Cette étape crée le lien entre l'événement et l'activité, ce qui permet de déterminer la durée des activités, grâce à l'horodatage des événements. Les données du log sont composées de traces, qui elles-mêmes sont composées d'événements. Il s'en suit qu'une trace (une séquence d'événements) correspond à un appel dans le centre d'appels. Tous les événements sont instantanés. En outre, selon la ressource liée à l'événement, il peut s'agir d'une activité différente dans le processus de gestion. Ici, l'objectif est de déterminer la durée de chaque activité pour chaque appel à l'aide du log des événements. Nous avons donc élaboré des règles métiers pour relier les événements et les activités.

Dans un appel on peut enregistrer plusieurs durées, ce qui permet de décrire le type d'appel en termes d'état. Un appel a deux états principaux :

- En attente : l'appel est dans une file d'attente, en attente d'une ressource pour le traiter.

- En ligne : l'appel est traité par une ressource.

Un processus d'appel est une séquence d'états - transitions, de type "en attente" - "en ligne". L'état d'attente est facultatif facultatif (ou équivalent potentiellement à une durée nulle) et deux états en ligne peuvent être consécutifs. Un appel peut être géré par une ou plusieurs ressources. Par exemple, s'il y a deux ressources qui gèrent l'appel, il y aura deux durées correspondant à un état en ligne, une pour chaque ressource concernée. La durée entre les événements peut alors être facilement calculée.

Dans notre cas, lorsque plusieurs actions sont liées par une seule ressource, nous considérons cette séquence d'action comme une seule activité. Ceci est possible parce que nous considérons que les actions ne peuvent pas être interrompues. Le constat effectué lors des observations terrains est que lorsqu'un appel commence à être traité, il n'est que très rarement interrompu avant la fin de son traitement. Il s'ensuit qu'un temps de communication dans le log serait lié à une activité (ou à un ensemble d'action) effectuée par une ressource. Pour lier les processus métier et le log des événements il faut identifier au moins quatre types d'événements :

- *E1-Ringing* : Le premier événement pour un appel, lorsqu'il arrive au centre d'appels et qu'il attend d'être pris en charge par une ressource.
- *E2-Established* : L'appel est décroché et la ressource est en ligne.
- *E3-Released* : L'appel est libéré.
- *E4-Held* : L'appel est suspendu.

À titre d'exemple, nous considérons deux modèles d'appels du point de vue de l'événement et du processus métier décrit à la figure 3.4. Dans cette figure, il y a deux types de formalisme, d'une part les enchaînements d'événements qui proviennent des données de téléphonie, d'autre part les enchaînements d'activités qui proviennent du processus métier. L'objectif est de réussir à faire le lien entre les deux modèles. Dans ce processus de gestion, deux séquences d'activités sont possibles : i) A0-A2-A3-A3-A4, avec toutes les actions exécutées par la ressource R1, ou ii) A0-A2-A3-A5, les trois premières actions étant effectuées par R1 et la dernière par R2. Puisque nous supposons que les activités ne peuvent pas être interrompues, la séquence A2-A3-A4 est équivalente à l'activité A1 et la séquence A2-A3 à l'activité A1'. Ces activités peuvent être décrites comme suit :

- A0 correspond au temps d'attente, c'est la présence d'un appel dans la file d'attente.
- A1 représente la réponse d'un appel par un opérateur ; la gravité de l'appel est évaluée et s'il n'a pas besoin d'être diagnostiqué par un médecin, l'opérateur donnera quelques conseils à l'appelant.
- A1' représente la réponse d'un opérateur à un appel avant qu'il ne soit transféré à un médecin.
- A5 représente le diagnostic d'un appel d'un médecin durant lequel il décide ou non d'envoyer des moyens au patient.

L'application de l'algorithme 1 implique que la séquence (i) est liée au modèle 1 (Pattern 1 - voir figure) et que la durée d0 d'un appel dans la file d'attente est la différence entre l'horodatage de l'événement *E2-Established* et le premier événement *E1-Ringing*. L'heure de communication est la différence entre l'horodatage du dernier

événement *E3-Released* et l'événement *E2-Established*. L'algorithme 1 est également appliqué au modèle 2 pour obtenir la durée qui nous permet de lier la durée des activités du processus métier au journal des événements.

Algorithme 1 : How to link events to duration of communication

Résultat : Obtain the duration of each communication time

R_j : $j = 1..n$, $n \in (N)$ the number of resources

d_j : the communication time of the resource j

$E_{i,j}$: $i = 1..4$ the timestamp of the event i realized by the resource j ;

si $n = 1$ **alors**

$d_0 = E_{2,1} - E_{1,1}$

$d_n = E_{4,1} - E_{2,1}$

sinon

$d_0 = E_{2,1} - E_{1,1}$

$d_j = E_{3,j} - E_{2,j}$

$d_n = E_{4,n} - E_{2,n}$

fin

Ainsi, cette étape, comme résumée dans la figure 3.5 prend en entrée, les données sous forme de log issue de la première étape et les données de processus formalisées issues de l'étape 2. Le fait de croiser les deux résultats et de fusionner les deux modèles nous permet d'obtenir un jeu de données contenant plus d'information et notamment les temps d'attente et de communication des différents appels.

2.b Extraire des connaissances

En utilisant les données précédemment générées on peut extraire de la connaissance de ces dernières en faisant usage du *Process Mining* et du *Data Mining* (voir Figure 3.6). L'intérêt ici est de récupérer les données qui vont nous permettre de paramétrer le modèle de simulation.

Identifier la gravité

Les appels peuvent être classés par ordre de priorité et il est possible d'établir un niveau de gravité pour chaque appel en fonction de l'ensemble des ressources qui le traite. Nous pouvons classer par ordre de priorité le type de ressources : les ressources essentielles traitent les appels les plus urgents, tandis que les ressources non essentielles traitent les appels les moins urgents.

Par exemple, dans l'algorithme 2 avec $p=3$, Rt1 est un MR et la ressource la plus critique, Rt2 est un médecin généraliste et une ressource intermédiaire, et Rt3 est ARM et une ressource non critique. Rt3 reçoit tous les appels. Il/elle traite les appels non critiques et transfère le reste soit vers Rt2, soit directement vers Rt1. Rt2 transfère à Rt1 les appels qu'il considère comme très critiques. Essentiellement, tous les appels transférés à Rt1 par Rt2 et Rt3 sont considérés comme les appels les plus

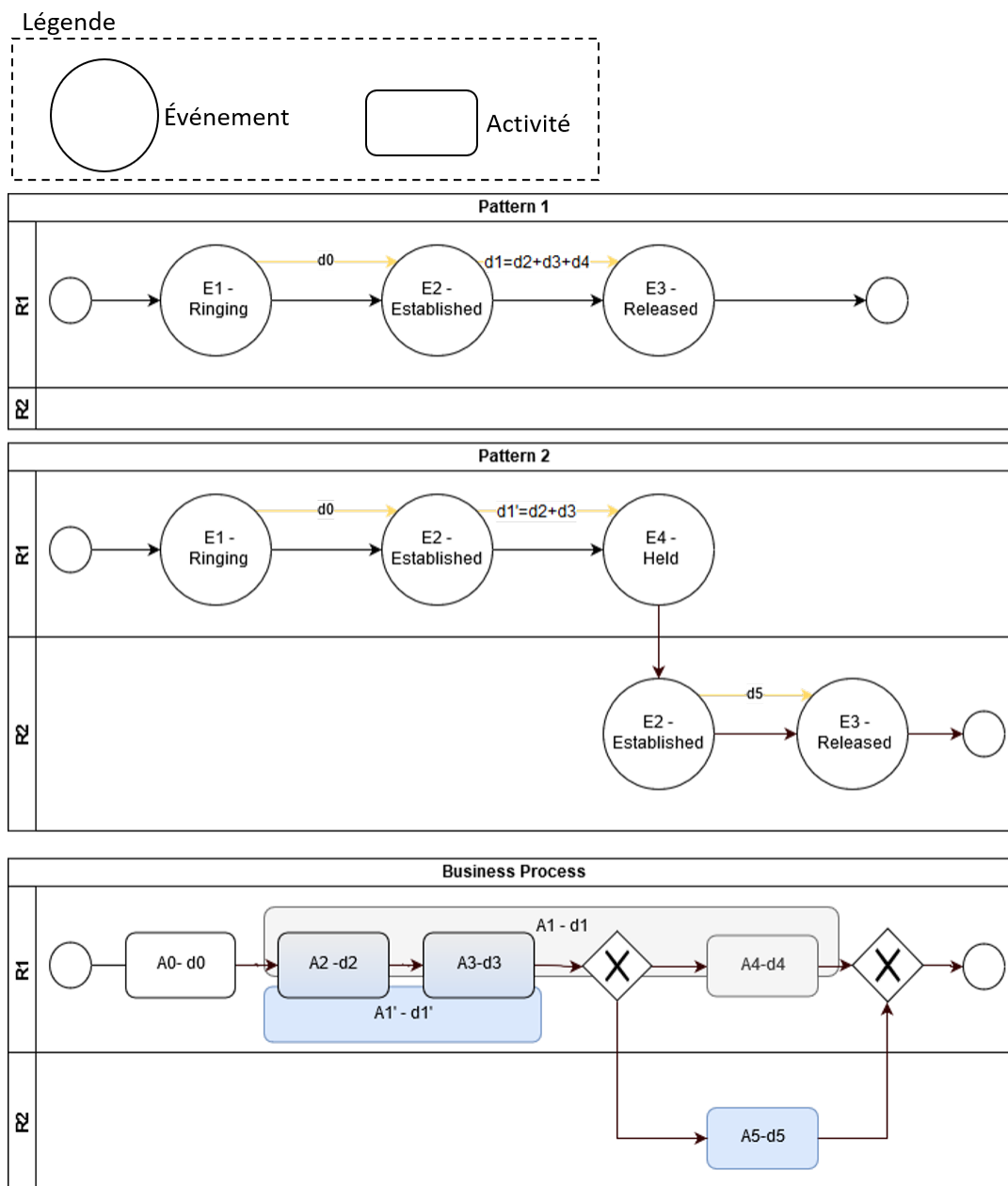


Figure 3.4 – Un processus métier et des exemples de patrons d'appels

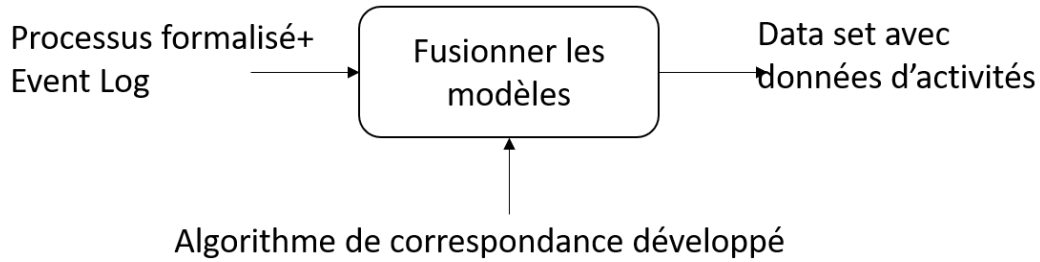


Figure 3.5 – Fusionner les modèles : entrées, sorties et outils

urgents (graves), tandis que ceux qui sont traités par Rt2 sont considérés comme étant d'un niveau de gravité intermédiaire.

Algorithme 2 : How to set the severity of a call

Résultat : Obtain the Severity of each Call

Rt_j : $j = 1..p$ the type of resource

S_i : $i = 1..m$, $m \in (N)$ The severity of the call i ;

$$A_{i,j} = \begin{cases} 1 & \text{if the Resource } Rt_j \text{ handled the call } i \\ 0 & \text{if not} \end{cases}$$

$\forall i, S_i = j$ pour $A_{ij} \neq 0$

Modéliser le temps de service pour les activités

Avec la durée des activités, on peut dès lors trouver la distribution de probabilité de chaque activité, ou du moins la distribution qui correspond le mieux aux données observées sur le terrain. Cela nous permet de calibrer le modèle de simulation avec précision. Dans la littérature existante, on suppose généralement que les temps de service suivent une distribution log-normale (Zuzakova et al., 2012)(Bhulai, 2009). Cette répartition peut convenir dans des cas particuliers où les ressources et les appels sont considérés comme homogènes et identiques.

Dans le présent document, nous faisons l'hypothèse que ce n'est pas nécessairement vrai puisque les ressources et les appels sont hétérogènes. Ainsi afin de modéliser les temps de service il y a deux possibilités : trouver une distribution théorique qui correspondent aux données réelles ou travailler directement avec les données réelles et la distribution empirique des données.

Dans le premier cas, une manière de faire est la suivante : en analysant les temps de service collectés à partir de données réelles, il est possible d'appliquer un algorithme de correspondance afin d'identifier les distributions théoriques les plus adaptées à notre cas. Cela signifie que pour chaque distribution théorique on va chercher le jeu de paramètres qui permet de minimiser l'écart entre les données réelles et la distribution théorique. Pour mesurer cet écart, on va utiliser la somme résiduelle des carrés (Dwars et al., 2013), également connue sous le nom de Somme des erreurs au carré (SSE). Une faible valeur de SSE indique un ajustement serré entre les données réelles et le modèle ce qui permet de choisir la meilleure distribution

Dans le deuxième cas, il s'agit de travailler avec les données réelles et d'utiliser la distribution empirique des données directement pour modéliser les temps d'activités.

Obtenir le profil des appels entrants

Le profil des appels peut être défini en fonction des données réelles, qui sont représentées sous la forme d'un histogramme avec un intervalle de classe de 30 minutes. En se basant sur les données réelles, l'objectif est de réussir à obtenir une distribution théorique de l'arrivée des appels. La plus grande difficulté sur cette partie est de disposer de suffisamment de données pour être capable d'identifier une distribution théoriques.

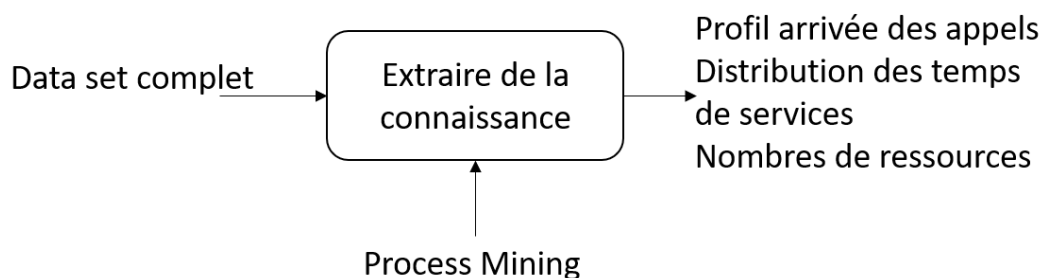


Figure 3.6 – Extraire les connaissances : entrées, sorties et outils

3.2.3 Étape 3 : Diagnostiquer

La première étape du diagnostic consiste à choisir la manière dont on va modéliser le réel. On considère deux grands types de modèles : les modèles analytiques et les modèles de simulation. Les modèles analytiques sont construits en définissant le modèle par des équations mathématiques. On lui donne en entrée des paramètres, le modèle effectue des calculs exacts et on obtient les résultats en sortie. Le modèle de simulation quant à lui va fonctionner avec un système de règles pour passer d'un état au suivant.

Lorsqu'il s'agit de décrire des systèmes dynamiques et hétérogènes, les modèles analytiques s'avèrent limités. Si on considère l'exemple de la théorie des files d'attente, on constate que si on complexifie la situation, assez rapidement nous ne disposons plus de modèle analytique permettant de résoudre le problème. (Borshchev, 2013b).

C'est pourquoi nous avons choisi d'utiliser un modèle de simulation dans notre étude. La simulation présente plusieurs avantages (Borshchev, 2013b) :

- Cela permet de tester et de trouver des solutions à des problèmes où d'autres méthodes n'ont pas fonctionné (comme la programmation linéaire ou les méthodes analytiques).
- Une fois le niveau d'abstraction établi, il est relativement facile de créer le modèle et de le faire évoluer de manière incrémentale et modulable. Cela apporte plus de flexibilité qu'un modèle analytique.
- Il est possible d'utiliser des éléments visuels de manière à ce que le modèle de simulation ressemble à la réalité. Cela rend le modèle plus accessible et il est plus aisé de communiquer avec les experts métiers.

- Il est possible de mesurer n'importe quelle valeur au sein du processus, il est facile de rajouter des calculs pour obtenir des statistiques sur le fonctionnement du modèle à n'importe quel moment dans le temps.
- On peut faire tourner le modèle en temps réel et voir son évolution au fil du temps, ce qui permet de voir l'état du système à chaque étape. Cette fonctionnalité est non seulement utile à des fins de communication mais aussi pour la construction et la vérification du modèle.
- Enfin l'aspect visuel est moins austère que peut l'être un modèle analytique avec uniquement des formules. Ce qui en fait un outil didactique très pertinent pour expliquer les résultats aux experts métiers.

3.a Paramétrer le double numérique

Cette partie consiste ainsi à construire le modèle de simulation (voir Figure 3.7). Il faut choisir le type de simulation que l'on souhaite utiliser puis le logiciel ou l'outil de programmation avec lequel on souhaite réaliser la modélisation. Le modèle est ensuite paramétré avec les données et distributions obtenues lors de l'étape précédente.

Sur la base du processus de gestion et des données, la simulation à événements discrets (SED) est utilisée pour créer un modèle du centre d'appels avec les paramètres extraits lors de l'étape précédente. La plupart du temps, les problèmes des centres d'appels sont modélisés en utilisant la théorie des files d'attente (Koole et al., 2002), qui est plus adaptée à un seul type d'appel et d'agent. Mais, les centres d'appels ont évolué et utilisent maintenant des ressources hétérogènes et reçoivent plusieurs types d'appels. C'est ce qu'on appelle des centres d'appels multi-sources avec des agents multicompetents, le SED semble donc être plus appropriée comme le démontrent Bhulai (Bhulai, 2009) et Roubos et al (Roubos et al., 2013). La simulation n'est pas seulement utilisée comme un outil d'aide à la décision, mais aussi comme un outil d'aide à la gestion (Mason, 2013).

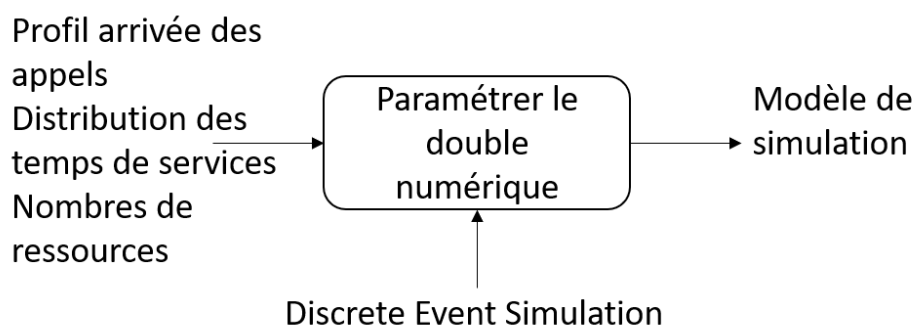


Figure 3.7 – Paramétrer : entrées, sorties et outils

3.b Evaluer l'AS-IS

Dans cette étape, il s'agit de définir les indicateurs de performance pour évaluer le centre d'appels. À l'aide des données, le temps d'attente pour accéder au centre d'appels est déterminé et utilisé pour évaluer la qualité de service (QoS). Le QoS est

une mesure de la réactivité d'un centre d'appels (voir chapitre 1.3) (Avramidis et al., 2009). Il s'agit habituellement de la fraction des appels desservis en deçà d'un seuil donné divisée par le nombre total d'appels entrants. Cet indicateur est généralement utilisé pour l'évaluation des centres d'appels et spécialement pour les centres d'appels d'urgence (Passmore et al., 2013)(Penverne et al., 2017). Dans notre étude, nous avons également tenu compte du temps d'attente maximal.

Par exemple, si on souhaite mettre l'accent sur la réactivité du centre d'appels. Il est naturel de se concentrer sur l'évaluation du temps d'attente avant qu'un appel ne soit traité. En générant l'histogramme des temps d'attente, nous pouvons rapidement évaluer le QoS et le temps d'attente maximum du centre d'appels.

En fonction du seuil défini par les experts du système réel, il est aisé de vérifier si les exigences de qualité sont respectées par le centre d'appels. Si ce n'est pas le cas, il faut alors chercher à améliorer la situation actuelle. On sera donc ensuite capable de tester de nouvelles configurations organisationnelles en vue de répondre aux exigences.

Le principe de cette étape (voir figure 3.8) est d'utiliser la SED afin de reproduire la réalité pour mieux comprendre où le système a des difficultés de fonctionnement, en identifiant les configurations dans lesquelles les temps d'attentes sont les plus longs.

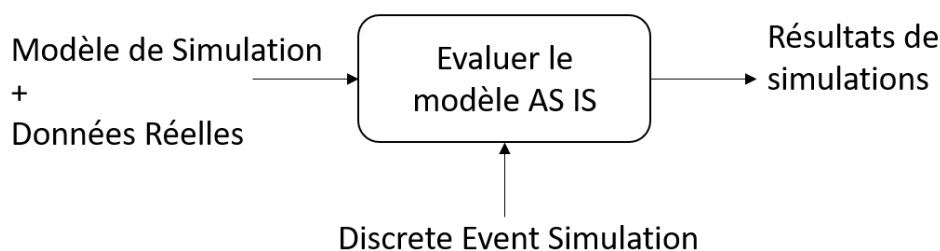


Figure 3.8 – Évaluer le AS-IS : entrées, sorties et outils

3.c Validation du modèle

Plus de 15 techniques de validation ont été identifiées dans la littérature par Robert G. Sargent (Sargent, 2013). Cet auteur indique que le choix d'une ou de plusieurs de ces techniques est fondé sur de multiples critères objectifs ou subjectifs. Nous avons donc examiné de nombreuses méthodes statistiques et opté pour l'approche de l'inspection corrélée (*correlated inspection approach*) proposée par (Law et al., 2000). Ici, nous souhaitons avoir une distribution qui permet de reproduire les temps de traitement des appels arrivant de manière totalement aléatoire suivant un profil d'arrivée prédéfini (défini par un nombre d'appels par tranche horaire).

La performance du centre d'appels réel doit être comparée à celle du centre d'appels simulé par rapport à l'indicateur de performance (KPI) le plus précis. Plus la performance du centre d'appels réel est proche du modèle de simulation, plus le niveau de confiance du modèle est élevé. Tukey (Tukey, 1977) a mis au point plusieurs méthodes qui sont utilisées pour comparer des séries univariées de données,

comme c'est le cas pour notre étude. Nous préconisons d'utiliser trois tests - deux graphiques et un analytique - pour évaluer la validité du modèle. Les tests graphiques doivent permettre de comparer la distribution des données à l'aide de leurs boîtes à moustaches (boxplots) et histogrammes. Cette méthode nous fournit des informations sur la distribution des données, la valeur moyenne et les valeurs extrêmes. Pour que le modèle soit valide, il doit avoir une forme similaire au modèle réel. Enfin, nous proposons d'utiliser le test analytique de Pearson pour évaluer la corrélation entre les deux séries. Si la valeur d'essai est supérieure ou égale à 95%, le modèle est considéré comme valide.

L'important dans cette étape est d'être capable d'attester que le modèle correspond à la réalité selon les hypothèses qui ont été formulées. (voir Figure 3.9). Cette étape est primordiale car si le modèle est trop éloigné de la réalité, toutes les préconisations qui pourront être faites en utilisant ce modèle seront caduques.

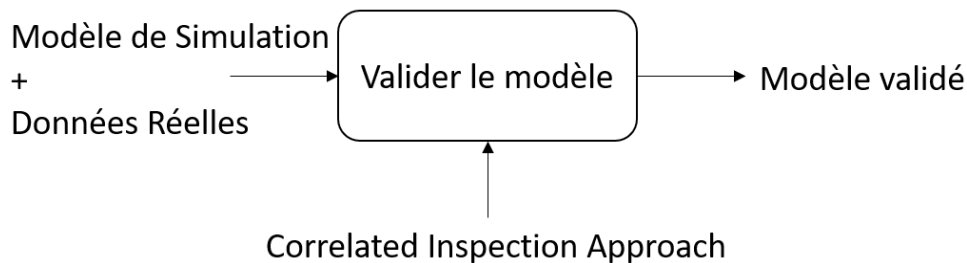


Figure 3.9 – Valider le modèle : entrées, sorties et outil

3.2.4 Étape 4 : Améliorer

4.a Simuler des scénarios d'amélioration

Avec la collecte des données et des informations extraites des journaux, une SED est mise en place pour tester plusieurs scénarios et évaluer l'impact sur la réactivité du centre d'appels. Le principe de cette étape est itératif (voir figure 3.10). Nous allons tester des scénarios d'amélioration en jouant sur différents paramètres et/ou variables illustrées par les onglets de la figure 3.1. Par exemple, il est possible de faire varier les règles des files d'attente : mettre une file d'attente par type d'appel ou une seule file d'attente pour tous les types d'appels ? Si tel est le cas est-ce que la file d'attente doit être gérée en FIFO ? (*First In, First Out*, le premier appel qui entre dans la file d'attente sera le premier appel décroché) Ou alors faut-il mettre une priorité en fonction de la provenance de l'appel (Urgence ou PDSA) ? Il est aussi possible de modifier la spécialité des ressources : tester l'effet qu'auraient des ARM spécialisés dans le décroché des appels d'urgence ou dans celui des appels de la PDSA. A partir de là, il est possible d'envisager des scénarios mixtes avec certaines ressources spécialisées et d'autres qui puissent répondre à tous les types d'appels. Un autre type de scénario envisageable est de modifier le processus de décroché des appels en mettant en place un niveau supplémentaire de décroché qui aurait pour but d'évaluer la gravité de l'appel le plus rapidement possible.

Une fois un scénario joué, il est analysé pour voir son impact sur les différents indicateurs de performance défini lors de l'étape 3b. De cette analyse des résultats obtenus, il est possible de définir un nouveau scénario afin d'affiner la préconisation. Plusieurs itérations peuvent être faites jusqu'à obtenir des résultats satisfaisants.

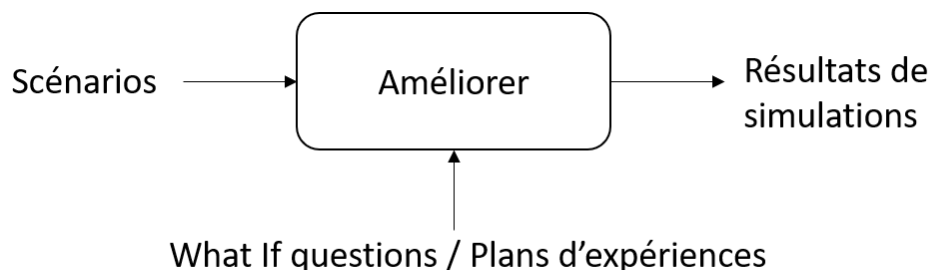


Figure 3.10 – Améliorer : entrées, sorties et outils

4.b Mettre en oeuvre

Si un ou plusieurs scénarios répondent aux objectifs des experts, ils pourraient être testés dans un laboratoire d'usages (*Living Lab*) ou par un jeu sérieux *Serious Game* dans une zone sécurisée pour évaluer le scénario en conditions réelles avec les ressources du métier. Cette étape permettra d'une part de présenter les scénarios retenus aux experts métiers (enjeux) et d'autre part d'évaluer la difficulté de mise en place de cette solution (accessibilité). Cette étape peut aussi s'avérer très intéressante pour obtenir les avis de l'équipe par rapport à des difficultés de mise en place qui pourraient être rencontrées et auxquelles nous n'aurions pas pensé (conduite du changement).

3.3 Synthèse

Face à la grande hétérogénéité des SAMU et notamment de leur composante aval, il semble pertinent de proposer un cadre méthodologique d'évaluation et d'amélioration de leurs organisations. A ce jour dans la littérature, il n'y a pas de travaux qui proposent une méthode complète, générique et outillée pour l'étude organisationnelle des centres d'appels d'urgence. La contribution proposée met donc en place un cadre méthodologique en deux temps qui s'appuie sur l'étude des données réelles pour soutenir la formulation d'un état des lieux objectif, puis sur la création d'un double numérique pour soutenir la conception et l'évaluation de propositions d'améliorations organisationnelles. Chaque partie est décomposée en étapes pour lesquelles nous proposons des outils opérationnels permettant d'atteindre les objectifs visés. Nous proposons aussi 6 dimensions organisationnelles (**la stratégie pré-hospitalière, l'accès au centre d'appels** et le modèle de décroché que l'on décline en quatre dimensions : **le type de plate-forme, la spécialité du centre d'appels, la gestion des files d'attente et les compétences en matière de ressources**). Ces six dimensions permettent de caractériser l'hétérogénéité des centres d'appels d'urgence. Elles sont autant de leviers sur lesquels il est possible d'agir afin de modifier les

organisations des centres d'appels d'urgence et mesurer l'effet de ceux ci sur les performances des centres d'appels.

Avec cette contribution nous souhaitons proposer un cadre d'étude structuré des centres d'appels associé à une boîte à outils. Les outils identifiés pour les différentes étapes n'ont pas la prétention d'être uniques et peuvent être modifiés ou adaptés. Nous ne prétendons pas non plus avoir choisi le meilleur outil pour chaque étape, mais simplement un outil qui permet de réaliser l'objectif de l'étape. Afin de faire progresser encore ce travail, il pourrait être utile de développer les transitions entre les étapes, de manière à ce que le format de sortie d'une étape corresponde exactement à celui d'entrée de l'étape suivante. L'idéal serait de pouvoir proposer une boîte à outils consolidée et unifiée qui utiliserait des données d'entrées harmonisées et réaliserait automatiquement toutes les étapes les unes à la suite des autres. Mais pour arriver à ce résultat, de nombreux verrous restent à lever. Notamment au niveau de la génération automatique du modèle de simulation à partir de la connaissance extraite des logs de données. Le tableau 3.1 propose une synthèse des contributions de ce chapitre et des principales limitations associées.

Dans la suite, nous proposons d'appliquer cette méthodologie pas à pas, afin d'évaluer dans le détail les bénéfices et les limites de la proposition faite dans un cadre d'application réelle. Le chapitre 4 correspond ainsi à l'étude et l'analyse des données réelles du SAMU associée aux étapes de collecte (section 3.2.1) et d'interprétation (section 3.2.2) des données. Le chapitre 5 présente quant à lui la contribution en lien avec les préconisations d'amélioration de l'organisation des SAMU. Ce chapitre 5 correspond à la partie gauche de la figure 3.1, donc celle qui traite du double numérique et fait référence aux étapes de diagnostic (section 3.2.3) et d'amélioration (section 3.2.4).

Tableau 3.1 – Contributions et Limites : Chapitre 3

Chapitre 3	
Contributions	Proposer des dimensions pour caractériser la diversité organisationnelle des SAMU
	Une méthodologie outillée pour le diagnostic et l'amélioration des centres d'appels d'urgence
Limitations	Une boîte à outils pouvant être adaptée et améliorée
	Les différentes étapes ne sont pas totalement consolidées.

4

Influence des types d'organisation sur les performances d'un SAMU

Dans ce chapitre, nous étudions les données des SAMU afin d'en extraire de l'information et de la connaissance, utiles à l'établissement d'un état des lieux organisationnel objectif. Ce chapitre correspond à la partie droite du cadre (cf. figure 4.1). Nous appliquons d'abord la première étape de la méthodologie présentée dans la section 3.2.1 à propos de la récupération des données et de leur pré-traitement. Nous présentons ensuite la deuxième étape de la méthodologie présentée dans la section 3.2.2 à propos de l'interprétation des données.



Figure 4.1 – Partie de la méthodologie traitée dans ce chapitre

4.1 Étape 1 : Collecter et préparer les données

Grâce à nos SAMU pilotes de l'Aveyron (SAMU 12), du Tarn (SAMU 81) et de la Haute-Garonne (SAMU 31), nous avons pu avoir accès à des données de téléphonie réelles à l'échelle de l'appel. Les données ont été récoltées entre le 1^{er} février 2018 et le 1^{er} mars 2018 pour l'Aveyron et le Tarn. Pour la Haute-Garonne nous disposons uniquement des appels du 25 mars 2018 qui est un dimanche. Bien qu'une journée soit un échantillon d'une taille très faible, le fait que ce soit un dimanche rend cette journée significative. En effet, le dimanche est la journée où les SAMU reçoivent le plus d'appels. Il existe une quantité considérable de données de téléphonie par exemple, un appel peut générer à lui seul plusieurs dizaines de lignes de logs. Toutefois ces données ne sont pas faciles à obtenir, et encore moins à exploiter. Les SAMU n'ont pas directement la main dessus, ils ne disposent généralement pas des compétences nécessaires pour extraire et traiter les registres de téléphonie. Il a donc fallu faire appel au prestataire du logiciel de téléphonie pour qu'il fasse l'extraction et le prétraitement du registre (notamment l'anonymisation des numéros de téléphone) de manière à ce qu'il soit possible de l'exploiter dans le cadre de cette recherche.

Quoi qu'il en soit, le fait d'avoir des données de février à mars est particulièrement intéressant. En effet, ce sont des mois d'hiver durant lesquels l'épidémie saisonnière de grippe traverse la France. Durant cette période, le nombre d'appels au SAMU augmente. Ainsi ce sont des mois difficiles où les performances des SAMU sont mises à rude épreuve.

4.1.1 Les approches pour la découverte et la cartographie des processus

Lorsqu'on cherche à modéliser un système, on va bien souvent passer par la découverte des processus de l'organisation afin de savoir qui fait quoi, quand ? Pour découvrir les processus, cela nécessite d'obtenir de l'information à propos des activités du système. Pour faire cela, (Dumas et al., 2013) considèrent trois techniques de récupération et de création de processus :

- Les approches *Interview-based* qui réfèrent à des méthodes basées sur des entretiens auprès des experts du domaine pour leur demander comment sont effectués les processus. La difficulté étant que les processus sont bien souvent multi acteurs, ainsi pour avoir une vision globale du système il faut croiser les vues de chacun de manière à ne pas perdre d'information. L'avantage de cette technique est qu'elle permet d'avoir une vue très détaillée des processus de l'organisation. Toutefois, l'humain étant faillible par nature, les experts auront tendance à expliquer les processus dans le fonctionnement normal lorsque tout va bien ou au contraire à se focaliser sur les cas exceptionnels qui ont marqué les esprits. Ainsi, il est possible que les cas particuliers qui peuvent poser problème soient occultés.

- Les approches *workshop-based process discovery*. Cette technique consiste à former des ateliers de travail afin de rassembler tous les acteurs pour qu'ils puissent confronter leurs visions du processus et leurs places dans ce dernier. Cette méthode demande à être capable de réunir l'ensemble des rôles à une même réunion. Pour avoir une vision complète et satisfaisante du processus, plusieurs ateliers de ce type seront nécessaires.
- Les approches *Evidence-based* qui se basent sur l'étude d'éléments de descriptions tangibles d'un processus. Cela peut prendre la forme de documents écrits, d'organigrammes, de procédures. Toutefois, les documents de description édités par des humains peuvent être limités et oublier une partie du processus. C'est pourquoi il est intéressant de les croiser et de les compléter avec d'autres méthodes. Parmi les approches *Evidence-based*, on trouve aussi la découverte automatique de processus (*Process Mining*), qui elle se base sur les enregistrements faits dans le système d'information. Si les registres sont complets et que chaque événement peut être relié à un cas individuel du processus, à une activité et à un moment précis dans le temps, alors on peut reconstruire le processus dans son intégralité. Cette méthode implique d'avoir accès à ces données, ce qui n'est pas toujours évident.

Tableau 4.1 – Forces et faiblesses de chaque méthode (d'après (Dumas et al., 2013))

Critère d'évaluation	<i>Evidence</i>	<i>Interview</i>	<i>Workshop</i>
Objectivité	fort	moyen fort	moyen fort
Richesse	moyen	fort	fort
Temps	moyen faible	moyen	fort
Retour des experts	faible	fort	fort

Chacune de ces méthodes a donc des avantages et des limites. Le tableau 4.1 résume les caractéristiques de chaque méthode selon plusieurs critères (vu par (Dumas et al., 2013)) : l'objectivité, la richesse en information, le temps passé, et le retour des experts. L'objectivité correspond à quel point on peut croire à la donnée récoltée comme étant la réalité. Il y aura plus de biais dans des données transmises par des humains que par des systèmes d'information. Par richesse, on entend la qualité et la précision de l'information que l'on va pouvoir avoir. La compréhension d'un système d'information peut être limitée là où un humain peut expliquer précisément un écart dans les données, car il sera capable de relier des données chiffrées à une situation complexe. Le temps d'une méthode à l'autre ne sera pas le même. Il sera plus long de réussir à réunir plusieurs acteurs à des ateliers que d'analyser des données, si on fait l'hypothèse que l'accès aux données est simple. Or dans notre domaine, nous l'avons vu, l'accès aux données est plutôt long et fastidieux. Selon la méthode il sera plus ou moins facile d'avoir un retour rapide des experts sur le modèle construit. Lors d'atelier ou d'entretien, le modèle de processus peut être construit de manière itérative rapidement, l'analyse des données rend l'échange avec les experts moins direct.

4.1.2 Sous-étape 1.a : Collecter et nettoyer les données

Format des données récupérées

Les données ainsi récoltées, nous ont été transmises sous forme de registres d'événements (logs). Avant d'exploiter ces données, un traitement est nécessaire pour enlever les doublons, les données incomplètes et obtenir un format qui puisse être exploité. La figure 4.2 présente les deux grandes étapes appliquées et des extraits des logs à chacune des étapes. Un premier travail de pré-traitement a été fait par le prestataire de téléphonie qui nous a fourni les données. En effet, sans cela les données se présentaient sous la forme d'un fichier de texte difficilement exploitable. Le premier traitement appliqué permet d'obtenir une table de données avec différentes colonnes (cf. figure 4.3). Nous présentons ci-dessous les significations des éléments que nous avons utilisés par la suite. Afin d'illustrer le contenu de chaque colonne nous proposons un exemple dans la figure 4.4. :

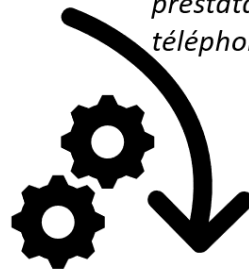
- **EventNo** : Le numéro de l'événement qui est un identifiant unique de l'événement. Dans l'exemple, nous avons 23 événements différents, chacun avec son propre identifiant.
- **EventName** : Le nom de l'événement. On constate qu'il y a deux types d'événements, ceux qui commencent par « Event », et ceux qui commencent par « Request ». Ces derniers sont des demandes d'information faites par le serveur, nous ne les utiliserons pas dans la suite, car ils ne nous apportent pas d'information pertinente. Nous reviendrons juste après sur les différents événements et leurs significations.
- **Localtime1** : Le jour de l'événement au format DD-MM-AAAA. Ici notre appel date du 31 janvier 2018.
- **LocalTime2** : L'heure de l'événement au format HH :mm :ss.ms. L'appel arrive enregistre son premier événement à 23 :43 :30 et il sera raccroché à 23 :46 :59.
- **ThisDN** : Le code du poste de travail, DN correspond à **Directory Number**. Dans notre exemple on constate 2 codes différents 25954 et 25009. Le code 25954 correspond à la file d'attente et 25009 correspond au code d'un poste de travail d'ARM.
- **OtherDN** : Le DN d'un autre poste de travail impliqué dans la gestion de l'événement courant. Cela peut par exemple être le numéro du poste du travail auquel l'appel est transféré. Typiquement sur l'événement numéro 107 **EventRouteCall** on a un **OtherDN** qui vaut 25009, cet événement nous indique que l'appel provient de la file d'attente **ThisDN** dont l'identifiant est 25954.
- **ThisQueue** : File d'attente dans laquelle est l'appel. Dans l'exemple on retrouve le code correspondant à la file d'attente soit 25954.
- **ThirdPartyDN** : Identifiant (DN) du poste de travail d'un troisième intervenant sur l'événement courant concernant l'appel.
- **ConnID** : Identifiant de l'appel, plusieurs événements vont être reliés à un même appel, afin de pouvoir regrouper ces événements l'appel à un identifiant

Données de Téléphonie Brutes

```

@18 24 01 07 10 [5] 8.1.004.01 distribute_event message EventReleased
AttributeTimeSecs 071000
AttributeTimeSecs 102862241 (18 24 01)
AttributeCustomerID "Resources"
AttributeCallID "4375A14C269F5LP09052LPR8000020"
AttributeConnID 006f02a790029226
AttributeCallID 19009
AttributePropagateCallType 0
AttributeCallType 0
AttributeThisDN "81111111"
@18 24 01 07 10 [5] 8.1.004.01 distribute_event message EventReleased generated
@18 24 01 07 10 [5] 8.1.004.01 distribute_event message EventReleased sent to [306] (00000005 StatusServer 172.18.15.118 62661)
-- call_id: connected destroyed
-- directed_call_id: app_006f02a790029226@01e431a0
@18 24 01 07 10 [5] [306] call [destroyed] [said 4375A14C269F5LP09052LPR8000020] [connection-id 006f02a790029226] [call-id 19009]
-- ChannelEvent Delete
@18 24 01 07 10 [5] 8.1.004.01 distribute_event message EventDelivered
AttributeTimeSecs 071000
AttributeCustomerID "Resources"
AttributeTimeSecs 102862241 (18 24 01)
AttributeAgentID "811111"
AttributeThisDN "811111"
    
```

Pré-traitement par le prestataire de téléphonie



Données de Téléphonie sous forme de table (log)

EventNo	EventName	LocalTime	ThisDN	OtherDN	ThisQueue	ThirdParty DN	ConnID	Transfer ConnID	Previous ConnID	CallID	ANI	DNIS	Agent ID
15527	EventACK	2018-02-02 15:49:07.537	AgentStateDN										
15528	EventUserEvent	2018-02-02 15:49:07.537	AgentStateDN	AgentStateDN									
15529	EventTreatmentApplied	2018-02-02 15:49:07.672	25901		25901		006f02a790029226			4185	81LLL3H	25601	
15530	EventUserEvent	2018-02-02 15:49:07.672	25901		25901		006f02a790029226			4185	81LLL3H	25601	19
15531	RequestUpdateUserData	2018-02-02	25901				006f02a790029226						
15532	RequestSetCallInfo	2018-02-02					006f02a790029226						
15533	EventAttachedDataChanged	2018-02-02 15:49:07.673	25901			25901	006f02a790029226			4185	81LLL3H	25601	
15534	RequestSetCallInfo	2018-02-02					006f02a790029226						
15535	RequestDistributeEvent	2018-02-02	VQ_G_CENTRE15	V81LLL3HL			V_G_CENTRE15			4185	81LLL3H	25601	
15536	EventACK	2018-02-02 15:49:07.673	VQ_G_CENTRE15										
15537	EventAttachedDataChanged	2018-02-02 15:49:07.673	VQ_G_CENTRE15	V81LLL3HL			V_G_CENTRE15			4185	81LLL3H	25601	
15538	RequestUpdateUserData	2018-02-02	25901				006f02a790029226						
15539	RequestSetCallInfo	2018-02-02					006f02a790029226						
15540	EventAttachedDataChanged	2018-02-02 15:49:07.674	25901			25901	006f02a790029226			4185	81LLL3H	25601	
15541	RequestSetCallInfo	2018-02-02					006f02a790029226						
15542	RequestDistributeEvent	2018-02-02	VQ_G_CENTRE15	V81LLL3HL			V_G_CENTRE15			4185	81LLL3H	25601	

Données triées et nettoyées

Unnamed: 0.1	EventName	LocalTime	ThisDN	OtherDN	ConnID	DNIS
128	EventRingin	2018-01-31 23:43:30.468	25009	HV3O1V2O1	006f02a790028c34	25601.0
137	EventEstablished	2018-01-31 23:43:34.442	25009	HV3O1V2O1	006f02a790028c34	25601.0
881	EventReleased	2018-01-31 23:46:59.336	25009	HV3O1V2O1	006f02a790028c34	25601.0
142567	EventRingin	2018-02-02 07:58:24.035	25009	VTTLH202H	006f02a79002902c	25601.0
142580	EventEstablished	2018-02-02 07:58:24.573	25009	VTTLH202H	006f02a79002902c	25601.0
...
409966	EventEstablished	2018-04-02 23:53:56.719	25008	VOL8180TH	006f02a790039ab0	25601.0
409994	EventReleased	2018-04-02 23:54:52.192	25008	VOL8180TH	006f02a790039ab0	25601.0
410255	EventRingin	2018-04-02 23:57:29.018	25007	HVH8V0V0	006f02a790039ab3	25601.0

Nettoyage et traitement des données

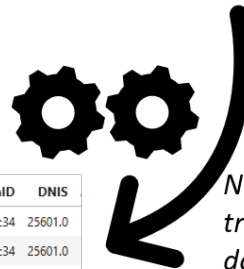


Figure 4.2 – Traitement des données pour obtenir un log exploitable.

EventNo	EventName	LocalTime	ThisDN	OtherDN	ThisQueue	ThirdParty DN	ConnID	Transfer Conn'n	Previous Conn'n	CallID	ANI	DNIS	Agent ID
15527	EventACK	2018-02-02 15:49:07.537	AgentStateDN										
15528	EventUserEvent	2018-02-02 15:49:07.537	AgentStateDN	AgentStateDN									
15529	EventTreatmentApplied	2018-02-02 15:49:07.672	25901		25901		006f02a790029226			4185	81LLL3H	25601	
15530	EventUserEvent	2018-02-02 15:49:07.672	25901		25901		006f02a790029226			4185	81LLL3H	25601	19
15531	RequestUpdateUserData	2018-02-02	25901				006f02a790029226						
15532	RequestSetCallInfo	2018-02-02					006f02a790029226						
15533	EventAttachedDataChanged	2018-02-02 15:49:07.673	25901			25901				4185	81LLL3H	25601	
15534	RequestSetCallInfo	2018-02-02					006f02a790029226						
15535	RequestDistributeEvent	2018-02-02	VQ_G_CENTRE15	V81LLL3HL	_G_CENTRE15		006f02a790029226			4185	81LLL3H	25601	
15536	EventACK	2018-02-02 15:49:07.673	VQ_G_CENTRE15										
15537	EventAttachedDataChanged	2018-02-02 15:49:07.673	VQ_G_CENTRE15	V81LLL3HL	_G_CENTRE15		006f02a790029226			4185	81LLL3H	25601	
15538	RequestUpdateUserData	2018-02-02	25901				006f02a790029226						
15539	RequestSetCallInfo	2018-02-02					006f02a790029226						
15540	EventAttachedDataChanged	2018-02-02 15:49:07.674	25901			25901				4185	81LLL3H	25601	
15541	RequestSetCallInfo	2018-02-02					006f02a790029226						
15542	RequestDistributeEvent	2018-02-02	VQ_G_CENTRE15	V81LLL3HL	_G_CENTRE15		006f02a790029226			4185	81LLL3H	25601	

Figure 4.3 – Extrait d'un log de données non traité.

unique. L'identifiant unique de notre appel est *006f02a790028c34*. C'est bien le même code qu'on retrouve pour chaque événement.

- **TransferConnID** : Identifiant de l'appel vers lequel la communication a été transférée. Lorsqu'un appel est transféré, par exemple, d'un ARM à un MR, il se peut que l'identifiant de l'appel change, car l'ARM peut mettre l'appel du patient en attente, composer un nouvel appel vers le MR puis transférer le patient sur ce nouvel appel avec le MR. Les deux appels ont alors des identifiants différents et ce champ permet de les relier. Il n'y a pas d'appels transférés dans notre exemple.
- **PreviousConnID** : Cet identifiant est utilisé dans le cas d'un transfert d'appel avec changement d'identifiant. Il correspond à l'identifiant de l'appel précédant depuis lequel le nouvel appel courant a été transféré (appel de consultation, par exemple lorsqu'un ARM appelle un MR afin de lui résumer l'appel). Dans notre exemple il n'y en a pas.
- **CallID** : Cet attribut contient l'identification de l'appel fournie par le commutateur, qui identifie un appel de manière unique. Contrairement au ConnID attribué par le serveur, le CallID est créé par le commutateur lorsque l'appel entrant arrive, ou lorsque les appels sortants de l'agent/du système sont créés. L'attribut doit être présent si le commutateur génère et distribue le paramètre correspondant au serveur (CallID est égal à zéro tant que le commutateur ne fournit pas cette information au serveur).
- **ANI** : Cela signifie *Automatic Number Identification* C'est le numéro de l'appelant. Dans notre exemple, les numéros des appelants ont été anonymisés c'est pourquoi nous avons un numéro sous la forme "HV3O1V2O1".
- **DNIS** : Cela signifie *Directory Number Identification System*, c'est le canal d'entrée utilisé par le numéro appelant, le numéro qu'il a composé pour joindre le centre d'appels. Ici le code est le 25601, qui correspond aux appels qui arrivent en ayant composé le 15.
- **AgentID** : L'identifiant de la ressource qui gère l'événement. L'appel a été géré par l'agent avec l'identifiant 25528. Si on fait le lien avec les tables de signification, on voit que cela correspond à un ARM.

Comme évoqué précédemment, dans cette table le champ **EventName** peut prendre plusieurs valeurs. De là, nous avons pu identifier les différents événements qui nous

EventNo	EventName	LocalTime	ThisDN	OtherDN	ThisQueue	ThirdPartyDN	ConnID	TransferConnID	PreviousConnID	CallID	ANI	DNIS	AgentID
103	RequestUpdateUserData	2018-01-31	25954										
104	RequestSetCallInfo	2018-01-31											
105	EventAttachedDataChanged	2018-01-31 23:43:30.266	25954			25954				1430	/3O1V2C	25601	
106	RequestSetCallInfo	2018-01-31											
107	RequestRouteCall	2018-01-31	25954	25009									
112	EventUserEvent	2018-01-31 23:43:30.266	25954			25528				1439			
114	EventRouteUsed	2018-01-31 23:43:30.266	25954	HV3O1V2O1	25954	25009				1439	/3O1V2C	25601	
115	EventDiverted	2018-01-31 23:43:30.266	25954	HV3O1V2O1	25954	25009				1439	/3O1V2C	25601	
123	RequestSetCallInfo	2018-01-31											
124	RequestSetCallInfo	2018-01-31											
125	RequestSetCallInfo	2018-01-31											
126	RequestSetCallInfo	2018-01-31											
127	RequestSetCallInfo	2018-01-31											
128	EventCallPartyState	2018-01-31 23:43:30.467											
129	EventRinging	2018-01-31 23:43:30.466	25009	HV3O1V2O1	25954					1439	/3O1V2C	25601	25528
138	EventEstablished	2018-01-31 23:43:34.442	25009	HV3O1V2O1	25954					1439	/3O1V2C	25601	25528
151	RequestUpdateUserData	2018-01-31	25954										
152	RequestSetCallInfo	2018-01-31											
153	EventAttachedDataChanged	2018-01-31 23:43:34.522	25009			25954				1430	/3O1V2C	25601	
154	RequestSetCallInfo	2018-01-31											
155	EventAttachedDataChanged	2018-01-31 23:43:34.522				25954				1430	/3O1V2C	25601	
882	EventReleased	2018-01-31 23:46:59.336	25009	HV3O1V2O1	25954					1439	/3O1V2C	25601	25528
883	EventReleased	2018-01-31 23:46:59.337AlcatelA4400::								1430	/3O1V2C	25601	

Figure 4.4 – Exemple de log avant traitement, sous forme de table pour un appel.

intéressent dans un log de téléphonie afin de donner un sens métier aux événements. Les informations sont tirées du document de références du logiciel de téléphonie¹.

- **Event Ringing** : l'évènement annonçant la survenue d'un appel (Sonnerie de l'appel).
- **Event Established** : l'évènement exprimant qu'un appel est décroché.
- **Event Abandoned** : l'évènement exprimant qu'un l'appel est abandonné.
- **Event PartyChanged** : L'objet de téléphonie spécifié par **OtherDN** a remplacé l'objet de téléphonie spécifié par **OtherDN** dans l'évènement reçu précédemment ; ou le **Previous ConnID** de l'appel s'est vu attribuer une nouvelle valeur de **ConnID**. Cela peut être le cas quand un appel est transféré, si le **Party** est modifié c'est que l'interlocuteur change. Cet évènement a lieu notamment quand un ARM transfère l'appel à un MR ou inversement.
- **Event Network Reached** : l'appel a rejoint le réseau distant. Cela concerne les appels sortants externes, c'est-à-dire les appels qui vont être passés en dehors du réseau du centre d'appels. Par exemple lorsqu'il faut contacter une agence d'ambulances ou encore les pompiers.
- **Event Held** : l'appel est mis en attente.
- **Event Retrieved** : l'appel est repris en ligne après avoir été mis en attente.
- **Event Party Added** : un ou plusieurs participants a été ajouté à l'appel à la suite d'une conférence. Si un seul participant est ajouté (comme dans le cas d'une simple conférence téléphonique), le nouvel interlocuteur correspondant est spécifié dans OtherDN.
- **Event Party Deleted** : l'objet de téléphonie spécifié par OtherDN a été supprimé de la conférence téléphonique en question. Cela signifie qu'une personne qui était dans la conversation téléphonique a quitté l'appel. C'est le cas lorsqu'un ARM à un patient en ligne, il crée une conversation téléphonique avec le MR. Il y a donc trois personnes dans la conversation, puis l'ARM se retire de la conversation laissant le patient avec le MR, c'est à ce moment ou on aura un **Event Party Deleted** avec en **OtherDN** le code de l'ARM.
- **Event Destination Busy** : l'appel sortant sonne occupé.

1. « Genesys Events and Models Reference, System Level Guides Current - 14-09-2018 »

- **Event Released** : l'appel est raccroché et terminé.
- **Event Error** : Une erreur est arrivée dans le log.

Une première étape qui concorde avec le nettoyage et la préparation des données et d'identifier les événements de début et les événements de fin possibles. En effet, une trace pour être dite complète doit avoir un début et une fin identifiables sinon c'est une trace incomplète qui ne sera pas prise en compte.

- **Event Ringing** est un événement de départ pour les appels entrants.
- **Event Dialing** est un événement de départ pour les appels sortants.
- **Event Released** est un événement de fin pour un appel où la communication a pu être établie.
- **Event Destination Busy** est un événement de fin pour les appels sortants qui ne sont pas décrochés, car l'appelé est occupé.
- Après un **Event Established** on peut avoir une **Event Released**.

Nettoyage et préparation des données

L'objectif de cette partie est d'effectuer un prétraitement afin d'avoir une base de données complétée qui nous permette facilement de faire des analyses. Pour ce faire, plusieurs étapes ont été suivies :

- Seulement un certaines colonnes ont été conservées (**EventNo**, **EventName**, **LocalTime**, **ThisDN**, **OtherDN**, **ConnID**, **DNIS**, **AgentID**, **ANI**). La colonne **CallID** n'a pas été conservée, car elle était redondante avec l'identifiant du **ConnID**. Les colonnes **ThisQueue** et **ThirdPartyDN** sont assez peu utilisées et ne nous apportent pas d'information que nous sommes en mesure d'exploiter. De ce fait ces colonnes ont été écartées.
- Les données sont triées par ordre chronologique.
- On nettoie les données, on supprime les lignes qui n'ont pas de **ConnID**, celles qui n'ont pas un **LocalTime** complet ainsi que les événements qui ont été identifiés comme non pertinents (**EventCallPartyState** et **EventAttachedDataChanged**) qui sont des événements que nous n'exploitons pas. Enfin on supprime les doublons. Les données brutes sont composées de 69360 appels et 9890976 événements sur les deux mois (pour le SAMU d'Albi). Après traitement des données, on obtient un jeu de données contenant 56918 appels soit 82% des données extraites initialement. Cela nous assure un nombre toujours significatif d'appels qui ne compromet pas la pertinence de notre analyse.
- On complète les données avec le jour de la semaine, cela nous permettra par la suite de faire rapidement un tri pour étudier des jours spécifiques de la semaine.
- On complète les données avec le rôle de la ressource qui effectue l'action. En effet pour chaque événement nous avons l'identifiant de l'agent qui effectue la ressource, on vient compléter cette information en renseignant si cette ressource est un ARM ou un MR, cela nous permet de donner plus de sens métier aux données.
- On transforme les données en log en les regroupant selon leur **ConnID**.

- On crée une nouvelle colonne *Type Appel* dans laquelle on va renseigner si l'appel est un **appel entrant**, un **appel interne** ou un **appel sortant**. Pour ce faire, on effectue un tri sur les appels entrants (uniquement les appels qui ne commencent par « EventRingring »). On affecte à ces derniers le type d'appel entrant. On tri les appels qui commencent par *Event Dialing*, ceux avec un événement *Event Network Reached* sont des appels sortants externes, les autres sont des appels sortants internes.

Ce premier traitement nous permet d'avoir une base de données structurée et exploitable sur laquelle il sera aisé d'appliquer d'autres algorithmes de traitement afin d'en extraire la connaissance utile à notre analyse. Cette étape est essentielle au bon déroulement des suivantes.

4.1.3 Sous-étape 1.b : Découvrir le processus métier

Cette étape correspond à la partie 3.2.1. Elle peut être effectuée en parallèle du traitement des données. L'objectif ici est d'obtenir de l'information sur les processus métiers. Des observations et interviews dans les centres d'appels du SAMU ont été réalisées à cette fin.

Méthodologie pour les entretiens

La phase d'étude terrain doit permettre de valider et préciser les enjeux métiers d'un point de vue opérationnel. De ce fait, des observations et entretiens ont été réalisés dans les SAMU partenaires : Rodez et Toulouse. Le SAMU de Carcassonne et celui de Nantes ont également fait l'objet d'une visite pour appréhender la diversité des organisations mises en place. Le planning suivant a été réalisé dans les SAMU de Rodez et de Toulouse :

- 8h pour la présentation du projet, des enjeux et du planning d'entretien.
- 8h d'observation et entretien avec un Médecin Régulateur (MR).
- 8h d'observation et entretien avec un Assistant de Régulation médicale (ARM).
- 4h d'observation et entretien avec le Gestionnaire des Moyens (GdM).
- 4h d'entretien avec le responsable du système d'information et de la téléphonie.

On peut rajouter 4h au SAMU de Nantes et 4h à celui de Carcassonne pour comprendre leur fonctionnement. Les observations sont plus succinctes et moins détaillées pour Nantes et Carcassonne. Ces visites ont tout de même permis de constater des modes de fonctionnement différents.

Particularité des systèmes d'information

Observations Toulouse

Le système de téléphonie du SAMU 31 dispose d'une interface informatisée. De cette manière les ARM peuvent décrocher, raccrocher, transférer les appels depuis l'écran de l'ordinateur. De plus, l'interface permet de visualiser les appels en attente

ainsi que de la disponibilité de chacun des acteurs (ARM et MR). Cela permet de rapidement voir qui est déjà en ligne et le nombre d'appels qu'ils ont dans leur salle d'attente personnelle. Les appels arrivent dans une salle d'attente commune et peuvent être décrochés par tous les ARM. Ces derniers choisissent les appels qu'ils souhaitent décrocher. Lorsqu'ils doivent transférer un appel à un MR, ils le choisissent de manière arbitraire. De la même manière, le MR choisit l'appel qu'il souhaite décrocher sans priorité établie par le système. De plus, il y a une visibilité sur la nature de l'appelant. Si c'est une personne dont le numéro est enregistré dans le système (médecin, ARM, SAMU de la région voisine, équipe pompier, etc.), l'identification apparaîtra en information complémentaire au numéro appelant. De même, si ce numéro a déjà appelé récemment et qu'un dossier médical a été créé, le numéro de dossier s'affiche à l'écran.

Observations Rodez

Le système de téléphonie de Rodez n'est pas interfacé informatiquement pour le décrocher des appels. MR et ARM décrochent les appels « à l'ancienne » avec le combiné. Il n'y a pas de visibilité sur les appels en attente. La seule information disponible est un voyant lumineux qui clignote s'il y a plus d'un appel en attente. Il n'y a pas de visibilité sur les numéros connus. Deux téléphones doivent être gérés, celui qui reçoit les appels externes et celui qui reçoit les appels internes. Les files d'attente sont gérées avec un mode FIFO (First In, First Out), cela signifie que les premiers appels qui entrent dans la file d'attente seront les premiers appels à être décrochés.

Observations Carcassonne

À Carcassonne la téléphonie est mutualisée avec celle des pompiers. Ainsi lorsqu'un appel entrant 15 n'est pas décroché par un des ARM, car ils ne sont pas disponibles, l'appel est décroché par un opérateur pompier, et inversement. Sinon le fonctionnement est le même qu'à Toulouse. À savoir, une salle d'attente commune dans laquelle les ARM viennent décrocher les appels.

Observations Nantes

Le SAMU de Nantes est équipé d'un système de téléphonie similaire à celui de Toulouse avec une interface informatisée. Les appels sont décrochés automatiquement et ce n'est pas l'ARM qui décroche. Donc cela signifie que les appels sont routés automatiquement vers des opérateurs suivant des règles prédéfinies (que nous ne connaissons pas). Ce qui n'est pas le cas à Toulouse ni à Rodez.

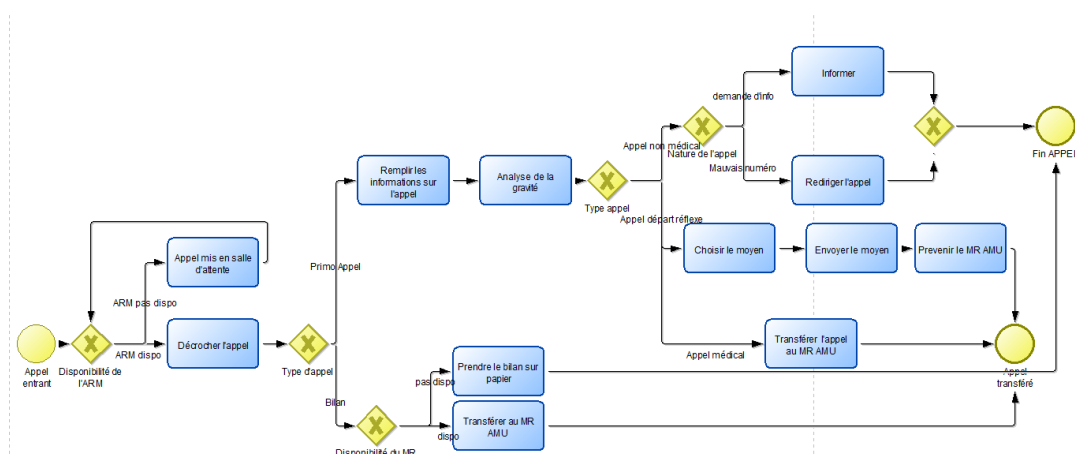
Dès cette première partie, on constate une hétérogénéité au niveau du matériel et de facto au niveau des pratiques de fonctionnement. Des méthodes de décrochés des appels différentes, automatisée à Nantes, en FIFO à Rodez, au choix à Toulouse. Une visibilité sur les appels en attente qui n'existe pas forcément. De plus les fournisseurs de téléphonie ne sont pas les mêmes d'un département à l'autre, et ils travaillent tous sur des logiciels différents.

Tableau 4.2 – Caractéristiques des systèmes d'information dans les SAMU visités

	Toulouse	Rodez	Nantes	Albi	Carcassonne
Logiciel téléphonie	MediQ V2	CCD Alcatel +	n.d	MediQ V2	Systel
Logiciel régulation	AppliSamu	Centaure 15	AppliSamu	Centaure 15	Systel
Règles files attentes	au choix.	FIFO	n.d	n.d	n.d
Message d'accueil	oui	oui	oui	oui	oui
Décroché automatique	non	non	oui	non	n.d

Processus de traitement des appels

Dans chacun des centres visités, lorsque cela était possible nous avons observé et modélisé les processus de réception et de traitement de l'appel. L'étude de ces processus a été effectuée dans les SAMU de Rodez et de Toulouse pour les comparer. Si l'on compare les deux processus globaux de Rodez (cf. Figure 4.5) et Toulouse (cf. Figure 4.6), on constate une première différence évidente qui est la présence d'une ressource supplémentaire dans l'organisation à Toulouse. En effet, un ARM est affecté à temps plein sur la gestion des moyens. C'est un ARM qui ne décroche pas d'appel, mais qui gère les réservations d'ambulances, l'envoi des équipes SMUR et le contact avec les pompiers pour envoyer leurs équipes.

**Figure 4.5** – Processus de la prise d'un appel d'urgence à Rodez

Si l'on regarde de plus près les processus d'évaluation de l'appel à Rodez et à Toulouse, il apparaît que les ARM de Toulouse décrochent aussi les appels de la permanence des soins afin de les transférer à un médecin généraliste en régulation, tandis qu'à Rodez une différence est faite entre les ARM de l'aide médicale urgente et ceux de la permanence des soins. Les deux sont gérés de manière indépendante. La manière de récupérer les bilans d'interventions est complètement informatisée à Toulouse, alors que l'utilisation du papier est encore très présente à Rodez. La gestion des appels en attente est différente. En effet, à Toulouse, il y a une visibilité sur les appels en attente. On sait depuis combien de temps ils attendent. Pour chaque appel, on sait aussi s'il est en lien avec un dossier médical déjà ouvert. Les ARM peuvent choisir l'appel qu'ils souhaitent décrocher. Il n'y a pas de critères officiels pour la priorisation du décroché. Intuitivement, les appels qui ont attendu le plus longtemps

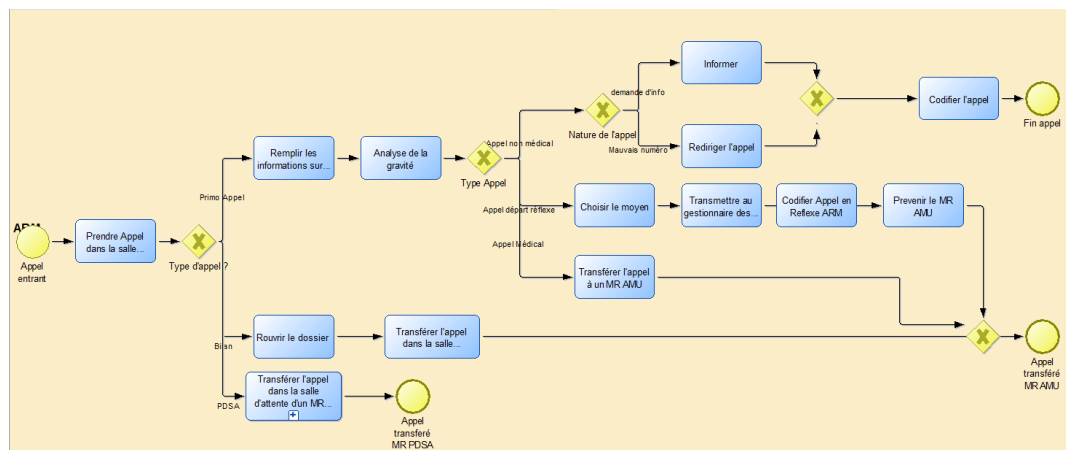


Figure 4.6 – Processus de la prise d'un appel d'urgence à Toulouse

sont décrochés en priorité, ce qui souligne l'intérêt de modéliser d'une manière fine les différents états d'un appel. Cela permet en effet de voir combien de temps un appel peut rester en attente pour définir un indicateur de temps maximum d'attente ou de nombre d'appels pris avant un certain temps. À Rodez, le logiciel de téléphonie ne permet pas la visualisation de l'attente des appels. On ne sait pas combien d'appels sont en attente d'être décrochés ni depuis combien de temps.

Bien que les grandes étapes de la régulation d'un appel soient les mêmes d'un SAMU à un autre, lorsqu'on les étudie individuellement, on constate qu'il existe des différences sensibles. En effet, en fonction des outils informatiques disponibles, la manière de décrocher l'appel ou suivant les spécialisations des ressources, les étapes et modalités associées peuvent varier.

4.2 Étape 2 : Interpréter et intégrer

Une fois que l'on dispose à la fois des données de téléphonie nettoyées ainsi que d'une vision précise des processus métier, il faut rassembler et lier ces informations afin d'en extraire des connaissances. Pour ce faire, nous préconisons l'utilisation d'une méthode d'exploration automatique de processus (*Process Mining*).

Le Process Mining

Le Process Mining consiste à extraire de la connaissance d'un registre d'événements appelé *Event Log*, sauvegardé dans un système d'information. Il se place dans les méthodes *Evidence-base* pour la découverte de processus. Le *Process Mining* se base sur trois activités afin de générer de la connaissance : la découverte du processus, le contrôle de conformité et l'amélioration du processus métier (Van Der Aalst, 2015).

L'histoire du *Process Mining* commence avec les travaux de Agrawal et al. (Agrawal et al., 1998). Ils développent un algorithme afin de générer des graphes de flux à

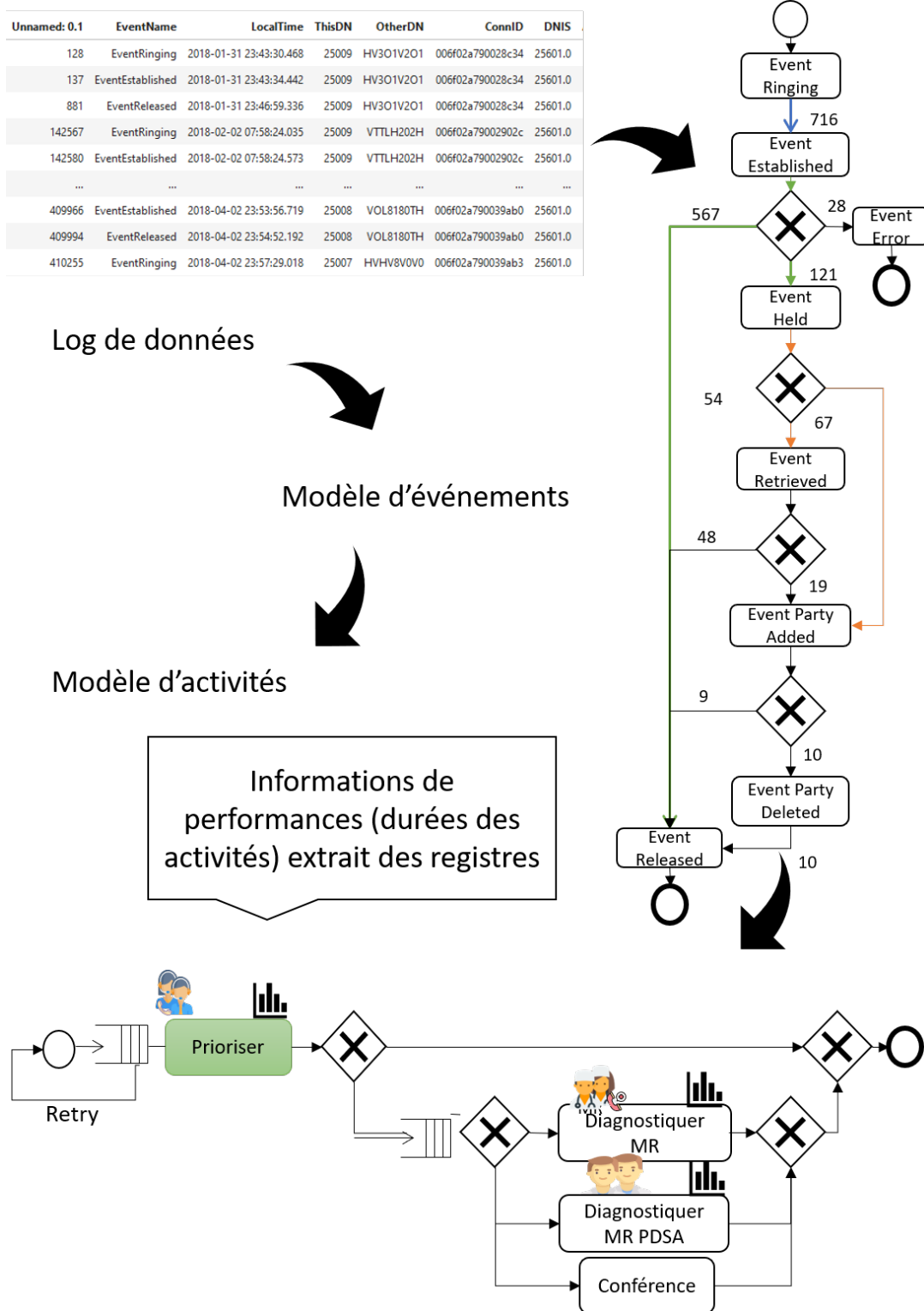


Figure 4.7 – Utilisation du Process Mining pour nos données

partir de l'exécution d'un processus enregistré sous la forme de logs de workflows. Ces travaux ont ensuite été repris et complétés par Weitjers et Van der Aalst qui ont proposé une technique de découverte et de validation de processus capable de gérer les cas particuliers et de valider le processus de flux de travail en mesurant l'écart entre le modèle et l'exécution réelle du processus.

Les processus dans le domaine de la santé sont connus pour être fortement dynamiques, complexes et faisant intervenir de nombreux acteurs de disciplines variées (Homayounfar, 2012). C'est pourquoi le *Process Mining* est de plus en plus utilisé pour analyser les processus en utilisant les données disponibles sur les systèmes d'information des établissements de santé. Cette approche présente l'avantage de pouvoir répondre à plusieurs types de questions tels que :

- Qu'est ce qu'il s'est passé ? : Permettre d'identifier le processus et ses activités. (par exemple, quel est le traitement typique d'un appel ?)
- Pourquoi est-ce que cela est arrivé ? : Comprendre les enchaînements d'activités et leurs conséquences (par exemple, pourquoi est-ce qu'un appel a attendu avant d'être décroché ?)
- Que va-t-il se passer ? : Identifier les circonstances dans lesquelles une activité arrive dans le but de la prévoir (par exemple, est-ce que cet appel va attendre avant d'être décroché ?)

Le *Process Mining* a été utilisé dans différents domaines au sein des systèmes de santé. C'est en oncologie et en chirurgie qu'on retrouve le plus grand nombre d'études à ce propos. Dans le domaine de l'urgence, Rojas et al. identifient 3 études seulement dans leur revue de littérature (Rojas et al., 2016).

4.2.1 Sous-étape 2.a : Fusionner les modèles

Les données de téléphonie nous procurent des informations sur les heures d'arrivée de chaque appel et les différents événements qui en découlent. Toutefois, nous n'avons pas directement l'information sur la durée de chaque activité. C'est pourquoi il faut établir une correspondance entre les suites d'événements et les activités qu'elles représentent.

Quelques exemples d'appels et d'enchaînements d'événements sont présentés dans les figures 4.8 et 4.9 pour chaque type de trace nous avons identifié le nombre de fois où elle est retrouvée dans le log de données ainsi que les enchaînements de temps d'attente ou de communication. On constate qu'il y a des briques constitutives qui se dessinent, des enchaînements qui se répètent et qui correspondent à la même chose.

- Une première brique est l'enchaînement *Event Ringing* \implies *Event Established* qui correspond à un appel qui contacte le centre d'appels et qui réussit à le joindre, le temps qui s'écoule entre les deux événements est un temps d'attente.
- Une deuxième est l'enchaînement *Event Ringing* \implies *Event Released* qui traduit une tentative d'appel non concluante. Soit l'appelant ne rappelle pas, et cela s'arrête là, soit l'appelant essaye de rappeler et on aura de nouveau un bloc *Event Ringing* \implies *Event Released*, dans le cas d'un échec, soit un bloc *Event Ringing* \implies *Event Established* dans un cas de réussite.

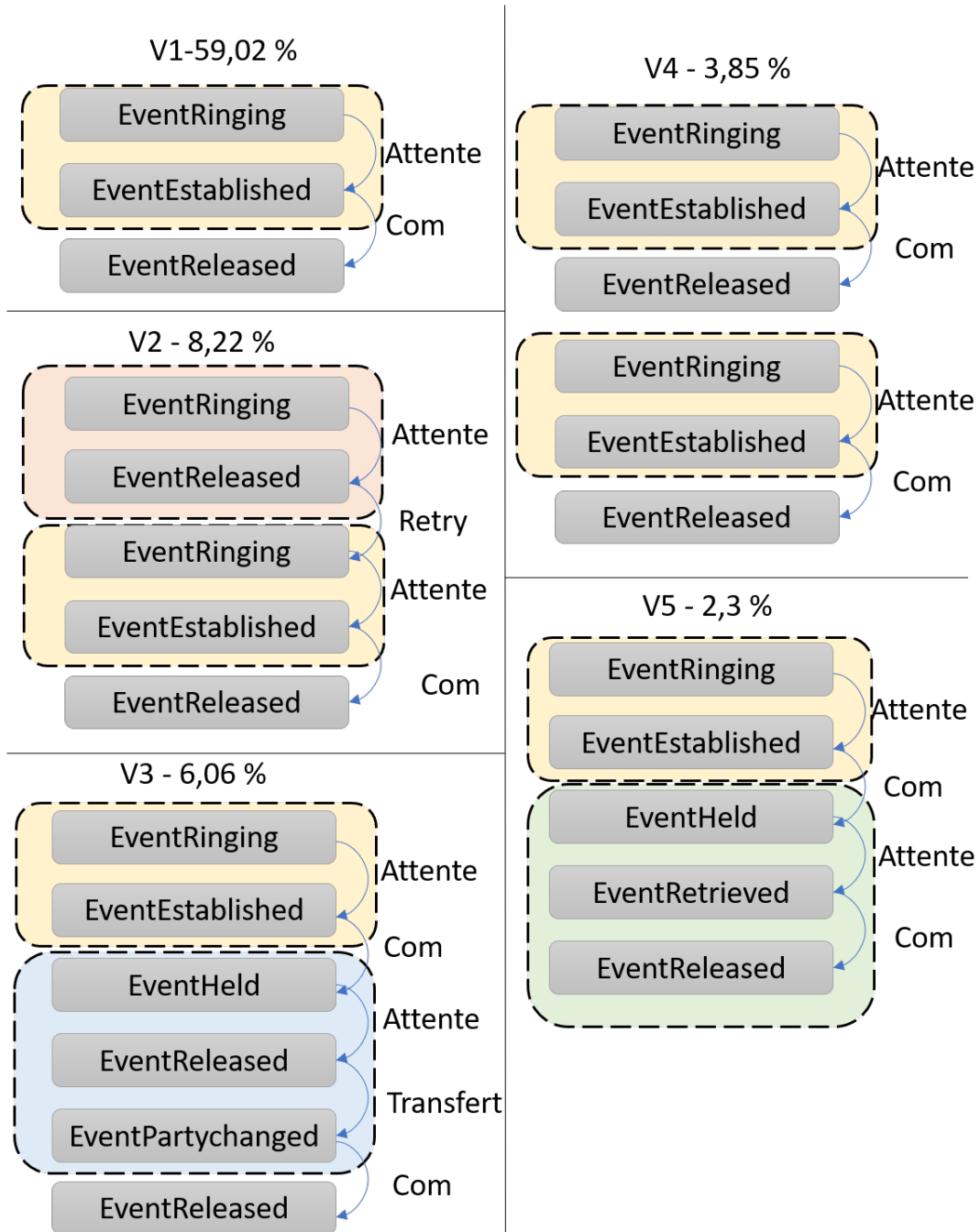


Figure 4.8 – Quelques exemples de traces (Partie 1)

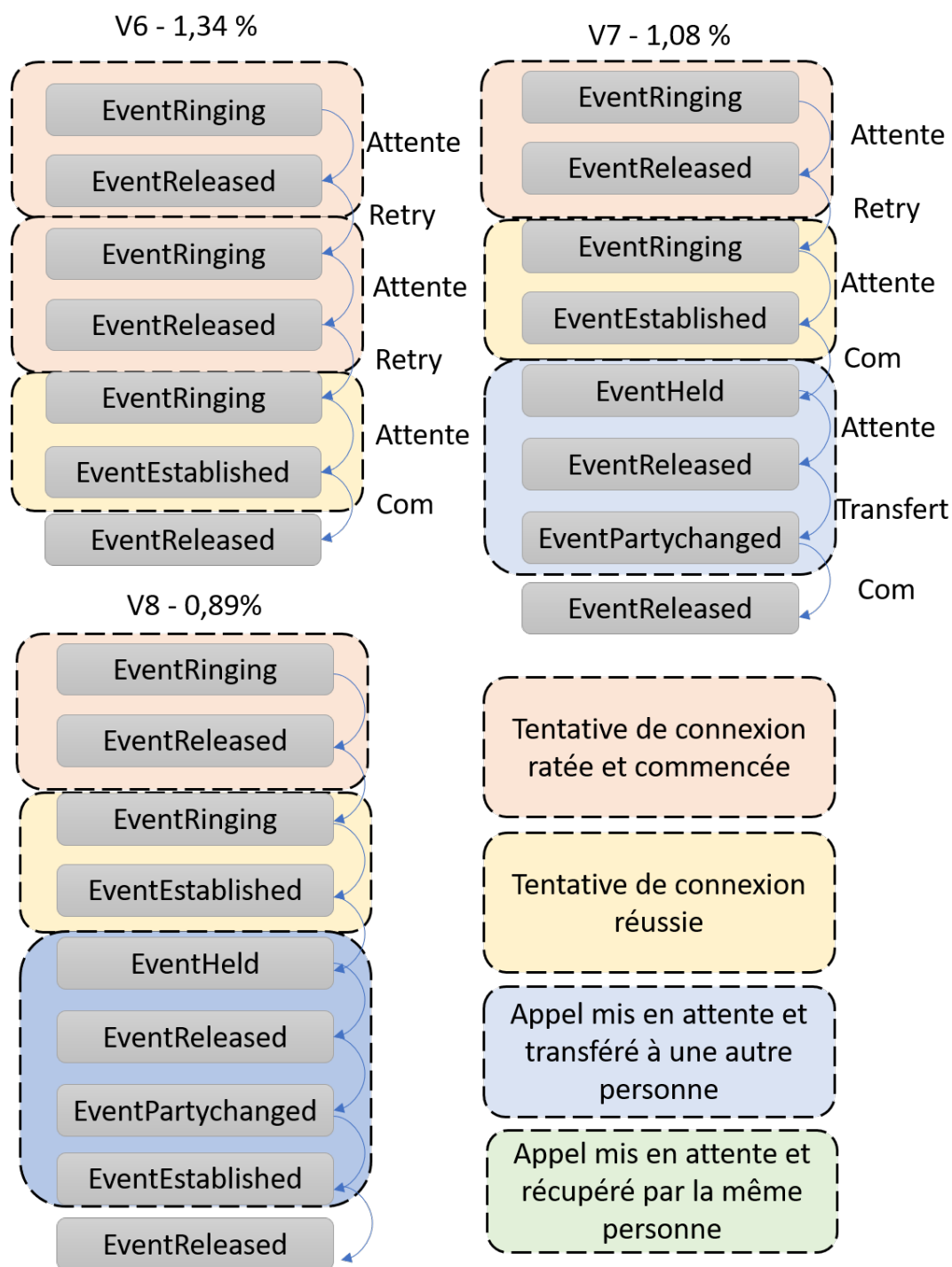


Figure 4.9 – Quelques exemples de traces (Partie 2)

- Lorsque les appels sont mis en attente il y a plusieurs cas possibles, soit l'appel est mis en attente et repris par la même personne plus tard alors on aura un comme dans le cas V5, *Event Held* \implies *Event Retrieved* \implies *Event Released*. Soit l'appel est transféré à une autre personne et on aura plutôt comme dans le V3, *Event Held* \implies *Event Released* \implies *Event Party Changed*, ce qui correspond à une mise en attente de l'appel, *EventReleased* suivi de *EventPartyChanged* implique que le premier interlocuteur raccroche, et que l'interlocuteur change. On notera que le fait de transférer l'appel à un interlocuteur est plus courant (6,06% des appels) que le fait de mettre en attente et de le reprendre plus tard (2,3% des appels).

Ces briques constitutives nous permettent de former des traces différentes, en effet il peut y avoir plusieurs tentatives de connexion qui échouent avant de réussir à joindre le centre d'appels, l'appel peut être mis en attente plusieurs fois. Ou encore, comme dans la trace V4 de la figure 4.8, un appel qui va être pris en charge une première fois par une ressource puis qui va rappeler et être pris en charge une deuxième fois.

Ces événements sont très précieux, car ils vont nous permettre d'identifier les différentes activités de nos processus métier. En fonction de l'enchaînement d'événements et du type de ressource qui est impliquée, ce ne sera pas la même activité. Par exemple, l'enchaînement d'un *Event Established* et d'un *Event Released* peut correspondre à un temps de communication avec l'ARM ou avec le MR, cela va dépendre du type de ressource impliquée. Si on a un enchaînement *Event Retrieved* et *Event Released* réalisé par un MR, alors c'est une activité de régulation par un MR après que l'appel ait été mis en attente. Ainsi, nous devons définir des règles métiers pour être capable de transformer nos enchaînements d'événements en activités qui ont du sens d'un point de vue métier.

4.2.2 Sous-étape 2.b : Extraire les connaissances

Les Centres d'appels peuvent être vus comme un ensemble de files d'attente.

Un centre d'appels va être constitué de k lignes de téléphones qui permettent à un appel de joindre le centre. Ces lignes vont être gérées par $w \leq k$ postes de travail sur lesquels vont être assignés $m \leq k$ agents (cf. figure 4.10). Lorsqu'un appel arrive, trois scénarios sont possibles :

1. L'appel est distribué directement si les agents sont disponibles.
2. L'appel est placé dans une file d'attente si les agents ne sont pas disponibles.
3. L'appel est bloqué s'il n'y a pas de lignes disponibles, toutes les lignes sont déjà occupées par un appel, la file d'attente ne peut pas accueillir plus d'appels.

Dans un tel système, les ressources sont immédiatement libérées dès qu'un appel a été servi. Il existe différents types d'arrivées (ou d'appelants), à savoir ceux qui établissent un premier contact, ceux qui recomposent en raison d'une tonalité d'occupation et ceux qui abandonnent la file d'attente en attendant. Pour caractériser ce système, la

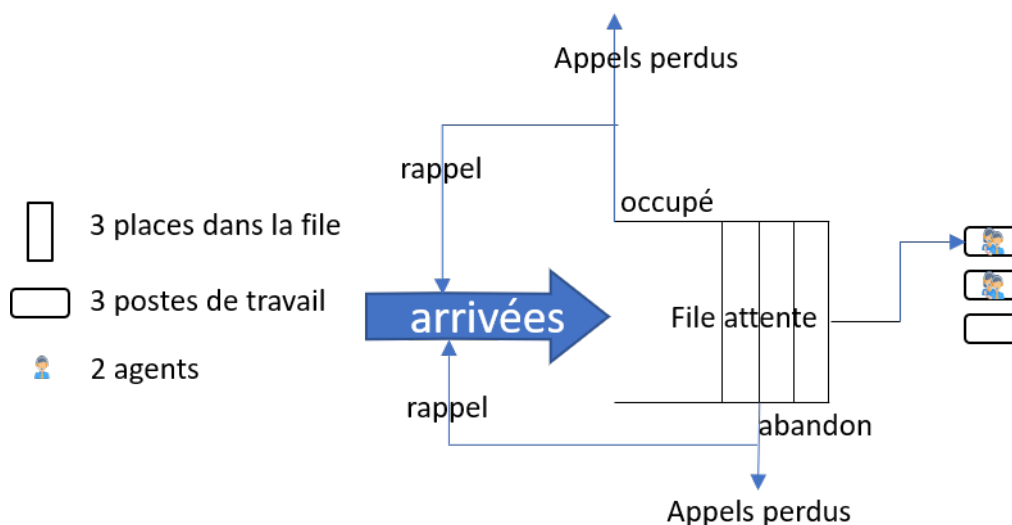


Figure 4.10 – Représentation d'un centre d'appels par file d'attente. (Gans et al., 2003)

notation de Kendall a été inventée (Kendall, 1953). Celle-ci décrit la file d'attente. La notation de Kendall est aujourd'hui généralement exprimée par les termes suivants 1/2/3/4/5/6, où chaque numéro représente :

1. Le processus d'arrivée des appels, c'est-à-dire la manière dont ils se répartissent dans le temps sur un créneau donné (généralement suivant une loi exponentielle, exprimé par un M).
2. La répartition du temps de service, la loi qui régit le temps de réponse de l'appel (généralement suivant une loi exponentielle, exprimé par un M).
3. Le nombre de serveurs, qui correspond au nombre de ressources disponibles pour décrocher.
4. Le nombre de troncs² (infini si l'on ne tient pas compte des troncs).
5. La taille de la population appelante (infinie si on l'ignore).
6. La modalité de gestion des files d'attente, généralement premier entré, premier sorti (FIFO).

La file d'attente la plus simple est la file M/M/1, qui est composée de l'arrivée des appels suivant une loi de Poisson, des temps de service suivant une distribution exponentielle aléatoire, et un serveur (soit une seule ressource pour décrocher). Cependant, un tel modèle est le plus souvent une simplification excessive, comme le montre le nombre des reprises possibles dans la figure 4.10 (Gans et al., 2003).

L'arrivée des appels est un grand vecteur d'aléatoire. Le processus d'arrivée des appels est généralement modélisé par une loi de Poisson (Saaty, 1961). La loi de Poisson est une loi de probabilité qui s'applique aux événements rares et indépendants les uns des autres. Ce qui est le cas pour les appels d'urgence. Si une personne passe un appel d'urgence, cela est généralement fait indépendamment de si son voisin vient d'appeler les urgences. Des cas existent néanmoins comme les attentats.

2. Cela correspond à la taille de la file d'attente et le nombre d'appels maximal qui peut être en simultané dans le centre d'appels.

Modéliser les temps de communication

Le temps de communication ou temps de service a été défini par Mandelbaum et Zeltyn (Zeltyn et al., 2005) Comme étant « le temps qu'un agent passe à gérer un appel ». Dans les faits, cela inclut le temps de communication passé avec l'appelant, mais aussi le temps de gestion pouvant être occasionné après l'appel. Ce temps de gestion peut correspondre à la complétion du dossier du patient une fois l'appel terminé. Dans la plupart des cas, ce temps de traitement moyen est une distribution exponentielle. Ce choix de modélisation présente certaines caractéristiques qui permettent de simplifier les calculs dans certains modèles de files d'attente.

Un problème commun dans les centres d'appels est d'anticiper le nombre d'agents qu'il faut planifier à chaque moment de la journée. Ce nombre d'agents doit permettre de satisfaire une attente minimale pour l'appelant. Ces décisions découlent directement du nombre d'appels et de la manière dont ils sont repartis durant la période à planifier, donc de la prévision du nombre d'appels. Une approche assez commune pour déterminer le nombre de ressources nécessaires pour atteindre une qualité de service cible est la formule d'Erlang. (Takacs, 1969)

Créé en 1917 par Agner Erlang, la formule d'*Erlang-C* est utilisée pour déterminer le nombre d'opérateurs nécessaire pour atteindre une certaine qualité de service dans les centres d'appels. Ses avantages découlent du fait qu'elle est facile à utiliser et à coder dans un logiciel ou même dans une feuille de calcul. Elle est très utile, car très peu gourmande en termes de ressources pour le calcul des plannings. Ce qui à l'époque n'était pas négligeable.

Toutefois cette formule fonctionne sous quelques hypothèses restrictives : tous les appels sont considérés comme décrochés, ainsi on ne prend pas en compte les appels abandonnés. On considère aussi que tous les agents peuvent décrocher tous les appels, donc on ne prend pas en compte les compétences de chacun.

Pour calculer la probabilité qu'un appel soit mis en attente la formule d'Erlang C s'appuie sur plusieurs données :

- Le nombre d'agents N .
- Le temps de service moyen .
- Le nombre d'appels par intervalle de temps.
- La charge A est définie comme le temps qu'il faudra pour traiter tous les appels sur l'intervalle donné. Mettons que sur un créneau de 30 minutes on reçoive 200 appels avec un temps moyen de prise en charge de 3 minutes, alors la charge vaudra 600 appels minute. La charge s'exprime habituelle en Erlang. Un Erlang correspond au temps nécessaire pour gérer les appels durant une heure, dans notre exemple, la charge vaut alors 10 Erlangs (600 *appels minutes*/60 *minutes*).

Ce qui nous permet de calculer la formule d'Erlang-C pour obtenir la probabilité qu'un appel entrant ne soit pas décroché immédiatement, car tous les agents sont déjà en ligne :

$$X(N, A) = \frac{\frac{A^N}{N!} \frac{N}{N-A}}{\frac{A^N}{N!} \frac{N}{N-A} + \sum_{m=0}^{N-1} \frac{A^m}{m!} + \frac{A^N}{N!}} \quad (4.1)$$

Une fois la probabilité pour qu'un appel ne soit pas décroché calculée, on peut l'utiliser pour calculer la qualité de service estimée dans le centre appel :

$$QualitédeService = 1 - [X(N, A) * e^{-[(N-A)*(\frac{QoS_{cible}}{AHT})}] \quad (4.2)$$

Si la qualité de service n'atteint pas la cible alors il faut rajouter une ressource et refaire le calcul pour voir si l'ajoute de cette ressource permet d'atteindre l'objectif de qualité de service. De cette méthode, il est possible de calculer le nombre d'agents nécessaires sur chaque intervalle j durant une journée i pour chaque créneau. Pour cela il faut avoir la charge $A_{i,j}$ et le temps moyen de communication $S_{i,j}$ sur chacun des créneaux pour faire le calcul.

Il s'agit d'une approche simplifiée qui prédit N dans une file d'attente $M/M/N$. Par conséquent, comme soulignée précédemment, la formule ne tient pas compte des appels abandonnés avant d'être décrochés ou encore des appels qui sont mis en attente durant un certain temps. Dans un tel cas, le processus d'arrivée est supposé être de Poisson avec des temps d'arrivée indépendants distribués de manière exponentielle négative. Avoir des temps d'arrivée indépendants signifie que chaque appel est indépendant du précédent est distribué aléatoirement sur le créneau défini. Les temps de service sont également supposés être distribués de manière exponentielle, et les agents N sont tous statistiquement identiques. Cela signifie que l'on considère que tous les ARM sont identiques et ont les mêmes performances. Outre la formule Erlang-C, Erlang a également produit la formule de la probabilité qu'un appel soit bloqué dans une file d'attente $M/M/N/b$, avec un nombre b de places dans la file d'attente comme la formule Erlang-B (Brockmeyer et al., 1948).

Ces modèles d'Erlang reposent sur des hypothèses fortes qui sont largement contestées. L'hypothèse de la distribution exponentielle notamment a été remise en question ; par exemple, certaines études ont montré que la distribution log-normale permet une bien meilleure approximation des temps de session (Brown et al., 2005 ; Bolotin, 2013).

Afin de modéliser notre centre d'appels, nous avons donc besoin d'identifier et de définir le modèle de file d'attente que nous souhaitons utiliser. C'est pourquoi dans les sections suivantes nous détaillerons nos choix de modélisation pour l'arrivée des appels, la distribution des temps de communication et la manière dont on paramètre la file d'attente.

Générer la distribution des temps de session des appels

L'objectif ici est d'avoir un modèle qui se rapproche le plus possible du comportement réel du SAMU. Comme nous disposons de 26000 appels pour établir les distributions réelles, c'est une quantité qui est statistiquement significative et qui nous permet de nous baser sur ces données pour notre modèle.

Les statistiques descriptives sur les durées d'activités sont des données importantes comme les durées minimums, moyennes, maximum, mais aussi les occurrences de chaque activité. Il est possible de calculer ces données grâce au *Process Mining*. En effet, une fois les activités identifiées, il est possible d'analyser la durée de chacune d'elle. De ce fait, que ce soit pour le temps d'attente des appels ou encore pour le temps de session des communications avec l'ARM, le MR ou le MR PDSA, il est possible de faire des statistiques sur la durée de ces appels et d'évaluer la performance actuelle des centres d'appels.

Toutefois, nous avons quand même voulu identifier la distribution théorique la plus proche de nos données réelles. Pour ce faire nous avons calculé les temps de communication de l'ARM et du MR grâce au *Process Mining*. Une fois les durées récupérées, nous construisons un histogramme pour chaque activité de communication du processus. Puis nous appliquons un test d'adéquation (*Goodness of Fit test*) (D'Agostino, 1986). C'est un algorithme de correspondance avec lequel on applique des distributions théoriques sur notre modèle réel. On cherche le meilleur jeu de paramètres pour minimiser la distance entre les données réelles et la courbe théorique.

Bien que cette procédure d'identification soit assez simple, elle soulève plusieurs questions :

- Comment choisit-on le nombre de classes de notre histogramme ? En effet, le nombre de classes aura un impact sur la forme de notre histogramme et sur le résultat de l'algorithme de correspondance.
- Comment mesure-t-on la distance entre le modèle théorique et nos données réelles pour évaluer « la meilleure distribution théorique » ?

Pour le choix du nombre de classes (et in fine de la largeur de celles-ci), il existe plusieurs règles dans la littérature scientifique. Dogan et al. détaillent les différentes méthodes et leurs usages dans leur revue de littérature sur le sujet (Dogan et al., 2010). Dans notre cas, comme nos données ne sont pas symétriquement et normalement distribuées, nous ne pouvons pas utiliser la règle de Sturge (Sturges, 1926) qui est la plus commune. Ainsi nous allons utiliser la règle de Rice³ qui consiste à prendre un nombre de classes égal à deux fois la racine cubique du nombre d'observations (cf. équation 4.3).

$$k = 2\sqrt[3]{n}, \text{ où } k \text{ est le nombre de classe, } n \text{ est le nombre d'observation} \quad (4.3)$$

Pour la distance entre le modèle théorique et la distribution réelle, nous avons utilisé la SSE (*Sum Squared Error*) soit la somme des erreurs au carré. Plus celle-ci est proche de zéro et plus la distribution théorique correspond à la distribution empirique. La SSE permet d'avoir une comparaison entre les distributions théoriques. Ainsi cela permet d'identifier la distribution qui est la moins mauvaise par rapport à d'autres, mais pas nécessairement la meilleure dans l'absolu.

3. Online Statistics Education : A Multimedia Course of Study (<http://onlinestatbook.com/>). Project Leader : David M. Lane, Rice University

La figure 4.11 présente la distribution réelle des appels (histogramme, en bleu) ainsi que les distributions théoriques log-normale et exponentielles dont les SSE sont respectivement $1,47e-05$ et $2,18e-05$ pour la communication des ARM au SAMU d'Albi. Le tableau 4.3 apporte des informations complémentaires sur les statistiques descriptives de ces données. Ainsi la distribution log-normale correspond bien mieux à nos données réelles que la distribution exponentielle ce qui corrobore les résultats de Brown et al. (L. Brown et al., 2005). De ce fait notre jeu de données ne valide pas les hypothèses de la formule d'Erlang concernant la distribution exponentielle de la durée des appels. Le choix pour modéliser la durée des appels va être important surtout dans le chapitre suivant dédié au double numérique pour que notre modèle soit le plus similaire possible à la réalité.

En conclusion, nous avons décidé de travailler avec la distribution réelle des appels que ce soit pour les communications avec les ARM ou avec les MR. Toutefois, des recherches supplémentaires pourraient être effectuées pour identifier les lois de distribution théoriques les plus appropriées pour les différentes étapes de communication d'un appel d'urgence, bien que la distribution log-normale semble déjà apporter un niveau de précision satisfaisant.

Profil d'arrivée des appels : Est-ce que l'arrivée des appels au SAMU est prévisible ?

Une des caractéristiques de l'urgence est qu'elle n'est pas planifiée. C'est pourquoi, on lui associe souvent un fort caractère stochastique, aléatoire et dépendant de la saisonnalité. Nous avons voulu vérifier ces hypothèses et voir si elles étaient valables pour nos cas d'études. Nous avons exploré les données de Rodez et Albi sur 8 semaines complètes (février et mars). La figure C.1 correspond au graphique de ce résultat pour Albi (le même graphique est disponible en annexe pour Rodez). Plusieurs constats peuvent être faits :

- On constate une différence en termes de volumétrie entre les jours de semaine et les fins de semaine. En effet, les samedi et dimanche se caractérisent par une augmentation du volume d'appels pouvant faire doubler, voir tripler le nombre d'appels entrants. Cette augmentation peut être due en partie à l'ouverture des lignes de la PDSA.

Tableau 4.3 – Informations sur la durée de communication des ARM

<i>SAMU</i>	Albi
<i>Période</i>	Février, Mars 2018
<i>Nbre appels</i>	26084
<i>Durée moyenne appel</i>	77 secondes
<i>Durée médiane</i>	67 secondes
<i>Déviat ion standard</i>	69.66 secondes
<i>Durée maximale</i>	1215 secondes

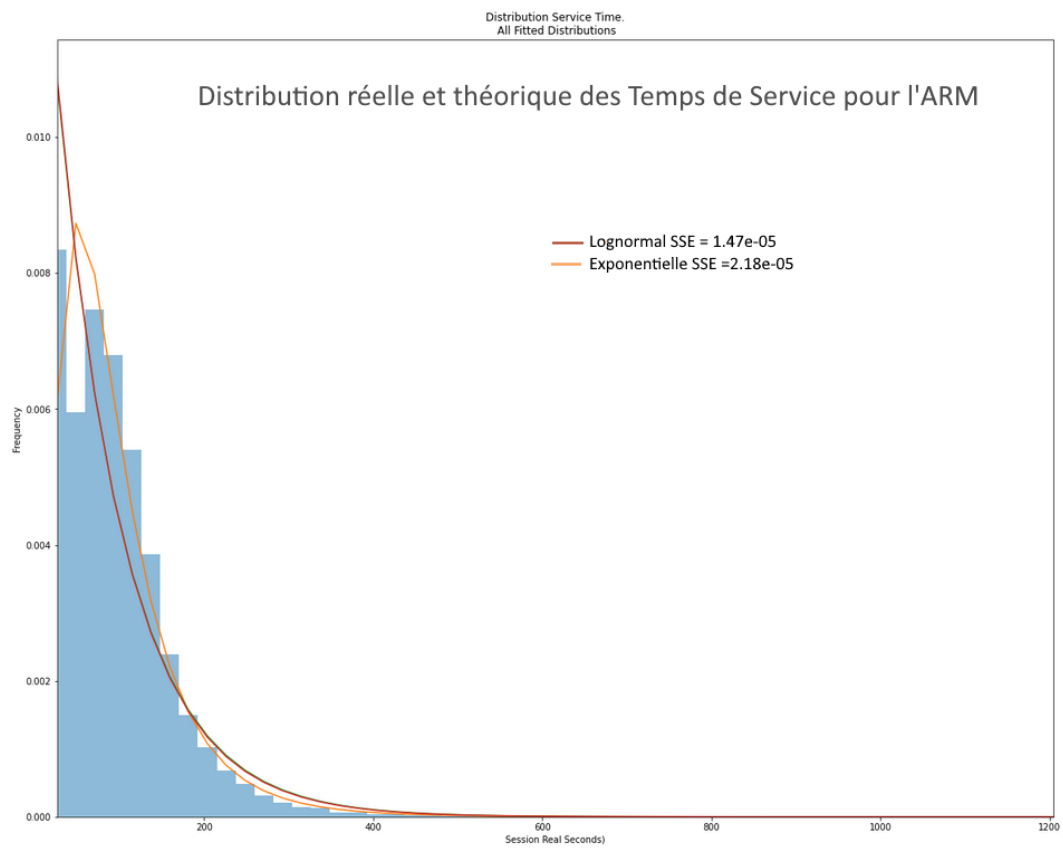


Figure 4.11 – Distribution théorique et empirique des temps de session pour l'ARM

- Le samedi et le dimanche n'ont pas tout à fait le même profil. On remarque que la journée du samedi présente un seul pic d'appels en milieu de journée, là où le dimanche présente deux pics d'appel un en début de matinée et un en fin d'après-midi.
- Au niveau des jours de la semaine, on constate des comportements similaires d'une semaine à l'autre. On retrouve deux pics d'appels journaliers en début de matinée et en fin de journée.
- Au sein d'une même journée on constate deux pics d'appels, un dans la matinée et un durant l'après-midi.

Si on regarde la provenance des appels par créneau horaire et par jour de la semaine sur ces mois là, on voit bien que l'augmentation du week-end est en grande partie due à l'ajout des appels de la PDSA.

Bien qu'il puisse y avoir une variabilité d'une semaine à l'autre, on retrouve néanmoins un profil similaire de l'arrivée des appels à l'échelle de la semaine. Les appels, bien que non programmés donc non prévisibles, se répartissent sur des horaires journaliers et hebdomadaires de façon assez récurrente avec peu de variabilité. Donc, aussi surprenant que cela puisse paraître, on peut prévoir avec une bonne fiabilité le nombre d'appels d'urgence sur chaque créneau horaire, chaque jour, hors considération des événements majeurs de type attentat ou pandémie par exemple. De ce fait, nous nous baserons sur les données historiques pour générer le profil des appels dans le cadre de nos simulations.

Le dimanche est une journée particulièrement critique en termes de volumétrie sur laquelle il n'y a pas forcément plus de ressources engagées pour le traitement des appels. De ce fait, nous allons cibler notre étude sur cette journée. En effet, si un centre d'appels est bien dimensionné pour tenir la charge du jour le plus critique, on peut faire l'hypothèse qu'il devrait être capable de gérer la charge des autres jours qui est moins importante.

Analyse des flux : Appels entrants et sortants

Les centres d'appels d'urgence ont la particularité d'être des centres d'appels mixtes, cela signifie qu'ils reçoivent des appels, mais qu'ils en émettent aussi vers l'extérieur. Toutefois la gestion des appels sortants est bien souvent négligée dans la littérature scientifique (Aboueljiane et al., 2014; M. v. Buuren et al., 2017). Comme nous avons la chance de disposer des données des centres d'appels à l'échelle de l'appel, il est possible de savoir la différence entre les appels sortants et entrants en étudiant les événements du registre de données.

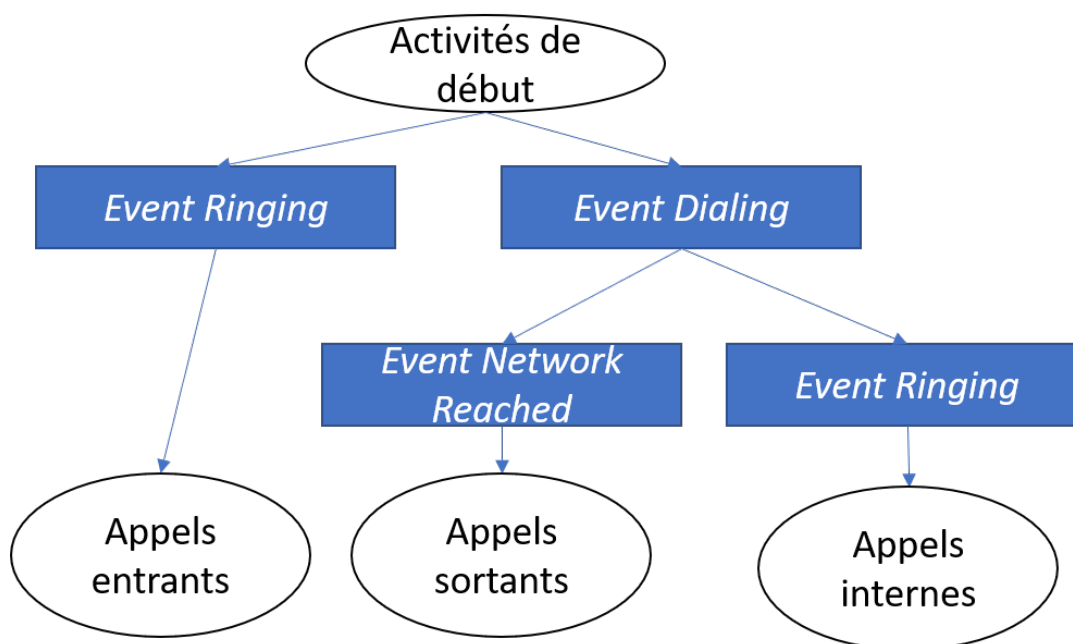


Figure 4.13 – Identifier les types d'appels par leur événement initial

L'analyse du premier événement dans la gestion d'un appel nous donne une information sur le type d'appel (cf. figure 4.13). Les traces des appels ont globalement deux types d'événements de début : *Event Ringing* ou *Event Dialing*. Une trace commençant par le premier correspondra à un appel entrant, car cet événement indique qu'un téléphone est en train de sonner. *Event Dialing* nous indique qu'un appel est en train d'être composé depuis un poste interne. Cet appel peut être interne ou sortant. Un appel interne correspond à une communication entre les acteurs du centre d'appels. Par exemple, un MR qui appelle un ARM. Un appel externe implique un appel issu du centre d'appels vers l'extérieur, on le reconnaît grâce à l'événement *EventNetworkReached* qui nous indique qu'un réseau distant a été contacté. C'est par exemple le cas lorsqu'un MR contacte un hôpital pour savoir s'il y a de la place dans un service particulier.

Si on regarde une journée de dimanche et que l'on compare le nombre d'appels sortants et le nombre d'appels entrants. On constate qu'il y a en moyenne quasiment autant d'appels sortants que d'appels entrants, car ils représentent respectivement 51% et 49% du total des appels (voir tableau 4.4). Ainsi, en termes de volumétrie, il ne semble pas acceptable de faire l'impasse sur les appels sortants.

Tableau 4.4 – Part des appels sortants et des appels entrants dans la volumétrie totale sur 8 semaines

	S 6	S 7	S 8	S 9	S 10	S 11	S 12	S 13	Moyenne
Appels entrants	46%	47%	49%	50%	48%	54%	47%	49%	49%
Appels sortants	54%	53%	51%	50%	52%	46%	53%	51%	51%

Il y a donc un nombre non négligeable d'appels qui sont émis par le centre d'appels. Il est intéressant de se demander quel est le temps passé sur ceux-ci par rapport au traitement des appels entrants. Par conséquent, nous avons regardé le temps total cumulé passé en ligne avec des appels entrants et sortant durant la journée. Les résultats sont contenus dans le tableau 4.5, on constate qu'en moyenne sur une journée il y a souvent deux fois plus de temps passé sur les appels sortants que sur les appels entrants.

Cela confirme notre intuition que les appels sortants ne peuvent pas être écartés de l'étude. La prise en compte de ces appels est une originalité nécessaire qui vient combler un manque vis-à-vis de l'état de l'art.

Tableau 4.5 – Sur quels appels les ARM et les MR passent-ils le plus de temps ?

	S 6	S 7	S 8	S 9	S 10	S 11	S 12	S 13	Moyenne
Appels entrants	36%	35%	38%	38%	38%	34%	28%	19%	33%
Appels sortants	28%	65%	62%	62%	62%	66%	72%	81%	62%

4.3 Synthèse

Dans ce chapitre, nous avons appliqué la première partie de la méthodologie développée dans le chapitre 3. La principale contribution de ce chapitre consiste donc à proposer une analyse détaillée des données de téléphonie des centres d'appels d'urgence du SAMU. Nous avons pu remarquer que les profils d'arrivées des appels présentaient une saisonnalité hebdomadaire avec une augmentation du nombre d'appels durant le week-end, mais aussi une saisonnalité journalière avec deux pics d'appels dans la journée. Nous avons développé plusieurs algorithmes pour l'étude des données afin de permettre l'évaluation des temps d'attentes et des temps de sessions.

En étudiant les temps de sessions des différents types d'appels, nous avons découvert que les appels sortants occupaient les ressources pendant 62% de leur temps en moyenne. La prise en compte des appels sortants dans cette analyse constitue une véritable originalité par rapport aux travaux menés jusqu'à présent sur le sujet.

Néanmoins, afin d'aller plus loin, il serait nécessaire d'avoir une quantité de données plus importante afin de mener des expérimentations supplémentaires concernant les distributions théoriques des temps de session. En effet, il serait intéressant d'étudier

l'effet des différentes lois pour voir laquelle se rapproche le plus du comportement réel du centre d'appels. Le tableau 4.6 propose une synthèse des contributions et des limitations de ce chapitre.

Tableau 4.6 – Contributions et Limites : Chapitre 4

	Chapitre 4
Contributions	Une analyse des données de téléphonie des centres d'appels d'urgence.
	L'inclusion des appels sortants dans notre étude.
	Proposition d'algorithmes pour l'étude des données de centres d'appels d'urgence.
Limitations	Un manque de données ne nous permet pas d'obtenir des distributions théoriques de temps de session.

5

Expérimentations SAMU

Une fois le système réel analysé, il est alors possible de chercher à améliorer la situation courante et de réduire le temps d'attente des appels ainsi que d'améliorer le QoS comme décrit au chapitre 1.3.

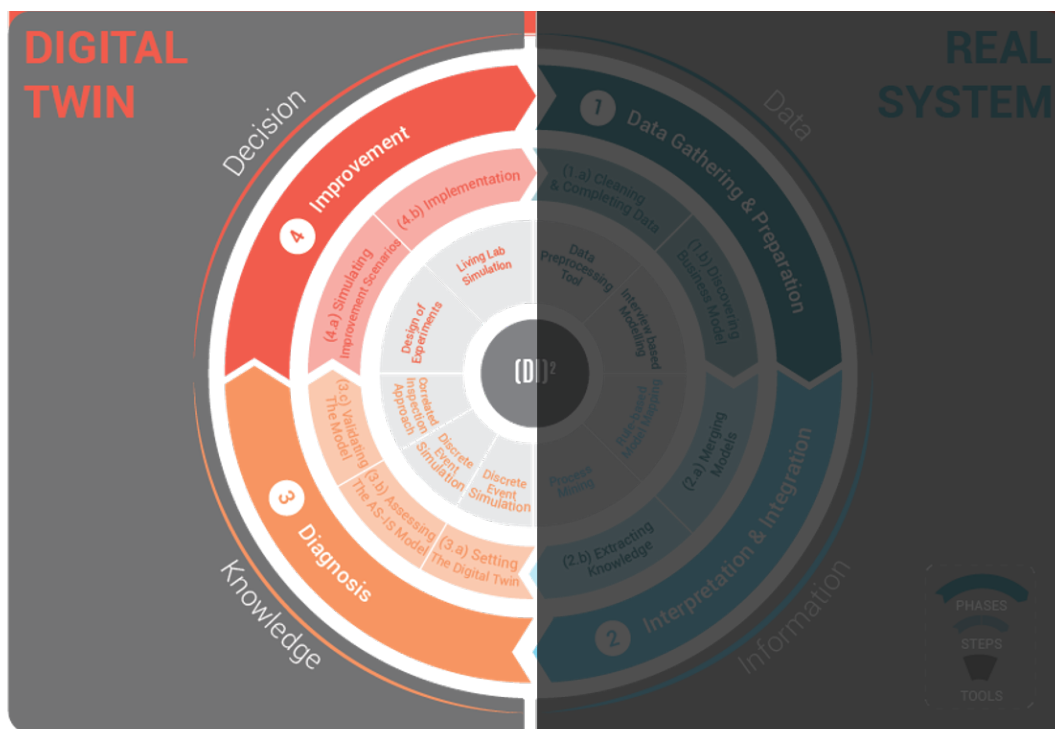


Figure 5.1 – Partie de la méthodologie traitée dans ce chapitre

Ce chapitre correspond à la partie gauche (cf. figure 5.1) de notre cadre conceptuel présenté au chapitre 3. Dans cette partie, on cherche de nouveaux « traitements » pour améliorer la performance du SAMU (QoS et Temps d'attente maximum). Plutôt que de tester ces traitements sur le processus réel (long, ou impossible, ou coûteux, ou risqué, etc.), on décide de développer un double numérique, véritable réplique virtuelle du processus de prise en charge des appels, grâce auquel il sera possible de reproduire l'organisation existante pour mieux la comprendre et tester d'autres

schémas organisationnels, avant une implémentation éventuelle dans le monde réel. La question de recherche abordée dans ce chapitre est finalement :

Considérant l'hétérogénéité des SAMU, quel processus de décroché des appels et de gestion des ressources est le plus à même de soutenir leur profil d'arrivée en temps normal ainsi que lors d'une augmentation soudaine du nombre d'appels ?

5.1 Étape 3 : Diagnostiquer

5.1.1 Sous-étape 3.a : Paramétrer le double numérique ; Hypothèses et Modèle conceptuel

Nous avons développé notre modèle de manière à ce qu'il soit le plus proche possible de la réalité. Toutefois, comme dans tous les modèles, il est nécessaire de faire des hypothèses, car nous ne disposons pas de l'information, ou parce que nous ne pouvons pas prendre en compte de l'intégralité de l'information. Ainsi, notre modèle sera valable sous certaines conditions.

Sélectionner l'outil support au double numérique

La première étape est de choisir l'outil support à la mise en oeuvre de ce double numérique.

Il existe plusieurs types de simulation : on trouve notamment la simulation à base d'agents, la simulation à événements discrets ainsi que la dynamique des systèmes.

La simulation à événements discrets

La simulation à événements discrets (SED) concerne la modélisation d'un système qui évolue dans le temps en prenant différents états. Les changements d'état se produisent à des instants correspondant à des événements (Stewart Robinson, 2004). Lors des événements les variables qui définissent le système changent, ce qui peut amener le système à un nouvel état.

En effet, un événement peut ne pas changer l'état du système et peut simplement correspondre à une récupération d'information, ou encore à une planification d'un événement futur dans le système, ce qui n'a pas d'impact direct sur l'état du système.

La SED est une technique de recherche opérationnelle qui est utilisée dans des domaines variés allant de l'industrie, à la santé en passant par la défense (Law et al., 2000). Elle offre plusieurs avantages :

- Elle permet de modéliser des systèmes complexes ayant de nombreux paramètres.
- Les scénarios simulés peuvent être observés dynamiquement et en temps réel via la représentation graphique du modèle. Cela permet de décrire plus visuellement le comportement du système.

- Il est assez aisé d'affiner le modèle de manière incrémentale et de voir l'effet des modifications sur le système.

Pour prendre en compte les phénomènes aléatoires, on lance plusieurs répliques. Ces dernières se basent sur des suites de nombres pseudo aléatoire présentes dans le simulateur. Elles permettent de prendre en compte les phénomènes aléatoires qui régissent le modèle et d'appliquer cette variabilité dans les simulations afin d'avoir un comportement plus proche de la réalité. De ce fait, les résultats obtenus seront plus fiables.

La SED est une technique qui est largement utilisée dans l'industrie depuis les années 1950.

La simulation à base d'agents

La simulation à base d'agents est une méthode qui est plus récente, elle est développée au début des années 2000. Son essor est motivé par une volonté de la communauté à aller plus loin dans la modélisation de système complexe. Cela est rendu possible grâce aux avancées techniques et technologiques qui permettent d'avoir plus de puissance de calcul car la simulation à base d'agents nécessite des ressources plus importantes que les autres types de simulation (Borshchev, 2013a). Comme soulevé par Law et al., (Law et al., 2000) il est difficile de trouver une définition claire pour la simulation à base d'agents. Tout d'abord, qu'est-ce qu'un agent ? C'est une entité autonome qui peut interagir avec son environnement (incluant les autres agents), et utiliser les informations recueillies pour prendre des décisions. Les agents ont des attributs, et un panel de règles qui définissent leurs comportements. Ces comportements peuvent évoluer dans le temps et s'adapter à l'environnement de manière dynamique. Si on suit cette définition, la simulation à base d'agents est de ce fait, un cas particulier de la SED, où les comportements sont définis directement au niveau de l'entité et non pas au niveau des blocs, par lesquels l'entité passe.

La modélisation basée sur les agents se concentre sur les composants actifs individuels d'un système. Elle s'oppose à la fois à l'approche plus abstraite de la dynamique des systèmes et à la méthode des événements discrets axée sur le processus et les enchainements d'activités du système¹.

Avec la modélisation basée sur les agents, les entités actives, les agents donc, doivent être identifiées et leurs comportements définis. Il peut s'agir de personnes, de ménages, de véhicules, d'équipements, de produits ou d'entreprises, selon ce qui est pertinent pour le système. Des connexions entre elles sont établies, des variables environnementales sont définies et des simulations sont effectuées. La dynamique globale du système émerge alors des interactions entre les nombreux comportements individuels.

Ainsi, la simulation à base d'agents est intéressante dans plusieurs cas (Borshchev, 2013b) :

- Si les entités du système interagissent entre elles et avec l'environnement (des colis, des personnes dans un aéroport, des patients dans un hôpital...).

1. <https://www.anylogic.com/use-of-simulation/agent-based-modeling/>

- Lorsqu'il est important pour les entités d'apprendre et d'adapter leurs comportements.
- Lorsque les mouvements de l'entité dépendent de sa perception de la situation et de son environnement plutôt que de règles fixées au préalable.

Les systèmes dynamiques

La simulation utilisant les systèmes dynamiques est une méthode créée dans le milieu des années 1950 par Jay Forrester (Forrester, 1971). Elle permet d'étudier des systèmes dynamiques en se basant sur des lois issues de la physique (Borshchev, 2013a). Ces méthodes se basent sur plusieurs principes concernant les stocks, les flux et les boucles de retours (Sterman, 2010) :

- Les stocks sont les endroits dans le système où il y a des accumulations du flux, ils sont source de déséquilibre dans le système.
- Les boucles de retour sont au coeur des systèmes dynamiques. Dans le cas d'un centre d'appels, cela correspondrait aux appelants qui rappellent parce qu'ils n'ont pas obtenu de réponse.
- les événements et les décisions sont vus comme un flux continu d'information qui vient alimenter le modèle.

Là où les systèmes à base d'agents agissent au niveau de l'agent ou la SED au niveau des événements, les systèmes dynamiques utilisent un niveau d'abstraction plus élevé.

Un intérêt visuel et didactique pour les experts qui rend la simulation attrayante en santé

L'usage de la simulation dans le domaine de la santé tend à aller en augmentant au fil des années (Salleh et al., 2017). Salleh et al. proposent une revue de littérature de l'usage de la simulation dans le domaine de la santé. Leur but est de présenter une vue d'ensemble des revues de littérature sur l'utilisation de la simulation entre 1990 et 2017. Les auteurs identifient quatre principaux domaines d'applications en santé :

1. Opération de santé et conception de système : Utilisation de la simulation pour la gestion des ressources, leur dimensionnement, pour l'optimisation des flux et la prévision de la demande de soin.
2. Aide à la décision médicale : Évaluer les effets à court et long terme de certains programmes de soin.
3. Modélisation de maladies infectieuses : Applications d'épidémiologie pour prédire la diffusion des épidémies et pandémies ainsi que leurs conséquences économiques.
4. Études diverses : Simulations d'événements ponctuels (coupe du monde, attaques terroristes) et de leurs effets sur le système de santé. On trouve aussi des comparaisons de diverses techniques de simulation.

La grande majorité des revues de littérature étudiées par ces auteurs se focalisent uniquement sur la gestion des opérations en santé, ce qui est aussi le domaine dans lequel nous travaillons. On trouve des travaux de simulation pour évaluer les services des urgences (Aboueljinane et al., 2013 ; Gul et al., 2015 ; Paul et al., 2010), pour l'amélioration du parcours patient et la gestion des flux de patients

(Crema et al., 2013; Sobolev et al., 2011; Van Sambeek et al., 2010), pour gérer la chaîne d’approvisionnement logistique des hôpitaux (Beliën et al., 2012; Kammoun et al., 2014), la gestion des lits d’hôpitaux (Baru et al., 2015), la gestion et le dimensionnement des ressources (Isern et al., 2016), ou plus spécifiquement concernant la planification des salles d’opération (Cardoen et al., 2010).

La simulation dans le domaine de la santé est donc très utilisée depuis de nombreuses années car sa capacité à permettre de traduire des échanges et des interactions complexes est très intéressante. En effet, le domaine médical est constitué de nombreux facteurs humains qui viennent générer beaucoup d’aléatoires. Ce qui vient rendre très pertinent l’usage de la simulation pour traiter les problèmes. De plus, la simulation a un intérêt visuel non négligeable pour l’explication des résultats auprès des experts. Le fait de pouvoir rejouer les situations pour en expliquer les tenants et les aboutissants présente un avantage didactique et pédagogique.

Comme nous l’avons expliqué dans la section 3.2.4 nous avons choisi de travailler avec des techniques de simulation. Comme vu ci-dessus, il en existe plusieurs types et pour cette étude il nous a paru pertinent d’utiliser la SED. Notamment la simulation à base d’agents a été écartée car les appels ne sont pas des entités qui interagissent entre elles.

Des hypothèses pour la modélisation

Les simulations ont été réalisées sur un ordinateur équipé d’un processeur Intel® Core™ i5-6200U CPU à 2,3 GHz et d’une mémoire RAM de 8 Go. Le logiciel utilisé dans ce manuscrit est le logiciel Witness®, développé par Lanner² et qui fait partie des leaders sur le marché des moteurs de simulation à événements discrets. Un visuel du modèle de simulation est disponible dans la figure 5.2.

Dans le chapitre 2 nous avons détaillé les différents types d’appels qui sont reçus par les centres d’appels d’urgence. Pour rappel, on distingue les appels de la population, les appels provenant des équipes dépêchées sur le terrain, ainsi que les appels des autres acteurs de la chaîne des urgences. Les données du fichier de téléphonie nous permettent de savoir quel est le numéro composé pour joindre le centre d’appels. Toutefois nous n’avons pas les informations nécessaires pour différencier clairement les appels de la population des appels des intervenants. Comme précédemment décrit dans le chapitre 1.2, il y a plusieurs types d’appels. Pour la suite de cette étude, nous faisons l’hypothèse que tous les appels entrants sont considérés comme étant des appels de la population entrés par le 15 ou la PDSA. Dans un premier temps, nous ne différencions pas les appels bilans des appels de la population. Ce choix peut avoir un impact sur la précision du modèle, car les appels bilans n’ont pas forcément le même profil d’arrivée que les appels entrants. Les appels bilans dépendent directement du choix d’envoyer des moyens et des intervenants lors d’un primo appel.

Une deuxième hypothèse concerne les appels sortants. Dans la littérature scientifique relative aux centres d’appels d’urgences, il n’y a pas à notre connaissance de travaux

2. <https://www.lanner.com/fr-fr/technologie/witness-simulation-software.html>

dans lesquels les appels sortants sont considérés. Ceci constitue une faiblesse importante dans les études passées des centres d'appels urgents comme démontré en section 4.2.2. Nous proposons dans notre travail d'en tenir compte et de les inclure dans notre modélisation. Bien qu'il existe plusieurs types d'appels sortants, vers les intervenants, vers les établissements de santé, etc. (voir figure 1.9), nous ne distinguons néanmoins pas la nature de l'appel sortant ni son destinataire. Nous avons utilisé les données de téléphonie pour extraire le profil des appels sortants ainsi que la distribution de leur durée. Toutefois, nous n'avons pas l'information de la nature du destinataire, donc nous n'avons pas pu spécifier la génération ou la durée de l'appel sortant en fonction de son destinataire. C'est pourquoi nous avons considéré les appels sortants comme une grande famille d'appels sans définir de spécificités en son sein.

Concernant les appels internes, soit les appels qui sont générés entre les opérateurs du centre d'appels, nous avons uniquement considéré les appels de l'ARM vers le MR. D'après les données dont nous disposons, ces appels correspondent à 74% de la totalité des appels internes. 13% des appels internes sont de l'ARM vers le MR PDSA, nous les avons aussi considérés, et nous avons négligé les 13% restants qui concernent des appels d'un ARM vers un autre ARM ou d'un MR vers d'autres ressources.

Réplifications, Horizon, Explications

Définir le nombre de réplifications

Comme nous avons déjà pu le mentionner, les réplifications permettent de modifier les suites pseudo-aléatoires du simulateur et ainsi d'avoir de nouveaux tirages pour les valeurs aléatoires du système. Dans notre cas, nous avons de l'aléatoire en ce qui concerne l'arrivée des appels : le nombre d'appels par tranches de 30 minutes. Il y a aussi de l'aléatoire sur le niveau de gravité de l'appel, sur sa durée de traitement, sur la possibilité de générer un appel interne. Les réplifications nous permettent de couvrir tout le spectre des valeurs possibles lorsqu'il y a de l'aléatoire. L'objectif est d'avoir un nombre significatif de 100 réplifications afin de couvrir tous les types de journées « probables » d'un centre d'appels, les meilleurs comme les pires.

Définir l'horizon de simulation

Comme nous avons déjà pu l'expliquer dans le chapitre précédent, nous avons limité le périmètre d'études aux dimanches, car ce sont les journées les plus chargées avec une augmentation significative du nombre d'appels reçus. De ce fait, nous avons fixé l'horizon de simulation à 24h. Chaque simulation correspond à un dimanche dans un centre d'appels. Lorsqu'on lance une nouvelle réplification, cela revient à simuler un nouveau dimanche. Dans notre cas, nous avons travaillé sur une base de 100 réplifications, donc de 100 dimanches, soit environ deux ans de dimanches simulés.

5.1.2 Sous-étape 3.b : Évaluation de l'AS-IS : Un Modèle conceptuel pour guider la construction du modèle

La figure 5.3 présente le processus complet suivi par notre modèle de simulation : une version agrandie est disponible en annexe C. Nous décrivons ci-dessous, pour chacune des activités, les données utilisées et hypothèses considérées. Afin d'être le plus lisibles possible, nous décomposons ces explications par macro-activités. (cf. figure 5.4)

Macro-activité 1 : La gestion des appels entrants

Cette activité ne concerne que l'acteur ARM. Dans cette étude nous considérons deux types d'appels entrants, les appels provenant du numéro d'urgence nationale français, le 15, et les appels provenant de la permanence des soins ambulatoires (39 66 en Occitanie, 116 117 pour d'autres régions comme en Pays de la Loire, Normandie et en Corse). L'arrivée des appels est paramétrée en se basant sur le profil des données historiques comme expliqué dans le chapitre 4. Nous avons segmenté la journée en créneaux de 30 minutes. Ce choix de segmentation a été fait, car il permet d'avoir une vision assez précise et détaillée de l'évolution de la performance du centre d'appels au cours de la journée. Habituellement, les études mesurent les performances d'un centre d'appels à l'échelle de la journée au niveau global, toutefois la variabilité du nombre d'appels au cours de la journée est assez importante. C'est pourquoi il nous a paru intéressant de se baser sur des créneaux de 30 minutes pour étudier les performances du SAMU.

Sur chaque créneau nous avons regardé le nombre d'appels arrivés sur les 8 dimanches disponibles. Cet échantillon de 8 jours n'étant pas statistiquement significatif, il n'est pas possible ici d'identifier une distribution théorique satisfaisante pour l'arrivée des appels. Pour modéliser le caractère aléatoire du nombre d'appels sur chaque créneau horaire, nous avons choisi d'utiliser une loi de distribution uniforme s'étendant entre le nombre minimum et le nombre maximum d'appels reçus durant les 8 jours étudiés. Cela nous permet d'avoir des profils d'appels simulés qui sont cohérents par rapport aux données réelles. Chaque simulation ayant un horizon de 24 heures (1 journée pleine), cela va aussi permettre de lancer un nombre de répliques significatif, car l'arrivée des appels sur chaque créneau, même si elle est aléatoire, ne se produira qu'une seule fois au cours d'une même simulation. L'ensemble des répliques permettra donc de tirer plusieurs valeurs aléatoires d'arrivée d'appels sur chaque créneau et ainsi permettra de présenter les résultats sous la forme d'un intervalle de confiance. À titre d'exemple, la figure 5.5 montre le profil des appels 15 réels sur lesquels nous nous sommes basés. On peut remarquer deux pics d'appels dans une journée de dimanche autour de 9h et un autre autour de 16h.

Lorsqu'un appel arrive au SAMU, il arrive dans une file d'attente, l'appel va sonner au SAMU et pendant ce temps là pour l'appelant, le disque de bienvenue au SAMU déroule durant plusieurs secondes. Ainsi, il y a un temps d'attente minimal avant que l'appel puisse être décroché. Ce temps a été paramétré à 3 secondes en se basant sur

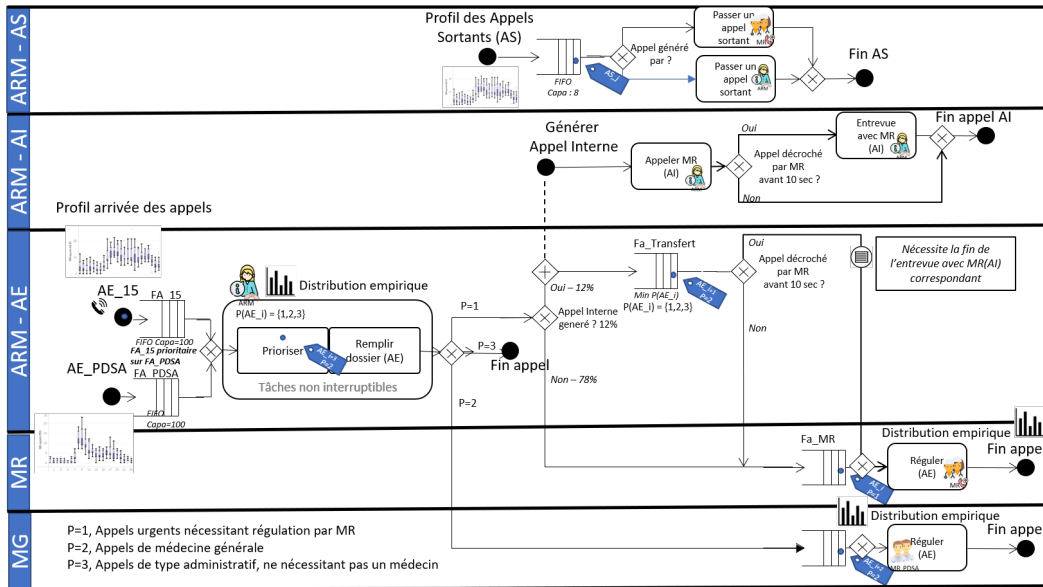


Figure 5.3 – Modèle conceptuel de la gestion des appels. (Voir version agrandie en annexe)

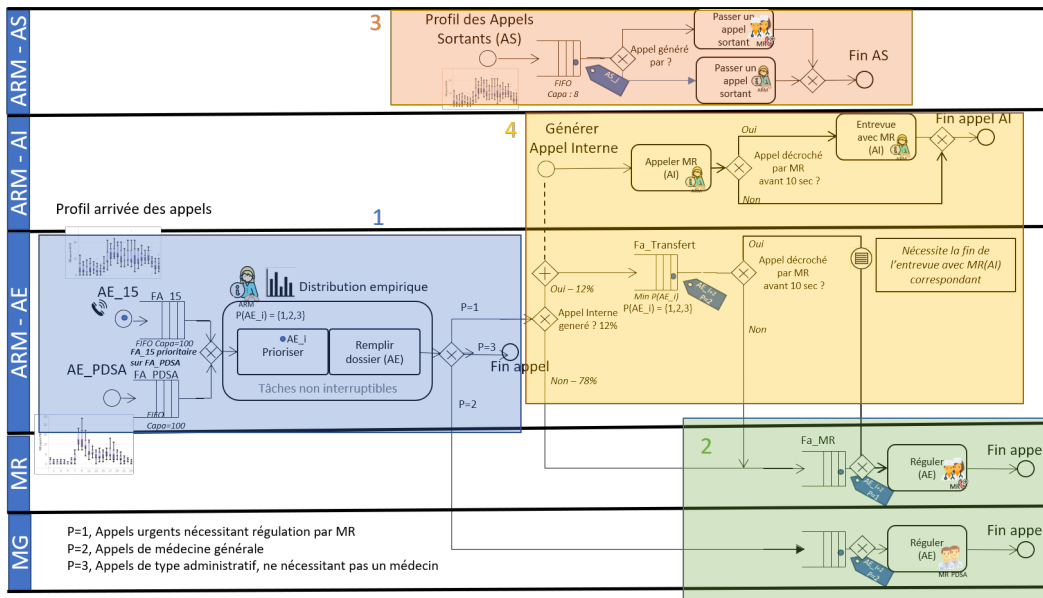


Figure 5.4 – Modèle conceptuel de la gestion des appels : Définition des macro-activités

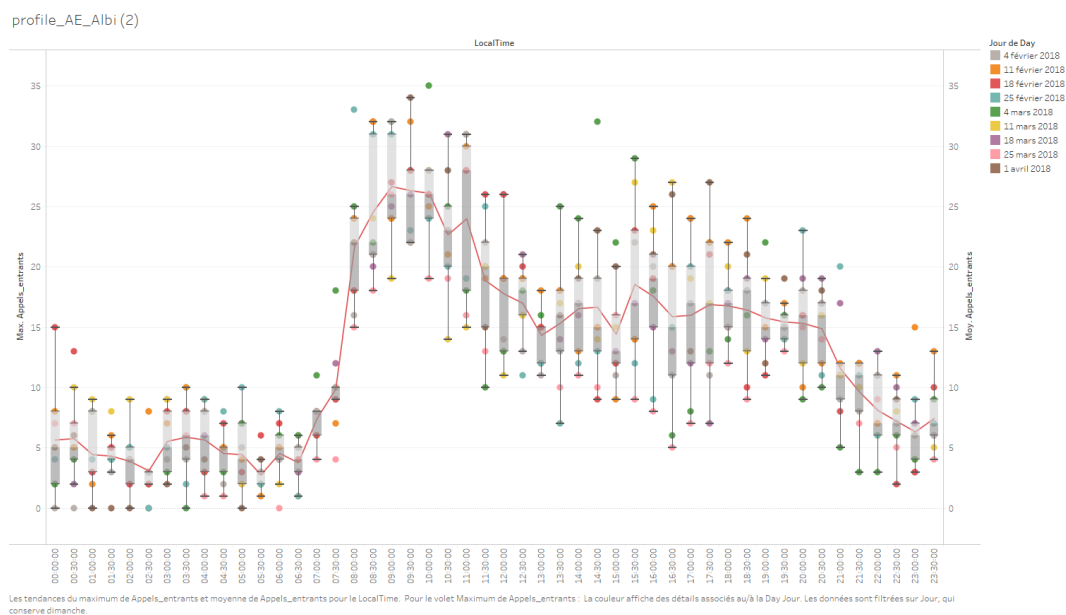


Figure 5.5 – Profil d'arrivée des appels entrants (un dimanche)

les données indiquées par les experts et notamment par les conclusions du rapport MARCUS (Moutard et al., 2019).

Concernant les files d'attente, il y a plusieurs questions qui se posent :

- Est-ce qu'il y a une file d'attente par canal d'entrée ? Dans ce cas, y a-t-il une file d'attente prioritaire par rapport à l'autre ?
- Est-ce qu'il y a une seule file d'attente gérée avec une loi de type « premier arrivé, premier servi » (FIFO) ?
- Est-ce qu'il y a une seule file d'attente avec une priorité pour les appels provenant du 15 par rapport à ceux provenant de la PDSA ?

Ces choix de modélisation sont à discuter avec les responsables des centres d'appels pour mettre en place ce qui correspond à l'usage dans le centre d'appels. Toutefois, il reste intéressant de s'interroger sur l'effet de ces dispositions de file d'attente sur les performances. Dans une étude préliminaire, nous avons testé différentes files d'attente (Petitdemange et al., 2018). Les résultats montrent que s'il peut paraître logique de prioriser les appels du 15 par rapport à ceux de la PDSA cela peut poser problème dans certains cas. En effet, bien qu'ils soient moins fréquents, des appels urgents qui nécessitent une régulation par un MR peuvent arriver par le canal de la PDSA. Ainsi la priorisation uniquement par la provenance de l'appel, n'est pas suffisante pour certains cas isolés.

Dans notre cas nous avons d'abord paramétré le modèle comme ayant deux files d'attente distinctes en fonction de la provenance de l'appel. Nous avons fait l'hypothèse qu'une priorité supérieure est appliquée pour la file d'attente des appels 15 (*FA_15*).

Une fois l'appel dans la file d'attente, il sera décroché dès lors qu'un ARM sera disponible. Durant ce premier temps de communication avec l'ARM, plusieurs tâches sont effectuées les unes à la suite des autres ces tâches ont déjà été présentée dans la section 1.2.5. Ces sous-tâches sont regroupées sous une macro-tâche qui correspond

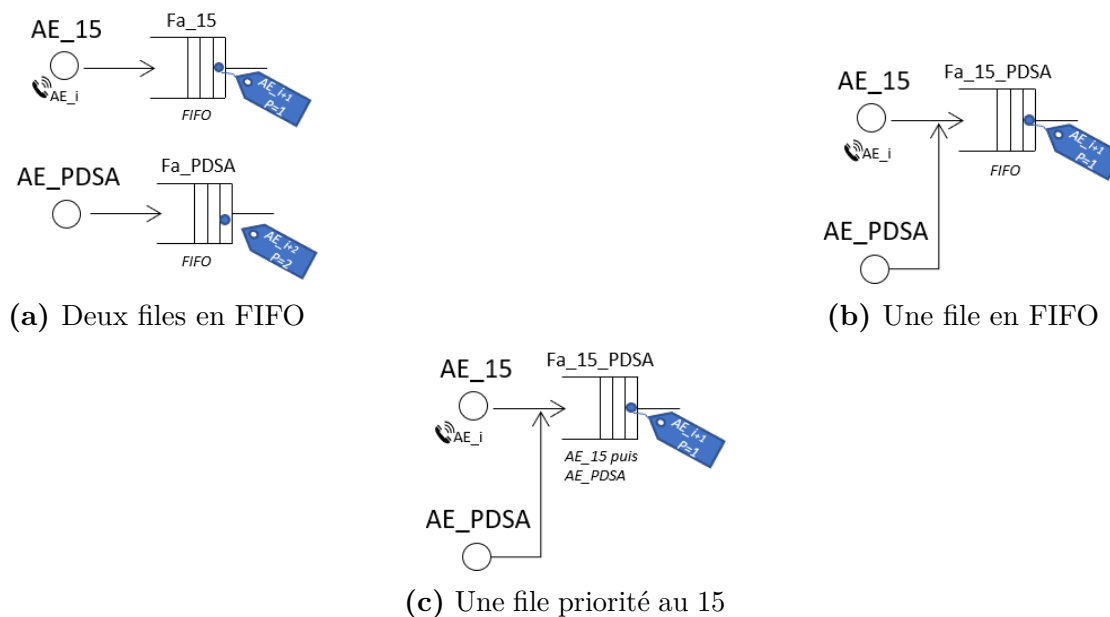


Figure 5.6 – différents types de file d’attentes

à ce premier temps de communication avec l’ARM. Dans nos données de téléphonie, nous n’avons pas d’indication pour identifier les temps de chaque sous-tâche. Cet enchaînement est bien souvent non interrompu, cela signifie que lorsqu’un ARM prend un appel, il effectue tout ou partie de la chaîne en fonction de la gravité de l’appel qui va donner lieu à sa priorisation (P1, P2 ou P3 comme présenté dans la section 1.2.5. C’est pourquoi nous avons modélisé le temps de communication de l’ARM d’un seul bloc.

Dans un premier temps, notre modèle est configuré pour que cet enchaînement d’activité corresponde à une seule macro-tâche qui ne peut être interrompue ou mise en pause. C’est-à-dire que lorsqu’un ARM décroche un appel, il le traite en entier avant d’être de nouveau disponible. C’est lors de cette étape que l’appel va être priorisé : L’ARM va lui attribuer un niveau de gravité qui va décider du devenir de l’appel. La Table 5.1 présente la répartition statistique de la gravité des appels par rapport à leur provenance. On constate que des appels de priorité 3 occupent 78% de la totalité des appels que ce soit pour les appels provenant de la PDSA ou du 15. Pour rappel, ce sont des appels de nature non médicale (adresse de la pharmacie de garde, coordonnées d’un médecin ou d’un dentiste, etc.), qui ne requièrent par de temps de communication avec un MR. Il serait alors utile d’assurer un traitement automatisé de ces appels, ce qui pourrait être un levier d’amélioration dans la gestion des appels d’urgence. Ensuite, d’après nos données, dans les appels provenant du 15, 18% sont des appels qui nécessitent une régulation auprès d’un MR et 2% auprès d’un généraliste. Les chiffres s’inversent pour les appels provenant de la PDSA. Cela peut paraître peu, toutefois parmi les appels entrants il y a aussi les appels bilans des équipes d’effecteurs (SMUR, pompiers, etc.) sur place qui parfois transmettent le bilan des interventions aux ARM plutôt qu’au MR. Ces données sont à nuancer : En effet, cela voudrait dire qu’1 appel sur 5 fait l’objet d’une régulation médicale, ce qui semble peu dans la mesure où tous les appels médicaux sont supposé faire l’objet

d'une régulation auprès du MR. Cette donnée est obtenue en analysant les données de téléphonie. Pour chaque identifiant de communication unique correspondant à l'appel d'un patient, nous identifions les ressources qui interviennent sur la gestion de cet appel. Toutefois, cela implique l'hypothèse suivante : tous les appels concernant un même dossier patient sont rassemblés sous le même identifiant de communication. Si ce n'est pas le cas, alors il se pourrait que la part des appels qui génère une communication avec le MR soit plus grande, mais pour vérifier cette hypothèse nous devons disposer du numéro de dossier patient et le relier aux données de téléphonies, ce qui à ce jour n'est pas le cas.

Tableau 5.1 – Répartition de la priorité des appels en fonction de la provenance des appels.

Priorité	P=1	P=2	P=3
	Régulation	Permanence des soins	Conseil
AE_15	18%	2%	62%
AE_PDSA	2%	22%	76%

Macro-activité 2 : La Régulation des appels

Cette activité de régulation concerne le MR pour les appels de priorité $P = 1$, et le MG pour les appels de priorité $P = 2$. Une fois les appels triés par l'ARM, ils sont envoyés dans la file d'attente idoine en attendant d'être régulés par un médecin. S'il n'y a pas de système de gravité codifié par l'ARM, les appels sont traités en FIFO dans leurs files d'attente respectives, à l'exception du MR lorsqu'il termine une communication en appel interne avec un ARM, il traite directement l'appel entrant en lien avec cet appel interne avant de continuer à gérer sa file d'attente.

Macro-activité 3 : La gestion des appels sortants externes

Les appels sortants sont des appels émis par les agents du centre d'appels vers l'extérieur. Ceux-ci peuvent correspondre à des appels vers les effecteurs pour envoyer une ambulance, le SMUR, une équipe des pompiers, ou encore un appel vers les établissements de santé pour savoir s'ils ont une place disponible pour un patient. Cela peut aussi être pour rappeler des patients (cette pratique est assez courante, notamment en ce qui concerne la PDSA). Plusieurs questions se posent en ce qui concerne les appels sortants :

- Qu'est-ce qui déclenche un appel sortant ?
- Est-ce que les appels sortants dépendent des appels entrants ?
- Est-ce que les appels sortants sont prioritaires par rapport aux appels entrants ?

Si on utilise uniquement les données de téléphonie, il est difficile d'identifier vers quels types de correspondants l'appel est passé. Pour cela il faudrait identifier les numéros récurrents, savoir à quels numéros correspondent les partenaires de l'urgence pour connaître la nature de l'appel sortant. Nous n'avons pas obtenu la liste des numéros connus des effecteurs, il nous a donc pas été possible de filtrer ces appels

dans la base de données. Une autre piste intéressante serait d'être capable de croiser les données d'intervention d'un appel entrant qui précède un appel sortant pour avoir plus d'information sur le moyen envoyé, savoir si cet appel sortant est relié au dossier de l'appel entrant. Ces informations permettraient d'estimer la probabilité qu'un appel entrant génère un appel sortant, mais aussi d'identifier les règles métiers qui font qu'un appel sortant est déclenché. Toutefois l'accès aux dossiers de régulation médicale est complexe et n'est pas possible en état pour des raisons de confidentialité. Ainsi, nous avons fait l'hypothèse que les appels sortants étaient générés en suivant leur profil d'émission historique. De la même manière que les appels entrants, nous avons analysé le nombre d'appels sortants émis sur chaque créneau de 30 minutes durant la journée. Afin d'avoir un profil type des émissions des appels sortants. La limite de cette technique est qu'elle ne prend pas en compte les interactions potentiellement existantes entre appels entrants et appels sortants. Néanmoins elle a l'avantage de considérer les appels sortants ce qui n'est pas le cas des études menées jusqu'à présent sur les centres d'appels d'urgence (cf section 5.1.1).

En conséquence, les appels sortants sont générés en suivant le profil présenté dans la figure 5.8. On constate que le profil des appels sortants a une forme similaire à celui des appels entrants, ce qui corrobore le fait que la génération des appels sortants dépend des appels entrants. Lorsqu'un appel est généré, il arrive dans une file d'attente jusqu'à ce qu'un agent puisse le composer. Le temps d'attente qui en découle est donc artificiel puisqu'il correspondrait au fait qu'un agent sait qu'il doit passer un appel, mais il gère d'autres activités au préalable. La durée des appels sortants est définie en suivant la distribution empirique issue des données historiques (cf. figure 5.7).

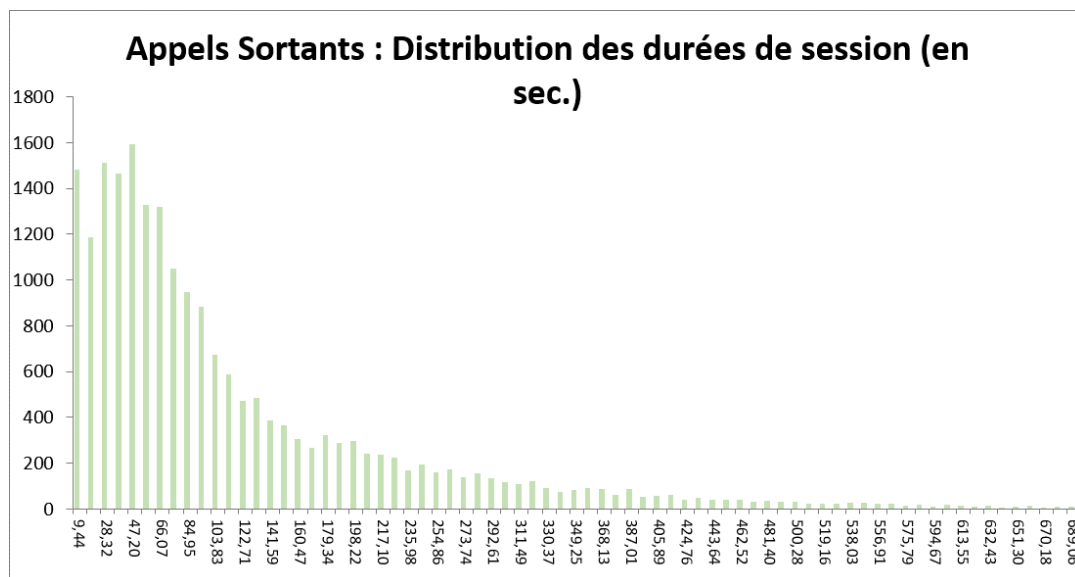


Figure 5.7 – répartition des durées de session pour les appels sortants

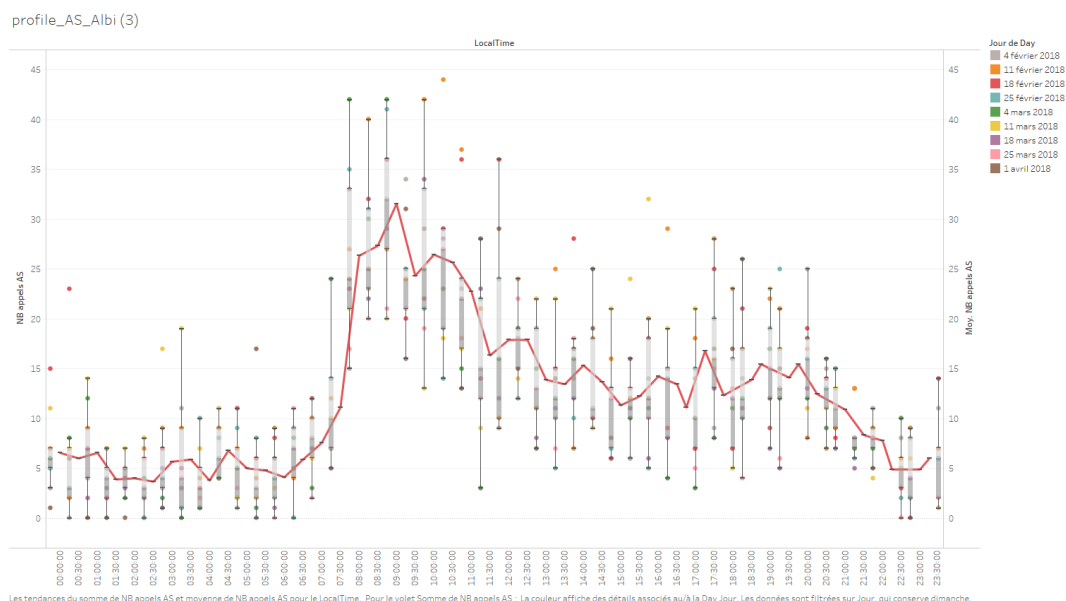


Figure 5.8 – Profil d'émission des appels sortants

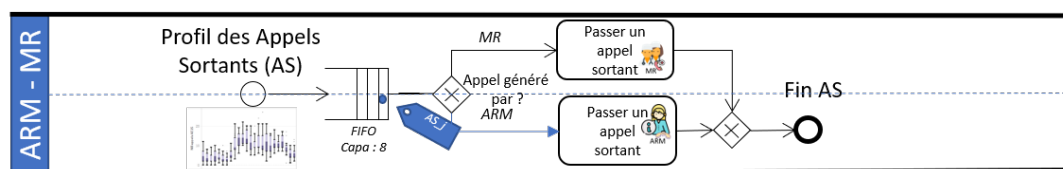


Figure 5.9 – Gestion des Appels sortants (AS)

Macro-activité 4 : La gestion des appels internes

Certains appels entrants peuvent donner lieu à un appel sortant interne (cf. figure 5.10) (une communication entre personnels du centre d'appels). Cela concerne 53% des appels de priorité P1, et 60% des appels de priorité P2. Dans ce cas, l'appel entrant est mis en attente (il est maintenu en ligne), l'ARM passe son appel interne. En accord avec les observations terrains réalisées, nous avons fait l'hypothèse que si le MR n'est pas disponible dans les 10 secondes, l'appel entrant est directement transféré dans la file d'attente du MR et l'appel interne est abandonné. L'ARM va alors soit passer un autre appel sortant s'il y a lieu, soit décrocher des appels.

Un temps de latence pour les agents

À travers nos observations sur le terrain et nos discussions avec les experts, nous avons constaté qu'après un appel, un agent peut prendre quelques secondes avant de décrocher directement un nouvel appel. Ce temps peut correspondre au temps nécessaire pour clôturer le dossier de régulation. Afin de reproduire cet effet dans notre modèle nous avons décidé d'ajouter un temps de latence après chaque appel dont le traitement vient de finir pour l'agent qui vient de raccrocher. Nous avons fixé ce temps en nous basant à la fois sur nos observations et sur l'avis des experts. Ce temps correspond à une loi uniforme qui s'étend entre 2 et 6 secondes. Ce temps

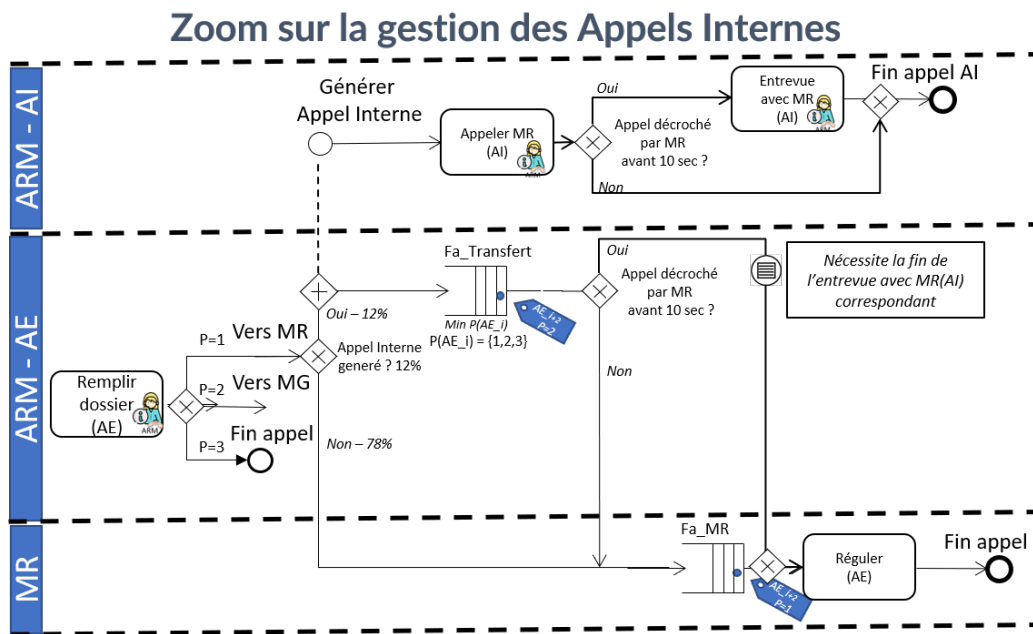


Figure 5.10 – Gestion des Appels internes (AI)

de latence est aussi appelé *wrap-up* il est d'ailleurs pris en compte dans les solutions de téléphonie avancée et ce temps est paramétrable. En effet l'agent peut prendre ce temps de latence ou l'annuler en fonction de son choix. Il serait toutefois intéressant de voir l'impact de la variation de ce temps de latence sur les performances du centre d'appels.

5.1.3 Sous-étape 3.c : Validation du modèle

Valider le profil d'arrivée des appels

Pour valider le profil d'arrivée, nous avons comparé nos données réelles à nos données simulées. Sur chaque créneau de 30 minutes, nous avons relevé le nombre d'appels entrants sur les 8 échantillons réels ainsi que sur chacune des 100 répliques. Pour chaque créneau horaire, nous avons mesuré l'écart sur le nombre minimum, la moyenne et le nombre maximum d'appels pour le profil réel et simulé. Les résultats sont présentés dans la figure 5.11. Au maximum on observe un écart de 7 appels sur le nombre d'appels maximum par jour. Nous pouvons donc conclure que le profil d'arrivée des appels simulés suit la même distribution que celui des appels réels.

Comparer les performances du modèle et de la réalité

Pour évaluer la justesse de notre modèle, nous avons comparé le temps d'attente des appels moyens, médians, et maximums ainsi que la déviation standard sur chacun des créneaux. Nous avons comparé les valeurs de notre échantillon de 8 dimanches avec celles de nos 100 répliques. On observe des écarts entre le réel et le simulé que nous allons expliquer et justifier. Les résultats sont présentés dans le tableau

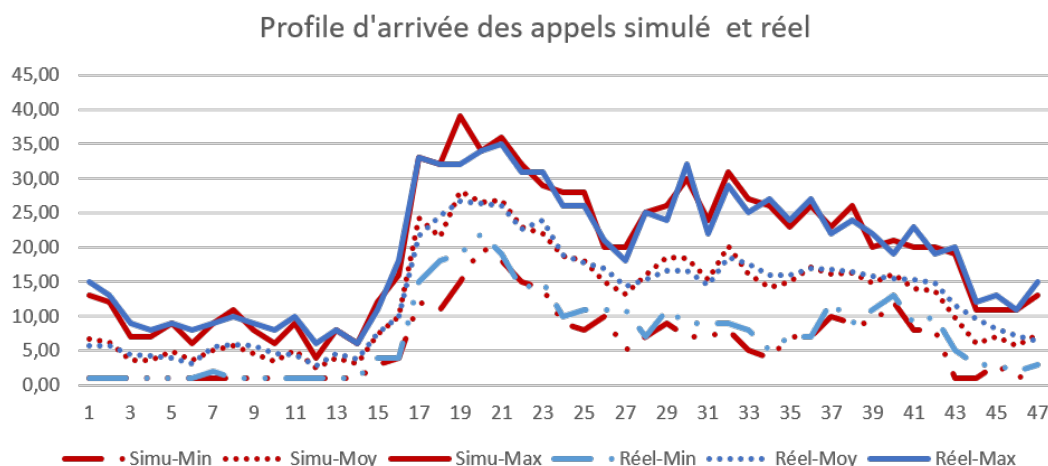


Figure 5.11 – Écarts entre les profils d'arrivées des appels réels et simulés

5.2. On note des écarts en moyenne qui varient entre 0,5 seconde et 6,7 secondes en fonction de l'heure de la journée. Ceci peut être problématique, car la moyenne de décrochée des appels réels varie entre 5 et 11,5 secondes, donc 6,7 secondes est un écart assez significatif. Les écarts au niveau de la médiane varient eux entre 0,5 et 1,5 seconde ce qui est tout à fait acceptable. C'est au niveau des temps d'attente maximums qu'on constate la plus grande variabilité. En effet, notamment durant les horaires de nuit profonde entre minuit et 7h00 les résultats de la simulation sont optimistes par rapport à la réalité. En effet le temps d'attente maximal est plus faible dans notre simulation qu'il n'a pu l'être en réalité sur les 8 dimanches étudiés. On obtient un temps maximal d'attente des appels simulé entre minuit et 07h00 qui est au maximum de 17 secondes. Dans les données réelles, les appels ayant attendu plus de 17 secondes correspondent à moins de 7% de la totalité des appels sur cet horizon. Notre modèle ne prend effectivement pas en compte ces appels qui doivent avoir une raison particulière d'avoir attendu parfois jusque 2 minutes 30. Typiquement, si on regarde le cas de l'appel qui a attendu 162 secondes, on constate que c'est un appel qui a appelé 2 fois avant d'être décroché par un ARM la troisième fois, or nous ne sommes pas en mesure d'expliquer pourquoi les ARM n'ont pas décrochés les deux premières fois, peut être ils n'étaient pas à leur poste de travail à ce moment-là. Nous pouvons encore améliorer la finesse de notre modèle afin qu'il soit en mesure de prendre en compte ces comportements particuliers qui sont à la marge du fonctionnement habituel du SAMU.

Les simplifications effectuées traduisent les comportements les plus fréquents d'un centre d'appels pour le traitement des appels entrants (voir figure 5.12). Il ne prend pas en compte les appels qui abandonnent, ou encore les personnes qui vont appeler, ne pas être décrochées, puis rappeler. Le fait de ne pas distinguer les appels entrants de la population et les appels entrants des équipes qui interviennent sur le terrain rajoute aussi de la variabilité. Ce sont pour toutes ces raisons notamment que l'on constate des écarts entre notre modèle et la réalité. Toutefois, bien que perfectible, ce modèle est une base de travail intéressante pour venir challenger le traitement des

appels entrants de la population. Pour la suite lors des évaluations des scénarios, les résultats seront comparés aux performances du modèle simulé.

Tableau 5.2 – Statistiques descriptives des performances du modèle comparées aux performances réelles

Créneau	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
Réal	Dev_std	18,0	6,1	5,0	8,3	12,1	4,2	8,3	19,8	18,2	11,0	14,6	16,5	12,8	17,8	14,3	9,9	15,1	15,6	14,4	10,6	27,0	26,6	18,0	12,5
	Max	152,0	34,0	31,0	76,0	90,0	17,0	56,0	163,0	220,0	84,0	181,0	248,0	115,0	151,0	112,0	96,0	134,0	119,0	128,0	57,0	249,0	245,0	102,0	109,0
	Moy	9,0	6,7	7,0	6,8	8,1	5,3	5,7	8,2	7,2	5,5	6,6	7,3	7,2	10,5	9,1	7,0	8,2	8,9	8,8	7,6	11,5	11,0	10,2	8,0
	Médiane	5,0	6,0	6,0	6,0	5,0	5,0	6,0	5,0	5,0	4,0	5,0	5,0	3,0	6,0	5,0	5,0	5,0	5,0	5,0	5,0	1,0	3,0	4,0	5,0
Simu	Dev_std	0,8	0,3	0,5	0,6	0,5	0,6	0,5	0,8	10,5	15,1	11,7	8,0	5,5	1,4	10,1	6,9	2,8	7,2	2,3	1,8	4,6	0,8	0,4	0,4
	Max	16,3	5,0	11,9	14,5	12,3	15,7	12,7	16,7	343,7	590,5	413,3	404,9	278,4	37,7	570,5	354,5	61,5	369,2	71,1	77,8	173,6	21,9	15,3	12,3
	Moy	4,6	4,5	4,5	4,5	4,5	4,5	4,6	5,8	6,5	6,1	5,1	4,9	4,7	5,1	5,0	4,8	4,9	4,7	4,6	4,8	4,6	4,6	4,5	4,5
	Médiane	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5	4,5
Écarts	Dev_std	17,2	5,8	4,4	7,7	11,6	3,6	7,9	19,0	7,7	-4,1	2,9	8,6	7,3	16,4	4,3	3,1	12,4	8,4	12,1	8,8	22,4	25,8	17,5	12,1
	Max	135,7	29,0	19,1	61,5	77,7	1,3	43,3	146,3	-123,7	-506,5	-232,3	-156,9	-163,4	113,3	-458,5	-258,5	72,5	-250,2	56,9	-20,8	75,4	223,1	86,7	96,7
	Moy	4,5	2,2	2,4	2,3	3,6	0,8	1,2	3,6	1,5	-1,0	0,5	2,2	2,3	5,9	4,1	2,0	3,4	4,0	4,1	3,0	6,7	6,5	5,7	3,5
	Médiane	0,5	1,5	1,5	1,5	0,5	0,5	1,5	0,5	0,5	-0,5	0,5	0,5	-1,5	1,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	-3,5	-1,5	-0,5

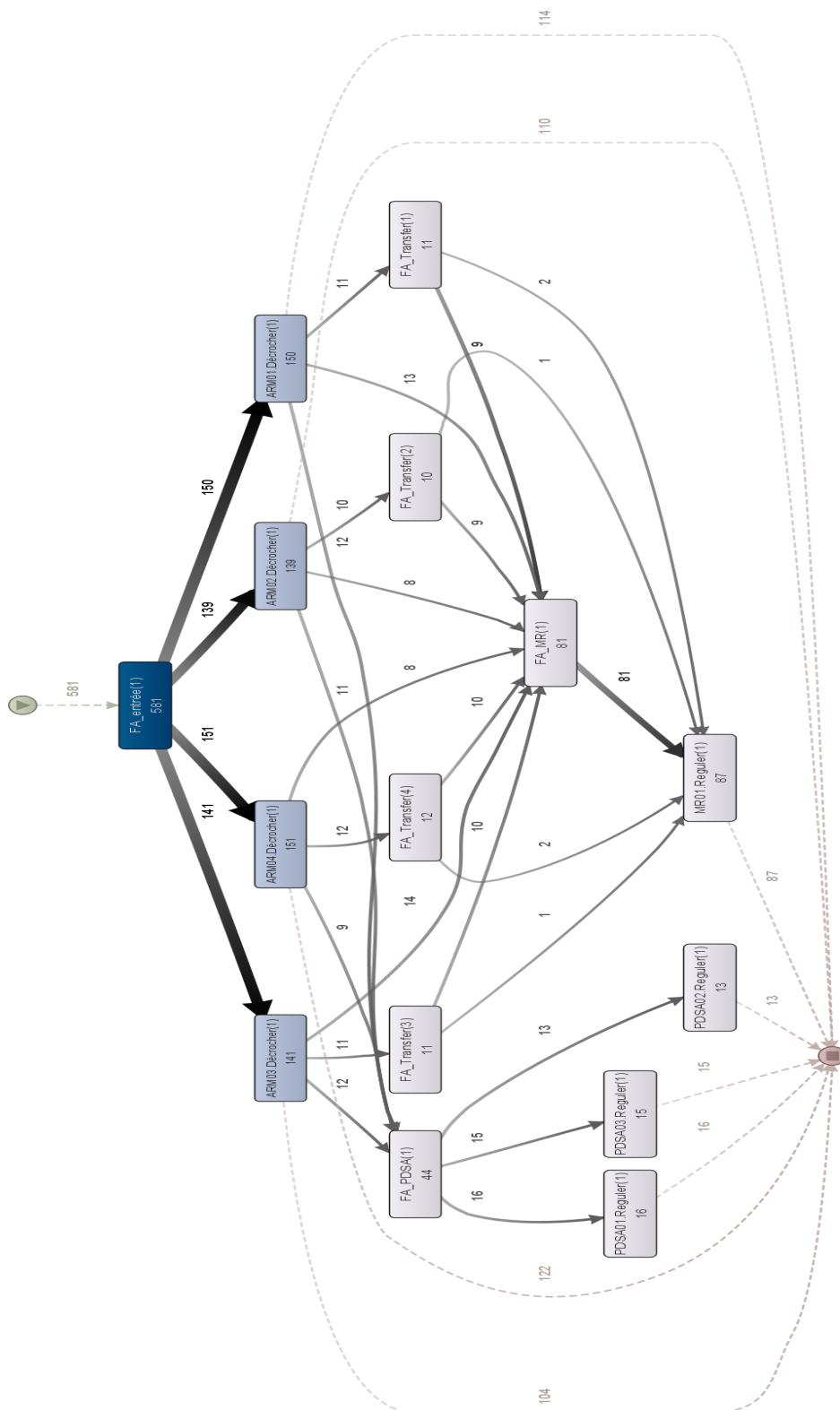


Figure 5.12 – Différents chemins pouvant être empruntés par un appel entrant dans le modèle de simulation

5.2 Étape 4 : Améliorer ; Scénarios de Simulation, Analyses et Résultats

5.2.1 Comment hiérarchiser les types d'appels, quel impact sur les performances ?

Nous avons identifié trois familles d'appels au sein d'un centre d'appels d'urgence : les appels entrants (AE), les appels sortants (AS) et les appels internes (AI). Nous nous sommes questionnés sur les règles de priorité qui peuvent exister entre ces appels. Comme il y a 3 types d'appels à trier et à prioriser, nous en avons donc 3!, soit 6 possibilités à envisager pour l'ARM et pour le MR, soit 12 cas possibles. Fixer ces priorités va permettre de définir les choix des ressources s'ils doivent choisir entre *Décrocher un appel entrant*, *Passer un appel interne* ou *Passer un appel sortant*. Concernant l'ARM 6 choix sont possibles :

- AS>AE>AI : Les appels sortants sont prioritaires sur les appels entrants, qui sont eux-mêmes prioritaires sur les appels internes. Cela implique qu'un ARM peut passer un appel interne uniquement s'il/elle n'a ni appels entrants à décrocher ni appels sortants à passer.
- AS>AI>AE : Les appels sortants sont prioritaires sur les appels internes, eux mêmes prioritaires sur les appels entrants. Cela implique qu'un ARM ne décroche un nouvel appel que s'il/elle a passé tous les appels sortants et appels internes en attente.
- AE>AI>AS : Les appels entrants sont prioritaires sur les appels internes, eux-mêmes prioritaires sur les appels sortants. Cela implique que le décroché des appels est prioritaire par rapport aux communications internes et externes. De plus si on ARM doit contacter un MR et passer un appel sortant, il/elle préférera d'abord contacter le MR.
- AE>AS>AI : Les appels entrants sont prioritaires sur les appels sortants, eux-mêmes prioritaires sur les appels internes. Cela implique que le décroché des appels est prioritaire par rapport aux communications internes et externes. De plus si on ARM doit contacter un MR et passer un appel sortant, il/elle préférera d'abord passer l'appel sortant.
- AI>AE>AS : Les appels internes sont prioritaires sur les appels entrants, eux mêmes prioritaires sur les appels sortants. Dans le cas où un ARM doit passer un appel interne et décrocher un nouvel appel, il/elle passera d'abord son appel interne avant de décrocher l'appel entrant et en dernier lieu il passera un appel sortant.
- AI>AS>AE : Les appels internes sont prioritaires sur les appels sortants eux-mêmes prioritaires sur les appels entrants.

5.2.2 Des propositions pour le traitement des appels : Le Triage et La Prépriorisation

Dans cette section nous proposons plusieurs modifications sur la manière de décrocher les appels. Nous présentons chacune des configurations, puis analysons les résultats

Scénario 0 : AS-IS

Samu Centre 15 référentiel et guide d'évaluation (2015)

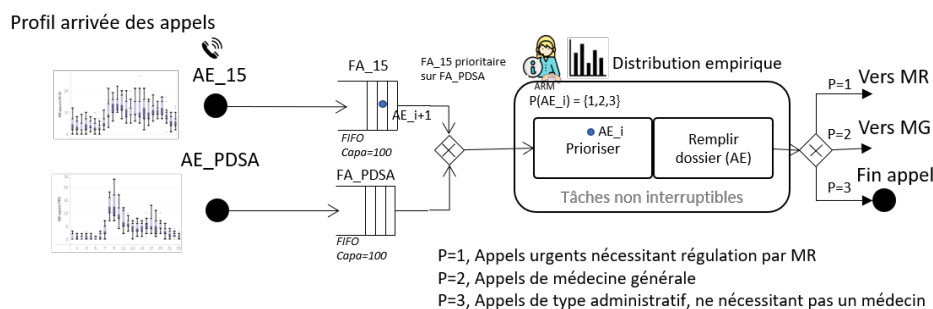


Figure 5.13 – Scénario AS-IS

donnés par chaque scénario. Le plan d'expériences suivi est présenté dans la figure 5.15. Deux nouvelles manières de tester les appels sont proposées. Le Triage et la Prépriorisation. Ces scénarios peuvent être combinés. Cela nous donne donc trois familles de scénarios différents. Pour les scénarios de Triage ainsi que pour le scénario combiné, nous proposons deux versions du scénario en faisant varier le paramètre de durée de la priorisation d'un appel. Cela nous fait donc un total de 6 scénarios à analyser. Afin de faciliter la lecture des résultats, nous proposons différents codes afin d'identifier les résultats :

- AS-IS : code pour un scénario qui correspond à la configuration initiale.
- PP : Scénario de Prépriorisation.
- T-1/3 : Scénario de Triage avec une durée de triage équivalente à un tiers du temps total de gestion de l'appel.
- T-2/3 : Scénario de Triage avec une durée de triage équivalente à deux tiers du temps total de gestion de l'appel.
- PP-T-1/3 : Scénario mixte de Prépriorisation et de Triage avec une durée de triage équivalente à un tiers du temps total de gestion de l'appel.
- PP-T-2/3 : Scénario mixte de Prépriorisation et de Triage avec une durée de triage équivalente à deux tiers du temps total de gestion de l'appel.

La Prépriorisation

En développant un scénario de prépriorisation, nous faisons l'hypothèse qu'à l'avenir il serait plausible d'avoir accès à de l'information sur la gravité de l'appel en amont. Certains pans de la recherche semblent très prometteurs à ce sujet : Que ce soit en utilisant un *Automatic Call Dispatcher* (ACD) pour permettre au patient de faire une première auto-évaluation de sa situation, ou encore du côté de la recherche sur l'analyse vocale des émotions à l'aide de l'intelligence artificielle (Blomberg et al., 2019). Blomberg et al, proposent par exemple d'utiliser un modèle de *Machine Learning* pour analyser la voix des patients afin d'identifier plus facilement les arrêts cardiaques. Leurs résultats montrent que l'algorithme de *Machine Learning* était en mesure de détecter les arrêts cardiaques plus rapidement que les opérateurs téléphoniques même s'il est encore nécessaire de l'améliorer pour augmenter sa précision. Nous pouvons espérer que le développement de ces techniques et de ces

technologies permettra aux centres d'appels de collecter plus de données sur l'appel au plus tôt dans le processus de gestion de l'urgence. Parmi ces nouvelles technologies on peut citer les objets connectés de santé qui peuvent aussi permettre de transmettre des informations permettant d'évaluer la gravité de l'état du patient (rythme cardiaque, hypoxie, pression artérielle, température, ECG etc.) sans oublier l'intégration de la télémédecine qui rajoutera la dimension visuelle au diagnostic.

Dans ce scénario, nous faisons donc l'hypothèse que l'information sur la gravité de l'appel est disponible lorsqu'il entre en file d'attente. De ce fait, l'ARM décrochera d'abord les appels selon leur priorité (cf. figure 5.14). Nous avons gardé une activité de priorisation de l'ARM suite au décroché de l'appel afin de vérifier et de valider la priorisation effectuée en amont.

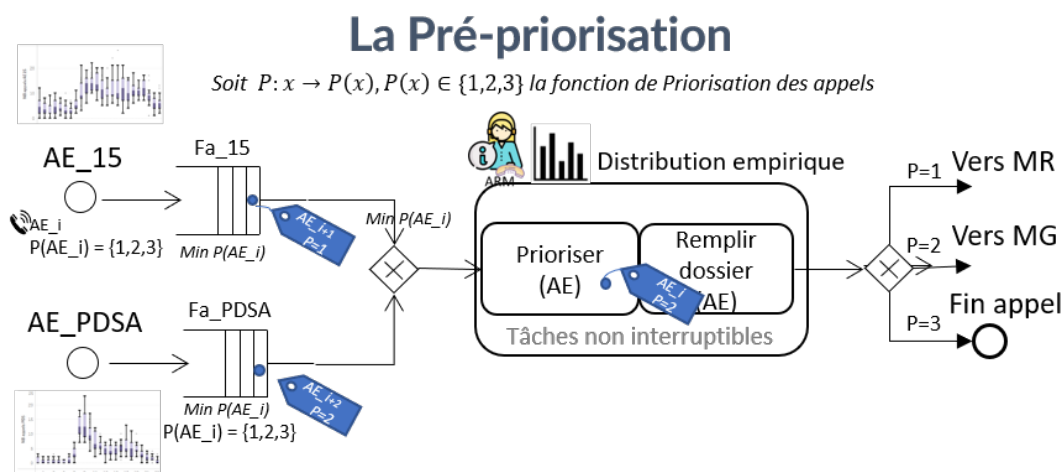


Figure 5.14 – Scénario de prépriorisation des appels

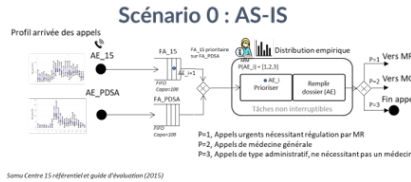
Le Triage

Une pratique qui va prochainement voir le jour suite au pacte de refondation des urgences avec la mise en place du SAS³ est le triage dans les services d'urgences physiques. Cela permet d'évaluer rapidement la gravité des patients pour les orientés du mieux possible et réduire l'engorgement (Storm-Versloot et al., 2011). Dans la continuité des travaux de Lamine et al. (Lamine et al., 2015). Nous proposons d'appliquer le même principe aux centres d'appels d'urgence en mettant la priorité sur le décroché des appels et l'évaluation de leur gravité avant de continuer dans le traitement de l'appel. Pour ce faire, nous scindons la macro activité 1 "Décrocher et prioriser" en deux étapes correspondant à *Prioriser* et *Remplir le dossier* (cf. figure 5.16). La première activité consiste à établir la gravité de l'appel, la deuxième concerne le reste du traitement, avec la création du dossier de régulation médicale et sa complétion. Nous ne disposons pas d'information précise sur la durée de l'activité de priorisation. Ainsi nous proposons de tester deux scénarios, un premier plutôt optimiste où on considère que le temps de priorisation d'un appel correspond à un tiers du temps total de la communication avec l'ARM. Et un deuxième scénario où ce temps de priorisation correspondant à deux tiers du temps total de la communication. À l'issue de cette première étape, l'ARM a plusieurs choix :

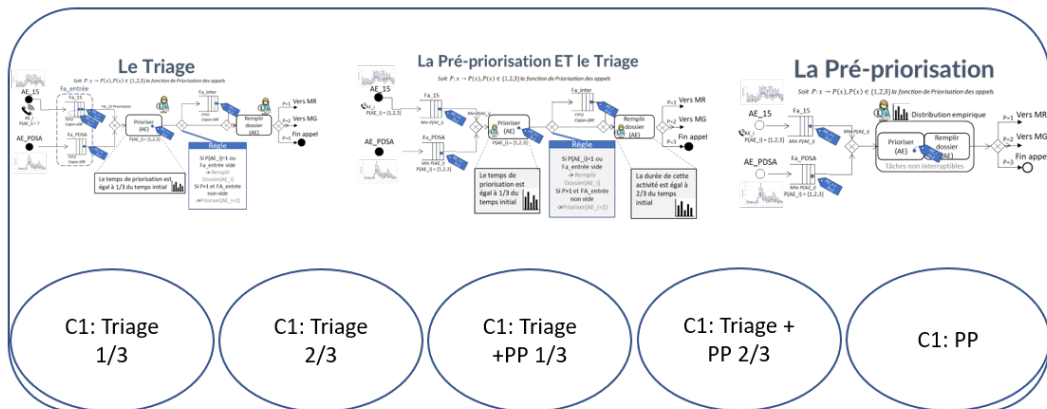
- Soit il n'y a pas d'appels dans la file d'attente d'arrivée des appels, alors on peut enchaîner sur la complétion du dossier de l'appel en cours.
- S'il y a un appel dans la file d'attente alors on regarde la priorité de l'appel en cours.
- Si l'appel en cours est de priorité 1 et est donc urgent, alors on continue de traiter cet appel-là.
- Si l'appel en cours est d'une priorité 2 ou 3 alors on le place dans une file d'attente intermédiaire et on priorise le nouvel appel.
- Si le nouvel appel est d'une priorité supérieure à l'appel mis en attente, on traite d'abord cet appel avant de finir par traiter l'appel mis en attente.

3. <https://www.gouvernement.fr/pacte-de-refondation-des-urgences-754-millions-et-des-mesures-perennes-pour-soutenir-les>

Plan d'expérience



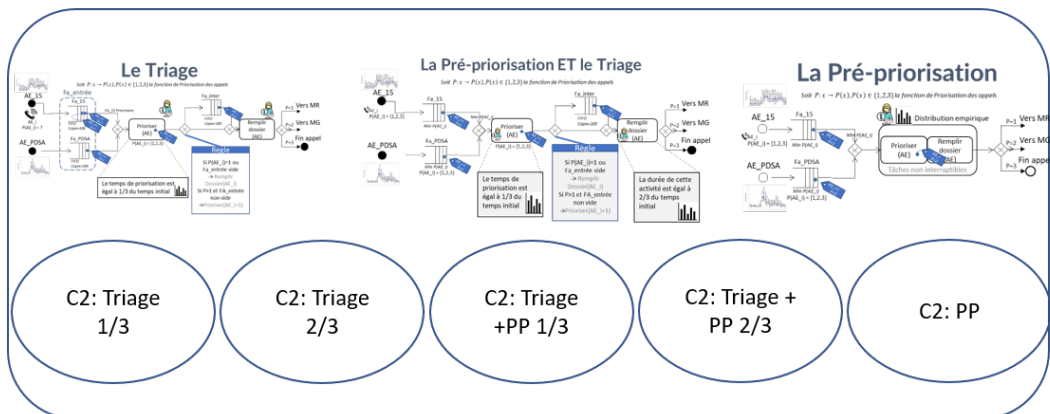
Configuration 1 : Les Appels Sortants sont prioritaires aux Appels Entrants



Analyse des résultats : Attente Maximale, QoS₂₀, QoS₆₀

Les ARM passent du temps sur les AS au détriment des AE donc nous testons une nouvelle configuration

Configuration 2 : Les Appels Sortants peuvent être préemptés au profit des Appels Entrants



36

Figure 5.15 – Plan d'expériences proposé

- Si le nouvel appel est d'une priorité inférieure ou égale à l'appel mis en attente, alors on met ce nouvel appel en attente et on reprend le traitement de l'appel précédent.

5.2.3 1ère Expérimentation : Quels sont les effets de la prépriorisation et du triage ?

Ainsi, nous avons fait une première expérimentation afin d'évaluer l'effet de la prépriorisation, du triage, des deux combinés par rapport au cas initial. Nous appellerons ce jeu de résultat : la configuration C1. La simulation est faite sur un horizon d'une journée, sur une base de 100 réplifications.

Effet sur la qualité de service

Pour chacun des scénarios nous avons calculé le QoS_{20} et le QoS_{60} moyenné sur les 100 réplifications sur chaque créneau de 30 minutes durant la journée. Nous avons détaillé sur chaque créneau de 30 minutes durant la journée, car les gains globaux à l'échelle de la journée pour le QoS est assez faible. Ces résultats sont présentés dans le tableau 5.3. Le scénario de triage permet au mieux un gain de 3% pour le T-1/3 et au pire un gain de 2,7% pour le T-2/3 sur le QoS_{60} . Pour le QoS_{20} les scénarios de triage permettent des gains respectivement de 3,7% et 3,6% pour les scénarios T-1/3 et T-2/3. Le scénario de prépriorisation permet un gain de 0,3% et de 0,5% respectivement pour QoS_{60} et pour QoS_{20} . La combinaison des scénarios de triage et de prépriorisation permettent respectivement des gains sur le QoS_{60} de 3,0% et de 2,7% pour le PP_T-1/3 et le PP_T-2/3. Pour le QoS_{20} c'est là où on note la meilleure amélioration avec des gains de 3,7% et 3,6% pour le PP_T-1/3 et le PP_T-2/3.

Bien que raisonnables, ces modifications présentent déjà des améliorations par rapport au scénario initial. Néanmoins les résultats présentés sont pour tous les appels confondus. Il est intéressant de regarder les gains obtenus par type d'appels. Ces résultats sont présentés dans le tableau 5.4. On constate que les gains obtenus sur chaque type d'appels peuvent varier entre -1,1% jusqu'à 7,1% en fonction du type d'appel et du scénario.

Or, si on calcule le QoS_{20} et le QoS_{60} sur chaque créneau de 30 minutes, on fait apparaître des gains nettement plus importants sur certains créneaux, notamment

Tableau 5.3 – Valeur du QoS au global sur la journée dans la première configuration.

	Moy. QoS_{60}	δ AS-IS	Moy. QoS_{20}	δ AS-IS
AS-IS	95,4%		86,8%	
T-1/3	98,5%	3,0%	90,2%	3,4%
T-2/3	98,2%	2,7%	90,4%	3,2%
PP	95,7%	0,3%	87,3%	0,5%
PP_T-1/3	98,4%	3,0%	90,2%	3,7%
PP_T-2/3	98,2%	2,7%	90,0%	3,6%

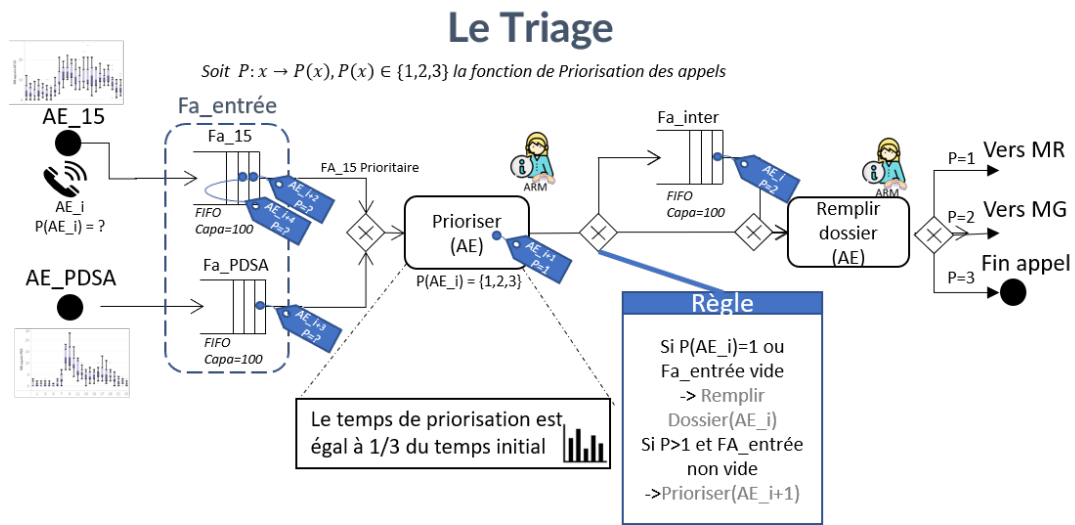


Figure 5.16 – Scénario de triage des appels

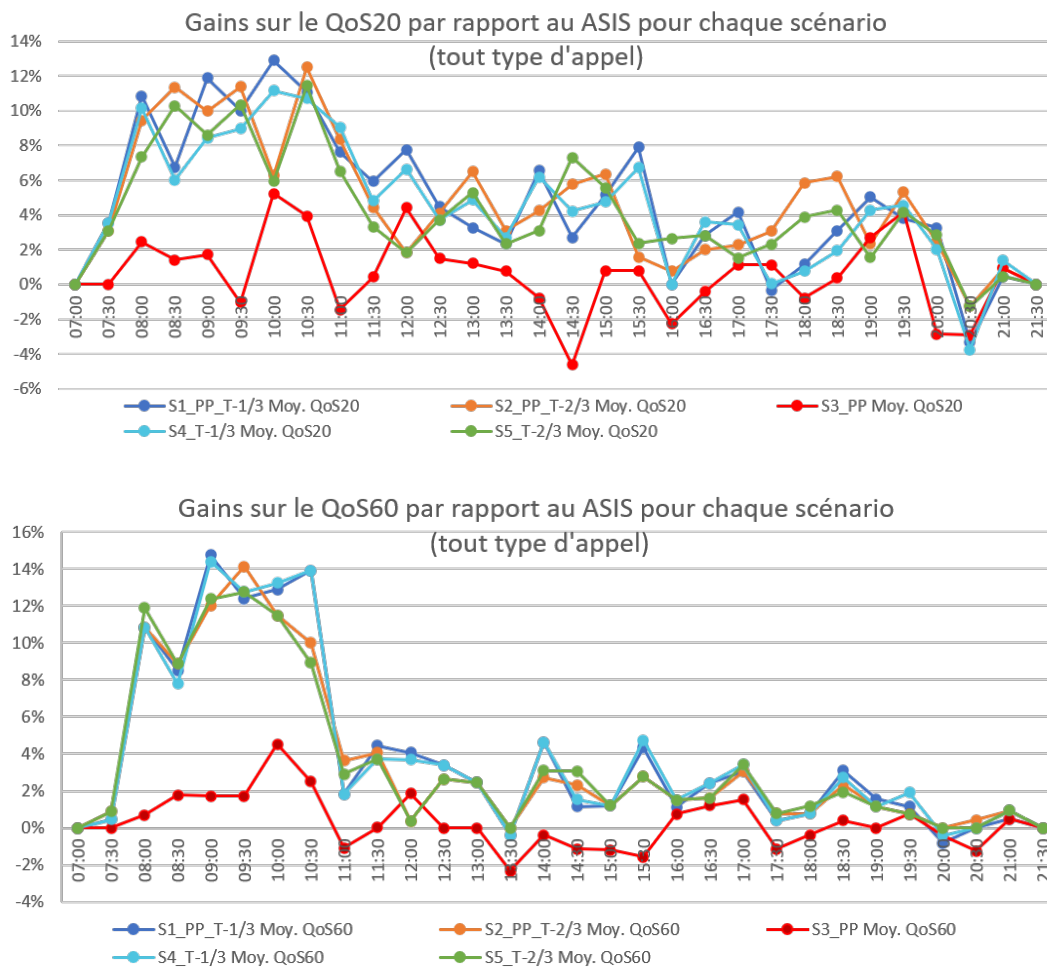


Figure 5.17 – Gains sur le QoS20 et QoS60 au global pour tous type d'appels, dans la configuration C1

Tableau 5.4 – Amélioration du QoS au global par priorité d'appels

		Moy. QoS20	δ AS-IS	Moy. QoS60	δ AS-IS
S1_PP_T-1/3	P1	93,6%	5,2%	99,3%	3,0%
	P2	94,5%	3,3%	99,3%	2,0%
	P3	76,1%	5,7%	95,6%	6,4%
S2_PP_T-2/3	P1	93,1%	4,8%	99,1%	2,7%
	P2	94,2%	3,0%	99,2%	1,9%
	P3	76,8%	6,3%	95,0%	5,8%
S3_PP	P1	90,3%	2,0%	97,6%	1,3%
	P2	92,5%	1,3%	98,3%	1,0%
	P3	69,4%	-1,1%	88,0%	-1,1%
S4_T-1/3	P1	92,8%	4,5%	99,0%	2,6%
	P2	94,3%	3,1%	99,1%	1,8%
	P3	76,1%	5,6%	96,2%	7,1%
S5_T-2/3	P1	92,2%	3,9%	98,6%	2,2%
	P2	93,5%	2,3%	99,0%	1,7%
	P3	76,7%	6,2%	95,5%	6,3%
S6_AS-IS	P1	88,4%	-	96,6%	-
	P2	91,2%	-	97,3%	-
	P3	70,5%	-	89,2%	-

Tableau 5.5 – Gains par rapport au AS-IS pour les appels P1 sur chaque créneau

Appels P1	S1_PP_T-1/3		S2_PP_T-2/3		S3_PP		S4_T-1/3		S5_T-2/3	
	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60
07 :30	3,0%	0,0%	3,0%	0,0%	0,0%	0,0%	3,0%	0,0%	3,0%	0,0%
08 :00	9,6%	4,3%	10,6%	5,3%	4,3%	2,1%	7,4%	4,3%	5,3%	3,2%
08 :30	8,9%	6,7%	14,4%	7,8%	4,4%	3,3%	6,7%	4,4%	13,3%	7,8%
09 :00	17,9%	10,5%	11,6%	8,4%	3,2%	4,2%	14,7%	7,4%	10,5%	6,3%
09 :30	9,3%	11,3%	12,4%	12,4%	2,1%	4,1%	7,2%	11,3%	8,2%	10,3%
10 :00	16,8%	11,6%	5,3%	10,5%	8,4%	9,5%	13,7%	10,5%	7,4%	10,5%
10 :30	10,5%	10,5%	12,6%	6,3%	6,3%	4,2%	12,6%	10,5%	10,5%	4,2%
11 :00	10,8%	0,0%	9,7%	1,1%	1,1%	-1,1%	10,8%	0,0%	6,5%	-1,1%
11 :30	3,4%	1,1%	2,2%	2,2%	0,1%	1,1%	-3,4%	1,1%	3,4%	1,1%
12 :00	4,4%	2,2%	2,2%	1,1%	4,4%	1,1%	3,3%	2,2%	1,1%	1,1%
12 :30	0,0%	3,4%	0,0%	2,3%	0,0%	3,4%	0,0%	2,3%	0,0%	2,3%
13 :00	7,7%	2,2%	6,6%	2,2%	2,2%	0,0%	7,7%	2,2%	6,6%	2,2%
13 :30	2,2%	-1,1%	5,6%	0,0%	5,6%	-1,1%	3,4%	-1,1%	5,6%	0,0%
14 :00	14,4%	4,4%	8,9%	2,2%	2,2%	1,1%	12,2%	4,4%	10,0%	2,2%
14 :30	8,4%	2,1%	6,3%	4,2%	-6,3%	-2,1%	7,4%	1,1%	8,4%	4,2%
15 :00	1,1%	1,1%	2,3%	1,1%	2,3%	0,0%	3,4%	1,1%	1,1%	1,1%
15 :30	6,3%	4,2%	0,0%	4,2%	2,1%	2,1%	7,4%	4,2%	-1,1%	3,2%
16 :00	0,0%	2,2%	2,2%	2,2%	-2,2%	1,1%	-1,1%	2,2%	2,2%	2,2%
16 :30	3,9%	2,6%	3,9%	2,6%	0,0%	1,3%	3,9%	2,6%	3,9%	2,6%
17 :00	4,5%	3,4%	2,2%	3,4%	2,2%	2,2%	3,4%	3,4%	0,0%	3,4%
17 :30	0,0%	0,0%	4,5%	-1,1%	5,6%	-1,1%	1,1%	-1,1%	3,4%	-1,1%
18 :00	1,1%	1,1%	1,1%	1,1%	2,2%	1,1%	2,2%	1,1%	1,1%	1,1%
18 :30	3,4%	1,1%	2,2%	0,0%	0,0%	-1,1%	0,0%	0,0%	0,0%	-1,1%
19 :00	3,3%	0,0%	2,2%	0,0%	3,3%	0,0%	1,1%	0,0%	1,1%	0,0%
19 :30	4,4%	1,1%	5,6%	0,0%	5,6%	1,1%	4,4%	1,1%	3,3%	0,0%
20 :00	-1,1%	-1,1%	0,0%	-2,2%	1,1%	0,0%	-1,1%	-1,1%	-2,2%	-2,2%
20 :30	-3,3%	1,1%	1,1%	1,1%	-3,3%	0,0%	-2,2%	1,1%	0,0%	0,0%
21 :00	1,3%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
21 :30	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

Tableau 5.6 – Gains par rapport au AS-IS pour les appels P2 sur chaque créneau

S1_PP_T-1/3		S2_PP_T-2/3		S3_PP		S4_T-1/3		S5_T-2/3	
Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60
0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00	0,02	0,00
0,11	0,05	0,09	0,05	0,03	0,01	0,10	0,02	0,05	0,07
0,04	0,03	0,03	0,03	0,03	0,02	0,03	0,01	0,02	0,02
0,09	0,08	0,08	0,06	0,02	0,07	0,04	0,07	0,04	0,06
0,06	0,09	0,12	0,11	0,01	0,04	0,06	0,08	0,10	0,10
0,10	0,04	0,00	0,04	0,04	0,03	0,08	0,05	0,00	0,03
0,09	0,08	0,11	0,07	0,07	0,04	0,09	0,06	0,04	0,05
0,05	-0,01	0,04	0,01	0,02	0,01	0,05	0,00	0,01	0,01
0,05	0,02	0,04	0,02	0,01	0,01	0,07	0,02	0,04	0,02
0,01	0,02	-0,02	0,01	0,01	0,02	0,01	0,02	0,00	0,00
0,08	0,03	0,02	0,01	0,03	0,00	0,06	0,03	0,02	0,01
-0,02	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00	0,00
0,03	0,00	0,04	0,00	0,03	-0,01	0,03	0,00	0,04	0,00
0,00	0,01	0,00	0,01	0,00	0,00	-0,01	0,01	0,00	0,01
0,03	0,00	0,05	0,00	0,00	-0,02	0,05	0,00	0,05	0,00
0,06	0,02	0,05	0,02	0,00	0,02	0,06	0,02	0,05	0,02
0,10	0,03	0,09	0,02	0,05	0,02	0,09	0,03	0,09	0,03
0,01	0,01	0,00	0,01	0,01	0,01	0,03	0,01	0,01	0,01
0,00	0,00	-0,03	0,00	-0,03	-0,01	0,01	0,00	-0,01	0,00
0,03	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00
-0,03	0,01	0,01	0,00	-0,03	0,01	-0,01	0,01	0,00	0,00
0,02	0,00	0,03	0,00	-0,02	0,00	0,02	0,00	0,00	0,00
0,00	0,01	0,03	0,01	0,00	0,01	-0,01	0,01	0,04	0,01
0,01	0,01	0,01	0,01	0,01	0,00	0,01	0,01	0,00	0,01
0,05	0,03	0,07	0,01	0,05	0,01	0,05	0,03	0,05	0,01
-0,04	0,00	0,00	0,00	-0,04	-0,02	-0,04	0,00	0,00	0,00
0,02	0,00	0,00	0,00	0,02	0,00	0,02	0,00	0,00	0,00
-0,03	-0,03	0,00	0,00	-0,03	0,00	0,00	0,00	0,00	0,00
0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Tableau 5.7 – Gains par rapport au AS-IS pour les appels P3 sur chaque créneau

S1_PP_T-1/3		S2_PP_T-2/3		S3_PP		S4_T-1/3		S5_T-2/3	
Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60	Moy. QoS20	Moy. QoS60
0,06	0,01	0,04	0,02	0,00	0,00	0,06	0,01	0,04	0,02
0,12	0,22	0,09	0,21	0,00	-0,01	0,13	0,25	0,11	0,25
0,07	0,15	0,16	0,15	-0,03	0,00	0,08	0,17	0,15	0,16
0,09	0,25	0,10	0,21	0,00	-0,06	0,07	0,28	0,11	0,24
0,14	0,17	0,10	0,19	-0,06	-0,03	0,13	0,19	0,13	0,18
0,12	0,22	0,13	0,19	0,03	0,01	0,12	0,23	0,10	0,20
0,13	0,22	0,14	0,16	-0,01	0,00	0,10	0,24	0,19	0,17
0,07	0,06	0,11	0,08	-0,07	-0,03	0,11	0,05	0,11	0,08
0,09	0,09	0,07	0,07	0,00	-0,02	0,10	0,07	0,03	0,07
0,16	0,07	0,05	-0,01	0,07	0,02	0,14	0,06	0,04	0,00
0,06	0,04	0,09	0,04	0,02	-0,03	0,05	0,05	0,08	0,04
0,02	0,04	0,10	0,04	0,00	0,00	0,05	0,04	0,07	0,04
0,02	0,00	0,00	0,00	-0,05	-0,04	0,02	0,00	-0,02	0,00
0,04	0,07	0,03	0,04	-0,04	-0,02	0,06	0,07	-0,01	0,05
-0,03	0,01	0,06	0,02	-0,06	0,00	0,01	0,03	0,08	0,04
0,08	0,01	0,11	0,01	0,00	-0,04	0,05	0,01	0,10	0,01
0,08	0,05	-0,01	0,02	-0,03	-0,07	0,05	0,06	0,02	0,02
-0,01	0,00	0,00	0,01	-0,05	0,00	-0,01	0,01	0,04	0,01
0,04	0,04	0,04	0,02	0,01	0,03	0,05	0,04	0,05	0,02
0,05	0,05	0,03	0,05	0,00	0,02	0,05	0,06	0,03	0,06
0,01	0,00	0,03	0,03	0,00	-0,03	0,00	0,01	0,03	0,03
0,01	0,01	0,12	0,01	-0,03	-0,02	-0,01	0,01	0,09	0,02
0,05	0,06	0,12	0,05	0,01	0,01	0,06	0,06	0,08	0,05
0,09	0,03	0,03	0,02	0,03	0,00	0,09	0,02	0,03	0,02
0,02	0,00	0,04	0,01	0,02	0,00	0,04	0,02	0,04	0,01
0,11	-0,01	0,06	0,02	-0,06	0,00	0,08	0,00	0,09	0,02
-0,06	-0,01	-0,04	0,00	-0,05	-0,03	-0,08	-0,01	-0,03	0,00
0,01	0,02	0,02	0,02	0,03	0,01	0,03	0,02	0,01	0,02
0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Tableau 5.8 – Qualité de service et gains par rapport à l’AS-IS pour chaque scénario

	AS-IS		Triage				Prériorisation				Triage + Prériorisation			
	QoS60	QoS20	QoS 20	δ AS-IS	QoS60	δ AS-IS	QoS20	δ AS-IS	QoS60	δ AS-IS	QoS20	δ AS IS	QoS60	δ AS IS
07 :00	1,00	0,99	1,00	0,5%	1,00	0,0%	0,99	0,0%	1,00	0,0%	1,00	0,5%	1,00	0,0%
07 :30	0,99	0,95	0,99	3,6%	1,00	0,4%	0,95	0,0%	0,99	0,0%	0,99	3,6%	1,00	0,4%
08 :00	0,84	0,63	0,73	9,8%	0,95	10,8%	0,66	3,1%	0,86	1,7%	0,74	10,5%	0,95	10,8%
08 :30	0,85	0,64	0,72	7,8%	0,95	9,6%	0,67	2,5%	0,88	2,5%	0,73	8,5%	0,95	10,3%
09 :00	0,76	0,50	0,57	7,4%	0,91	15,1%	0,51	1,0%	0,79	3,1%	0,61	10,9%	0,91	15,4%
09 :30	0,82	0,59	0,70	10,7%	0,94	12,4%	0,61	2,4%	0,83	1,0%	0,71	11,7%	0,94	12,1%
10 :00	0,82	0,56	0,68	11,8%	0,93	11,5%	0,58	2,1%	0,85	3,8%	0,69	13,6%	0,93	11,1%
10 :30	0,81	0,58	0,65	6,4%	0,95	13,2%	0,60	1,8%	0,85	3,2%	0,65	6,8%	0,95	13,2%
11 :00	0,91	0,67	0,79	11,6%	0,96	4,7%	0,68	0,4%	0,89	-1,4%	0,78	10,1%	0,96	4,7%
11 :30	0,94	0,83	0,91	8,5%	0,99	4,8%	0,86	2,9%	0,95	1,1%	0,92	9,6%	1,00	5,5%
12 :00	0,96	0,82	0,88	6,3%	0,99	3,3%	0,85	3,0%	0,97	1,8%	0,89	7,4%	0,99	3,7%
12 :30	0,97	0,89	0,92	2,3%	0,99	1,9%	0,89	0,0%	0,97	0,4%	0,92	3,0%	0,99	1,9%
13 :00	0,99	0,93	0,96	3,3%	1,00	1,2%	0,93	0,0%	0,98	-0,4%	0,94	1,6%	1,00	1,2%
13 :30	0,99	0,90	0,95	4,7%	0,99	0,4%	0,91	0,8%	0,97	-1,9%	0,94	4,3%	0,99	0,4%
14 :00	0,93	0,81	0,89	8,1%	1,00	6,6%	0,81	-0,4%	0,94	0,8%	0,90	8,5%	1,00	6,6%
14 :30	0,96	0,82	0,85	3,1%	0,97	0,8%	0,80	-2,3%	0,95	-1,2%	0,84	1,5%	0,97	0,4%
15 :00	0,98	0,89	0,93	4,0%	1,00	1,2%	0,90	0,8%	0,97	-1,6%	0,94	4,4%	1,00	1,2%
15 :30	0,96	0,79	0,89	9,5%	1,00	3,6%	0,82	3,2%	0,95	-1,6%	0,90	10,7%	1,00	3,2%
16 :00	0,97	0,90	0,91	1,1%	0,99	1,9%	0,89	-1,5%	0,97	-0,4%	0,91	1,1%	0,99	1,5%
16 :30	0,98	0,89	0,94	5,2%	1,00	2,0%	0,89	0,0%	0,98	-0,4%	0,93	4,4%	1,00	2,0%
17 :00	0,98	0,89	0,95	6,1%	1,00	2,3%	0,92	2,7%	0,98	0,0%	0,96	6,8%	1,00	1,9%
17 :30	0,98	0,89	0,88	-1,1%	0,98	0,4%	0,89	0,8%	0,98	0,0%	0,87	-1,5%	0,98	0,4%
18 :00	0,99	0,91	0,93	2,3%	1,00	0,8%	0,91	0,8%	0,97	-1,6%	0,93	2,7%	1,00	0,8%
18 :30	0,97	0,87	0,91	4,3%	0,99	1,9%	0,90	3,5%	0,98	0,4%	0,92	5,4%	1,00	2,3%
19 :00	0,99	0,95	0,96	0,4%	0,99	0,4%	0,95	-0,8%	0,98	-0,4%	0,97	1,2%	1,00	0,8%
19 :30	0,97	0,86	0,94	7,2%	0,99	2,7%	0,87	0,8%	0,97	0,4%	0,93	6,4%	0,98	1,9%
20 :00	0,97	0,92	0,92	0,4%	0,98	0,4%	0,93	1,6%	0,98	1,2%	0,93	1,6%	0,97	0,0%
20 :30	0,99	0,96	0,94	-2,1%	1,00	0,4%	0,96	0,0%	0,99	0,0%	0,94	-1,7%	1,00	0,4%
21 :00	0,99	0,98	0,99	0,9%	1,00	1,4%	0,97	-0,5%	1,00	0,9%	0,98	0,0%	1,00	0,9%
21 :30	1,00	1,00	1,00	0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%

ceux qui présentent un nombre d’appels élevés. Un premier constat est qu’entre 00h00 et 07h00 ainsi qu’entre 21h30 et 23h30, le QoS_{20} est de 100%, tous les appels sont décrochés en moins de 20 secondes. C’est pourquoi dans le tableau 5.8 nous ne présentons que les créneaux entre 07h00 et 21h30. La mise en place du triage permet d’obtenir un gain pouvant aller jusqu’à 15% par rapport au scénario AS-IS. La figure 5.17 présente les gains en pourcentage par rapport au scénario de base pour chacun des scénarios. Par exemple, sur le créneau de 09h00, le QoS_{60} passe de 0,76 à 0,91. Ce sont sur les créneaux où il y a le plus d’appels, soit entre 7h30 et 11h00 où on observe les améliorations les plus sensibles, entre 8% et 15% pour les scénarios de triage et les scénarios mixtes. Ce qu’il est intéressant de noter c’est que le scénario de triage permet de passer au-dessus des 90% d’appels décrochés en moins de 60 secondes entre 7h30 et 11h00.

Une autre chose que l’on peut voir est que sur certains créneaux, la prériorisation ou le triage a un impact légèrement négatif (au plus -5%) sur les performances. De plus, on constate que le scénario PP est le moins performant de tous. Cette perte de performance se doit d’être expliquée. Pour ce faire, il est nécessaire d’avoir plus de précisions. Si on détaille les gains par niveau de gravité des appels (voir figures 5.25, 5.27 et 5.20) on constate que les gains sont bien meilleurs pour les appels de priorité P1 et que la perte de performance a lieu sur les appels de priorité P3.

Pour les appels P1, le scénario de triage seul améliore jusque 18% du QoS_{60} sur certains créneaux, comme celui de 9h00, entre un temps de triage à 1/3 ou 2/3 on note une différence de 3% entre les deux scénarios. Donc, même avec un temps de triage très important on note des gains non négligeables en matière de décroché

des appels. Le scénario de PP permet des gains pouvant aller jusque 8% (comme à 10h00). Néanmoins on note une baisse en moyenne pour les appels P1 notamment sur le créneau de 14h30. Cette diminution de production est due au fait que lorsqu'un appel est décroché, les ARM ne vont pas interrompre l'appel en cours même s'il y a de nouveaux appels à décrocher. De ce fait, si les 4 ARM décrochent des appels qui sont des P3 et qu'un appel P1 arrive en file d'attente, ce dernier va attendre qu'un des ARM ait fini de traiter son appel courant. Ce cas de figure n'est pas possible dans les autres scénarios, c'est pourquoi les résultats sont globalement meilleurs.

Pour les appels P2, les gains vont jusqu'à 12% d'amélioration pour le QoS_{20} , ce sont encore une fois les scénarios proposant du triage seul ou mixé à la prépriorisation qui permettent d'obtenir les meilleurs résultats. Le QoS_{60} est amélioré de 8 à 10 % lors du premier pic d'appels au niveau des scénarios de triage et des scénarios mixtes, le scénario de prépriorisation ne permet une amélioration que de 4%.

On constate enfin une légère baisse de productivité sur les appels P3 : en mettant en place du triage ou de la prépriorisation, on accepte de dégrader les performances de décroché pour les appels de faible gravité afin d'augmenter la performance pour les appels de forte gravité. C'est pourquoi le scénario seul de prépriorisation a un effet globalement négatif sur les appels P3, car ils ne sont pas prioritaires dans ce cas de figure. On constate néanmoins l'intérêt des scénarios de triage et des scénarios mixtes, car ils permettent tout de même d'aller améliorer aussi le décrocher des appels P3. En effet, la particularité du triage est de mettre l'accent sur le décroché des appels pour qu'il soit décroché puis évalué le plus rapidement possible. On améliore ainsi les performances globales du système en termes de décroché des appels.

Effet sur le temps d'attente maximum

Pour chaque scénario, nous avons regardé le temps maximum d'attente sur l'intégralité des répliques afin de voir l'effet des modifications organisationnelles sur les points les plus extrêmes. La figure 5.21 présente les résultats obtenus pour tous les appels confondus. On a un maximum de temps d'attente qui est à 274,8 secondes sur le créneau de 9h00. C'est entre 07h00 et 21h30 où les temps d'attentes maximums sont élevés et globalement supérieurs à 1 minute. Au niveau des scénarios, on remarque que la prépriorisation a des performances globalement équivalentes au AS-IS. Les scénarios de triage permettent de réduire le temps d'attente maximal à 200 secondes. Les scénarios mixtes ont des résultats qui dépendent de la durée du triage, ils sont sensiblement meilleurs si ce dernier est à 1/3, car on observe une réduction du maximum à 200 secondes contre une amélioration bien moins nette et plus proche du cas initial pour le 2/3.

Lorsqu'on étudie les appels de gravité P1, on constate que les scénarios de triage permettent une amélioration globale des performances. On obtient une réduction du temps maximal d'attente pouvant aller jusqu'à 50% par rapport au scénario AS-IS. Sur le créneau de 09h00, on passe d'un maximum de 192 à 102 secondes d'attente au maximum, ce qui est très significatif pour un appel d'urgence. En revanche, plusieurs points ont un maximum qui est supérieur au scénario principal sur les créneaux de 11h00, 13h30, 14h30, 17h30, 18h30 et 20h30. Nous avons donc identifié les répliques

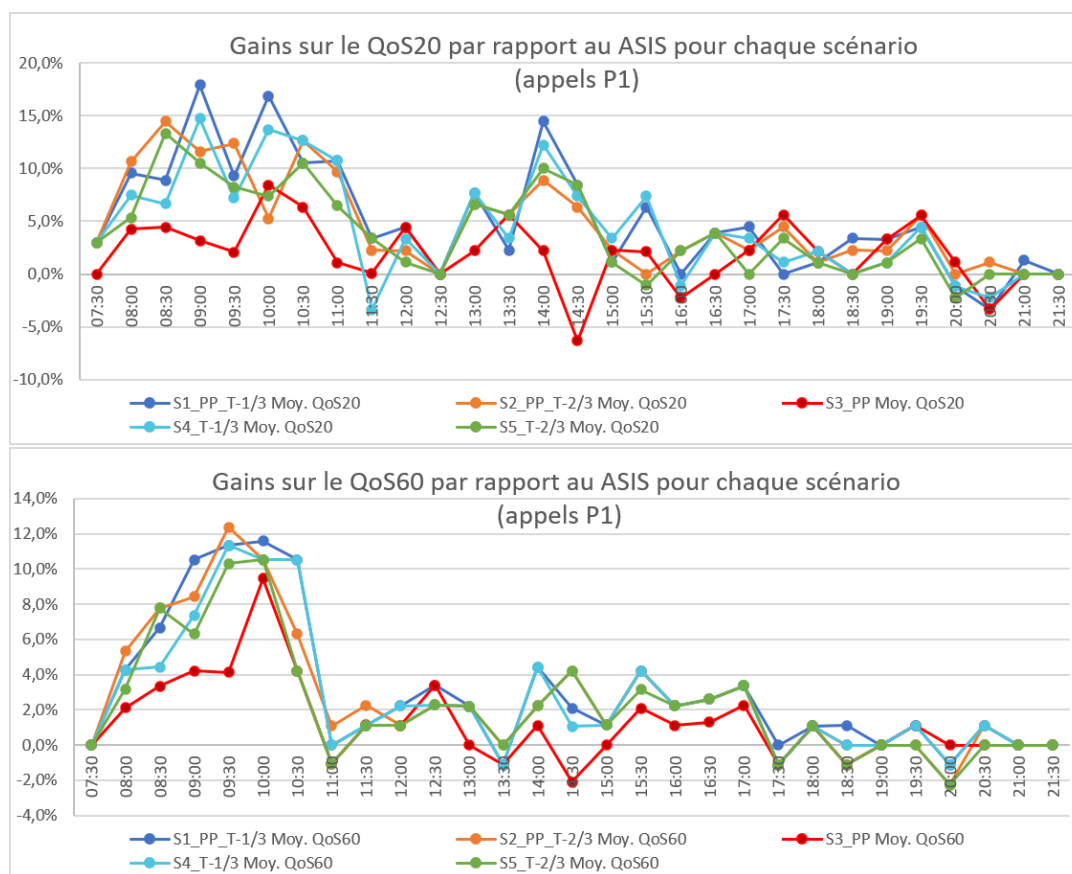


Figure 5.18 – Gains sur les QoS20 et QoS60 pour les appels P1 dans la configuration C1

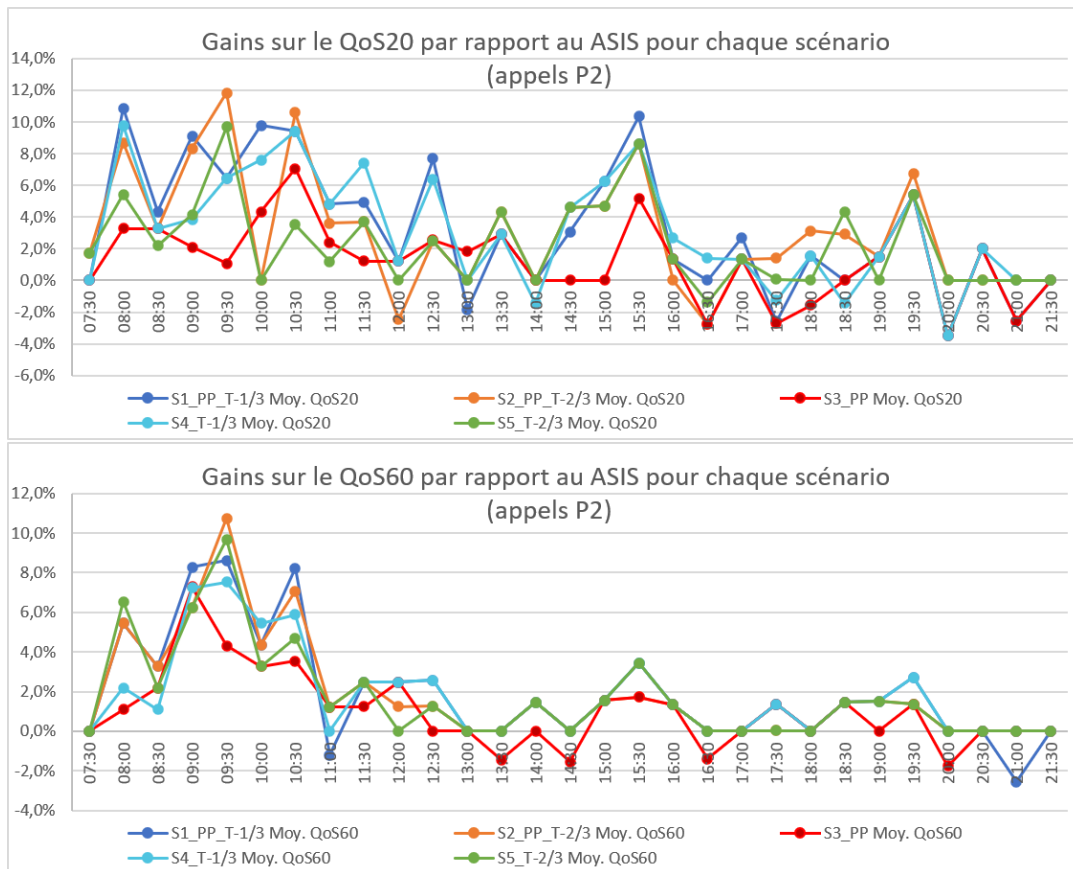


Figure 5.19 – Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C1

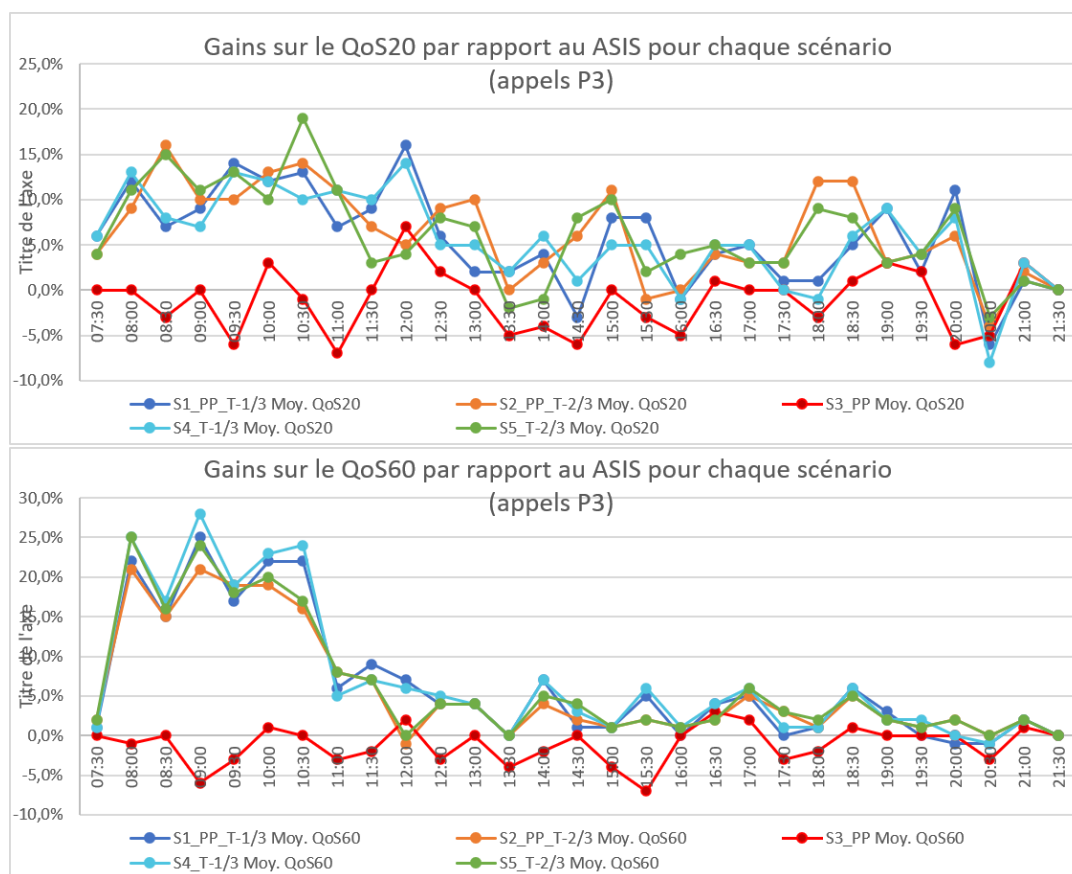


Figure 5.20 – Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C1

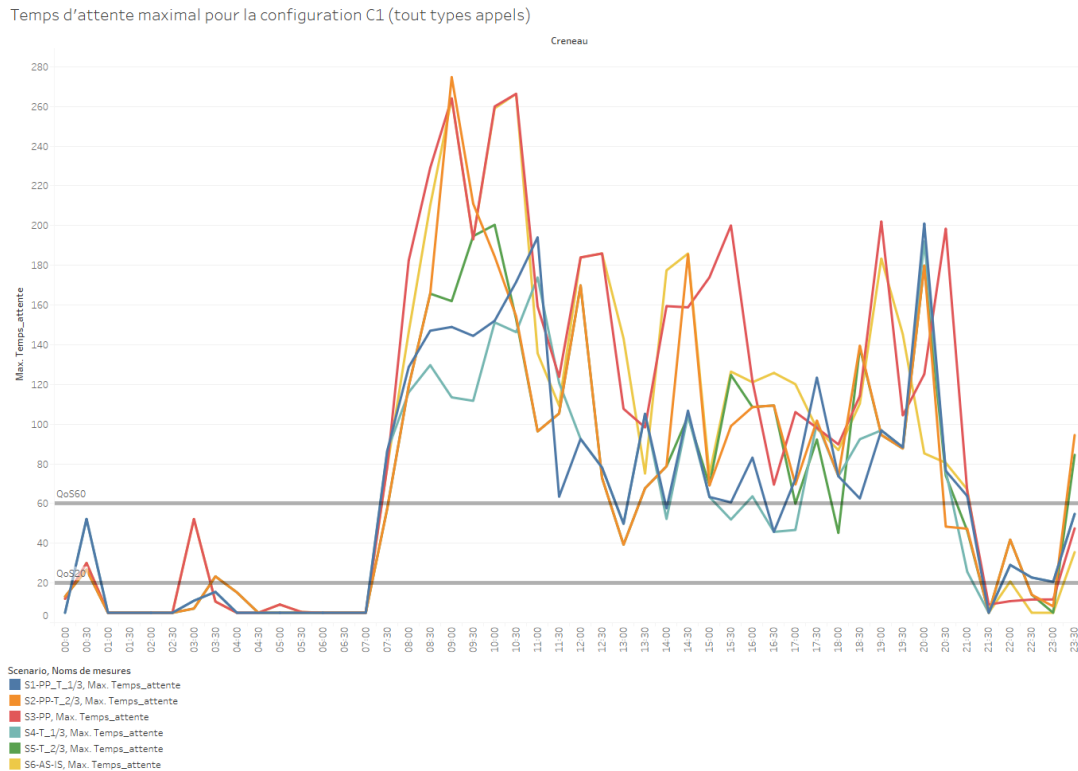


Figure 5.21 – Temps d’attente maximum pour chaque scénario de la configuration C1

et les appels avec un temps d’attente aussi long pour expliquer la cause de cette attente. Cette vérification directement dans le log de données de la simulation nous permet d’aller vérifier le comportement du modèle et de voir si cet écart au niveau des temps d’attente est une erreur inhérente au modèle ou si les temps d’attentes sont longs, car les agents étaient déjà en ligne. A titre d’exemple, nous allons détailler un cas ci-dessous :

Sur le créneau de 18h30 de la réplcation 12, on constate que l’appel 531 attend 92,4 secondes avant d’être décroché. Il arrive dans le centre d’appels à 18h42m48 et il n’est pas décroché avant 18h44m20. Pour comprendre pourquoi il n’est pas décroché, nous regardons ce que font les ARM durant cette période.

- L’ARM 1 est en train de remplir le dossier de l’appel 528 depuis 18h42m28 jusque 18h44m20 puis il priorise l’appel 532 qui est de priorité 1 aussi.
- L’ARM 2 est en train de remplir le dossier de l’appel 530 depuis 18h42m46 jusque 18h46m00.
- L’ARM3 passe un appel sortant externe de 18h40m35 jusque 18h44m18 puis il décroche l’appel 531 à 18h44m20.
- L’ARM 4 remplis le dossier de l’appel 527 de 18h42m38 jusque 18h44m05.

Donc il n’y a pas d’ARM disponible pour décrocher l’appel avant 18h44m20, car ils sont déjà tous soit en train de traiter un appel entrant, soit en train de passer un appel sortant. Pour tous les autres appels, on retrouve le même cas, les appels attendent, car les ressources sont déjà en ligne sur d’autres appels. On constate

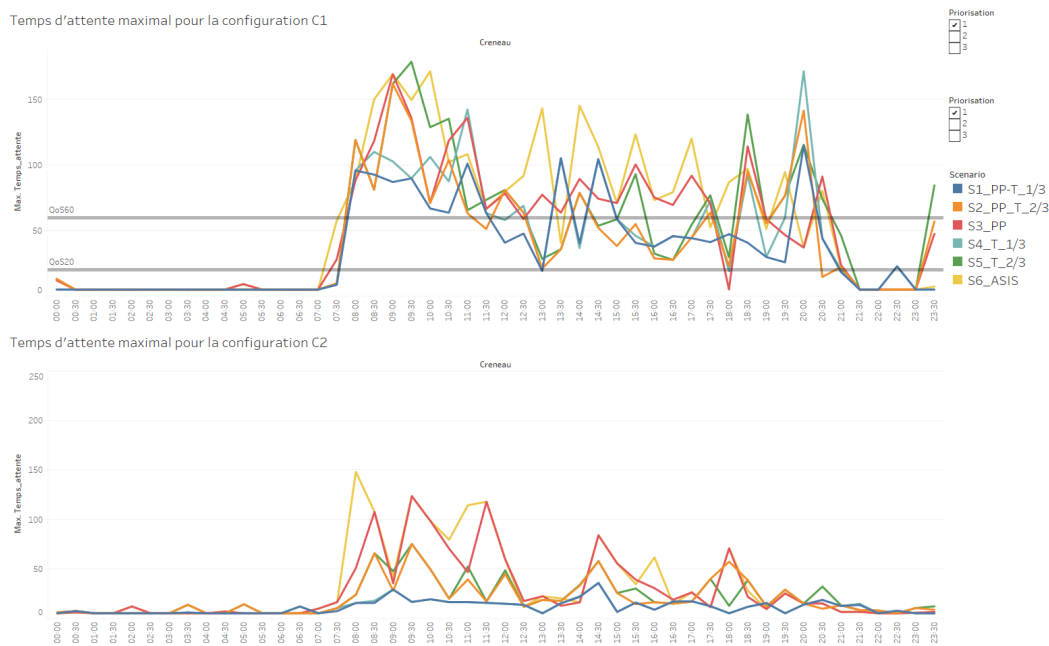


Figure 5.22 – Temps d’attente maximum pour chaque scénario de la configuration C1 : appels P1

que les ressources sont souvent occupées avec des appels sortants. Le traitement prioritaire des appels sortants semble donc avoir un impact fort sur la performance des centres d’appels. En effet, que ce soit pour les appels de 11h, 13h30, 14h30, ou 20h il y a trois ARM sur quatre qui sont en communication sur un appel sortant. Donc les temps d’attente longs ne sont pas dû à une erreur inhérente au modèle mais sont le résultat d’une réalité organisationnelle associée au scénario étudié.

Nous avons effectué le même travail sur le scénario de prépriorisation. Au global, la prépriorisation permet une amélioration sur le temps d’attente maximum de 11,8 secondes en moyenne avec un écart type de 35,6 secondes. Nous avons donc certains cas où les performances sont réduites avec l’utilisation de la prépriorisation. En effet, la pire performance est une augmentation du temps d’attente de 84,4 secondes par rapport au pire cas du scénario initial. Par contre, la meilleure amélioration est une diminution de 103,2 secondes par rapport au cas initial. Dans le tableau 5.9 nous détaillons les activités courantes des ARM dans les cas où le scénario de prépriorisation ne donne pas de meilleurs résultats que le scénario initial. On observe que les ARM sont plus de la moitié du temps en train de passer un appel sortant plutôt que de décrocher un appel entrant. Comme les appels sortants ne sont pas interruptibles même si des appels entrants sont en attentes d’être décrochés, cela engendre des temps d’attentes plus longs. Les temps d’attentes ne sont donc pas dus à la nature de la gestion des files d’attente, mais à la limite de la contrainte des activités non interruptibles.

De ce constat, nous avons envisagé un nouveau scénario d’amélioration dans lequel nous relaxons la priorité sur les appels sortants.

Tableau 5.9 – Activité courante des ARM plutôt que de décrocher l'appel entrant

	ARM 1	ARM 2	ARM 3	ARM 4
09h00	AS	AS	AS	AS
11h00	Décrocher 223	AS	Décrocher 221	AS
13h30	AS	AS	Décrocher 316	Décrocher 315
15h30	AS	AS	Décrocher 423	AS
19h00	AS	Décrocher	AS	AS
20h30	Décrocher 551	AS	AS	AS
22h30	Décrocher	Décrocher	Décrocher	AS

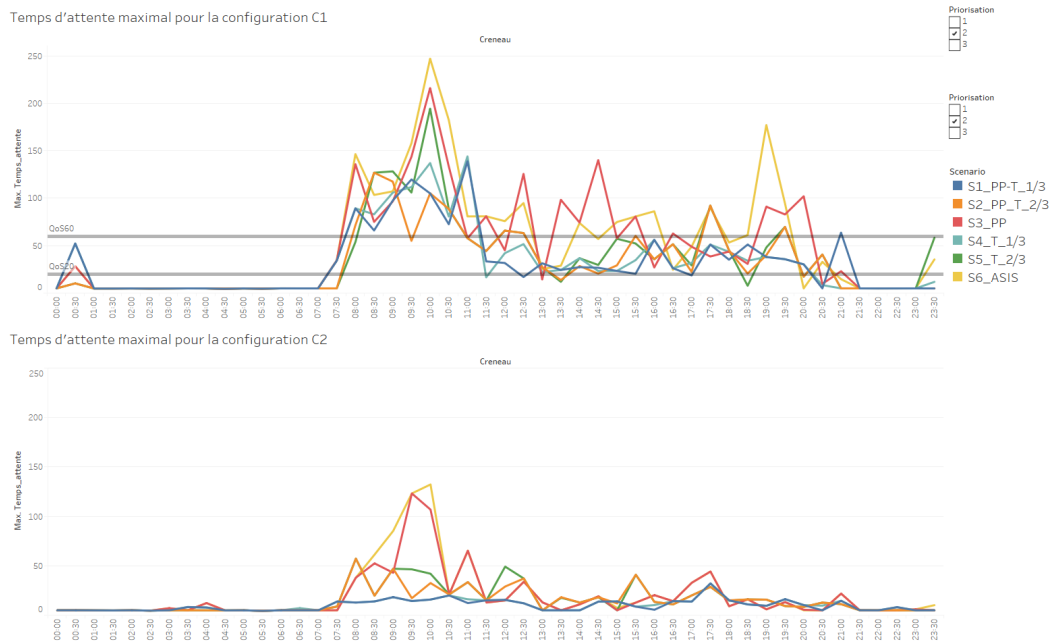


Figure 5.23 – Temps d'attente maximum pour chaque scénario de la configuration C1 : appels P2

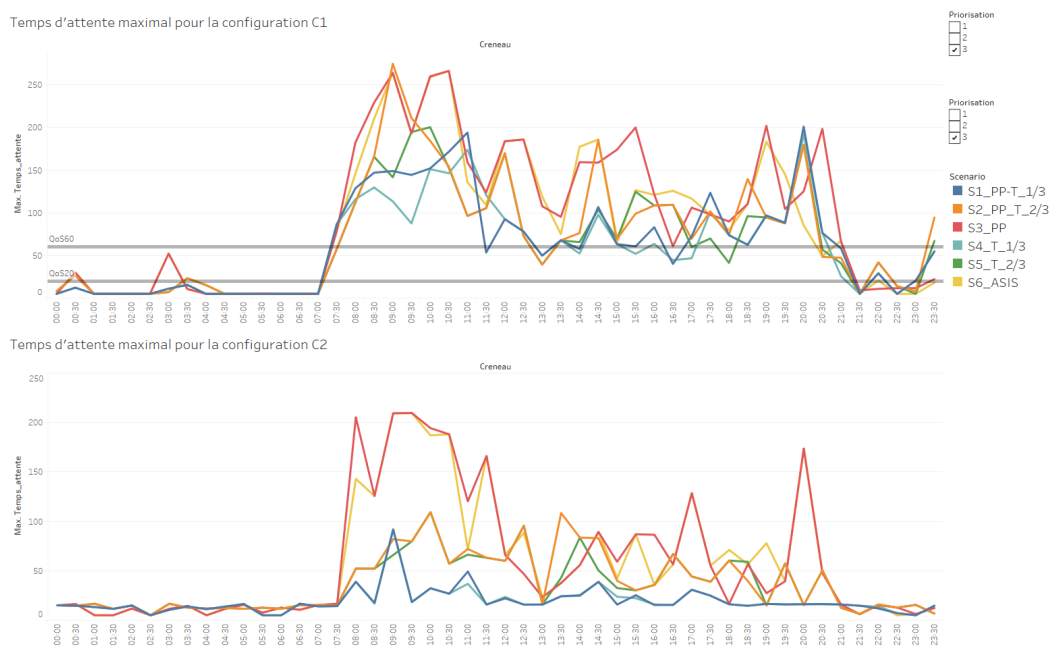


Figure 5.24 – Temps d’attente maximum pour chaque scénario de la configuration C1 : appels P3

La mise en place d’un système de triage ou de prépriorisation permet de gagner du temps de façon significative sur le décroché des appels. Concernant les appels les plus urgents de priorité P1 : Les scénarios mixtes de triage permettent d’améliorer le QoS_{20} de 3 à 5% en moyenne et de 12 à 18% au maximum. Le triage mixte améliore également le QoS_{60} de 3% en moyenne et de 12% au maximum. Les scénarios de prépriorisation seuls permettent une amélioration de 1 à 2% des QoS en moyenne et une amélioration au maximum de 8 à 9 %. Les scénarios de triage seuls permettent une amélioration de 4% en moyenne sur le QoS_{20} et de 13 à 15% au maximum et de 3% en moyenne et de 11% au maximum sur le QoS_{60} . Toutefois on observe des pertes de temps non négligeables dus à la non-disponibilité des ressources. Ainsi, nous souhaitons tester l’hypothèse suivante : *Autoriser l’interruption des appels sortants au profit du décroché des appels entrants peut-il avoir un impact positif sur la performance du décroché ?*

5.2.4 Une piste d’amélioration : modifier la priorisation des appels sortants

Dans cette configuration nous proposons d’autoriser la préemption des ressources pour le décroché des appels. Prenons un exemple pour détailler ce que cela signifie : jusqu’à présent, si un ARM commence un appel sortant et qu’un appel entrant arrive pendant ce temps, l’ARM termine son appel sortant puis va décrocher l’appel entrant. Dans ce nouveau scénario, si un appel entrant arrive alors qu’un ARM

est en train de passer un appel sortant, il peut mettre en pause son appel sortant pour décrocher l'appel entrant afin de s'en occuper. Dans ces conditions nous avons relancé une simulation des scénarios dans cette configuration que nous nommerons : Configuration C2. Les résultats de ces simulations sont comparés au scénario AS-IS de la configuration C1.

Effet sur la qualité de service

Dans la table 5.10 nous comparons les QoS_{20} et les QoS_{60} des différents scénarios à ceux du AS-IS modifié avec la règle des appels sortants pouvant être préemptés. Tous les types d'appels sont concernés et indifférenciés dans un premier temps. Plusieurs constats peuvent être faits :

- Bien qu'on observe toujours de légères variations négatives par rapport à l'AS-IS, celles-ci sont maintenant inférieures à 1% pour les appels de catégorie P3.
- Le scénario de triage améliore le QoS_{20} jusqu'à 21% du AS-IS sur certains créneaux. En comparaison, dans la première configuration le gain maximum avec le scénario de Triage pour le QoS_{20} était de 15%.
- Le scénario de prépriorisation a des gains moindres (au maximum 2,8% pour le QoS_{20} , et 2,1% pour le QoS_{60}).
- Le scénario de triage permet d'atteindre l'objectif de 99% des appels décrochés en moins de 60 secondes sur l'intégralité des créneaux. De plus, il permet d'avoir plus de 90% des appels décrochés en moins de 20 secondes sur l'ensemble des créneaux.
- Les organisations combinées du Triage et de la prépriorisation permettent une amélioration des QoS_{20} et QoS_{60} pouvant aller respectivement de 22,1% et 13,4%. Les gains sont les mêmes que pour le scénario de triage. La combinaison des deux est un peu meilleure sur certains créneaux comme à 12h00 où 98% des appels sont décrochés en moins de 20 secondes pour 97% avec seulement le Triage. Le triage seul semble suffisant pour améliorer les performances du centre d'appels.

Dans la table 5.11 nous comparons les qualités de service pour les différents scénarios avec la nouvelle configuration au scénario AS-IS de la configuration initiale. Cela permet de se rendre compte de la totalité des gains acquis. On observe des améliorations non négligeables. Dans le cas initial, on pouvait avoir jusqu'à 50% des appels qui n'étaient pas décrochés en moins de 20 secondes sur certains créneaux. La prépriorisation permet de passer à un minimum de 71% des appels décrochés en moins de 20 secondes. Le triage améliore encore ce résultat avec un minimum de 90% des appels décrochés en moins de 20 secondes. Cela correspond donc à une amélioration pouvant aller de 21 à 40% selon le scénario.

Effet sur le temps d'attente maximum

Cette nouvelle configuration permet aussi de réduire le temps maximal avant le décroché des appels. En effet, préempter les ressources sur les appels sortants au profit

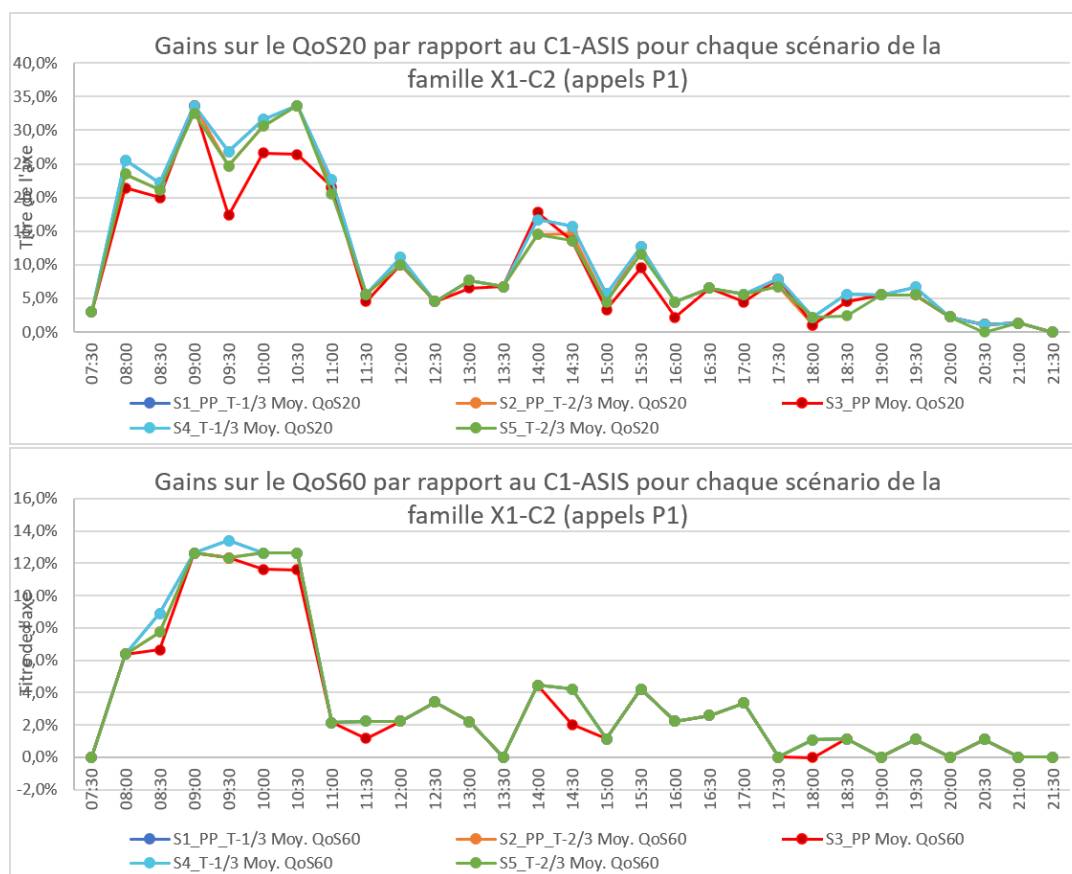


Figure 5.25 – Gains sur les QoS20 et QoS60 pour les appels P1 dans la configuration C2

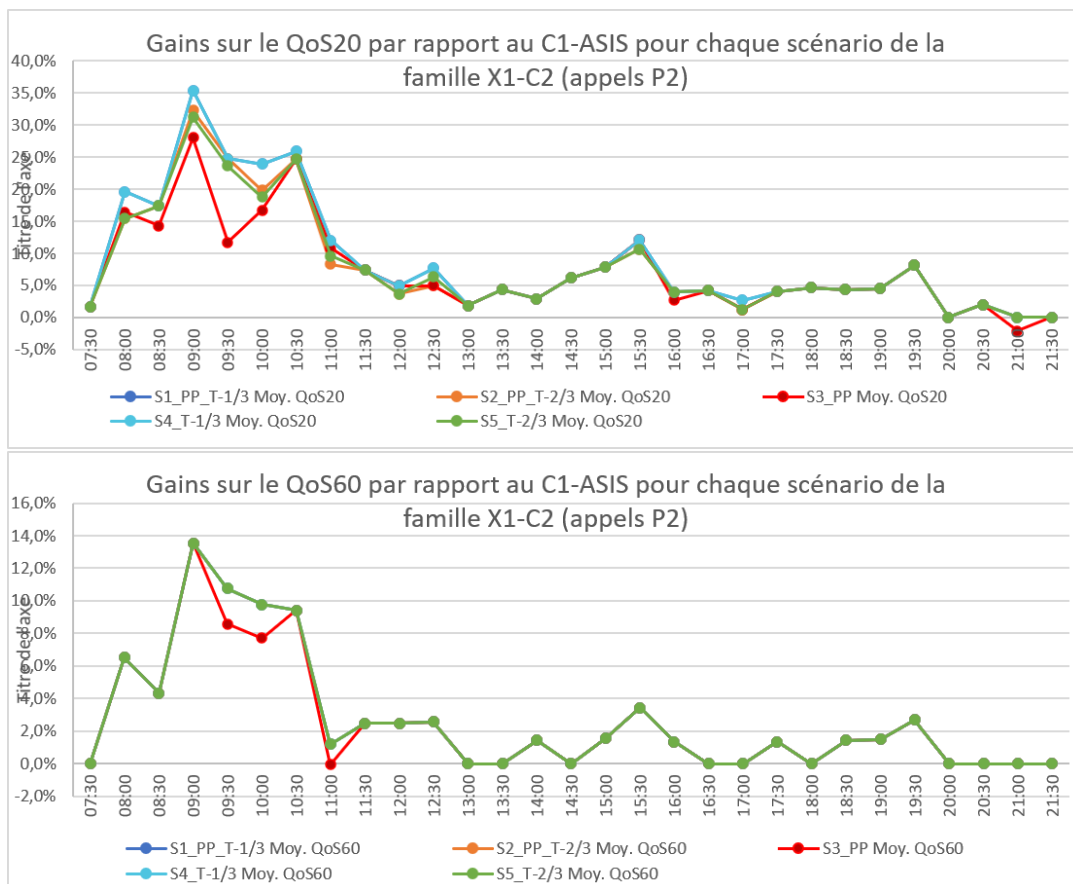


Figure 5.26 – Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C2

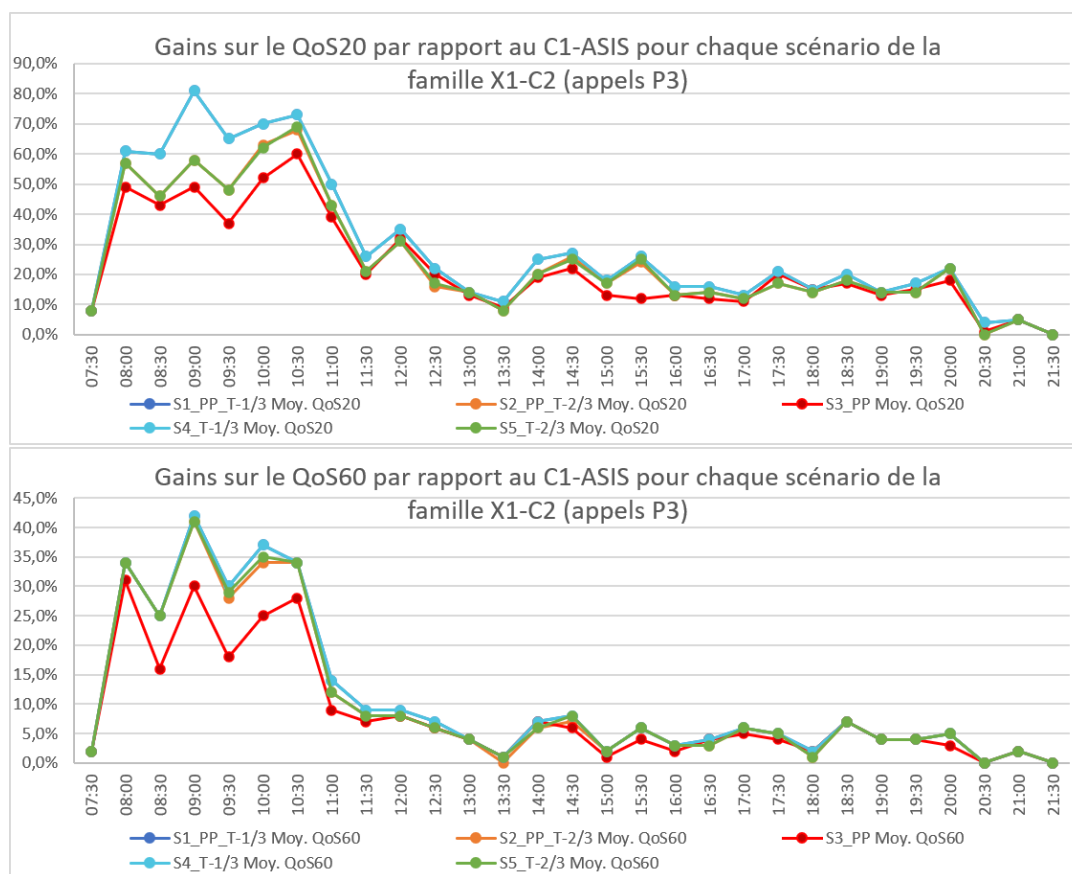


Figure 5.27 – Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C2

Tableau 5.10 – Configuration 2 : Comparaison du QoS entre les scénarios et l’AS-IS.

AS préemptés - Scénarios comparés entre eux														
Creneau	S4-AS-IS		S1 - PT&PP				S2-Pré-Priorisation				S3 - Triage			
	QoS20	QoS60	QoS20	δ AS-IS	QoS60	δ AS-IS	QoS20	δ AS-IS	QoS60	δ AS-IS	QoS20	δ AS-IS	QoS60	δ AS-IS
07 :00	1,00	1,00	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%
07 :30	0,99	1,00	1,00	0,9%	1,00	0,0%	0,99	0,0%	1,00	0,0%	1,00	0,9%	1,00	0,0%
08 :00	0,82	0,95	0,98	15,7%	1,00	4,9%	0,82	0,0%	0,95	0,0%	0,98	15,7%	1,00	4,9%
08 :30	0,87	0,96	0,99	12,1%	1,00	3,9%	0,87	0,0%	0,96	0,0%	0,99	11,7%	1,00	3,9%
09 :00	0,70	0,86	0,90	20,3%	0,99	13,4%	0,71	1,4%	0,88	1,7%	0,90	20,3%	0,99	13,4%
09 :30	0,72	0,87	0,94	22,1%	1,00	12,4%	0,74	2,4%	0,89	2,1%	0,93	21,7%	1,00	12,4%
10 :00	0,77	0,87	0,92	15,7%	1,00	12,2%	0,79	2,8%	0,89	1,7%	0,92	15,7%	1,00	12,2%
10 :30	0,81	0,91	0,97	16,1%	1,00	8,9%	0,82	1,1%	0,93	1,4%	0,98	16,4%	1,00	8,9%
11 :00	0,86	0,95	0,98	11,9%	1,00	4,7%	0,88	2,2%	0,96	0,4%	0,98	11,9%	1,00	4,7%
11 :30	0,92	0,97	0,97	5,6%	1,00	2,6%	0,92	0,0%	0,97	-0,7%	0,97	5,6%	1,00	2,6%
12 :00	0,94	0,97	0,99	4,4%	1,00	3,3%	0,94	0,4%	0,97	0,7%	0,98	4,1%	1,00	3,3%
12 :30	0,96	0,98	0,99	3,4%	1,00	1,5%	0,96	0,4%	0,98	0,0%	0,99	3,0%	1,00	1,5%
13 :00	1,00	1,00	0,99	-0,4%	1,00	0,0%	0,99	-0,4%	1,00	0,0%	0,99	-0,4%	1,00	0,0%
13 :30	0,98	0,99	0,98	0,0%	1,00	0,8%	0,98	0,4%	1,00	0,4%	0,98	0,0%	1,00	0,8%
14 :00	0,95	1,00	0,99	4,2%	1,00	0,4%	0,94	-0,8%	0,99	-0,8%	0,99	4,2%	1,00	0,4%
14 :30	0,95	0,98	0,98	2,7%	1,00	1,5%	0,96	0,8%	0,98	0,0%	0,98	2,7%	1,00	1,5%
15 :00	0,96	0,99	0,98	2,4%	1,00	1,2%	0,97	0,8%	1,00	0,8%	0,98	2,4%	1,00	1,2%
15 :30	0,94	0,98	0,96	2,4%	1,00	2,4%	0,93	-0,8%	0,98	0,0%	0,96	2,0%	1,00	2,4%
16 :00	0,97	0,99	1,00	2,6%	1,00	1,1%	0,96	-0,8%	0,99	0,0%	0,99	2,3%	1,00	1,1%
16 :30	0,98	1,00	0,99	0,4%	1,00	0,4%	0,98	0,0%	1,00	0,0%	0,99	0,4%	1,00	0,4%
17 :00	0,96	0,98	1,00	3,8%	1,00	1,9%	0,96	-0,4%	0,97	-0,8%	1,00	3,8%	1,00	1,9%
17 :30	0,97	0,99	0,98	1,9%	1,00	0,8%	0,96	-0,8%	0,99	-0,4%	0,98	1,9%	1,00	0,8%
18 :00	0,95	1,00	0,99	3,9%	1,00	0,4%	0,96	1,2%	0,99	-0,4%	0,99	3,9%	1,00	0,4%
18 :30	0,96	0,98	0,98	2,3%	1,00	1,6%	0,95	-0,8%	0,98	0,0%	0,98	1,9%	1,00	1,6%
19 :00	0,99	1,00	1,00	0,4%	1,00	0,0%	0,99	0,0%	1,00	0,0%	1,00	0,4%	1,00	0,0%
19 :30	0,96	0,99	0,98	1,9%	1,00	0,8%	0,96	0,0%	0,99	0,0%	0,98	1,9%	1,00	0,8%
20 :00	0,96	0,98	0,99	2,8%	1,00	2,0%	0,96	0,0%	0,98	0,0%	0,99	2,8%	1,00	2,0%
20 :30	0,99	1,00	1,00	1,3%	1,00	0,4%	0,99	0,0%	1,00	0,4%	1,00	1,3%	1,00	0,4%
21 :00	0,99	0,99	1,00	0,5%	1,00	0,9%	0,99	0,0%	0,99	0,0%	1,00	0,5%	1,00	0,9%
21 :30	1,00	1,00	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%

Tableau 5.11 – Comparaison des QoS entre les scénarios de la nouvelle configuration et l’AS-IS de la configuration originale.

AS préempté - Scénario comparé à l’AS-IS sans préemption														
	S4-AS-IS_1		S1 - PT&PP				S2-Pré-Priorisation				S3 - Triage			
	QoS20	QoS60	QoS20	δ AS-IS-1	QoS60	δ AS-IS-1	QoS20	δ AS-IS-1	QoS60	δ AS-IS-1	QoS20	δ AS-IS-1	QoS60	δ AS-IS-1
07 :00	0,99	1,00	1,00	0,5%	1,00	0,0%	1,00	0,5%	1,00	0,0%	1,00	0,5%	1,00	0,0%
07 :30	0,95	0,99	1,00	4,9%	1,00	0,9%	0,99	4,0%	1,00	0,9%	1,00	4,9%	1,00	0,9%
08 :00	0,63	0,84	0,98	34,3%	1,00	15,7%	0,82	18,5%	0,95	10,8%	0,98	34,3%	1,00	15,7%
08 :30	0,64	0,85	0,99	34,8%	1,00	14,9%	0,87	22,7%	0,96	11,0%	0,99	34,4%	1,00	14,9%
09 :00	0,50	0,76	0,90	40,2%	0,99	23,4%	0,71	21,3%	0,88	11,7%	0,90	40,2%	0,99	23,4%
09 :30	0,59	0,82	0,94	34,8%	1,00	17,6%	0,74	15,2%	0,89	7,2%	0,93	34,5%	1,00	17,6%
10 :00	0,56	0,82	0,92	36,6%	1,00	18,1%	0,79	23,7%	0,89	7,7%	0,92	36,6%	1,00	18,1%
10 :30	0,58	0,81	0,97	38,9%	1,00	18,6%	0,82	23,9%	0,93	11,1%	0,98	39,3%	1,00	18,6%
11 :00	0,67	0,91	0,98	30,4%	1,00	9,1%	0,88	20,7%	0,96	4,7%	0,98	30,4%	1,00	9,1%
11 :30	0,83	0,94	0,97	14,8%	1,00	5,9%	0,92	9,2%	0,97	2,6%	0,97	14,8%	1,00	5,9%
12 :00	0,82	0,96	0,99	17,0%	1,00	4,4%	0,94	12,9%	0,97	1,8%	0,98	16,6%	1,00	4,4%
12 :30	0,89	0,97	0,99	9,8%	1,00	3,0%	0,96	6,8%	0,98	1,5%	0,99	9,4%	1,00	3,0%
13 :00	0,93	0,99	0,99	6,5%	1,00	1,2%	0,99	6,5%	1,00	1,2%	0,99	6,5%	1,00	1,2%
13 :30	0,90	0,99	0,98	8,1%	1,00	1,2%	0,98	8,5%	1,00	0,8%	0,98	8,1%	1,00	1,2%
14 :00	0,81	0,93	0,99	18,1%	1,00	6,6%	0,94	13,1%	0,99	5,4%	0,99	18,1%	1,00	6,6%
14 :30	0,82	0,96	0,98	15,8%	1,00	3,5%	0,96	13,8%	0,98	1,9%	0,98	15,8%	1,00	3,5%
15 :00	0,89	0,98	0,98	9,1%	1,00	1,6%	0,97	7,5%	1,00	1,2%	0,98	9,1%	1,00	1,6%
15 :30	0,79	0,96	0,96	17,0%	1,00	3,6%	0,93	13,8%	0,98	1,2%	0,96	16,6%	1,00	3,6%
16 :00	0,90	0,97	1,00	9,4%	1,00	2,6%	0,96	6,0%	0,99	1,5%	0,99	9,1%	1,00	2,6%
16 :30	0,89	0,98	0,99	10,0%	1,00	2,0%	0,98	9,6%	1,00	1,6%	0,99	10,0%	1,00	2,0%
17 :00	0,89	0,98	1,00	10,6%	1,00	2,3%	0,96	6,5%	0,97	-0,4%	1,00	10,6%	1,00	2,3%
17 :30	0,89	0,98	0,98	9,8%	1,00	2,3%	0,96	7,2%	0,99	1,1%	0,98	9,8%	1,00	2,3%
18 :00	0,91	0,99	0,99	8,6%	1,00	1,2%	0,96	5,8%	0,99	0,4%	0,99	8,6%	1,00	1,2%
18 :30	0,87	0,97	0,98	11,2%	1,00	2,7%	0,95	8,1%	0,98	1,2%	0,98	10,9%	1,00	2,7%
19 :00	0,95	0,99	1,00	4,3%	1,00	1,2%	0,99	3,9%	1,00	1,2%	1,00	4,3%	1,00	1,2%
19 :30	0,86	0,97	0,98	11,7%	1,00	3,4%	0,96	9,8%	0,99	2,7%	0,98	11,7%	1,00	3,4%
20 :00	0,92	0,97	0,99	7,3%	1,00	2,8%	0,96	4,5%	0,98	0,8%	0,99	7,3%	1,00	2,8%
20 :30	0,96	0,99	1,00	4,2%	1,00	0,8%	0,99	2,9%	1,00	0,8%	1,00	4,2%	1,00	0,8%
21 :00	0,98	0,99	1,00	1,9%	1,00	1,4%	0,99	1,4%	0,99	0,5%	1,00	1,9%	1,00	1,4%
21 :30	1,00	1,00	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%	1,00	0,0%

des appels entrants permet de réduire le temps d'attente maximal à 58,3 secondes avec le scénario de triage et à 116,6 secondes avec le scénario de prépriorisation. Cette amélioration engendre un gain de temps pouvant aller respectivement jusqu'à 109 et 152 secondes pour la prépriorisation et le Triage. C'est une amélioration de plus de deux minutes par rapport à l'AS-IS de la configuration initiale. Les scénarios de triage et de prépriorisation apportaient déjà une amélioration significative pouvant aller jusque 108 secondes. Avec la nouvelle règle de gestion des appels sortants il est possible de gagner 48 secondes supplémentaires. La figure 5.28 représente la comparaison du temps maximum d'attente avant décroché entre les deux configurations, pour chaque scénario, pour les appels de gravité supérieure. Encore une fois, nous ne présentons les résultats qu'à partir de 07h00 jusque 21h30, car les créneaux précédents ne sont pas significatifs. Sur cette figure nous avons fait apparaître deux lignes de références correspondant au QoS_{20} et au QoS_{60} . Si la courbe est au-dessus de cette ligne alors cela signifie qu'il y a des appels qui sont décrochés en plus de 20 ou 60 secondes. Plusieurs résultats ressortent de ce graphique :

- Dans la deuxième configuration, il y a deux scénarios qui permettent de faire en sorte que tous les appels soient décrochés en moins de 60 secondes sur chaque créneau de la journée. Ce sont les scénarios de triage et celui combinant triage et prépriorisation.
- Le scénario de prépriorisation seul donne des résultats assez peu satisfaisants, car il donne des résultats assez similaires à l'AS-IS de la deuxième configuration, sur certains créneaux on observe une diminution du temps maximum pouvant aller jusque 68% de celui du AS-IS. À 12h00 on passe d'un maximum à 121,6 secondes à 38,6 secondes. On note toutefois une dégradation du temps sur le créneau de 13h30 où le temps maximal du AS-IS est de 20,9 secondes, et avec le scénario de prépriorisation il passe à 86,7 secondes. Ce résultat étant surprenant, nous allons étudier le log de données plus en détail afin de comprendre ce qu'il s'est passé pour avoir une telle baisse de performance.
- Concernant le scénario de triage, il y a quelques créneaux où le temps d'attente est supérieur à celui du scénario AS-IS, lorsque ce temps d'attente était supérieur à 20 secondes nous avons aussi vérifié dans le log la raison de cette modification.

Prenons l'appel 304 de la réplique 89. Il arrive dans le centre d'appels à 13h03m15 et ne sera décroché que 40,7 secondes plus tard. La figure 5.29 détaille les activités par lesquelles passe chaque appel. Il est ainsi facile de retracer ce qu'il s'est passé et pourquoi l'appel 304 n'a pas pu être décroché plus tôt. On voit sur la figure que l'ARM 1 commence à remplir le dossier de l'appel 302 à 13h03m11 et il finit à 13h03m53. L'ARM 2 priorise l'appel 303 de 13h02m59 à 13h03m21, il termine son activité après le décroché du 304. L'ARM 3 est en train de remplir le dossier du 301 entre 13h02m48 et 13h04m04. L'ARM 4 remplit le dossier du 300 entre 13h02m28 et 13h03m26 appel de priorité 1 qui génère ensuite un appel interne que l'ARM 4 continue de traiter. C'est donc l'ARM 1 qui après avoir traité l'appel 302 va décrocher l'appel 304 pour le traiter. Ce temps d'attente long n'est pas dû à la nature du scénario, mais à la disponibilité des ressources. Nous sommes typiquement dans un cas où beaucoup d'appels arrivent en même temps. Pour palier ces attentes on pourrait envisager d'autoriser l'interruption de l'activité *Remplir le dossier* pour les appels de gravité 3.

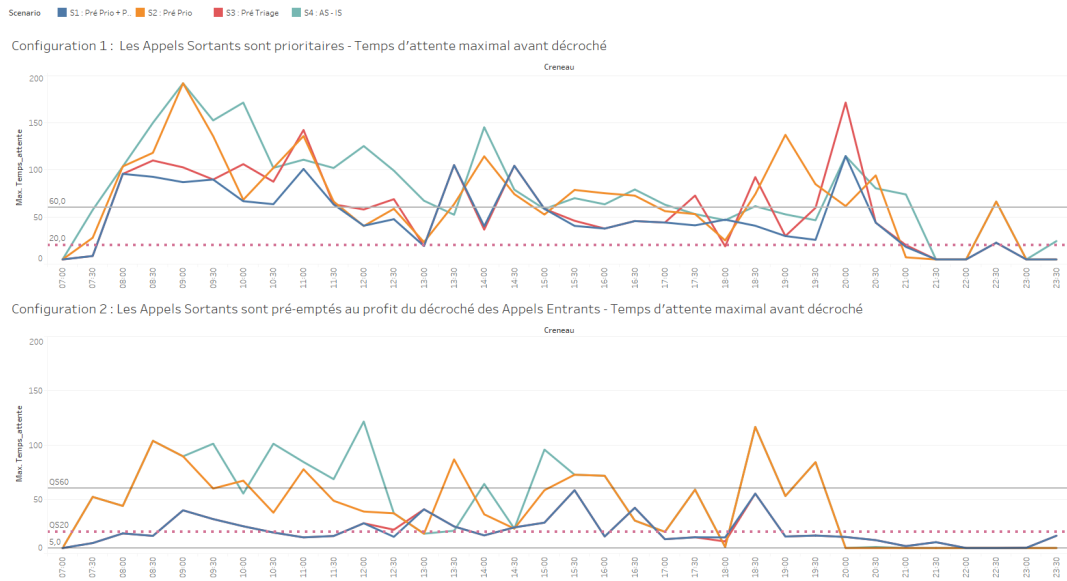


Figure 5.28 – Comparaison des temps d’attentes maxima pour les appels de gravité P1 entre les deux configurations

Conclusions sur les expérimentations

La modification de l’organisation du décroché de l’appel permet des améliorations sensibles sur la réduction du temps d’attente. Bien que ces gains doivent être mis en regard des hypothèses faites lors de la création du modèle, ils restent significatifs.

On identifie notamment une réduction du temps maximal d’attente pour les appels les plus urgents (P1) pouvant aller jusque 40% sur certains créneaux. Ce qui correspond à une diminution pouvant aller jusqu’à 1min30 sur certains appels. De plus les modifications permettent de respecter les critères nationaux de décrochés des appels et faire en sorte que tous les appels soient décrochés en moins de 60 secondes. Ainsi une combinaison idéale serait de mettre en place un triage des appels combiné à la prépriorisation de ceux-ci et à la possibilité d’interrompre les appels sortants au profit des appels entrant. Cette combinaison permet des gains significatifs notamment pour les appels de la priorité P1 qui nous le rappelons sont les plus critiques

5.2.5 Gérer une crise : comment le centre d’appels réagit-il confronté à une augmentation soudaine et significative des appels ?

Dans cette section, nous souhaitons évaluer la réaction du centre d’appels face à une augmentation significative et brutale des appels tout au long de la journée. Nous proposons de doubler le profil d’arrivée des appels ce qui pourrait correspondre à

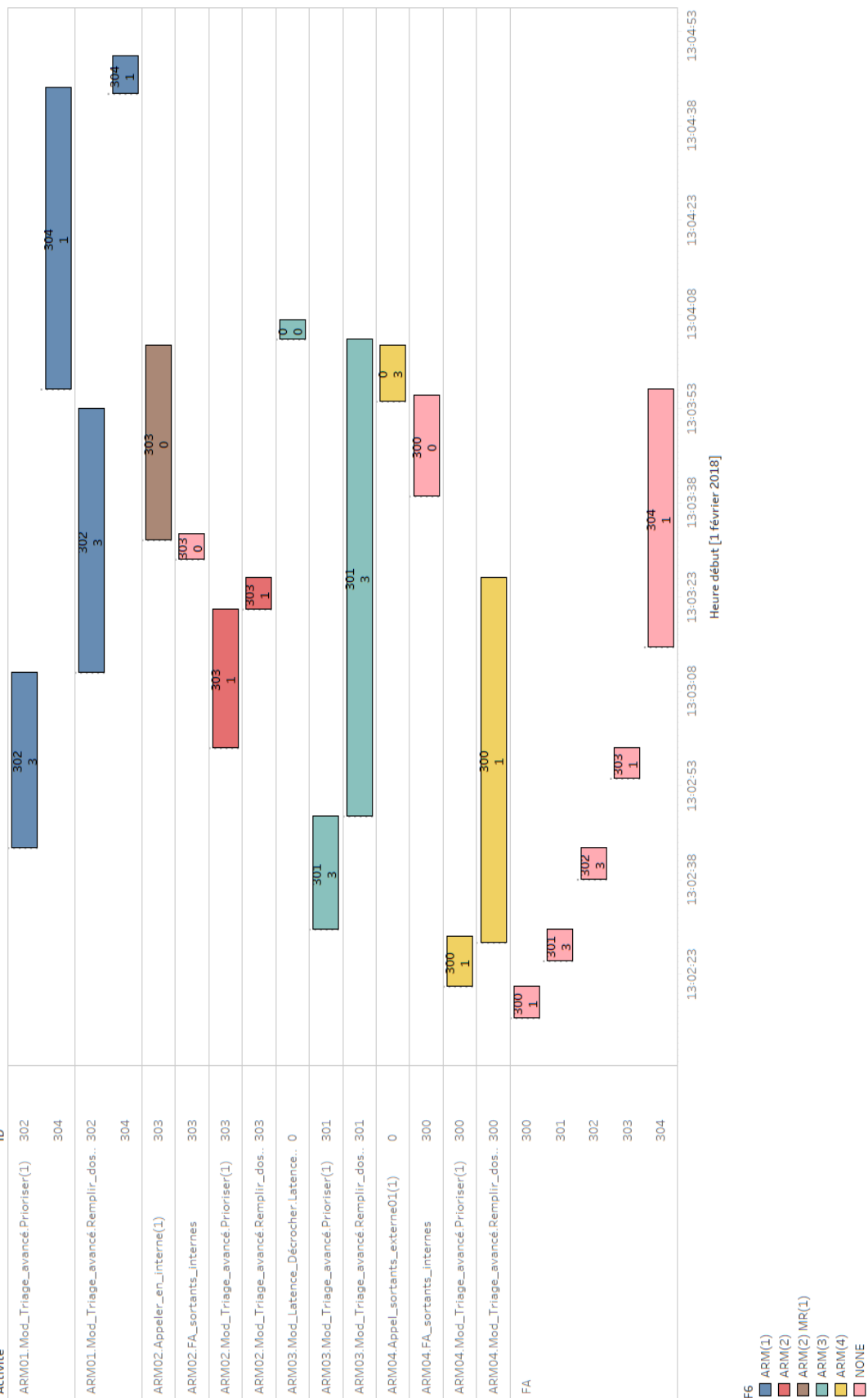


Figure 5.29 – Détail de la réplication 89 pour le scénario de Triage dans la configuration 2

un scénario d'épidémie par exemple de grippe saisonnière. Nous regardons l'effet de cette augmentation sur chacun des créneaux horaires en termes de temps d'attente maximum, de QoS_{20} et de QoS_{60} . Nous étudions dans un premier temps l'effet sur la configuration C1, puis nous appliquons la configuration C2. Les données sont analysées pour chacun des scénarios de prépriorisation, de triage ainsi que pour les scénarios mixtes. Si on regarde la courbe d'arrivée du nombre d'appels par créneau (voir figure 5.30) on constate comme précédemment qu'on a plusieurs périodes. De 00h00 à 7h30, durant la nuit profonde il y a assez peu d'appels et la qualité de service est plutôt bonne. Puis, il y a un premier pic d'appels entre 8h00 et 10h00 et un second pic d'appels entre 14h et 16h. Suite à cela, la situation se stabilise vers 21h30 avec une diminution du nombre d'appels. Pour la suite, et comme précédemment nous réduisons l'horizon d'étude aux créneaux entre 7h30 et 21h30 qui sont donc les plus intéressants, car c'est là où le centre d'appels montre des difficultés de décroché des appels.

Doubler la demande sur la configuration C1

Effet sur la qualité de service

Le fait de doubler le nombre d'appels reçus a un effet très négatif sur la qualité de service. La figure 5.30 présente les valeurs de QoS pour chaque scénario ainsi que le nombre moyen d'appels sur chaque créneau. On constate que la qualité de service diminue lorsque le nombre d'appels augmente. Que ce soit le QoS_{20} ou le QoS_{60} , dès 8h00 ils valent respectivement 5,7% et 22,8% (voir figure 5.30). Cela implique que 81,2% des appels attendent plus de 60 secondes dès 8h00 si le nombre d'appels double. Le centre d'appels est donc rapidement saturé si on reste avec un nombre de ressources fixes. Le pire créneau est celui de 10h00 où le QoS_{20} et le QoS_{60} valent respectivement 1,6% et 8,7%. On constate que tous les scénarios testés engendrent une amélioration globale de la qualité de service et permettent de réduire l'effet de l'augmentation de la durée d'attente des appels.

La figure 5.31 présente les gains par rapport au scénario AS-IS pour les appels de priorité 1. On constate que les scénarios mixtes qui combinent prépriorisation et triage permettent d'aller chercher des gains supérieurs à 50% du scénario initial pour le QoS_{60} , cela implique qu'après l'application de ces scénarios, dans le pire des cas 40% des appels sont décrochés en moins de 60 secondes. Même si ce n'est pas suffisant vis-à-vis des objectifs de performance en régime nominal, cela permet d'atténuer significativement l'effet néfaste de l'augmentation du nombre d'appels sur les performances du centre d'appels en situation de crise.

Ensuite, c'est le scénario de prépriorisation qui permet les meilleurs gains sur le QoS_{60} avec jusqu'à 50% d'amélioration sur les créneaux les plus chargés. Il est intéressant de noter que ce n'est pas le cas pour le QoS_{20} où la prépriorisation s'avère moins performante avec des gains moindres et une amélioration au maximum de 15%. Ce qui le rend moins performant que les scénarios mixtes qui vont chercher un gain pouvant aller jusque 25%.

Les scénarios de triage seuls permettent d'améliorer le QoS_{60} de 30%. On note que lorsqu'on augmente le nombre d'appels, le fait d'avoir un triage plus long (ce qui est

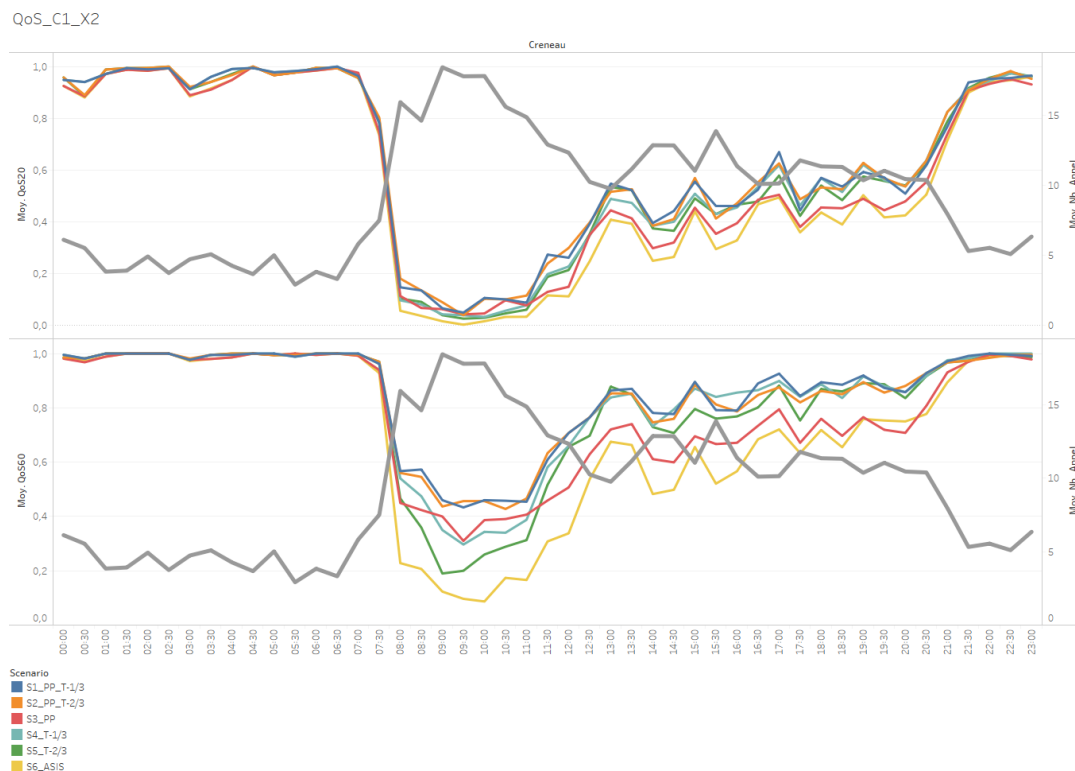


Figure 5.30 – Qualité de Service pour chacun des scénarios. Demande double, Configuration C1

le cas pour le scénario $S5_T - 2/3$) donne des écarts de performance plus importants que ce que l'on avait pu constater avec un nombre d'appels normal. Concernant le QoS_{20} , les scénarios de triage sont meilleurs que celui de prépriorisation avec des gains supérieurs à 10% après 12h30.

Concernant les appels de priorité P2, les gains sur la qualité de service sont représentés dans la figure 5.32. On constate que c'est encore une fois les scénarios mixtes qui permettent d'obtenir les meilleurs résultats et qui réduisent le QoS_{60} de 50% au maximum lors du premier pic d'appels. et lors du deuxième pic d'appels, on gagne 20% par rapport au scénario AS-IS. Le QoS_{20} est amélioré de 30% sur le premier pic d'appels ainsi que sur le second pic d'appels.

Le scénario de prépriorisation permet une amélioration au global du QoS_{60} avec des gains qui vont jusque 35% sur le premier pic d'appels et des gains de 10% sur le deuxième pic d'appels. Concernant le QoS_{20} on observe des gains de 15% lors du premier pic d'appels et des gains de 10% lors du deuxième pic d'appels. Le scénario de prépriorisation est plus efficace que le scénario T-2/3 sur le premier pic d'appels, sur le deuxième pic les performances sont équivalentes.

Les scénarios de triage T-1/3 et T-2/3 améliorent le QoS_{60} entre 20% et 30% durant le premier pic, on note une différence de performance de 20% entre T-1/3 et T-2-3, le fait d'augmenter le temps de triage va effectivement avoir un effet plutôt important sur la qualité de service lorsque le centre d'appels est saturé. Lors du deuxième pic d'appels T-1/3 améliore QoS_{60} de 20% et T-2/3 l'améliore de 10%. Concernant

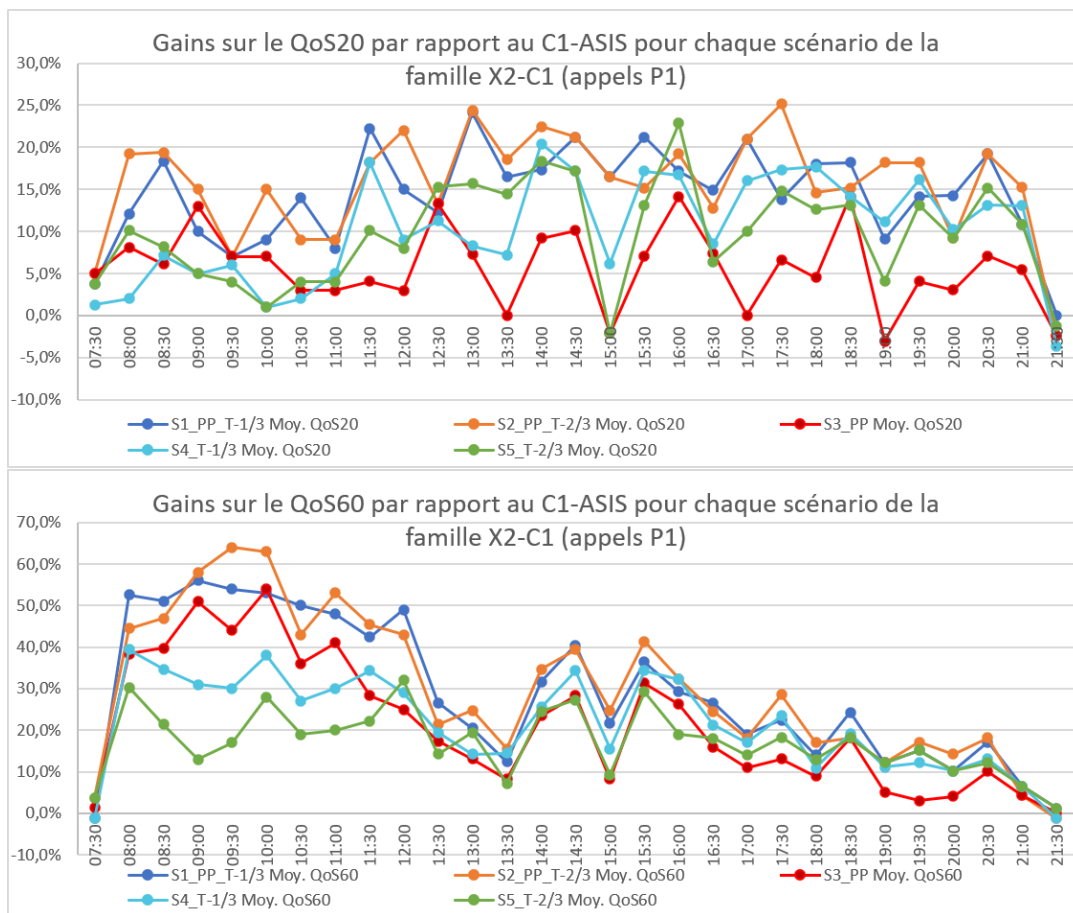


Figure 5.31 – Demande doublée : Gains sur les QoS_{20} et QoS_{60} pour les appels P1 dans la configuration C1

QoS_{20} les scénarios de triage permettent d'obtenir 5% d'amélioration lors du premier pic et au maximum 15% lors du deuxième pic. Le triage seul reste moins efficace que les scénarios mixtes.

La figure 5.20 montre les gains pour les appels de priorité 3. Pour le QoS_{60} , le premier constat est que le scénario de prépriorisation a un effet négatif sur le décroché des appels P3 (au bénéfice des appels de priorité supérieur). Le constat est similaire pour le QoS_{20} : il y a quelques gains qui allant jusque 5%, mais il y a des effets négatifs pouvant aller jusque -5%.

Les scénarios de triage et les scénarios mixtes permettent une amélioration du QoS_{60} surtout après le premier pic d'appels et durant le deuxième. Durant le premier pic d'appels il n'y a pas vraiment de scénario qui permettent d'améliorer le décrocher des appels P3, car la priorité est mise sur les appels P1 et P2. Toutefois on note quand même des améliorations pouvant aller jusque 20% pour le QoS_{20} avec le scénario T-1/3, et jusque 30% sur le QoS_{60} avec les deux scénarios de triage.

Effet sur le temps d'attente maximum

Le temps d'attente maximum si on double les appels monte jusque 800 secondes lors du premier pic d'appels, puis 400 secondes lors du second pic d'appels. Le temps d'attente maximum est globalement le même que ce soit pour les appels P1, P2 ou P3 concernant le AS-IS. Ce qui est logique puisqu'il n'y a pas de priorisation entre les types d'appels.

Pour les appels P1 et P2, ce sont les scénarios mixtes et le scénario de prépriorisation qui permettent d'obtenir la meilleure réduction du temps d'attente maximal dans la première configuration. Ils permettent notamment de maintenir l'attente en dessous de 200 secondes. Les scénarios de triage seuls sont peu performants et on note des temps d'attente pouvant aller jusque 600 secondes lors du premier pic d'appels.

Pour les appels P3, on constate que le scénario de prépriorisation engendre un temps d'attente plus important que les autres scénarios. Lors du pic d'appels, ce sont les scénarios PP-T-1/3 et T1/3 qui permettent la meilleure réduction du temps d'appels.

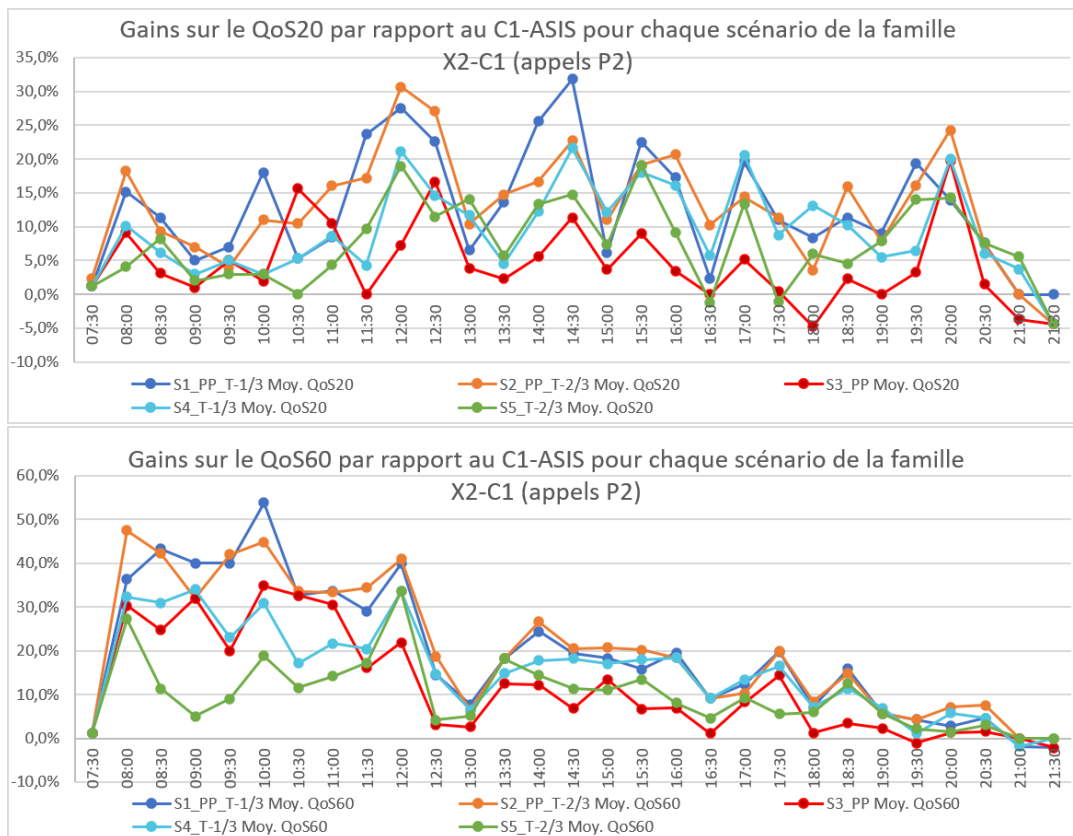


Figure 5.32 – Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C1

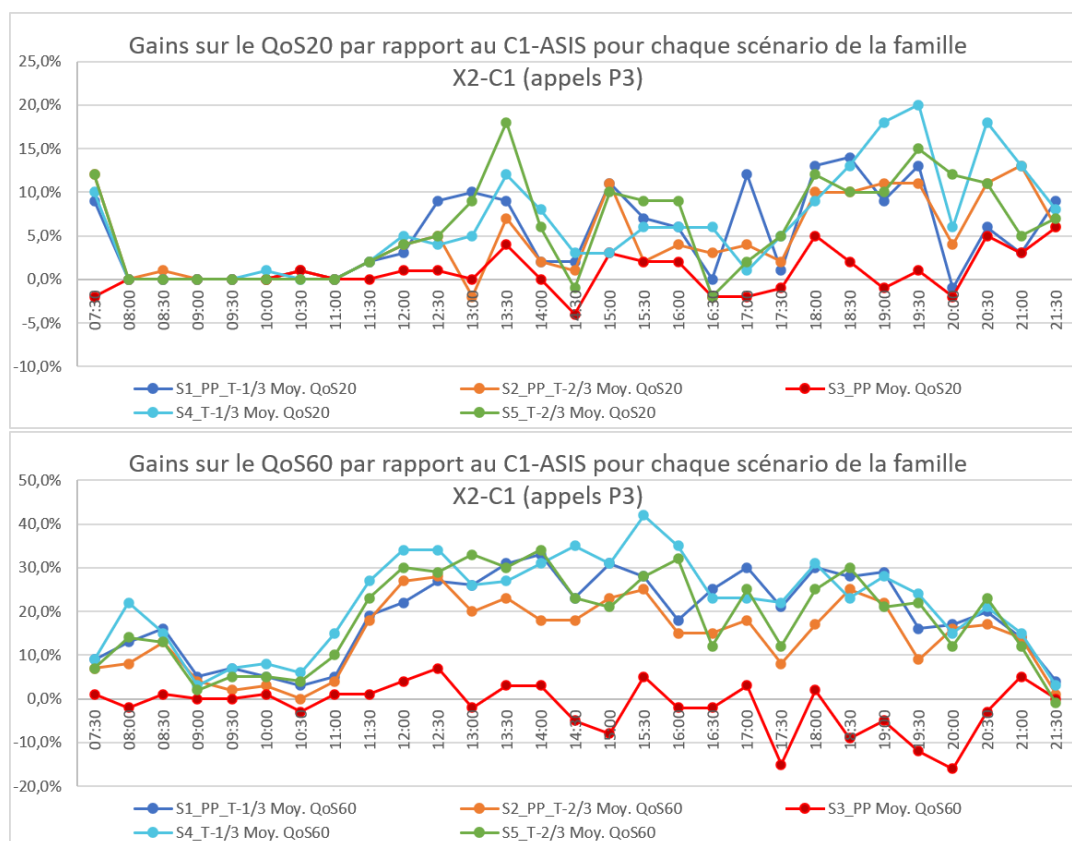


Figure 5.33 – Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C1

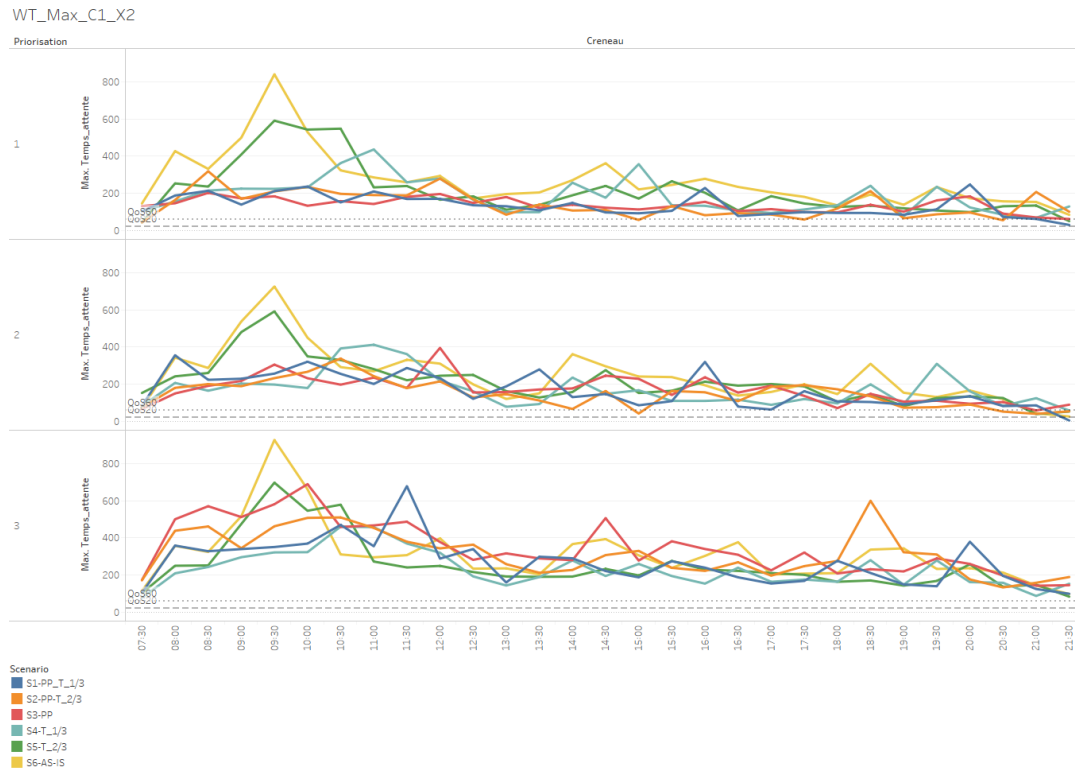


Figure 5.34 – Effet de l’augmentation du nombre d’appels sur le temps d’attente maximum en configuration C1

Doubler la demande sur la configuration C2

Nous avons déjà pu constater que les scénarios proposés permettent d’améliorer la prise en charge des appels lorsque le nombre d’appels double. Même si cela n’est pas suffisant pour faire en sorte que le centre d’appels ne soit pas saturé. Nous souhaitons maintenant tester l’effet des scénarios dans la configuration C2 (cf. section 5.2.4) et voir la réaction du système à l’augmentation du nombre d’appels. La figure 5.35 montre la valeur du QoS_{20} et du QoS_{60} pour chaque scénario. On observe qu’au global deux types de scénarios se démarquent : le scénario mixte PP_T-1/3 et le scénario de triage T-1/3 et les autres scénarios. Au niveau du QoS_{20} les deux premiers scénarios permettent de le maintenir au-dessus de 60% là où l’AS-IS descend jusqu’à 20% au plus haut du premier pic d’appels. Les autres scénarios de PP, et de T-2/3 améliorent la situation, mais uniquement d’une dizaine de %. Les différences entre les scénarios sont plus nettes au niveau du QoS_{60} . En effet, au niveau du premier pic d’appels, on constate que le scénario de prépriorisation améliore le QoS_{60} de 16% , le scénario de T-2/3 l’améliore de 32%, le scénario mixte PP-T-2/3 l’améliore de 37 % et les deux scénarios PP-T1/3 et T1/3 permettent de conserver un QoS_{60} supérieur à 90%.

Effet sur la qualité de service

Les figures 5.36, 5.37 et 5.38 présentent les gains sur chaque type d’appels comparé au scénario AS-IS initial, soit celui de la première configuration.

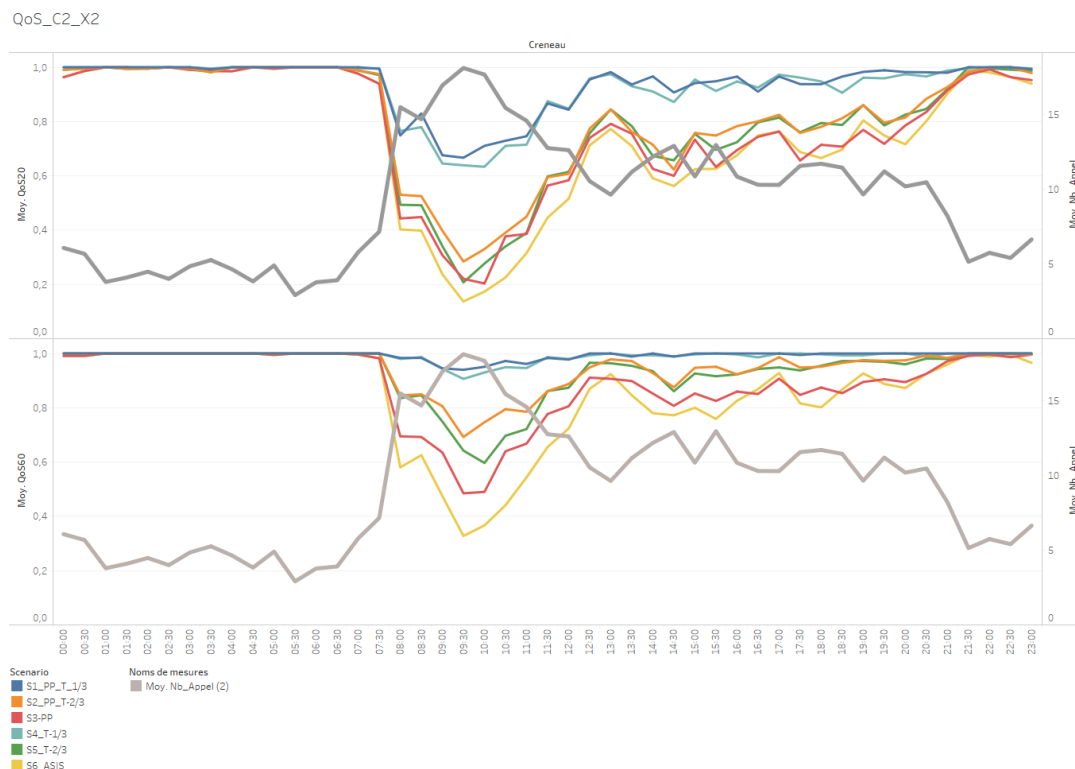


Figure 5.35 – Qualité de Service pour chacun des scénarios. Demande double, Configuration C2

Pour les appels de type P1 et P2, on retrouve les deux scénarios T-1/3 et PP-T-1/3 qui permettent une amélioration de 90% par rapport au scénario initial. Ces scénarios permettent de faire en sorte que même si on double le nombre d'appels, 95% des appels P1 sont décrochés en moins de 20 secondes. Les scénarios de PP-T-2/3, T-2/3 et PP permettent aussi une amélioration par rapport au AS-IS qui permet de gagner 50% d'appels décrochés en moins de 20 secondes lors du pic d'appels. Au niveau du QoS_{60} ce sont les scénarios mixtes et le scénario T-1/3 qui sont les plus efficaces avec une amélioration très sensible de 80 à 90 % lorsque le pic d'appels est le plus important.

Pour les appels de type P3, on améliore un peu moins le QoS_{20} car la priorité est une fois encore mise sur les appels P1 et P2. Toutefois ce sont quand même des gains de près de 50% qui peuvent être observés durant le premier pic d'appels pour les scénarios PP-T1/3 et T-1/3. Lors du deuxième pic d'appels, on observe des gains de 80% qui permettent de faire en sorte que 95% des appels soient décrochés en moins de 20 secondes. Pour le QoS_{60} on note la faiblesse du scénario de prépriorisation qui ne permet pas d'améliorer autant les performances des appels P3 par rapport aux autres scénarios.

Effet sur le temps d'attente maximum

Avec la mise en place de la configuration C2, il est intéressant de noter que le scénario PP-T-1/3 permet de réduire le temps d'attente maximal à 162,5 secondes. De plus, si ce n'est pour trois créneaux, on note que le temps d'attente maximal est inférieur

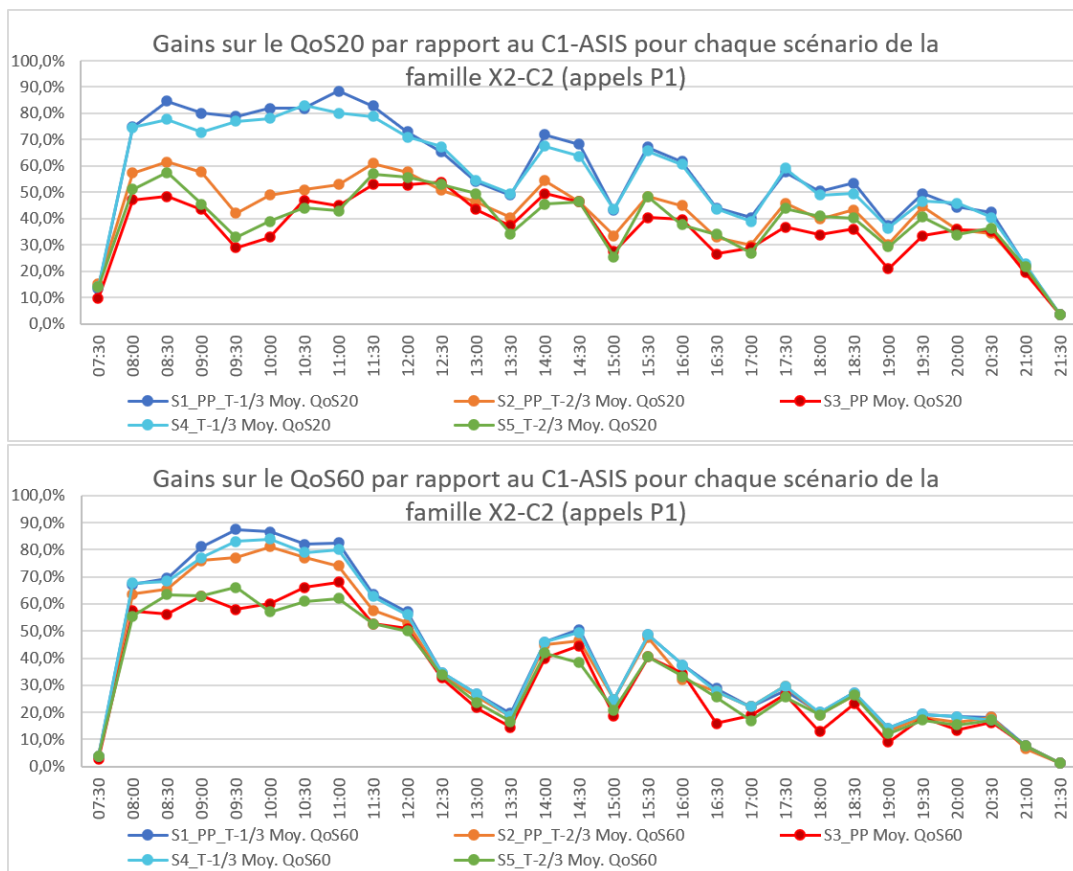


Figure 5.36 – Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P1 dans la configuration C2

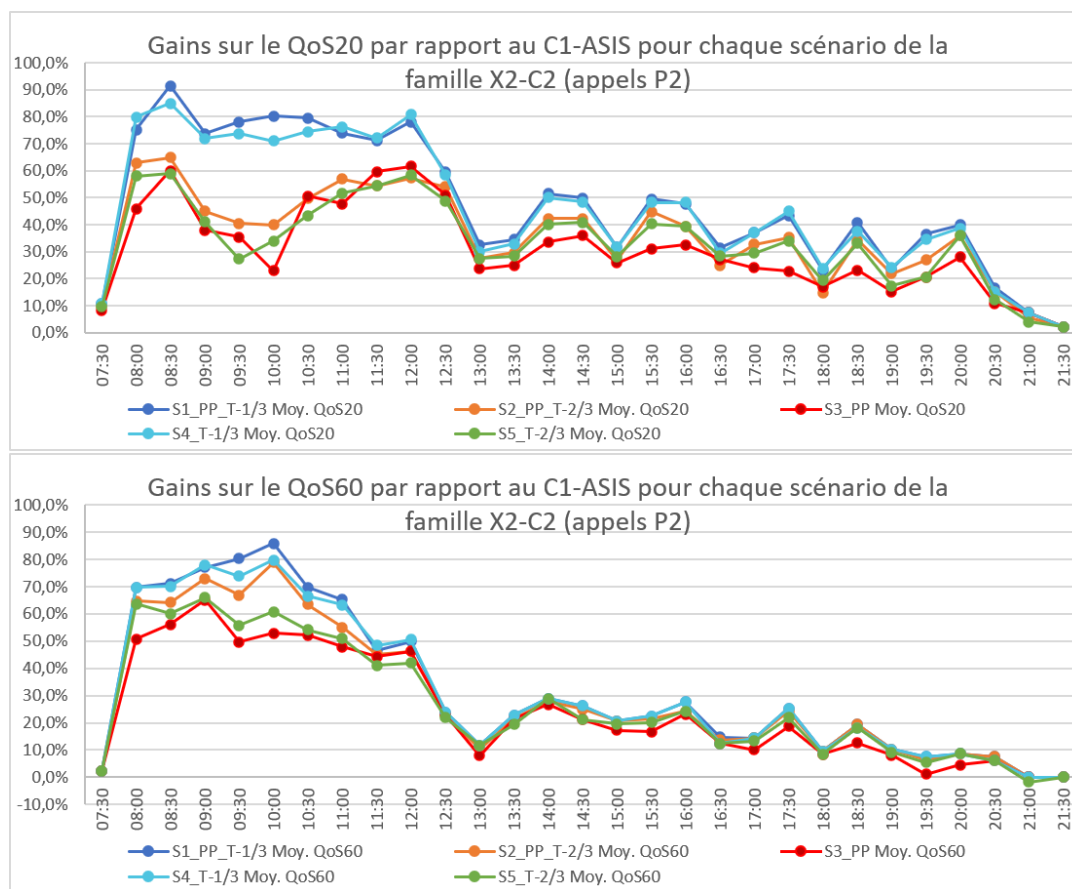


Figure 5.37 – Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C2

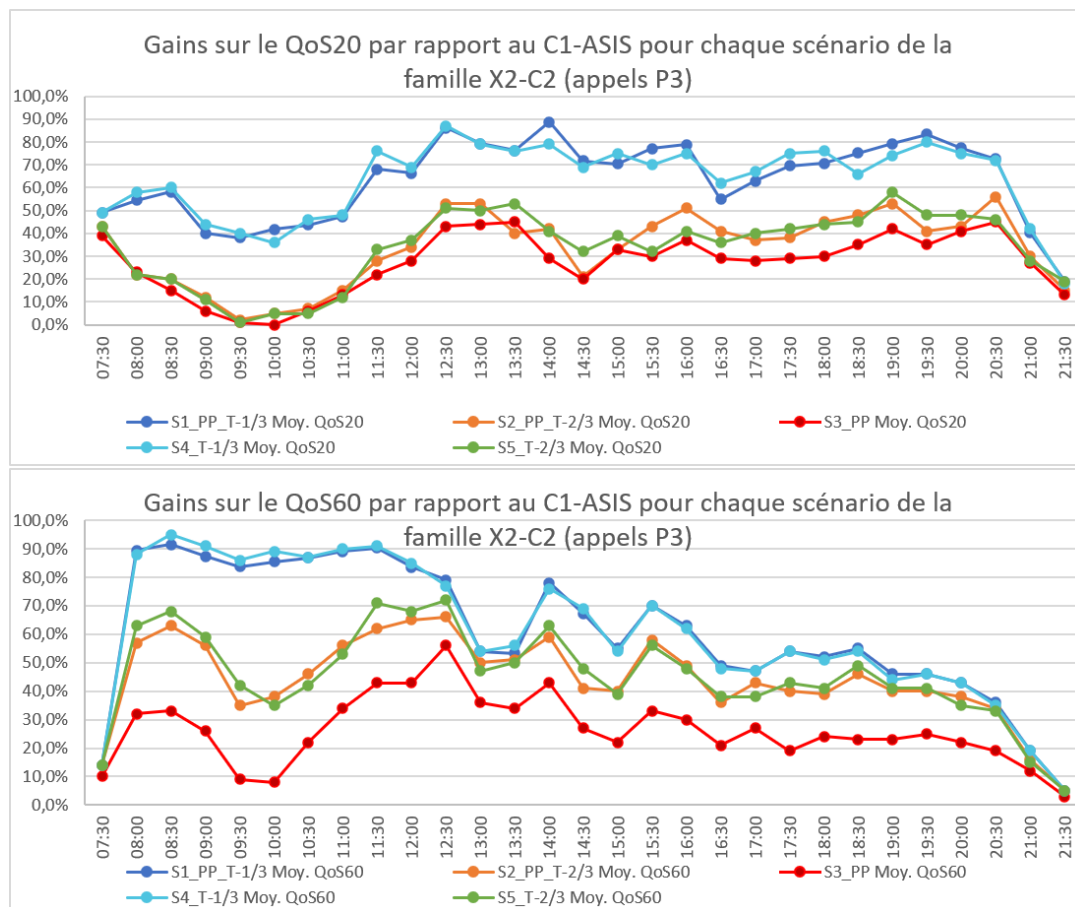


Figure 5.38 – Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C2

à 60 secondes, donc le QoS_{60} est de 100% sur ces créneaux pour les appels P1 et P2. Le scénario de T-1/3 maintient temps d'attente maximum à 244,6 secondes lors du premier pic d'appels puis dès 12h, le temps d'attente reste inférieur à 60 secondes pour les appels P1 et P2. Le scénario mixte PP-T-2/3 permet de réduire le temps d'attente maximum durant le premier pic à moins de 200 secondes. On note néanmoins un pic d'attente à 326 secondes à 11h30.

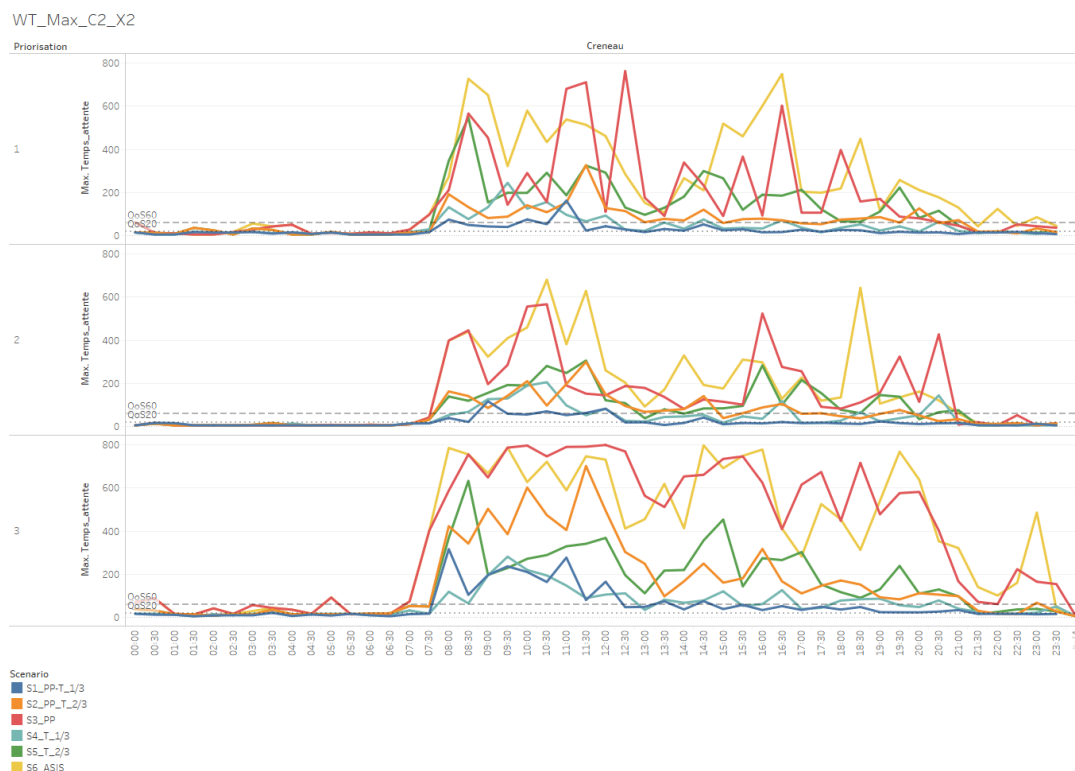


Figure 5.39 – Effet de l'augmentation du nombre d'appels sur le temps d'attente maximum en configuration C1

Dans le cas où le nombre d'appels double, on constate que la mise en place de scénarios de modification de l'organisation a un effet positif et tend à améliorer significativement les performances du système.

Le premier constat est qu'il semble évident que la priorité doit être mise sur le décroché des appels entrants, le fait de pouvoir interrompre un appel sortant afin de décrocher un appel entrant améliore le QoS_{20} moyen sur la journée de 51% à 71% et le QoS_{60} de 66% à 83%.

Le scénario de prépriorisation permet de réduire le temps d'attente des appels au global. Toutefois, tout comme pour l'AS-IS, il a tendance à légèrement augmenter le temps d'attente des appels P3. On note néanmoins une légère amélioration du QoS_{20} de 5% et du QoS_{60} de 7%. La faiblesse principale de ce scénario réside dans la difficulté de sa mise en oeuvre car il s'appuie sur des nouvelles technologies qui permettent d'obtenir de l'information supplémentaire

sur la gravité de l'appel en amont du décroché. Ce scénario serait donc un objectif à atteindre en complément du triage.

Le scénario de triage présente par contre un réel intérêt pour les centres d'appels d'urgence, il est plutôt facile à mettre en place, car il ne nécessite pas d'information supplémentaire en amont de l'arrivée de l'appel dans le centre. On peut soit envisager d'affecter un ARM spécifiquement à ce poste, soit demander aux ARM d'interrompre les appels courants pour décrocher les nouveaux, ce qui peut être humainement plus éprouvant. Si on arrive à garder le temps de triage en deçà d'un certain seuil (qui n'est pas identifié pour le moment) on peut prétendre à maintenir un QoS_{20} à 95% et un QoS_{60} supérieur à 95% même si on double le nombre d'appels. Il serait donc intéressant d'identifier le temps optimal de priorisation de l'appel pour lequel on observe une réduction de la qualité de service tout en laissant un temps suffisant à l'ARM pour identifier la gravité de l'appel.

Les scénarios mixtes de prépriorisation et de triage sont également intéressants et obtiennent en moyenne sur la journée les mêmes résultats que le triage seul. Dans la mesure où la prépriorisation n'est pas aisée à mettre en place, il serait plus intéressant pour les centres d'appels de mettre en place du triage qui en moyenne sur la journée permet d'obtenir un QoS_{60} de 100% ainsi qu'un QoS_{20} de 99%. Ainsi, les objectifs de performances fixés par la HAS seraient respectés.

5.3 Synthèse

Dans ce chapitre, nous avons proposé des scénarios de nouvelles organisations pour les SAMU. Ces scénarios ont été évalués au regard de plusieurs indicateurs de performance le QoS_{20} et le QoS_{60} ainsi que le temps d'attente maximal. Les contributions principales ainsi que les limites de ce chapitre sont présentées dans le tableau 5.12. La simulation de scénarios de triage, de prépriorisation et de scénarios mixtes en régime nominal et lors d'une augmentation soudaine du nombre d'appels constitue la contribution principale de ce chapitre. Plusieurs résultats intéressants doivent être soulignés :

- La mise en place d'un triage rapide des appels permet de respecter les objectifs gouvernementaux pour la QoS.
- La mise en place de prépriorisation améliore le décroché pour les appels de priorité P1 au détriment des appels de priorité P3.
- Le triage des appels permet de réduire le temps d'attente maximal de 152 secondes pour les appels P1 (soit 97% de gains), 230 secondes pour les P2 (soit 95% de gains) et 239 secondes pour les P3 (soit 92% de gains).
- La prépriorisation permet de réduire le temps d'attente maximal de respectivement 134, 171 et 156 secondes (soit, 89, 97, et 85% de gains) pour les appels P1, P2, P3.
- Même lorsque le nombre d'appels double, la mise en place du triage permet de maintenir un QoS_{20} supérieur à 70% et un QoS_{60} supérieur à 95%.

Ainsi, la mise en place du triage rapide des appels dans les SAMU semble une bonne piste pour améliorer les performances de décroché des centres d'appels d'urgence. Toutefois, ces conclusions sont toujours à nuancer vis-à-vis des hypothèses sous-jacentes au modèle. Notre modèle pourrait être ajusté, notamment en améliorant sa capacité à prendre en compte les appels abandonnés. De plus, nous n'avons pas différencié les appels de la population et les appels bilans des équipes d'effecteurs. Ces derniers ne suivent pas tout à fait le même processus et les temps de sessions seraient légèrement différents par rapport aux appels de la population. Qui plus est, notre modèle pourrait gagner en précision concernant la définition des temps de sessions qui sont considérés comme indépendants du profil d'arrivée. Or, la tendance montre que plus il y a d'appels dans la file d'attente plus les temps de sessions des appels ont tendance à diminuer. Enfin, il faudrait aussi étudier plus en profondeur le lien entre les appels entrants et l'émission d'appels sortants pour identifier s'il existe une relation entre ces deux types d'appels.

Tableau 5.12 – Contributions et Limitation : Chapitre 5

Chapitre 5	
Contributions	Évaluation de la mise en place d'un scénario de triage et/ou de prépriorisation
	Évaluation de ces scénarios lors d'une augmentation soudaine du nombre d'appels
	Préconisations organisationnelles pour les SAMU
Limitations	Les appels abandonnés ne sont pas considérés
	Les appels bilans et les appels des autres praticiens ne sont pas différenciés des appels entrants
	Les temps de sessions sont considérés comme indépendants du profil d'arrivée
	Les émissions des appels sortants sont considérées comme indépendants des appels entrants

6

Conclusion et perspectives

6.1 Bilan du projet de recherche

L'objectif de ce manuscrit était donc de proposer et d'évaluer des outils qui pourraient permettre d'améliorer la performance des activités aval d'un centre d'appels d'urgence, et notamment d'un SAMU, en proposant de nouvelles manières de fonctionner et en évaluant de nouveaux schémas organisationnels.

Ce chapitre propose un bilan de nos apports vis-à-vis du problème posé, en reprenant la question de recherche générale et les principales contributions apportées dans chaque chapitre. La partie 6.2 introduit les nombreuses perspectives possibles qui pourraient faire l'objet de recherches futures.

Dans le chapitre 1, nous avons fait le constat qu'à l'échelle internationale il est nécessaire de parler des systèmes d'aide médicale urgente, car ils sont multiples et hétérogènes d'un pays à un autre, et même au sein d'un même pays. D'un point de vue métier, la modernisation des SAMU (pour ce qui est de la France) est un enjeu majeur. Parmi les défis à relever on trouve :

1. Objectiver et harmoniser l'activité opérationnelle des centres d'appels.
2. Proposer de nouveaux modèles cibles organisationnels.
3. Mettre en place une stratégie d'entraide et de collaboration entre les centres d'appels d'urgence.

Notre travail s'est focalisé sur les deux premiers items uniquement. Pour réaliser ces objectifs, de nombreux travaux seront nécessaires pour développer de nouveaux modèles pertinents sur les plans techniques, économiques et politiques. L'étude de la littérature a montré que des verrous scientifiques doivent encore être levés. Il s'agit, aux travers d'indicateurs de performance d'être capable de modéliser, d'évaluer et d'améliorer l'activité complexe, soumise à une grande incertitude des centres d'appels d'urgences tant dans leur régime nominal que lors d'événements de crises.

La question de recherche du manuscrit qui en découle était la suivante :

Comment structurer et outiller une démarche de progrès permanent des centres d'appels d'urgence ?

Dans cette logique d'amélioration continue et de structuration des centres d'appels d'urgence, la démarche à définir devait être en mesure de prendre en compte les diversités structurelles et l'hétérogénéité des centres d'appels. Ainsi découle la question de recherche adressée au chapitre 3 :

Quelle méthodologie instrumentée développer pour diagnostiquer et améliorer le fonctionnement d'un centre d'appels d'urgence ?

Un cadre méthodologique outillé a donc été proposé, représenté par la figure 3.1 qui se décline selon plusieurs volets (présenté dans la section 3.1.2) qui permettent de caractériser l'organisation d'un centre d'appels d'urgence : la stratégie préhospitalière, l'accès au centre d'appels et le modèle de décrochés que l'on décline en quatre dimensions (le type de plateforme, la spécialité du centre d'appels, la gestion des files d'attente et les compétences en matière de ressources).

La proposition se décompose en deux parties : *le système réel* et *le double numérique*. Chaque partie est segmentée en étapes auxquelles nous associons des outils pour réaliser le diagnostic organisationnel et l'amélioration des centres d'appels. La première partie du cadre a pour objectif de collecter des données afin d'en extraire de l'information pour diagnostiquer les centres d'appels. Cette partie du cadre nous permet de répondre à la question suivante :

Quelle influence les organisations en place dans les centres d'appels d'urgence ont-elles vis-à-vis de leur niveau de performance ?

Dans cette partie, nous avons proposé des algorithmes pour l'étude des données de téléphonies des centres d'appels. La prise en compte des appels entrants mais aussi des appels internes et des appels sortants constitue une nouveauté par rapport aux études menées jusqu'à présent. Une fois l'organisation actuelle analysée et évaluée nous souhaitons chercher à l'améliorer afin de répondre à la question suivante :

Quels schémas organisationnels préconiser pour des centres d'appels d'urgence en situation nominale et en situation de crise ?

Grâce à la connaissance extraite des données réelles, nous avons donc défini un processus de prise en charge de l'appel que nous avons confronté à différents scénarios organisationnels à l'aide d'un plan d'expériences en faisant varier :

- Le traitement prioritaire entre les appels sortants et les appels entrants.
- La mise en place d'un triage des appels.
- Le temps de triage des appels.
- La mise en place d'une prépriorisation des appels.

- L'intensité du profil d'arrivée des appels.

Afin de répondre à la question de recherche, nous concluons ci-dessous sur la performance en termes de décroché des appels des différents scénarios testés. Ces résultats sont bien sûr à nuancer au regard des hypothèses considérées dans le modèle.

- Il est essentiel de prioriser le décroché des appels entrants par rapport au traitement des appels sortants et faire en sorte que les ressources qui décrochent ne soient pas indisponibles à cause du flux sortant.
- Le fait d'avoir de l'information sur la gravité de l'appel en amont de son décroché n'est pertinent que s'il est possible d'interrompre le traitement d'appels moins graves au profit d'appels plus prioritaires présents dans la file d'attente.
- La mise en place d'un triage rapide des appels permet de répondre aux objectifs de qualité de service.
- Dans le cas d'une augmentation doublée du nombre d'appels, la mise en place du triage permet de prendre en charge 99% des appels en moins de 60 secondes là où le centre d'appels était saturé dans la configuration initiale.

6.2 Perspectives

Nous détaillons dans les paragraphes suivants les perspectives applicatives et académiques, en prenant en compte les limites de l'étude menée.

6.2.1 Perspectives applicatives

Plusieurs axes de travail sont à envisager sur le plan opérationnel à plus ou moins long terme. Dans notre modèle, le premier résultat est que la priorité doit être mise sur les appels entrants. En effet, si ce n'est pas déjà le cas, il est nécessaire que les ressources dédiées au décroché des appels mettent la priorité sur les appels entrants. À court terme, la mise en place d'un triage rapide des appels est une mesure relativement simple à mettre en oeuvre qui permet un gain important sur les performances du centre d'appels. Ce triage pourrait être envisagé de deux manières. La première est celle présentée dans ce manuscrit, où est conservée la même allocation de ressources, mais en autorisant l'interruption des appels durant leur traitement pour décrocher et prioriser les suivants. Une autre solution, qui mériterait d'être testée, est d'affecter une ressource particulière pour le décroché et la priorisation rapide des appels. C'est d'ailleurs ce qui a été mis en place durant la pandémie de COVID-19 par le SAMU de Nantes (Penverne et al., 2020). Leurs résultats viennent corroborer les nôtres : la mise en place d'un triage permet de décrocher plus d'appels et de prioriser rapidement les urgences vitales.

Tout au long de ce manuscrit, nous avons mis l'accent sur l'importance de l'analyse de l'existant. Il est primordial que les SAMU se dotent de moyens pour récolter, agréger et exploiter les données événementielles en lien avec les appels d'urgence. Normalement, le système SI-SAMU devrait permettre un accès simplifié et unifié à ces données. Ce sera un très bon point de départ pour permettre des projets de progrès permanent qui se basent sur des données objectives.

Utiliser le dossier patient et faire le lien avec les types d'interventions

Dans cette étude, nous n'avons travaillé qu'avec des données de téléphonie. Or, dans le système d'information des centres d'appels, on constate deux types de bases de données, celles de téléphonie et celles de dossier de régulation médicale. Ces dossiers de régulation médicale contiennent des données complémentaires pertinentes, comme la classification précise du motif de l'appel ainsi que le choix du type d'intervention envoyé. En disposant de ces données, il serait alors possible de faire un lien entre l'appel et le choix effectué sur celui-ci. Cela ouvre des perspectives intéressantes en termes de diagnostic pour établir une classification des types d'appels et des types de moyens envoyés. Sur cette base, il serait possible d'envisager une prédiction de moyens nécessaires afin d'aider à mieux dimensionner les ressources sur la chaîne des urgences.

Mettre en place de la collaboration sur plusieurs niveaux

En plus d'être une piste intéressante et envisagée d'un point de vue métier, la collaboration au sein des services d'urgence est un sujet de recherche en plein développement. On note, notamment les travaux de (Belaidi, 2009) qui proposent de voir la chaîne des urgences comme un réseau d'acteurs pouvant collaborer. Ils mettent en place un système d'indicateurs pour mesurer la collaboration dans le réseau des urgences. Ces travaux sont un excellent point de départ pour aborder la collaboration au niveau de la chaîne des urgences. Celle-ci peut être envisagée suivant plusieurs axes :

- au niveau de la coordination de tous les professionnels de la chaîne des urgences (pompiers, policiers, urgentistes).
- au niveau territorial, entre les mêmes acteurs d'un département à un autre.

La collaboration dans les centres d'appels pourrait alors être envisagée de différentes manières. On peut imaginer centraliser le premier niveau de décroché à l'échelle régionale. Si un département sature en appels, le décroché est basculé sur d'autres départements. De cette manière, la régulation par un médecin de l'appel peut rester à l'échelle du département, mais l'appel sera décroché et priorisé plus rapidement, ce qui permet d'identifier en amont les appels les plus prioritaires. On peut aussi envisager de centraliser le décroché et la régulation des appels à l'échelle régionale.

6.2.2 Des perspectives académiques

Éprouver la méthodologie sur les différents volets

La méthodologie que nous avons proposée a pour but de couvrir un large panel d'organisations dans les centres d'appels d'urgence. Néanmoins dans notre étude, nous n'avons pas pu faire d'expérimentations sur des paramètres comme la stratégie préhospitalière, l'accès au centre d'appels ou le type de plateforme. Il pourrait être extrêmement intéressant de comparer des modèles de gestion de l'urgence entre

différents pays, notamment le modèle américain qui fonctionne avec des centres d'appels centralisés, un numéro unique pour tout type d'urgence et une stratégie hospitalière de type *Scoop And Run*. En disposant de données de téléphonie issue de différents pays, de différents modèles, ainsi que leurs processus de gestion de l'appel et de leur politique d'évaluation de la performante : on pourrait comparer ces modèles afin d'évaluer factuellement leurs différences et leurs niveaux de performance suivant un référentiel commun qui resterait à définir.

De plus, il est tout à fait envisageable d'étendre cette méthodologie à d'autres cas d'applications au sein de la chaîne des urgences avec les centres d'appels des pompiers ou de la police par exemple.

Évaluer le centre d'appels selon d'autres indicateurs de performance

Dans notre étude, nous avons évalué l'accessibilité du centre d'appels en termes de qualité de service et de temps d'attente maximum. Toutefois, il serait important d'étendre cet ensemble d'indicateurs à d'autres dimensions de performance comme le taux d'utilisation des ressources pour avoir une mesure de l'efficacité du modèle et de son dimensionnement. Afin d'avoir une évaluation complète des centres d'appels d'urgence, il faudrait pouvoir les évaluer sur d'autres critères que le temps.

Aller plus loin sur la définition des temps de session

Nous avons vu dans le chapitre 4 que nous pouvions définir les temps de session en utilisant les données historiques ou une distribution théorique. Seulement, une hypothèse qu'il serait judicieux d'étudier concerne les temps de sessions qui peuvent dépendre du nombre d'appels présents dans le système : si un agent sait qu'il y a d'autres appels à traiter, il va probablement passer moins de temps sur l'appel en cours. A contrario, si l'ARM sait qu'il n'y a pas d'autres appels dans la file d'attente, alors il/elle sera plus à même de prendre son temps. Ainsi, les temps de session pourraient être définis en fonction de la taille de la file d'attente ou de la période de la journée. Avec une étude plus poussée des données des centres d'appels, il paraît probable de trouver une modélisation de ce phénomène et ainsi affiner notre modèle, et les conclusions afférentes.

De plus, il serait opportun d'analyser la relation entre le type d'appel et sa durée. Par exemple, est-ce que les appels pour AVC sont traités dans un temps qui suit le même ordre de grandeur que d'autres types d'urgence ? Est-ce qu'il y a des différences de temps de traitement entre les types d'interventions nécessaires ?

Prendre en compte la nature des appels entrants

Dans notre modèle, nous avons choisi de ne considérer que les appels entrants provenant de la population. Ainsi, nous avons mis de côté tous les appels des professionnels ainsi que les appels bilans des équipes qui vont sur le terrain. L'intégration de ces appels permettrait d'affiner le modèle et de prendre en compte leur gestion particulière qui diffère de celle des appels entrants issus de la population.

Continuer à travailler sur les appels sortants

L'inclusion des appels sortants dans notre modèle est une nouveauté qui n'avait pas été réalisée jusqu'à présent dans la littérature scientifique. Il reste néanmoins du travail à fournir pour être en mesure de les modéliser correctement, et de mieux comprendre les interactions entre appels sortants et appels entrants au sein du centre d'appels. En effet, nous faisons l'hypothèse que l'émission d'un appel sortant va dépendre de la nature de l'appel entrant. Ainsi, le fait de disposer des informations du dossier médical pourrait nous donner des informations supplémentaires sur quand et pourquoi des appels sortants sont générés, mobilisant ainsi des ressources pour autre chose que le décroché des appels.

Élargir l'horizon d'étude

Puisque le dimanche a été identifié comme le jour critique, notre étude était basée uniquement sur ce jour. Il serait intéressant de pouvoir simuler des semaines entières d'appels et de prendre en compte la saisonnalité : d'abord sur les variations à l'échelle de la semaine, puis sur l'année en considérant les effets de saisons, des vacances scolaires, des épidémies de gripes, etc.

En outre, l'analyse des situations de crises mérite d'être approfondie. Il faudrait tester des situations bien plus extrêmes, telles que celle du COVID-19 et expérimenter des organisations à mettre en place pour être en mesure de gérer ces crises. Pour cela, de nouveaux scénarios et de nouveaux indicateurs de performance sont à développer pour évaluer notamment la résilience du SAMU. La collaboration à l'échelle d'une région pourrait être un premier axe de travail intéressant.

Dans le cadre d'un projet PREPS financé par la Direction Générale de l'Offre de Soins (DGOS) débuté en mars 2020, en partenariat avec le CHU de Nantes et la région des Pays de la Loire, les perspectives de collaboration vont être explorées. En effet, dans le cadre de ce projet nous travaillerons sur la collaboration de 5 SAMU à l'échelle régionale.



Recommandations du rapport Marcus

Ci-dessous nous détaillons toutes les recommandations du rapport MARCUS (Moutard et al., 2019).

Une volumétrie des appels d'urgence difficile à objectiver, qui valide tout de même la nécessité de renforcer l'accessibilité aux centres d'appels

1. Diffuser et généraliser l'emploi d'une sémantique homogène du traitement de l'appel d'urgence entre les métiers, afin d'harmoniser la description opérationnelle
2. Fixer un seuil d'abandon des appels à 10 secondes

Une activité opérationnelle efficace, mais perfectible à court terme

3. Définir des indicateurs clefs de performance de l'activité opérationnelle
4. Historiser les données en perspective de la réalisation de modélisations d'activité
5. Harmoniser les axes organisationnels principaux des centres d'appels d'urgence
6. Adapter le dimensionnement des centres d'appels d'urgence pour optimiser le traitement des appels et répondre aux objectifs de performance établis
7. Tendre vers une meilleure flexibilité des cycles de travail des opérateurs pour garantir l'adéquation entre planification des ressources et prévision de l'activité opérationnelle
8. Renforcer les stratégies d'articulation et d'entraide entre les centres d'appels d'urgence pour gérer les pics d'appels
9. Faire évoluer les modalités de distribution des flux téléphoniques grâce à la virtualisation pour permettre l'interconnexion des centres d'appels d'urgence
10. Structurer les processus et les outils de pilotage opérationnel des centres d'appels d'urgence afin de suivre l'activité en temps réel aux niveaux local et national

11. Professionnaliser des agents aux fonction de supervision et de pilotage de l'activité opérationnelle
12. Encourager le développement et l'utilisation des nouvelles technologies afin d'améliorer le traitement des appels d'urgence

Une offre de formation des opérateurs des centres d'appels d'urgence insuffisante et hétérogène intra et inter-métiers

13. Faire de la formation un prérequis à la prise de poste en centre d'appels d'urgence
14. Faire converger les forces sur la définition des axes et des objectifs pédagogiques transverses à l'ensemble des centres d'appels d'urgence
15. Proposer une offre de formation continue permettant le maintien et le développement des acquis afin de renforcer l'amélioration continue au sein et entre les centres d'appels d'urgence

Une démarche qualité à structurer pour améliorer la qualité du service rendu aux usagers et renforcer l'attractivité du métier d'opérateur"

16. Renforcer et structurer davantage les dispositifs qualité existants
17. Approfondir et diffuser les initiatives qualité et les bonnes pratiques initiées sur l'ensemble du territoire pour étudier leur généralisation
18. Encourager la certification qualité de la relation usager pour valoriser l'activité d'accueil aux appels d'urgence
19. Piloter et suivre la démarche qualité des centres d'appels d'urgence tant au niveau local qu'au niveau national
20. Améliorer et harmoniser l'organisation du travail au sein des centres d'appels d'urgence
21. Moderniser l'environnement de travail des opérateurs
22. Former et informer les opérateurs sur les risques professionnels en centres d'appels d'urgence
23. Encourager la certification qualité des centres d'appels d'urgence pour harmoniser les pratiques et les organisations

Le Cadre d'Interopérabilité des Services d'Urgence (CISU) au service de l'amélioration de la gestion des appels d'urgence"

24. Mettre en oeuvre le CISU de manière effective et inter-applicative dans le cadre de l'expérimentation MARCUS

Un manque de lisibilité et/ou d'éducation de la population sur le bon usage des numéros d'urgence

25. Associer un flux urgent et un flux non urgent répondant à des demandes spécifiques

-
26. Repenser les stratégies de communication autour des numéros d'urgence en garantissant une communication d'ensemble, ou à minima coordonnée entre les différentes forces, et une récurrence permettant d'ancrer durablement les messages et de les pérenniser auprès de la population
 27. Tirer pleinement parti de l'ensemble des vecteurs éducatifs pour permettre une meilleure lisibilité des numéros d'urgence
 28. Porter un effort ciblé sur des populations clefs, adossé aux campagnes de formation aux gestes de premiers secours ou à des vecteurs de communication spécifiques
 29. Assurer un service respectant les exigences de la charte Marianne
 30. Encourager le développement et l'utilisation des nouveaux modes de saisine des services d'urgence afin d'améliorer l'accessibilité des appels d'urgence, tout en garantissant l'interopérabilité avec les systèmes d'information existants

Dénominateurs communs pour répondre à l'amélioration du traitement des communications d'urgence

31. Mettre en place un premier niveau de décroché assurant les fonctions de filtrage et/ou d'orientation des appels d'urgence
32. Confirmer la nécessité de prioriser les appels d'urgence par le niveau 1, selon un référentiel co-construit par les différents services métiers
33. Placer les fonctions de filtrage et/ou d'orientation de l'appel d'urgence (N1) à un niveau supra-départemental couvrant un bassin populationnel de 3 à 6 millions d'habitants et recevant entre 10 à 18 000 appels/jour (hors territoires ultramarins)
34. Maintenir la qualification métier (N2) à l'échelle locale existante, très majoritairement départementale, à ce stade
35. Institutionnaliser et mettre en place une gouvernance nationale partagée des communications d'urgence

B

Profil des appels entrants à Rodez

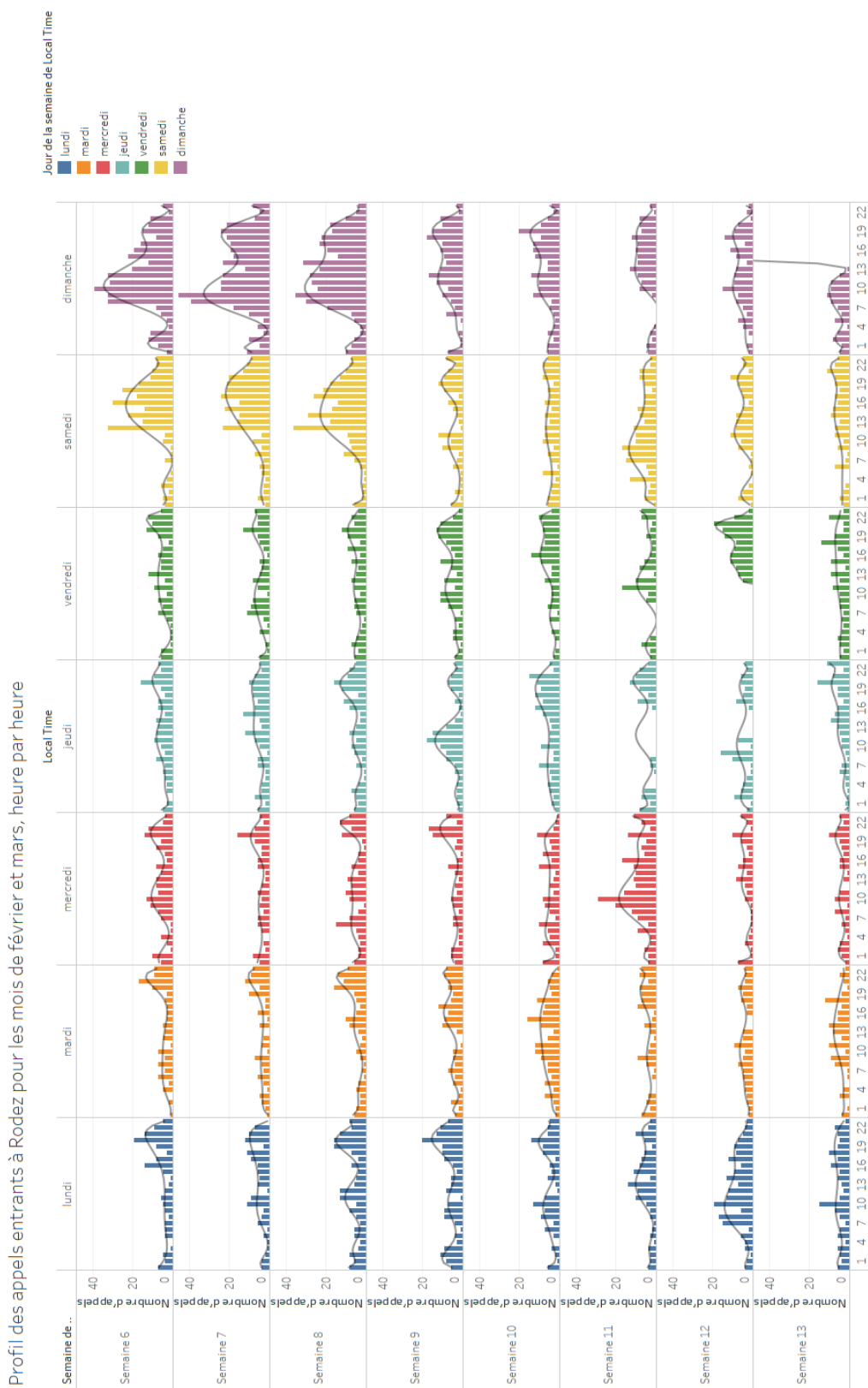
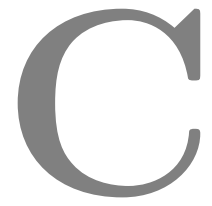


Figure B.1 – Profil des appels entrants sur huit semaines, heure par heure à Rodez



Modèle de simulation agrandi

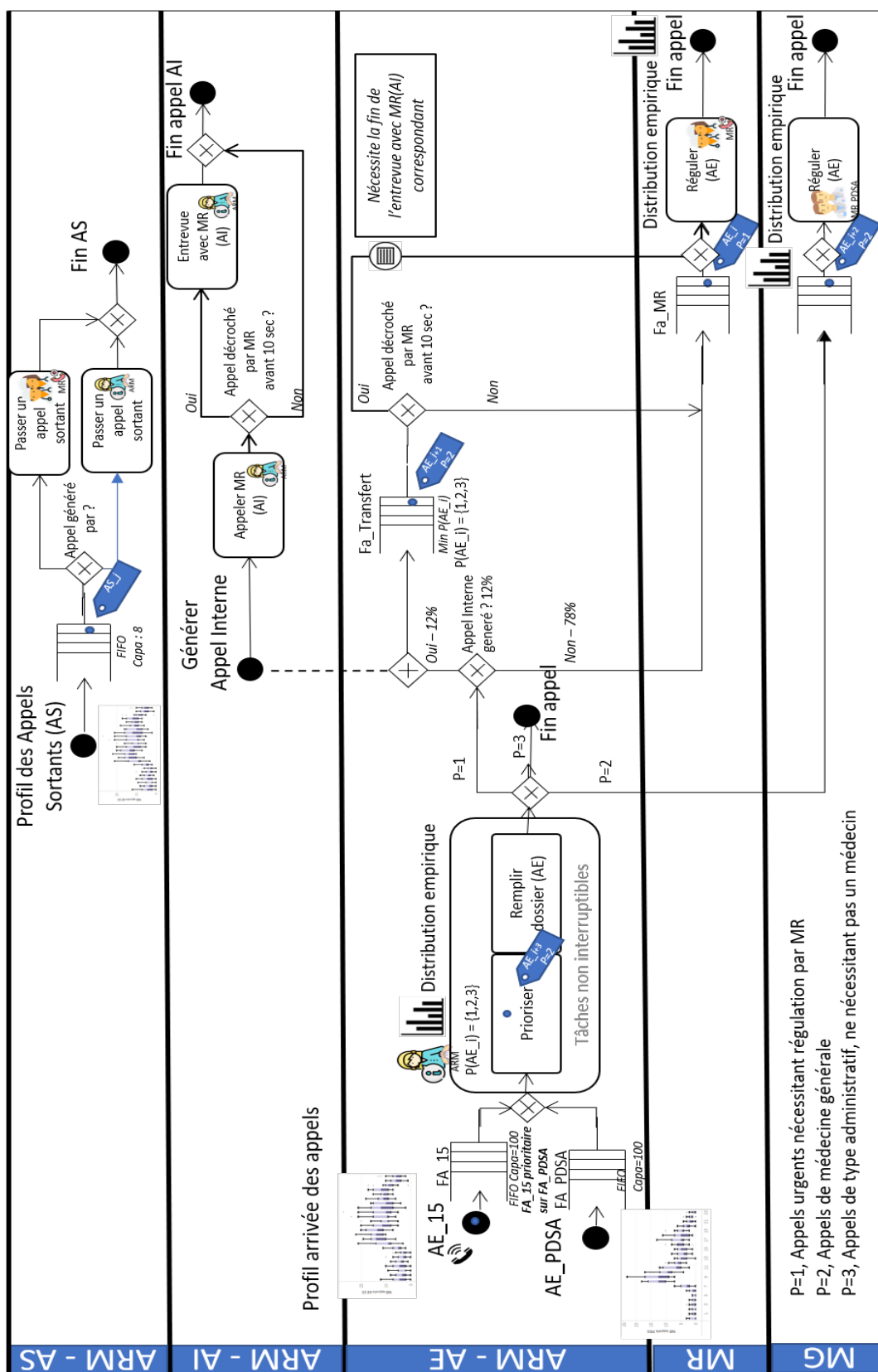


Figure C.1 – Modèle de simulation AS-IS

Table des figures

1	Méthodologie de recherche adaptée de Hevner et al. (2004)	10
2	Plan du manuscrit	12
1.1	La chaîne des urgences	17
1.2	Le modèle de décroché décentralisé	19
1.3	Le modèle de décroché centralisé : première version	20
1.4	Le modèle de décroché centralisé : deuxième version	20
1.5	Le modèle de décroché centralisé et unifié	21
1.6	Le modèle de décroché généralisé	21
1.7	Le modèle de décroché interconnecté	22
1.8	Missions et structure du SAMU	24
1.9	Flux et partenaires du SAMU	26
1.10	Cycle de vie d'un appel : les étapes.	28
1.11	Indicateurs de performance	38
1.12	Positionnement des types de sujets en fonction de l'activité de la chaîne des urgences	52
1.13	Résultats de recherche du 02/06/2020 sur Web of Science avec la requête :(("Emergency Medical Service*" OR EMS) AND((Call OR Contact OR Communication) AND Center))	53
1.14	Résultats de recherche du 02/06/2020 sur Web of Science avec la requête : "Emergency Medical Service*" OR EMS) AND (Ambulance* OR transport*	53
2.1	Présentation des données nationales	59

2.2	Extrait de la base de données	62
2.3	Exemple de clustering hiérarchique sur deux critères avec une distance simple.	64
2.4	Clustering proposé des départements français	64
2.5	Dendrogramme des départements français	66
2.6	Nombre d'interventions en fonction du nombre de moyens	68
3.1	La méthodologie outillée proposée (Petitdemange et al., 2019)	79
3.2	Collecter et nettoyer : entrées, sorties et outils	80
3.3	Découvrir le processus : entrées, sorties et outils	81
3.4	Un processus métier et des exemples de patrons d'appels	84
3.5	Fusionner les modèles : entrées, sorties et outils	85
3.6	Extraire les connaissances : entrées, sorties et outils	86
3.7	Paramétrer : entrées, sorties et outils	87
3.8	Évaluer le AS-IS : entrées, sorties et outils	88
3.9	Valider le modèle : entrées, sorties et outil	89
3.10	Améliorer : entrées, sorties et outils	90
4.1	Partie de la méthodologie traitée dans ce chapitre	93
4.2	Traitement des données pour obtenir un log exploitable.	97
4.3	Extrait d'un log de données non traité.	98
4.4	Exemple de log avant traitement, sous forme de table pour un appel.	99
4.5	Processus de la prise d'un appel d'urgence à Rodez	103
4.6	Processus de la prise d'un appel d'urgence à Toulouse	104
4.7	Utilisation du Process Mining pour nos données	105
4.8	Quelques exemples de traces (Partie 1)	107
4.9	Quelques exemples de traces (Partie 2)	108
4.10	Représentation d'un centre d'appels par file d'attente. (Gans et al., 2003)	110
4.11	Distribution théorique et empirique des temps de session pour l'ARM	115
4.12	Profil des appels entrants sur huit semaines, heure par heure	117
4.13	Identifier les types d'appels par leur événement initial	118
5.1	Partie de la méthodologie traitée dans ce chapitre	121

5.2	Extrait du modèle de simulation	126
5.3	Modèle conceptuel de la gestion des appels. (Voir version agrandie en annexe)	129
5.4	Modèle conceptuel de la gestion des appels : Définition des macro-activités	129
5.5	Profil d'arrivée des appels entrants (un dimanche)	130
5.6	différents types de file d'attentes	131
5.7	répartition des durées de session pour les appels sortants	133
5.8	Profil d'émission des appels sortants	134
5.9	Gestion des Appels sortants (AS)	134
5.10	Gestion des Appels internes (AI)	135
5.11	Écarts entre les profils d'arrivées des appels réels et simulés	136
5.12	Différents chemins pouvant être empruntés par un appel entrant dans le modèle de simulation	138
5.13	Scénario AS-IS	140
5.14	Scénario de prépriorisation des appels	142
5.15	Plan d'expériences proposé	143
5.16	Scénario de triage des appels	145
5.17	Gains sur le QoS20 et QoS60 au global pour tous type d'appels, dans la configuration C1	145
5.18	Gains sur les QoS20 et QoS60 pour les appels P1 dans la configuration C1	150
5.19	Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C1	151
5.20	Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C1	152
5.21	Temps d'attente maximum pour chaque scénario de la configuration C1	153
5.22	Temps d'attente maximum pour chaque scénario de la configuration C1 : appels P1	154
5.23	Temps d'attente maximum pour chaque scénario de la configuration C1 : appels P2	155
5.24	Temps d'attente maximum pour chaque scénario de la configuration C1 : appels P3	156
5.25	Gains sur les QoS20 et QoS60 pour les appels P1 dans la configuration C2	158

5.26	Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C2	159
5.27	Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C2	160
5.28	Comparaison des temps d'attentes maxima pour les appels de gravité P1 entre les deux configurations	163
5.29	Détail de la réplication 89 pour le scénario de Triage dans la configuration 2	164
5.30	Qualité de Service pour chacun des scénarios. Demande double, Configuration C1	166
5.31	Demande doublée : Gains sur les QoS_{20} et QoS_{60} pour les appels P1 dans la configuration C1	167
5.32	Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C1	169
5.33	Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C1	170
5.34	Effet de l'augmentation du nombre d'appels sur le temps d'attente maximum en configuration C1	171
5.35	Qualité de Service pour chacun des scénarios. Demande double, Configuration C2	172
5.36	Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P1 dans la configuration C2	173
5.37	Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P2 dans la configuration C2	174
5.38	Demande doublée : Gains sur les QoS20 et QoS60 pour les appels P3 dans la configuration C2	175
5.39	Effet de l'augmentation du nombre d'appels sur le temps d'attente maximum en configuration C1	176
B.1	Profil des appels entrants sur huit semaines, heure par heure à Rodez	190
C.1	Modèle de simulation AS-IS	192

Liste des tableaux

1.1	Synthèse des défis à relever	50
2.1	Statistiques descriptives sur les départements pour chacun des critères	62
2.2	Tableau récapitulatif des données pour les Bouches-du-Rhône, le Nord et le Pas-de-Calais comparé à la moyenne nationale	69
3.1	Contributions et Limites : Chapitre 3	91
4.1	Forces et faiblesses de chaque méthode (d'après (Dumas et al., 2013))	95
4.2	Caractéristiques des systèmes d'information dans les SAMU visités	103
4.3	Informations sur la durée de communication des ARM	114
4.4	Part des appels sortants et des appels entrants dans la volumétrie totale sur 8 semaines	119
4.5	Sur quels appels les ARM et les MR passent-ils le plus de temps ?	119
4.6	Contributions et Limites : Chapitre 4	120
5.1	Répartition de la priorité des appels en fonction de la provenance des appels.	132
5.2	Statistiques descriptives des performances du modèle comparées aux performances réelles	137
5.3	Valeur du QoS au global sur la journée dans la première configuration.	144
5.4	Amélioration du QoS au global par priorité d'appels	146
5.5	Gains par rapport au AS-IS pour les appels P1 sur chaque créneau	146
5.6	Gains par rapport au AS-IS pour les appels P2 sur chaque créneau	147

5.7	Gains par rapport au AS-IS pour les appels P3 sur chaque créneau	147
5.8	Qualité de service et gains par rapport à l'AS-IS pour chaque scénario	148
5.9	Activité courante des ARM plutôt que de décrocher l'appel entrant	155
5.10	Configuration 2 : Comparaison du QoS entre les scénarios et l'AS-IS.	161
5.11	Comparaison des QoS entre les scénarios de la nouvelle configuration et l'AS-IS de la configuration originale.	161
5.12	Contributions et Limitation : Chapitre 5	178

Bibliographie

- (Aboueljinane et al., 2013) L. Aboueljinane, E. Sahin, and Z. Jemai. “A review on simulation models applied to emergency medical service operations.” In: *Computers & Industrial Engineering* 66.4 (Dec. 2013), pp. 734–750 (cf. p. 37, 40, 124).
- (Aboueljinane, 2014) L. Aboueljinane. “Evaluation et amélioration des performances des Systèmes d’Aide Médicale Urgente : application au SAMU du département du Val de Marne.” fr. PhD thesis. Ecole Centrale Paris, June 2014 (cf. p. 34, 35, 40, 45, 47).
- (Aboueljinane et al., 2014) L. Aboueljinane, E. Sahin, Z. Jemai, and J. Marty. “A simulation study to improve the performance of an emergency medical service: Application to the French Val-de-Marne department.” In: *Simulation Modelling Practice and Theory* 47.Supplement C (Sept. 2014), pp. 46–59 (cf. p. 37, 48, 118).
- (Addis et al., 2012) B. Addis, R. Aringhieri, G. Carello, A. Grosso, and F. Maffioli. “Workforce management based on forecasted demand.” en. In: *Advanced Decision Making Methods Applied to Health Care*. International Series in Operations Research & Management Science. Springer, Milano, 2012, pp. 1–11 (cf. p. 37, 39).
- (Agrawal et al., 1998) R. Agrawal, D. Gunopulos, and F. Leymann. “Mining process models from workflow logs.” In: *International Conference on Extending Database Technology*. Springer. 1998, pp. 467–483 (cf. p. 104).
- (Aldrich et al., 1971) C. A. Aldrich, J. C. Hisserich, and L. B. Lave. “An analysis of the demand for emergency ambulance service in an urban area.” In: *American Journal of Public Health* 61.6 (1971), pp. 1156–1169 (cf. p. 43).
- (Aringhieri et al., 2017) R. Aringhieri, M. E. Bruni, S. Khodaparasti, and J. T. van Essen. “Emergency medical services and beyond: Addressing new challenges through a wide literature review.” In: *Computers & Operations Research* 78 (Feb. 2017), pp. 349–368 (cf. p. 37, 39, 45).
- (Avramidis et al., 2009) A. N. Avramidis, W. Chan, and P. L’Ecuyer. “Staffing multi-skill call centers via search methods and a performance approximation.” In: *IIE Transactions* 41.6 (Apr. 2009), pp. 483–497 (cf. p. 88).

- (Baker et al., 1986) J. Baker and K. Fitzpatrick. “Determination of an optimal forecast model for ambulance demand using goal programming.” In: *Journal of the Operational Research Society* 37.11 (1986), pp. 1047–1059 (cf. p. 43).
- (Bandara et al., 2012) D. Bandara, M. E. Mayorga, and L. A. McLay. “[Optimal dispatching strategies for emergency vehicles to increase patient survivability.](#)” In: *International Journal of Operational Research* 15.2 (Jan. 2012), pp. 195–214 (cf. p. 48).
- (Baru et al., 2015) R. A. Baru, E. A. Cudney, I. G. Guardiola, D. L. Warner, and R. E. Phillips. “Systematic review of operations research and simulation methods for bed management.” In: *IIE Annual Conference. Proceedings*. Institute of Industrial and Systems Engineers (IISE). 2015, p. 298 (cf. p. 125).
- (Başar et al., 2012) A. Başar, B. Çatay, and T. Ünlüyurt. “A taxonomy for emergency service station location problem.” In: *Optimization letters* 6.6 (2012), pp. 1147–1160 (cf. p. 40).
- (Belaidi, 2009) A. Belaidi. “[Organisation et pilotage de la prise en charge des patients dans le cadre du réseau des urgences.](#)” fr. PhD thesis. Université Jean Monnet - Saint-Etienne, July 2009 (cf. p. 37, 182).
- (Belaidi et al., 2007) A. Belaidi, B. Besombes, E. Marcon, and A. Guinet. “[Identifying and Modeling decision problems for emergency network in France: a literature analysis.](#)” In: *ORAHS’2007 (Operational Research Applied to Health Services), 33th International Conference*. St-Etienne, France: Publications de l’Université de Saint-Etienne, July 2007, pp. 353–366 (cf. p. 36).
- (Bélanger et al., 2016) V. Bélanger, Y. Kergosien, A. Ruiz, and P. Soriano. “An empirical comparison of relocation strategies in real-time ambulance fleet management.” In: *Computers & Industrial Engineering* 94 (2016), pp. 216–229 (cf. p. 37, 49).
- (Beliën et al., 2012) J. Beliën and H. Forcé. “Supply chain management of blood products: A literature review.” In: *European Journal of Operational Research* 217.1 (2012), pp. 1–16 (cf. p. 125).
- (Beuran et al., 2012) M. Beuran, S. Paun, B. Gaspar, M. Vartic, S. Hostiuc, A. Chiotoroiu, and I. Negoii. “Prehospital trauma care: a clinical review.” In: *Chirurgia* 107 (2012), pp. 564–570 (cf. p. 20).
- (Bhulai, 2009) S. Bhulai. “[DYNAMIC ROUTING POLICIES FOR MULTISKILL CALL CENTERS.](#)” en. In: *Probability in the Engineering and Informational Sciences* 23.1 (Jan. 2009), pp. 101–119 (cf. p. 85, 87).
- (Bieler et al., 2017) D. Bieler, A. Franke, R. Lefering, S. Hentsch, A. Willms, M. Kulla, E. Kollig, T. DGU, et al. “Does the presence of an emergency physician influence pre-hospital time, pre-hospital interventions and the mortality of severely injured patients? A matched-pair analysis based on the trauma registry of the German Trauma Society (TraumaRegister DGU®).” In: *Injury* 48.1 (2017), pp. 32–40 (cf. p. 21).
- (Blomberg et al., 2019) S. N. Blomberg, F. Folke, A. K. Ersboll, H. C. Christensen, C. Torp-Pedersen, M. R. Sayre, C. R. Counts, and F. K. Lippert. “[Machine learning as a supportive tool to recognize cardiac arrest in emergency calls.](#)” In: *RESUSCITATION* 138 (May 2019), pp. 322–329 (cf. p. 47, 140).

-
- (Bohm et al., 2018) K. Bohm and L. Kurland. “The accuracy of medical dispatch—a systematic review.” In: *Scandinavian journal of trauma, resuscitation and emergency medicine* 26.1 (2018), p. 94 (cf. p. 47).
- (Borshchev, 2013a) A. Borshchev. “[Multi-method Modeling](#).” In: *Proceedings of the 2013 Winter Simulation Conference: Simulation: Making Decisions in a Complex World*. WSC ’13. Piscataway, NJ, USA: IEEE Press, 2013, pp. 4089–4100 (cf. p. 123, 124).
- (Borshchev, 2013b) A. Borshchev. *The big book of simulation modeling: multimethod modeling with AnyLogic 6*. AnyLogic North America, 2013 (cf. p. 86, 123).
- (Bradbeer et al., 2000) P. Bradbeer, C. Findlay, and T. C. Fogarty. “An ambulance crew rostering system.” In: *Workshops on Real-World Applications of Evolutionary Computation*. Springer. 2000, pp. 267–279 (cf. p. 44).
- (L. H. Brown et al., 2007) L. H. Brown, E. B. Lerner, B. Larmon, T. LeGassick, and M. Taigman. “[Are EMS Call Volume Predictions Based on Demand Pattern Analysis Accurate?](#)” In: *Prehospital Emergency Care* 11.2 (Jan. 2007), pp. 199–203 (cf. p. 42).
- (L. Brown et al., 2005) L. Brown, N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, and L. Zhao. “Statistical Analysis of a Telephone Call Center: A Queueing-Science Perspective.” en. In: (2005), p. 61 (cf. p. 114).
- (M. v. Buuren et al., 2015) M. v. Buuren, G. J. Kommer, R. v. d. Mei, and S. Bhulai. “[A simulation model for emergency medical services call centers](#).” In: *2015 Winter Simulation Conference (WSC)*. Dec. 2015, pp. 844–855 (cf. p. 45, 48).
- (M. v. Buuren et al., 2012) M. v. Buuren, R. v. d. Mei, K. Aardal, and H. Post. “[Evaluating dynamic dispatch strategies for emergency medical services: TIFAR simulation tool](#).” In: *Proceedings of the 2012 Winter Simulation Conference (WSC)*. Dec. 2012, pp. 1–12 (cf. p. 49).
- (M. v. Buuren et al., 2018) M. van Buuren, C. Jagtenberg, T. van Barneveld, R. van der Mei, and S. Bhulai. “[Ambulance Dispatch Center Pilots Proactive Relocation Policies to Enhance Effectiveness](#).” In: *INTERFACES* 48.3 (June 2018), pp. 235–246 (cf. p. 47).
- (M. v. Buuren et al., 2017) M. van Buuren, G. J. Kommer, R. van der Mei, and S. Bhulai. “[EMS call center models with and without function differentiation: A comparison](#).” In: *OPERATIONS RESEARCH FOR HEALTH CARE* 12 (Mar. 2017), pp. 16–28 (cf. p. 46, 77, 118).
- (Caillaud et al., 2015) E. Caillaud, B. Rose, and V. Goepp. “Comment valider les recherches en génie industriel et en ingénierie système?” In: (2015) (cf. p. 9).
- (Cardoen et al., 2010) B. Cardoen, E. Demeulemeester, and J. Beliën. “Operating room planning and scheduling: A literature review.” In: *European journal of operational research* 201.3 (2010), pp. 921–932 (cf. p. 125).
- (Channouf et al., 2007) N. Channouf, P. L’Ecuyer, A. Ingolfsson, and A. N. Avramidis. “[The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta](#).” en. In: *Health Care Management Science* 10.1 (Feb. 2007), pp. 25–45 (cf. p. 43).

- (Cheang et al., 2003) B. Cheang, H. Li, A. Lim, and B. Rodrigues. “Nurse rostering problems—a bibliographic survey.” In: *European journal of operational research* 151.3 (2003), pp. 447–460 (cf. p. 43).
- (Cnossen et al., 2019) M. C. Cnossen, R. Van Der Brande, H. F. Lingsma, S. Polinder, F. Lecky, A. I. Maas, C. T. Investigators, and Participants. “Prehospital trauma care among 68 European neurotrauma centers: Results of the CENTER-TBI Provider Profiling Questionnaires.” In: *Journal of neurotrauma* 36.1 (2019), pp. 176–181 (cf. p. 22).
- (Crema et al., 2013) M. Crema and C. Verbano. “Guidelines for overcoming hospital managerial challenges: a systematic literature review.” In: *Therapeutics and clinical risk management* 9 (2013), p. 427 (cf. p. 125).
- (D’Agostino, 1986) R. B. D’Agostino. *Goodness-of-fit-techniques*. Vol. 68. CRC press, 1986 (cf. p. 113).
- (Dami et al., 2015) F. Dami, C. Golay, M. Pasquier, V. Fuchs, P.-N. Carron, and O. Hugli. “Prehospital triage accuracy in a criteria based dispatch centre.” In: *BMC Emergency Medicine* 15 (Oct. 2015), p. 32 (cf. p. 47).
- (Dogan et al., 2010) N. Dogan and I. Dogan. “Determination of the number of bins/classes used in histograms and frequency tables: a short bibliography: A SHORT BIBLIOGRAPHY.” In: *TurkStat, Journal of Statistical Research* 07 (Jan. 2010), pp. 77–86 (cf. p. 113).
- (Dongen, van B.F. et al., 2005) Dongen, van B.F., Aalst, van der W.M.P., Castro, J., Tentiento, E., and Information Systems IE&IS. “A meta model for process mining data.” en. In: *2nd Open INTEROP Workshop on Enterprise Modelling and Ontologies for Interoperability (EMOI - INTEROP ’05)*. FEUP, 2005 (cf. p. 78).
- (Dumas et al., 2013) M. Dumas, M. La Rosa, J. Mendling, H. A. Reijers, et al. *Fundamentals of business process management*. Vol. 1. Springer, 2013 (cf. p. 94, 95).
- (Dwars et al., 2013) R. Dwars, R. D. van der Mei, and S. Bhulai. “Capacity planning of emergency call centers.” en. Doctoral dissertation, MSc thesis. Amsterdam: VU University, 2013 (cf. p. 85).
- (Erdoğan et al., 2010) G. Erdoğan, E. Erkut, A. Ingolfsson, and G. Laporte. “Scheduling ambulance crews for maximum coverage.” In: *Journal of the Operational Research Society* 61.4 (2010), pp. 543–550 (cf. p. 44).
- (Ernst et al., 2004a) A. T. Ernst, H. Jiang, M. Krishnamoorthy, B. Owens, and D. Sier. “An annotated bibliography of personnel scheduling and rostering.” In: *Annals of Operations Research* 127.1-4 (2004), pp. 21–144 (cf. p. 43).
- (Ernst et al., 2004b) A. T. Ernst, H. Jiang, M. Krishnamoorthy, and D. Sier. “Staff scheduling and rostering: A review of applications, methods and models.” In: *European journal of operational research* 153.1 (2004), pp. 3–27 (cf. p. 43).
- (Fitzsimmons, 1971) J. A. Fitzsimmons. “An Emergency Medical System Simulation Model.” In: *Proceedings of the 5th Conference on Winter Simulation*. WSC ’71. New York, NY, USA: ACM, 1971, pp. 18–25 (cf. p. 49).

-
- (Forrester, 1971) J. W. Forrester. *World dynamics*. Wright-Allen Press, 1971 (cf. p. 124).
- (Française, 2020) O. québécois de la langue Française. *Le grand dictionnaire terminologique*. 2020 (cf. p. 73).
- (France et al., 2015) S. U. de France and S. F. de médecine d’Urgence. *Samu Centre 15 référentiel et guide d’évaluation*. 2015 (cf. p. 5, 16, 25, 29, 30, 33, 44).
- (Fujiwara et al., 1987) O. Fujiwara, T. Makjamroen, and K. K. Gupta. “Ambulance deployment analysis: A case study of Bangkok.” In: *European Journal of Operational Research* 31.1 (1987), pp. 9–18 (cf. p. 50).
- (Gans et al., 2003) N. Gans, G. Koole, and A. Mandelbaum. “Telephone Call Centers: Tutorial, Review, and Research Prospects.” In: *Manufacturing & Service Operations Management* 5.2 (Apr. 2003), pp. 79–141 (cf. p. 110).
- (Gardett et al., 2016) I. Gardett, J. Clawson, G. Scott, T. Barron, B. Patterson, and C. Olola. *Past, present, and future of emergency dispatch research: a systematic literature review*. 2016 (cf. p. 77).
- (Gendreau et al., 1997) M. Gendreau, G. Laporte, and F. Semet. “Solving an ambulance location model by tabu search.” In: *Location science* 5.2 (1997), pp. 75–88 (cf. p. 39).
- (Gendreau et al., 2001) M. Gendreau, G. Laporte, and F. Semet. “A dynamic model and parallel tabu search heuristic for real-time ambulance relocation.” In: *Parallel computing* 27.12 (2001), pp. 1641–1653 (cf. p. 48).
- (Gill et al., 2008) P. Gill, K. Stewart, E. Treasure, and B. Chadwick. “Methods of data collection in qualitative research: interviews and focus groups.” En. In: *British Dental Journal* 204.6 (Mar. 2008), p. 291 (cf. p. 81).
- (Goldberg, 2004) J. B. Goldberg. “Operations research models for the deployment of emergency services vehicles.” In: *EMS management Journal* 1.1 (2004), pp. 20–39 (cf. p. 34).
- (Grekousis et al., 2019) G. Grekousis and Y. Liu. “Where will the next emergency event occur? Predicting ambulance demand in emergency medical services using artificial intelligence.” In: *COMPUTERS ENVIRONMENT AND URBAN SYSTEMS* 76 (July 2019), pp. 110–122 (cf. p. 43).
- (Gul et al., 2015) M. Gul and A. F. Guneri. “A comprehensive review of emergency department simulation applications for normal and disaster conditions.” In: *Computers & Industrial Engineering* 83 (2015), pp. 327–344 (cf. p. 124).
- (Gunes et al., 2005) E. Gunes and R. Szechtman. “A simulation model of a helicopter ambulance service.” In: *Proceedings of the Winter Simulation Conference, 2005*. IEEE. 2005, 7–pp (cf. p. 39).
- (Gupta et al., 2010) N. K. Gupta, R. Dantu, H. Schulzrinne, A. Goulart, and W. Magnussen. “Next Generation 9-1-1: Architecture and Challenges in Realizing an IP-Multimedia-Based Emergency Service.” In: *Journal of Homeland Security and Emergency Management* 7.1 (2010) (cf. p. 41).
- (Hall, 1971) W. K. Hall. “Management science approaches to the determination of urban ambulance requirements.” In: *Socio-economic Planning Sciences* 5.5 (1971), pp. 491–499 (cf. p. 43).

- (Hand et al., 2001) D. J. Hand, H. Mannila, and P. Smyth. *Principles of data mining (adaptive computation and machine learning)*. MIT Press, 2001 (cf. p. 78).
- (Hannay et al., 2014) R. S. Hannay, A. Wyrzykowski, C. Ball, K. Laupland, and D. Feliciano. “Retrospective review of injury severity, interventions and outcomes among helicopter and nonhelicopter transport patients at a Level 1 urban trauma centre.” In: *Canadian Journal of Surgery* 57.1 (Feb. 2014), pp. 49–54 (cf. p. 21, 75).
- (Hevner et al., 2004) A. R. Hevner, S. T. March, J. Park, and S. Ram. “Design Science in Information Systems Research.” In: *MIS Quarterly* 28.1 (2004). Publisher: Management Information Systems Research Center, University of Minnesota, pp. 75–105 (cf. p. 9).
- (Homayounfar, 2012) P. Homayounfar. “Process mining challenges in hospital information systems.” In: *2012 Federated Conference on Computer Science and Information Systems (FedCSIS)*. IEEE. 2012, pp. 1135–1140 (cf. p. 106).
- (Hospitalier (MeaH)., 2008) M. d’Audit d’Expertise et d’Audit Hospitalier (MeaH). *Organisation des Centres 15 : "chaque appel compte, chaque seconde compte"*. Tech. rep. MEAH, 2008 (cf. p. 36, 44).
- (Inakawa et al., 2010) K. Inakawa, T. Furuta, and A. Suzuki. “Effect of ambulance station locations and number of ambulances to the quality of the emergency service.” In: *The Ninth International Symposium on Operations Research and Its Applications (ISORA’10)*. 2010, pp. 340–347 (cf. p. 36, 40).
- (Ingolfsson et al., 2003) A. Ingolfsson, E. Erkut, and S. Budge. “Simulation of single start station for Edmonton EMS.” en. In: *Journal of the Operational Research Society* 54.7 (July 2003), pp. 736–746 (cf. p. 34, 40, 49).
- (Isern et al., 2016) D. Isern and A. Moreno. “A systematic literature review of agents applied in healthcare.” In: *Journal of medical systems* 40.2 (2016), p. 43 (cf. p. 125).
- (Iskander, 1989) W. H. Iskander. “Simulation Modeling For Emergency Medical Service Systems.” In: *1989 Winter Simulation Conference Proceedings*. Dec. 1989, pp. 1107–1111 (cf. p. 39).
- (Javaudin et al., 2020) F. Javaudin, Y. Penverne, and E. Montassier. “Organisation of prehospital care: the French experience.” In: *European Journal of Emergency Medicine* 27.6 (2020), pp. 404–405 (cf. p. 21).
- (Jones, 1999) M. Jones. “An introduction to qualitative methods for health professionals.” In: *Journal of the Royal Sociandy of Medicine* 92.12 (1999), p. 657 (cf. p. 81).
- (Jutel, 2009) A. Jutel. “Sociology of diagnosis: a preliminary review.” In: *Sociology of health & illness* 31.2 (2009), pp. 278–299 (cf. p. 74).
- (Kammoun et al., 2014) A. Kammoun, T. Loukil, and W. Hachicha. “The use of discrete event simulation in hospital supply chain management.” In: *2014 International Conference on Advanced Logistics and Transport (ICALT)*. IEEE. 2014, pp. 143–148 (cf. p. 125).

-
- (Kaufman et al., 2009) L. Kaufman and P. J. Rousseeuw. *Finding groups in data: an introduction to cluster analysis*. Vol. 344. John Wiley & Sons, 2009 (cf. p. 61).
- (Kendall, 1953) D. G. Kendall. “Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain.” In: *The Annals of Mathematical Statistics* (1953), pp. 338–354 (cf. p. 110).
- (Koch et al., 2003) O. Koch and H. Weigl. “Modeling ambulance service of the Austrian Red Cross.” In: *Winter Simulation Conference*. Vol. 2. 2003, pp. 1701–1706 (cf. p. 37, 48).
- (Koole et al., 2002) G. Koole and A. Mandelbaum. “[Queueing Models of Call Centers: An Introduction](#).” en. In: *Annals of Operations Research* 113.1-4 (July 2002), pp. 41–59 (cf. p. 87).
- (Lamine et al., 2015) E. Lamine, F. Fontanili, M. Di Mascolo, and H. Pingaud. “[Improving the Management of an Emergency Call Service by Combining Process Mining and Discrete Event Simulation Approaches](#).” en. In: *Risks and Resilience of Collaborative Networks*. Ed. by L. M. Camarinha-Matos, F. Bénaben, and W. Picard. IFIP Advances in Information and Communication Technology. Springer International Publishing, 2015, pp. 535–546 (cf. p. 142).
- (Law et al., 2000) A. M. Law and D. w. Kelton. *Simulation Modeling and Analysis*. en. McGraw · Hill, 2000 (cf. p. 88, 122, 123).
- (Le Gallais, 1997) J. Le Gallais. “Systèmes numériques de contrôle-commande (SNCC).” In: *Techniques de l’ingénieur. Informatique industrielle* 2.R7505 (1997), R7505–1 (cf. p. 73).
- (Lee et al., 2012) T. Lee, S.-H. Cho, H. Jang, and J. G. Turner. “A simulation-based iterative method for a trauma center—Air ambulance location problem.” In: *Proceedings of the 2012 Winter Simulation Conference (WSC)*. IEEE. 2012, pp. 1–12 (cf. p. 34).
- (Leflar et al., 2013) J. J. Leflar and M. H. Siegel. *Organizational Resilience: Managing the Risks of Disruptive Events-A Practitioner’s Guide*. CRC Press, 2013 (cf. p. 35).
- (Lerner et al., 2007) E. B. Lerner, G. Nichol, D. W. Spaite, H. G. Garrison, and R. F. Maio. “A comprehensive framework for determining the cost of an emergency medical services system.” In: *Annals of emergency medicine* 49.3 (2007), pp. 304–313 (cf. p. 37).
- (Levy et al., 2007) G. Levy, L. Goldstein, Y. Erez, R. Levite, U. Bar, M. Marmor, G. Linn, E. Onn, Y. Levi, and Y. Bar-Dayan. “[Physician versus Paramedic in the Setting of Ground Forces Operations: Are They Interchangeable?](#)” en. In: *Military Medicine* 172.3 (Mar. 2007). Publisher: Oxford Academic, pp. 301–305 (cf. p. 21, 75).
- (Lewis et al., 2007) B. G. Lewis, R. D. Herbert, P. F. Summons, and W. J. Chivers. “Agent-based simulation of a multi-queue emergency services call centre to evaluate resource allocation.” en. In: *MODSIM 2007, International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand* (2007), p. 7 (cf. p. 45).
- (Li et al., 2009) Y. Li and E. Kozan. “Rostering ambulance services.” In: *Industrial engineering and management society* (2009), pp. 795–801 (cf. p. 44).

- (Ma, 2000) R. S. B. Ma. “The role of qualitative research in broadening the ‘evidence base’ for clinical practice.” In: *Journal of evaluation in clinical practice* 6.2 (2000), pp. 155–163 (cf. p. 81).
- (Markakis et al., 2017) E. K. Markakis, A. Lykourgiotis, I. Politis, A. Dagiuklas, Y. Rebahi, and E. Pallis. “[EMYNOS: Next Generation Emergency Communication.](#)” In: *IEEE Communications Magazine* 55.1 (Jan. 2017), pp. 139–145 (cf. p. 41).
- (Mason, 2013) A. J. Mason. “[Simulation and Real-Time Optimised Relocation for Improving Ambulance Operations.](#)” en. In: *Handbook of Healthcare Operations Management*. International Series in Operations Research & Management Science. Springer, New York, NY, 2013, pp. 289–317 (cf. p. 87).
- (McConnel et al., 1998) C. E. McConnel and R. W. Wilson. “The demand for prehospital emergency services in an aging society.” In: *Social Science & Medicine* 46.8 (1998), pp. 1027–1031 (cf. p. 43).
- (McCoy et al., 2013) C. E. McCoy, M. Menchine, S. Sampson, C. Anderson, and C. Kahn. “[Emergency Medical Services Out-of-Hospital Scene and Transport Times and Their Association With Mortality in Trauma Patients Presenting to an Urban Level I Trauma Center.](#)” en. In: *Annals of Emergency Medicine* 61.2 (Feb. 2013), pp. 167–174 (cf. p. 21, 75).
- (Miller et al., 2020) K. E. M. Miller, H. J. James, G. M. Holmes, and C. H. Van Houtven. “[The effect of rural hospital closures on emergency medical service response and transport times.](#)” In: *HEALTH SERVICES RESEARCH* 55.2 (Apr. 2020), pp. 288–300 (cf. p. 49).
- (Mitchell, 1999) T. M. Mitchell. “Machine learning and data mining.” In: *Communications of the ACM* 42.11 (1999), pp. 30–36 (cf. p. 61).
- (Mock et al., 2002) C. N. Mock, M. Tiska, M. Adu-Ampofo, and G. Boakye. “Improvements in prehospital trauma care in an African country with no formal emergency medical services.” In: *Journal of Trauma and Acute Care Surgery* 53.1 (2002), pp. 90–97 (cf. p. 16).
- (Montassier et al., 2015) E. Montassier, J. Labady, A. Andre, G. Potel, F. Berthier, J. Jenvrin, and Y. Penverne. “[The Effect of Work Shift Configurations on Emergency Medical Dispatch Center Response.](#)” In: *Prehospital Emergency Care* 19.2 (Apr. 2015), pp. 254–259 (cf. p. 44, 46).
- (Moutard et al., 2019) L.-c. R. Moutard and D. Y. Penverne. “Des défis à relever, des perspectives à concrétiser.” fr. In: (Dec. 2019), p. 116 (cf. p. 5, 25, 44, 55, 130, 185).
- (Müllner, 2011) D. Müllner. “Modern hierarchical, agglomerative clustering algorithms.” In: *arXiv preprint arXiv:1109.2378* (2011) (cf. p. 61, 63).
- (National Academies of Sciences et al., 2015) E. National Academies of Sciences, Medicine, et al. *Improving diagnosis in health care*. National Academies Press, 2015 (cf. p. 74).
- (Nikolaou et al., 2017) N. Nikolaou, M. Castrén, K. G. Monsieurs, D. Cimpoesu, M. Georgiou, V. Raffay, R. Koster, S. Hunyadi-Antičević, A. Truhlář, and L. Bossaert. “[Time delays to reach dispatch centres in different regions in Europe. Are we losing the window of opportunity? — The EUROCALL study.](#)” In: *Resuscitation* 111 (2017), pp. 8–13 (cf. p. 18).

-
- (OCDE, 2019) OCDE. *Panorama de la santé 2019*. 2019 (cf. p. 4).
- (Olsen, 1971) R. P. Olsen. “Can project management be defined?” In: Project Management Institute. 1971 (cf. p. 33).
- (Oorni et al., 2017) R. Oorni and A. Goulart. “In-Vehicle Emergency Call Services: eCall and Beyond.” In: *IEEE Communications Magazine* 55.1 (Jan. 2017), pp. 159–165 (cf. p. 24, 42).
- (Passmore et al., 2013) C. M. Passmore and J. Zhan. “Determining Appropriate Staffing Adjustments in a Call Center Staff Group.” In: *2013 International Conference on Social Computing*. Sept. 2013, pp. 1046–1053 (cf. p. 37, 39, 88).
- (Paul et al., 2010) S. A. Paul, M. C. Reddy, and C. J. DeFlitch. “A systematic review of simulation studies investigating emergency department overcrowding.” In: *Simulation* 86.8-9 (2010), pp. 559–571 (cf. p. 124).
- (Penn et al., 2017) C. Penn, T. Koole, and R. Natrass. “When seconds count: A study of communication variables in the opening segment of emergency calls.” en. In: *Journal of Health Psychology* 22.10 (Sept. 2017), pp. 1256–1264 (cf. p. 46).
- (Penverne et al., 2009) Y. Penverne, J. Jenvrina, L. Labastire, N. Tassou, A. Felix, G. Potel, and F. Berthier. “Management et coaching : perspective d’avenir dans les Samu centre 15 ?” In: *Journal Européen des Urgences* 22.4 (Dec. 2009), pp. 104–109 (cf. p. 36, 58).
- (Penverne et al., 2017) Y. Penverne, B. Leclere, J. Labady, F. Berthier, J. Jenvrin, F. Javaudin, E. Batard, and E. Montassier. “Key performance indicators’ assessment to develop best practices in an Emergency Medical Communication Centre.” eng. In: *European journal of emergency medicine : official journal of the European Society for Emergency Medicine* (May 2017) (cf. p. 35, 37, 44, 88).
- (Penverne et al., 2020) Y. Penverne, B. Leclere, J. Labady, F. Berthier, J. Jenvrin, F. Javaudin, and E. Montassier. “Impact of two-level filtering on emergency medical communication center triage during the COVID-19 pandemic: an uncontrolled before-after study.” In: *Scandinavian journal of trauma, resuscitation and emergency medicine* 28.1 (2020), pp. 1–6 (cf. p. 181).
- (Penverne et al., 2019) Y. Penverne, M. Terre, F. Javaudin, J. Jenvrin, F. Berthier, J. Labady, B. Leclere, and E. Montassier. “Connect dispatch centers for call handling improves performance.” In: *SCANDINAVIAN JOURNAL OF TRAUMA RESUSCITATION & EMERGENCY MEDICINE* 27 (Feb. 2019) (cf. p. 80).
- (Petitdemange et al., 2018) E. Petitdemange, F. Fontanili, E. Lamine, and M. Lauras. “Skill-based routing problem in Emergency Call Centers: toward an improvement of response time.” en. In: Sept. 2018, p.43–51 (cf. p. 130).
- (Petitdemange et al., 2019) E. Petitdemange, F. Fontanili, E. Lamine, M. Lauras, and U. Okongwu. “A Tool-Based Framework to Assess and Challenge the Responsiveness of Emergency Call Centers.” In: *IEEE Transactions on Engineering Management* (2019) (cf. p. 73, 79).
- (Rajagopalan et al., 2011) H. K. Rajagopalan, C. Saydam, H. Setzler, E. Sharer, et al. “Ambulance deployment and shift scheduling: An integrated approach.” In: *Journal of Service Science and Management* 4.01 (2011), p. 66 (cf. p. 44).

- (Riou et al., 2017) M. Riou, S. Ball, T. A. Williams, A. Whiteside, K. L. O'Halloran, J. Bray, G. D. Perkins, K. Smith, P. Cameron, D. M. Fatovich, M. Inoue, P. Bailey, D. Brink, and J. Finn. “[‘Tell me exactly what’s happened’: When linguistic choices affect the efficiency of emergency calls for cardiac arrest.](#)” English. In: *Resuscitation* 117 (Aug. 2017), pp. 58–65 (cf. p. 47).
- (Rojas et al., 2016) E. Rojas, J. Munoz-Gama, M. Sepúlveda, and D. Capurro. “Process mining in healthcare: A literature review.” In: *Journal of biomedical informatics* 61 (2016), pp. 224–236 (cf. p. 106).
- (Roubos et al., 2013) A. Roubos and O. Jouini. “[Call centers with hyperexponential patience modeling.](#)” In: *International Journal of Production Economics*. Meta-heuristics for manufacturing scheduling and logistics problems 141.1 (Jan. 2013), pp. 307–315 (cf. p. 87).
- (Saaty, 1961) T. L. Saaty. *Elements of queueing theory: with applications*. Vol. 34203. McGraw-Hill New York, 1961 (cf. p. 110).
- (Sacco et al., 2005) W. J. Sacco, D. M. Navin, K. E. Fiedler, R. K. Waddell II, W. B. Long, and R. F. Buckman Jr. “Precise formulation and evidence-based application of resource-constrained triage.” In: *Academic emergency medicine* 12.8 (2005), pp. 759–770 (cf. p. 36).
- (Salleh et al., 2017) S. Salleh, P. Thokala, A. Brennan, R. Hughes, and A. Booth. “[Simulation Modelling in Healthcare: An Umbrella Review of Systematic Literature Reviews.](#)” en. In: *PharmacoEconomics* 35.9 (Sept. 2017), pp. 937–949 (cf. p. 124).
- (Santé, 2011) H. A. de Santé. “Modalités de prise en charge d’un appel de demande de soins non programmés dans le cadre de la régulation médicale.” In: *HAS Paris* (2011) (cf. p. 30).
- (Sargent, 2013) R. G. Sargent. “[Verification and validation of simulation models.](#)” In: *Journal of Simulation* 7.1 (Feb. 2013), pp. 12–24 (cf. p. 88).
- (Sariyer, 2018) G. Sariyer. “[Sizing capacity levels in emergency medical services dispatch centers: Using the newsvendor approach.](#)” In: *AMERICAN JOURNAL OF EMERGENCY MEDICINE* 36.5 (May 2018), pp. 804–815 (cf. p. 42, 44).
- (Savas, 1969) E. S. Savas. “[Simulation and Cost-Effectiveness Analysis of New York’s Emergency Ambulance Service.](#)” In: *Management Science* 15.12 (Aug. 1969), B–608 (cf. p. 37, 39, 40).
- (Schmid et al., 2010) V. Schmid and K. F. Doerner. “Ambulance location and relocation problems with time-dependent travel times.” In: *European journal of operational research* 207.3 (2010), pp. 1293–1303 (cf. p. 48).
- (Setzler et al., 2009) H. Setzler, C. Saydam, and S. Park. “[EMS call volume predictions: A comparative study.](#)” In: *Computers & Operations Research* 36.6 (June 2009), pp. 1843–1851 (cf. p. 39, 43).
- (Sikka et al., 2005) N. Sikka and G. Margolis. “[Understanding diversity among prehospital care delivery systems around the world.](#)” English. In: *Emergency Medicine Clinics* 23.1 (Feb. 2005), pp. 99–114 (cf. p. 6, 15, 16).
- (Silva et al., 2010) P. M. S. Silva and L. R. Pinto. “Emergency medical systems analysis by simulation and optimization.” In: *Proceedings of the 2010 winter simulation conference*. IEEE. 2010, pp. 2422–2432 (cf. p. 35).

-
- (Smith et al., 2009) R. M. Smith and A. K. Conn. “Prehospital care Scoop and run or stay and play?” en. In: *Injury*. Current concepts in polytrauma management 40 (Nov. 2009), S23–S26 (cf. p. 6, 21, 75).
- (Snooks et al., 2011) H. Snooks, W. Y. Cheung, S. M. Gwini, I. Humphreys, A. Sánchez, and N. Siriwardena. “09 Can older people who fall be identified in the ambulance call centre to enable alternative responses or care pathways?” en. In: *Emergency Medicine Journal* 28.3 (Mar. 2011), e1–e1 (cf. p. 47).
- (Sobolev et al., 2011) B. G. Sobolev, V. Sanchez, and C. Vasilakis. “Systematic review of the use of computer simulation modeling of patient flow in surgical care.” In: *Journal of medical systems* 35.1 (2011), pp. 1–16 (cf. p. 125).
- (Sterman, 2010) J. Sterman. *Business dynamics*. Irwin/McGraw-Hill c2000., 2010 (cf. p. 124).
- (Stewart Robinson, 2004) S. Stewart Robinson. *The practice of model development and use*. 2004 (cf. p. 122).
- (Storm-Versloot et al., 2011) M. N. Storm-Versloot, D. T. Ubbink, J. Kappelhof, and J. S. Luitse. “Comparison of an informally structured triage system, the emergency severity index, and the manchester triage system to distinguish patient priority in the emergency department.” In: *Academic Emergency Medicine* 18.8 (2011), pp. 822–829 (cf. p. 142).
- (Sturges, 1926) H. A. Sturges. “The choice of a class interval.” In: *Journal of the american statistical association* 21.153 (1926), pp. 65–66 (cf. p. 113).
- (Su et al., 2003) S. Su and C.-L. Shih. “Modeling an emergency medical services system using computer simulation.” In: *International Journal of Medical Informatics* 72.1 (Dec. 2003), pp. 57–72 (cf. p. 34, 35, 39).
- (Swoveland et al., 1973) C. Swoveland, D. Uyeno, I. Vertinsky, and R. Vickson. “A simulation-based methodology for optimization of ambulance service policies.” In: *Socio-Economic Planning Sciences* 7.6 (Dec. 1973), pp. 697–703 (cf. p. 39).
- (Takacs, 1969) L. Takacs. “On Erlang’s formula.” In: *The annals of mathematical statistics* 40.1 (1969), pp. 71–78 (cf. p. 111).
- (Trudeau et al., 1989) P. Trudeau, J.-M. Rousseau, J. A. Ferland, and J. Choquette. “An Operations Research Approach For The Planning and Operation Of An Ambulance Service.” In: *INFOR: Information Systems and Operational Research* 27.1 (Jan. 1989), pp. 95–113 (cf. p. 39).
- (Tukey, 1977) J. W. Tukey. *Exploratory Data Analysis*. Anglais. 1st ed. Reading, Mass: Pearson, Jan. 1977 (cf. p. 88).
- (Van Der Aalst, 2015) W. M. P. Van Der Aalst. “Extracting Event Data from Databases to Unleash Process Mining.” en. In: *BPM - Driving Innovation in a Digital World*. Management for Professionals. Springer, Cham, 2015, pp. 105–128 (cf. p. 104).
- (Van der Aalst et al., 2004) W. Van der Aalst, T. Weijters, and L. Maruster. “Workflow mining: Discovering process models from event logs.” In: *IEEE transactions on knowledge and data engineering* 16.9 (2004), pp. 1128–1142 (cf. p. 80).

- (Van Sambeek et al., 2010) J. Van Sambeek, F. Cornelissen, P. Bakker, and J. J. Krabbendam. “Models as instruments for optimizing hospital processes: a systematic review.” In: *International journal of health care quality assurance* (2010) (cf. p. 125).
- (Viglino et al., 2017) D. Viglino, A. Vesin, S. Ruckly, X. Morelli, R. Slama, G. Debaty, V. Danel, M. Maignan, and J.-F. Timsit. “Daily volume of cases in emergency call centers: construction and validation of a predictive model.” In: *SCANDINAVIAN JOURNAL OF TRAUMA RESUSCITATION & EMERGENCY MEDICINE* 25 (Aug. 2017) (cf. p. 42).
- (Vile, 2013) J. Vile. “Time-dependent stochastic modelling for predicting demand and scheduling of emergency medical services.” en. phd. Cardiff University, 2013 (cf. p. 43).
- (Wang et al., 2007) T. Wang, A. Belaidi, A. Guinet, and B. Besombes. “Modeling and Simulation of emergency services with ARIS and Arena.” In: *International Conference on Industrial Engineering and Systems Management IESM 2007*. BeiJing, China, May 2007, p. 239 (cf. p. 49).
- (Weinlich et al., 2018) M. Weinlich, P. Kurz, M. B. Blau, F. Welcher, and S. Piatek. “Significant acceleration of emergency response using smartphone geolocation data and a worldwide emergency call support system.” In: *PLOS ONE* 13.5 (May 2018) (cf. p. 41).
- (Yersin, 2006) V. R. P. S. B. Yersin. “Accès aux soins d’urgence: organisation et utilité de la chaîne des secours.” In: *Rev Med Suisse* 2 (2006), p. 31574 (cf. p. 16).
- (Zeltyn et al., 2005) S. Zeltyn and A. Mandelbaum. “Call centers with impatient customers: Many-server asymptotics of the M/M/n+ G queue.” In: *Queueing Systems* 51.3-4 (2005), pp. 361–402 (cf. p. 111).
- (Zuzakova et al., 2012) B. B. Zuzakova and G. J. Kommer. “Optimal emergency medical service system design.” en. MA thesis. Amsterdam: VU University, May 2012 (cf. p. 85).

Résumé

SAMUFLUX : une démarche outillée de diagnostic et d'amélioration à base de doubles numériques : application aux centres d'appels d'urgence de trois SAMU

L'engorgement des structures d'urgence médicale est une réalité à laquelle les SAMU (Service d'Aide Médicale Urgente) n'échappent pas (5 à 15% d'augmentation des appels par an). De récents événements comme la pandémie de Covid-19 montrent les limites du système actuel pour faire face à des situations de crise. Le renfort en moyens humains ne peut être la seule solution face à ce constat et il devient incontournable de construire de nouveaux modèles organisationnels tout en visant une qualité de service permettant de répondre à 99% des appels entrants en moins de 60 secondes. Toutefois, ces modèles doivent prendre en compte la grande hétérogénéité des SAMU et de leur fonctionnement. À la lumière de ces constats, le travail de recherche présenté dans ce manuscrit a pour but *d'évaluer l'efficacité organisationnelle et la résilience des SAMU dans la prise en charge des flux d'appels téléphoniques d'urgence pour faire face au quotidien et aux situations de crise*. Cette évaluation nous permet de proposer et de tester de nouveaux schémas organisationnels afin de faire des préconisations adaptées aux particularités des centres d'appels d'urgence. Dans une première partie,

nous proposons une méthodologie outillée pour le diagnostic et l'amélioration des centres d'appels d'urgence. Elle se décompose en deux grands volets : l'étude des données issues des centres d'appel d'urgence puis la conception et l'utilisation d'un double numérique. Pour chaque étape de cette méthodologie nous proposons un outil associé. Dans une seconde partie, nous appliquons le premier volet de la méthodologie aux données de nos SAMU partenaires. Il s'agit d'être capable d'extraire de l'information et des connaissances des données de téléphonies ainsi que des processus métiers pour le traitement des appels d'urgence. La connaissance ainsi extraite permet de concevoir un double numérique qui soit proche du comportement réel des SAMU. Enfin dans une troisième partie, nous utilisons la matière produite précédemment pour modéliser et paramétrer un double numérique déployé sur un moteur de simulation à événements discret. Il nous permet de tester plusieurs scénarios en jouant sur différentes organisations de gestion de l'appel. Grâce à cela, nous faisons des préconisations sur les types d'organisations à adopter afin d'améliorer les performances des centres d'appel.

MOTS-CLÉS : Double numérique, Process Mining, Organisations, Centre d'Appel, Simulation à événements discrets, SAMU

Abstract

SAMUFLUX : a tool-based approach for diagnosis and improvement using digital twin : application to a cluster of three emergency call centers (SAMU)

The demand for emergency medical services has been significant and increasing over the last decade. In a constrained medico-economic context, the maintenance of operational capacities represents a strategic stake in front of the risk of congestion and insufficient accessibility for the population. Recent events such as the Covid-19 pandemic show the limits of the current system to face crisis situations. Reinforcement in human resources cannot be the only solution in front of this observation and it becomes unavoidable to build new organizational models while aiming at a quality of service allowing to answer 99% of the incoming calls in less than 60 seconds (90% in 15s and 99% in 30s MARCUS report and HAS recommendation October 2020). However, these models must take into account the great heterogeneity of EMS and their operation. In the light of these findings, the research work presented in this manuscript aims to evaluate the organizational effectiveness and resilience of EMS in managing the flow of emergency telephone calls to deal with daily life and crisis situations. This evaluation allows us to propose and test new organizational schemes in order

to make recommendations adapted to the particularities of emergency call centers. In a first part, we propose a methodology equipped for the diagnosis and improvement of emergency call centers. It can be broken down into two main parts: the study of data from emergency call centers, and then the design and use of a dual digital system. For each step of this methodology, we propose an associated tool. In a second part, we apply the first part of the methodology to our partner EMS data. The aim is to be able to extract information and knowledge from the telephony data as well as from the business processes for handling emergency calls. The knowledge thus extracted makes it possible to design a digital duplicate that is close to the real behavior of the EMS. Finally, in a third part, we use the material produced previously to model and parameterize a digital duplicate deployed on a discrete event simulation engine. It allows us to test several scenarios by playing on different call management organizations. Thanks to this, we make recommendations on the types of organizations to adopt in order to improve the performance of call centers.

KEYWORDS: Digital Twin, Process Mining, Organizations, Call Center, Discrete Event Simulation, Emergency Medical Services