



Département génie industriel et maintenance

Mémoire de fin d'étude en vue de l'obtention du diplôme

D'INGENIEUR d'État

- Filière -

Génie industriel

- Spécialité -

Management et Ingénierie de la Maintenance Industrielle

- Thème -

Optimisation de la chaîne d'approvisionnement avec le process mining

Cas d'étude : Hamoud Boualem

Réalisé par

BRAHIM MAZARI Ridha

Les membres de Jury :

BOUGCHICHE	Fazia	MAA	Présidente
BELAYADI	Djahida	MCB	Examinatrice
GHOMARI	Leila	MCA	Encadrante

Alger, le 30/06/2024.

Année universitaire 2023-2024

Dédicace

“

*À nos plus chers, martyrs des champs d'honneur...
Aux libres derrière les barreaux, fermes comme des
montagnes, et à tous ceux qui ont pu dire non aux ténèbres
et ont sacrifié leur vie pour la vérité.*

*À mes très chers parents,
sources de vie, d'amour et d'affection. Votre soutien est le
vent qui me porte vers mes rêves, et votre foi en moi est la
force qui me pousse à toujours me dépasser.*

*A mes frères bien-aimés,
Ilyas, Tarek et Mohamed, Votre être confère à l'expérience
de la vie une délicatesse gustative subtile et insaisissable.
Et à toute ma famille.*

*À vous, mes amis inspirants,
sources de tant de moments précieux partagés à l'ENP
(Surtout Bouraoui Land et G.H), au club VIC, à la
mosquée de Bouraoui et aux associations dans lesquelles
j'étais actif.*

*À tous ceux que j'ai connus et aimés,
je vous dédie ce travail.*

”

- *Ridha*

Remerciements

Tout d'abord, je remercie Allah le tout puissant de m'avoir donné le courage et la patience nécessaires à mener ce travail à son terme.

Je tiens à exprimer mes sincères remerciements à mon encadrante et ma deuxième mère, **Mme. GHOMARI Leila**, pour son soutien, ses conseils précieux et son expertise qui ont été essentiels à la réussite de ce travail.

Que les membres de jury trouvent, ici, l'expression de mes sincères remerciements pour l'honneur qu'ils me font en prenant le temps de lire et d'évaluer ce travail.

Je tiens à exprimer ma reconnaissance à l'entreprise SPA Hamoud Boualem pour m'avoir offert l'opportunité de réaliser ce projet. Un grand merci à Madame KEMOUCHE Meriem, Manager des Ressources Humaines, pour son soutien et sa confiance. Je remercie aussi Monsieur BENARBA Zakaria, mon encadrant au sein de l'entreprise, pour son accompagnement et ses conseils avisés, et ses collègues Adel, Bilel et Abderrahmane.

Je remercie également mes enseignants pour leur dévouement et leur passion pour l'enseignement, qui m'ont guidé tout au long de mon parcours, surtout Mme, KACHTOULI, Mme. MEZIANE-TANI, Mme. BOUAYAD, M. MAHDI, M. BENZAZZOUZ, M. TAIBI, Mme. BENALIA, Mme. HACHEMI, M. BOUSSOUFA, M. AMRANI.

Un grand merci à mes amis pour leur camaraderie et leurs encouragements.

Je tiens à remercier particulièrement ma famille pour leur encouragement sans faille, qui m'ont donné la force de persévérer dans les moments difficiles.

Enfin, je remercie toutes les personnes, de près ou de loin, qui ont contribué à l'aboutissement de ce projet. Merci à tous pour votre confiance et votre soutien.

Table des matières

Dédicace	I
Remerciements	II
Introduction générale	1
1 Étude bibliographique sur le Process Mining	3
1.1 Introduction	4
1.2 Analyse des traces	4
1.2.1 Objectifs de l'utilisation des Traces	5
1.2.2 Système à base de traces (SBT)	5
1.3 Principe général du process mining	9
1.4 Avantages et défis du process mining	12
1.5 Motifs courants dans les processus	14
1.6 Modélisation des processus	15
1.6.1 Réseaux de Petri (RdP)	16
1.6.2 Workflow Net	17
1.6.3 Process Tree	18
1.6.4 Business Process Model and Notation (BPMN)	19
1.7 Gestion des processus métier	20
1.7.1 Concepts clés de l'optimisation des processus	23
1.8 Propriétés du modèle miné	24
1.8.1 Alignement du modèle et des traces	25
1.8.2 Justesse	26
1.8.3 Précision	26
1.8.4 Généralisation	27
1.8.5 Simplicité	28
1.9 Techniques et algorithmes de fouille de processus	28
1.10 Conclusion	35
2 Implémentation : Étude de Cas Hamoud Boualem	36
2.1 Introduction	37
2.2 Présentation de l'organisme d'accueil	37
2.2.1 Historique de l'entreprise	37
2.2.2 Unité de Boufarik	38
2.2.3 Structure organisationnelle	38
2.2.4 Concurrents et fournisseurs	39
2.2.5 Direction Supply Chain	39

2.2.6	Identification des processus	41
2.2.7	Sélection du processus pour appliquer le Process mining	45
2.3	Analyse du processus choisi	47
2.3.1	Étapes du processus	47
2.3.2	Parties prenantes du processus	48
2.3.3	Documentation du processus	49
2.3.4	Modélisation du processus	51
2.3.5	Simulation du processus	52
2.3.6	Problèmes rencontrés dans l’approvisionnement	53
2.4	Préparation au process mining et analyse des données	53
2.4.1	Sources de données	53
2.4.2	Nettoyage et prétraitement des données	55
2.4.3	Description du journal d’événements	56
2.4.4	Logiciels utilisé pour l’analyse	57
2.5	Découverte du processus	58
2.6	Analyse de Conformité	62
2.7	Recommandations et optimisation	65
2.8	Plan de la mise en œuvre des recommandations	69
2.8.1	Étapes	69
2.8.2	Ressources nécessaires	70
2.8.3	Calendrier de mise en œuvre	71
2.8.4	Suivi et Évaluation	71
2.9	Conclusion	74
	Conclusion générale	75
	Annexes	82
	Résumé	93
	Abstract	94
	95	ملخص

Table des figures

1.1	Principe général d'un SBT [33]	8
1.2	Le pont entre la science des processus et la science des données	9
1.3	Les quatre relations principales entre deux activités	12
1.4	Choix non libre	13
1.5	Motifs élémentaires de processus	14
1.6	Exemple de boucles	15
1.7	Évolution du marquage d'un RdP	17
1.8	Exemple de Workflow-net	17
1.9	Exemple de Process Tree [10]	19
1.10	Évolution du développement de BPMN [11]	19
1.11	Exemple d'un modèle en BPMN [37]	20
1.12	Cycle de vie du BPM	21
1.13	Modèle de processus en fleur ??	27
2.1	Lignes de conditionnement de Hamoud Boualem - Boufarik	38
2.2	Organigramme de SPA Hamoud Boualem	39
2.3	Étapes de planification de la Demande	40
2.4	Étapes de planification de l'approvisionnement	41
2.5	Diagramme d'Ishikawa du processus de production	42
2.6	Modèle de processus d'approvisionnement en BPMN	52
2.7	Résumé du journal d'évènements	57
2.8	Modèle découvert avec alpha++	58
2.9	Modèle découvert avec Inductive Miner	58
2.10	Modèle découvert avec Disco	59
2.11	Boucle identifiée 1	61
2.12	Boucle identifiée 2	61
2.13	Point de décision identifié	62
2.14	Analyse de justesse	63
2.15	Modèle optimisé du processus d'approvisionnement	66
2.16	Diagramme de Gantt du plan de mise en oeuvre	71
2.17	Les changements après l'optimisation	73

Liste des tableaux

1.1	Comparaison des modèles de traces	7
2.1	Analyse fonctionnelle du processus d'approvisionnement	45
2.2	Table comparative entre les processus	46
2.3	Documentation du processus d'approvisionnement	51
2.4	Simulation du processus d'approvisionnement	52
2.5	Exemple du journal d'évènements	57
2.6	Simulation du processus optimisé	68
2.7	Comparaison des durées moyennes des étapes entre le modèle original et le modèle optimisé	72

Liste des sigles et acronymes

ABPD Automated Business Process Discovery

API Application Programming Interface

BPM Business Process Management

BPMN Business Process Model and Notation

BPCC Business Process Conformance Checking

DLUO Date Limite d'Utilisation Optimale

DSP Demand & Supply Planning

KPI Key Performance Indicators

MOQ Minimum Order Quantity

PA Performance Analysis

PM Process Mining

RdP Réseau de Petri

SCM Supply Chain Management

SBT Système à Base de Traces

SI Systèmes d'Information

SPA Société par Actions

TRS Taux de Rendement Synthétique

XES eXtensible Event Stream

Introduction générale

Dans un environnement économique en constante évolution, les entreprises modernes font face à des défis croissants pour maintenir leur compétitivité et leur succès. La gestion de la chaîne logistique (SCM) joue un rôle central dans cet effort, impactant directement la performance opérationnelle et la satisfaction des clients. Cependant, la complexité croissante des chaînes d'approvisionnement, caractérisées par des flux de matériaux et d'informations multiples et interdépendants, rend l'optimisation des processus logistiques un défi majeur.

Face à ces défis, le Process Mining se présente comme une approche innovante et prometteuse pour l'optimisation des processus logistiques. Le process mining combine des techniques de data science et de business process management pour analyser les données des systèmes d'information des entreprises, permettant de visualiser, comprendre et améliorer les flux de travail existants. Cette approche permet de découvrir des inefficacités, des goulots d'étranglement et des opportunités d'optimisation qui pourraient autrement rester cachés dans la complexité des processus.

Ce projet de fin d'études s'inscrit dans cette problématique en explorant le potentiel du process mining pour améliorer l'efficacité et réduire les inefficacités au sein d'une chaîne logistique réelle.

L'objectif principal de ce travail est de démontrer l'applicabilité du process mining pour l'optimisation des processus logistiques en s'appuyant sur un cas d'étude concret : l'entreprise SPA Hamoud Boualem, entreprise Algérienne leader dans le secteur des boissons gazeuses, avec un focus particulier sur le site de Boufarik, d'une importance stratégique dans la chaîne d'approvisionnement de l'entreprise. En effet, cette dernière, assure la production et la distribution de ses produits sur l'ensemble du territoire national. L'analyse des processus logistiques de cette entreprise offre une opportunité unique de mettre en œuvre le process mining dans un contexte réel et de démontrer ses avantages concrets.

Le choix du process mining comme outil d'optimisation des processus logistiques se justifie par plusieurs raisons :

- **Son caractère innovant** : Le process mining combine des techniques de pointe de data science et d'intelligence artificielle pour offrir une approche nouvelle et prometteuse à l'optimisation des processus.
- **Sa capacité à fournir des solutions concrètes** : Cette technique permet d'identifier des inefficacités et de proposer des recommandations d'amélioration basées sur des données réelles, permettant aux entreprises de prendre des décisions éclairées.

- **Son applicabilité à différents secteurs d'activité :** Le process mining peut être appliqué à une large gamme de secteurs, y compris la fabrication, la logistique, les services financiers et la santé.

La méthodologie adoptée comprend la préparation des données, l'application d'algorithmes de process mining pour modéliser les flux de travail existants, et l'analyse des résultats pour identifier les inefficacités et proposer des recommandations d'amélioration. La contribution principale de ce mémoire réside dans la fourniture d'un cadre méthodologique pour l'application du process mining dans un contexte de chaîne logistique.

Le mémoire est structuré en deux chapitres principaux.

1. Le premier chapitre présente une revue de littérature exhaustive sur le process mining, couvrant ses principes fondamentaux, ses avantages et défis, ainsi que les techniques et algorithmes utilisés.
2. Le deuxième chapitre se concentre sur l'implémentation pratique du process mining dans le cas de SPA Hamoud Boualem. À travers une étude de cas détaillée, nous démontrons les avantages du process mining pour l'analyse et l'optimisation des processus logistiques.

Chapitre 1

Étude bibliographique sur le Process Mining

1.1 Introduction

Au cœur des organisations, les processus métier déterminent le succès. Leur complexité croissante rend leur compréhension et leur optimisation un défi. La fouille de processus, ou "process mining", est une discipline émergente qui permet d'analyser et d'optimiser ses processus métier à partir des traces d'exécution enregistrées dans les systèmes d'information (SI). En exploitant des techniques avancées d'apprentissage automatique et d'analyse de données, la fouille de processus offre des insights précieux pour améliorer l'efficacité organisationnelle et promouvoir l'innovation.

Ce chapitre explore en profondeur le domaine de la fouille de processus. Nous commencerons par analyser les traces d'exécution et leur importance, puis nous aborderons les différentes techniques de fouille de processus ainsi que leurs applications concrètes. Enfin, nous discuterons des défis, des avantages et des outils disponibles pour la fouille de processus, offrant ainsi une vue d'ensemble complète de cette discipline révolutionnaire.

1.2 Analyse des traces

Cette section débute par une définition des traces, puis elle présente divers concepts associés, comme les traces brutes, les sources de traçage et les systèmes basés sur les traces.

- **Trace :**

La notion de trace, comme le soulignent Settouti et al. (2006) [44], revêt de multiples significations selon les contextes. De manière générale, elle désigne le vestige d'une interaction entre deux entités, qu'elles soient tangibles ou non, résultant d'un contact physique ou non. Sa forme et sa nature varient en fonction de son environnement.

Dans le contexte de notre étude, on la définit comme séquence ordonnée d'événements correspondant à une exécution particulière d'un processus. Les traces d'exécution sont généralement stockées dans des fichiers journaux ou des bases de données. [8]

- **Trace numérique :**

La trace numérique, comme définie par Settouti et al. (2006) [44], représente l'interaction entre une entité et un système informatique, enregistrée sous forme numérique. Par la suite, nous utiliserons le terme "trace" pour désigner cette trace numérique.

- **Trace brute ou première :**

Une trace brute est générée dans le système à la suite d'une activité ou d'une interaction, représentant ainsi la trace dans sa forme initiale sans aucune interprétation ni transformation [32]. Elle peut adopter un format spécifié par le modèle de collecte des traces du système ou du composant considéré, et est parfois appelée trace première. Les logs et les journaux d'événements systèmes en sont des exemples courants, collectés généralement à partir d'une source de traçage.

- **Source de traçage :**

Une source de traces, comme les fichiers journaux d'applications ou les flux de données, génère des informations sur les activités d'un système, conformément à un modèle de collecte défini [17]. Ces traces sont essentielles pour diverses applications telles que l'analyse du comportement du système, le débogage, la sécurité, l'apprentissage automatique et l'analyse de données. Le modèle de collecte de traces dicte la manière dont celles-ci sont capturées, stockées et formatées, tandis qu'un système à base de traces regroupe les outils nécessaires à leur gestion et leur analyse.

1.2.1 Objectifs de l'utilisation des Traces

Les traces, vestiges d'événements et d'interactions au sein d'un système, révèlent une mine d'informations pour l'analyse et l'optimisation. Elles permettent de :

- **Détection et correction des anomalies :** En repérant des erreurs, ralentissements ou comportements inhabituels, l'analyse des traces permet de corriger les défaillances sous-jacentes, renforçant ainsi la stabilité et l'expérience utilisateur.
- **Adaptation personnalisée de l'expérience utilisateur :** En examinant les expériences passées des utilisateurs, les traces facilitent l'adaptation du système selon leurs préférences, offrant des recommandations personnalisées et une présentation ajustée.
- **Optimisation de l'exécution en fonction du comportement utilisateur :** L'analyse des traces individuelles éclaire sur les habitudes et objectifs des utilisateurs, permettant une optimisation anticipative du système en prévoyant leurs actions futures.
- **Identification des modèles de comportement et amélioration des processus métiers :** L'examen des schémas d'activités révèle des modèles de comportement, identifiant ainsi des opportunités d'optimisation des workflows et des processus métiers.
- **Analyse statistique et comptable des activités et interactions :** Les traces fournissent une base pour l'analyse statistique, permettant aux gestionnaires de suivre les tendances, d'évaluer l'efficacité des fonctionnalités et de prendre des décisions informées pour l'optimisation continue.

1.2.2 Système à base de traces (SBT)

Notre approche s'appuie sur l'exploitation des traces générées par les systèmes informatiques, comme le propose Settouti et al. (2009) [43]. Pour ce faire, nous avons recours aux Systèmes à Base de Traces (SBT), qui offrent une solution générique à la modélisation et à la manipulation des traces.

Un modèle de trace est défini comme une représentation formelle d'une trace, incluant ses propriétés et attributs tels que la date, l'heure, l'identifiant de l'utilisateur et l'action

réalisée [30, 51]. Son objectif principal est d'expliciter de manière abstraite les éléments constitutifs de la trace. Par exemple, le format standardisé Common LogFile Format est utilisé dans les fichiers de journalisation du serveur Apache pour définir la sémantique de la trace.

Une trace modélisée, ou m-trace, est associée à un modèle de traces [42], définissant les éléments observés et les relations qui la composent. Bien qu'un modèle de traces définisse les données à tracer via ses attributs, il ne spécifie pas les données réellement collectées ; il impose plutôt des contraintes aux traces brutes. Ainsi, les systèmes basés sur ce modèle seront conformes aux exigences définies par le modèle de traces.

Application et choix du modèle de traces

Pour transformer des traces brutes en m-traces conformes au modèle choisi, plusieurs outils peuvent être utilisés. Parmi les options possibles figurent :

- **Extracto-chargeurs de type ETL** : Ces outils permettent d'extraire les données des sources brutes, de les transformer selon le modèle de traces choisi et de les charger dans un environnement d'analyse [26]. (Extract-Transform-Load est un module logiciel permettant d'effectuer des synchronisations d'information d'une source de données vers une autre)
- **API de modélisation spécifiques au modèle choisi** : Des API comme XESame pour le modèle d'événements permettent de manipuler et de modéliser les traces en suivant les règles du modèle choisi [46]. (Application Programming Interface est une bibliothèque de fonctionnalités par laquelle un logiciel offre des services à d'autres logiciels)

Avant de modéliser les traces, il est crucial de définir clairement les besoins du cas d'utilisation. Par exemple, dans le cadre de la fouille de processus, il est essentiel d'observer au moins l'activité effectuée et l'exécution à laquelle elle appartient [6].

À l'heure actuelle, deux modèles de traces principaux se distinguent pour la fouille de processus :

- **Modèle d'événements** : Ce modèle représente les traces comme une séquence d'événements ordonnés dans le temps, avec des attributs tels que l'identifiant de l'activité et l'horodatage [4].
- **Modèle de traces numériques** : Ce modèle offre une approche plus générale, permettant de représenter divers types de traces et d'associer des attributs définis par l'utilisateur [15].

Modèle de traces	Description	Avantages	Inconvénients
Modèle d'événements	Séquence d'événements ordonnés dans le temps	Simplicité, Compatibilité logicielle, Contrainte des traces, Organisation des activités	Manque de flexibilité pour représenter des traces complexes
Modèle de traces numériques	Représentation flexible de divers types de traces	Flexibilité, Adaptabilité, Contrainte des traces	Nécessite une définition claire des attributs, Non adapté à l'organisation des activités

TAB. 1.1 : Comparaison des modèles de traces

Comme le résume la table 1.1, le modèle de traces numériques peut représenter les traces des événements comme l'exécution d'une activité. Cependant, les algorithmes, logiciels et méthodes qui utiliseront ces traces doivent être compatibles avec le modèle choisi, ce qui n'est pas le cas ici. La plupart des logiciels et algorithmes se basent plutôt sur le modèle d'événements pour récupérer une représentation de l'organisation d'un processus.

Structure et stockage des traces

Que ce soit la trace brute (trace première) ou la trace transformée (m-trace), chacune respecte un schéma de données prédéfini par le modèle de traces utilisé.

- **Trace première** : Le schéma est dicté par le modèle de traces de la collecte, définissant les informations capturées lors de la collecte des données.
- **M-trace** : Le schéma est défini par le modèle des traces issues de la transformation, spécifiant les informations résultant de la transformation de la trace brute.

La trace résultante est souvent stockée sous forme d'enregistrements dans une table unique, avec des champs représentant les différents états des informations observées ou transformées, selon le modèle de traces utilisé. Le stockage peut se faire dans des fichiers texte pour les traces brutes, où les données sont séparées par un caractère spécifique. Des normes telles que XES (eXtensible Event Stream) basées sur XML émergent pour faciliter la structuration et la portabilité des traces, améliorant ainsi la communication entre les outils dédiés à la fouille de processus.

Principe d'un système à base de traces modélisées

Un système à base de traces modélisées (SBT) est un système qui utilise des traces modélisées pour son fonctionnement. Ces traces, qui représentent des séquences d'évé-

nements ordonnés, sont gérées, transformées et visualisées à différents niveaux selon les besoins du système [42]. La figure 1.1 illustre le principe général d'un SBT.

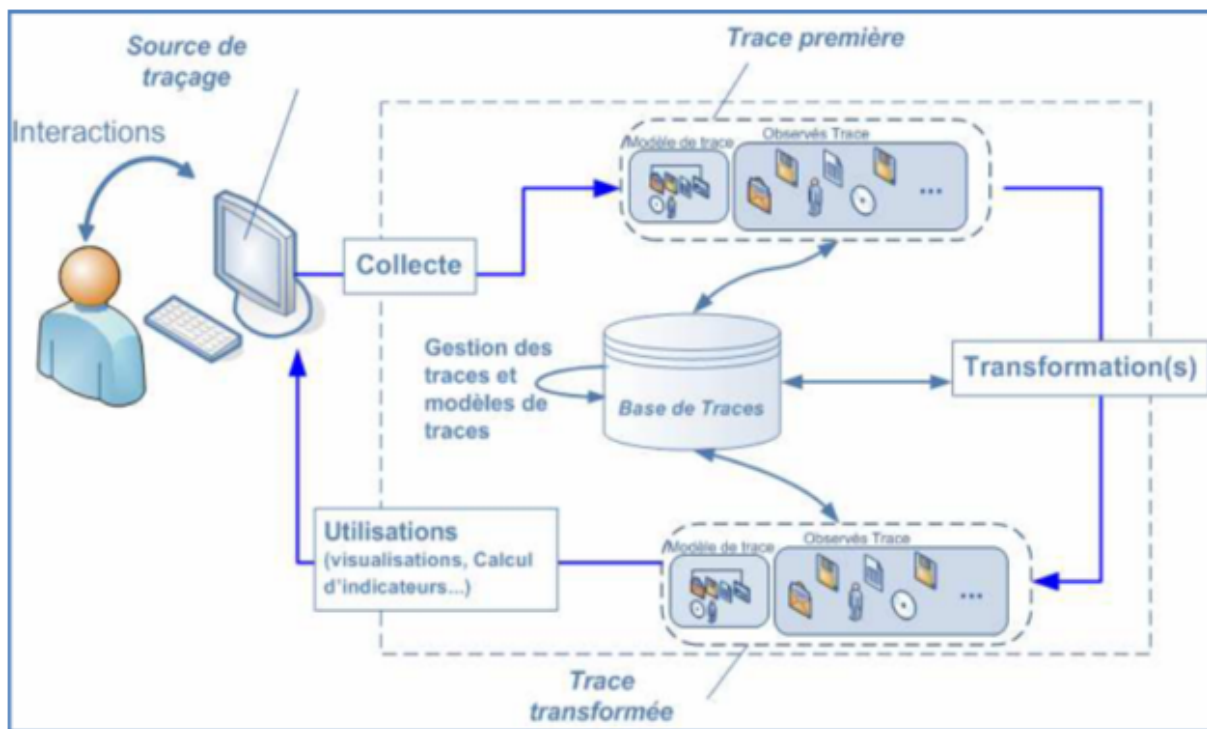


FIG. 1.1 : Principe général d'un SBT [33]

Les systèmes à base de traces, bien qu'adoptant des approches variées, partagent généralement des fonctionnalités principales réparties en trois phases distinctes.

La première phase, la collecte de traces, implique la définition des sources de traçage dans l'application interactive, la spécification du modèle de traces à utiliser, et enfin, la récupération des traces elles-mêmes.

La deuxième phase, la transformation des traces, comprend la définition des modèles de transformation et l'application de ces modèles aux traces obtenues lors de la phase de collecte.

Enfin, la troisième phase, l'analyse de traces, est chargée d'analyser et d'exploiter les traces en fonction des objectifs d'utilisation définis au préalable.

Chaque phase est généralement assurée par un composant distinct du SBT :

- **Composant de collecte de traces** : Responsable de la définition des sources de traçage, du modèle de traces et de la récupération des traces.
- **Composant de transformation de traces** : Responsable de la définition des modèles de transformation et de la transformation des traces.
- **Composant d'analyse de traces** : Responsable de l'analyse et de l'exploitation des traces selon les buts d'utilisation définis.

Dans le cadre de la recherche d'un compagnon pour l'exécution de processus métiers, nous nous concentrons sur les systèmes à base de traces utilisés pour l'apprentissage. L'objectif est d'obtenir une vue d'ensemble des systèmes existants et de les comparer en fonction de nos besoins spécifiques pour la fouille de processus (Annexe A).

1.3 Principe général du process mining

C'est à l'Université de Technologie d'Eindhoven que la fouille de processus a vu le jour, avec l'ambition d'exploiter les traces d'exécution pour la modélisation des processus [4]. Cette méthodologie se définit comme l'art de "découvrir une définition de processus à partir d'un ensemble d'exécutions de processus réels, en s'appuyant sur les journaux d'activités générés par les systèmes d'information impliqués dans ces processus" [1].

Notre recherche s'inscrit dans une démarche d'amélioration continue des processus métiers. En exploitant les données des systèmes d'information, le process mining nous permet de comprendre les pratiques des utilisateurs en identifiant les corrélations entre les activités et en analysant les écarts par rapport aux modèles attendus. Ainsi, nous proposons des suggestions concrètes pour optimiser les processus, réduire les erreurs et atteindre efficacement les objectifs.

Comme illustré par la figure 1.2, la fouille de processus établit un pont entre la science des processus et la fouille de données [1, 38, 13].

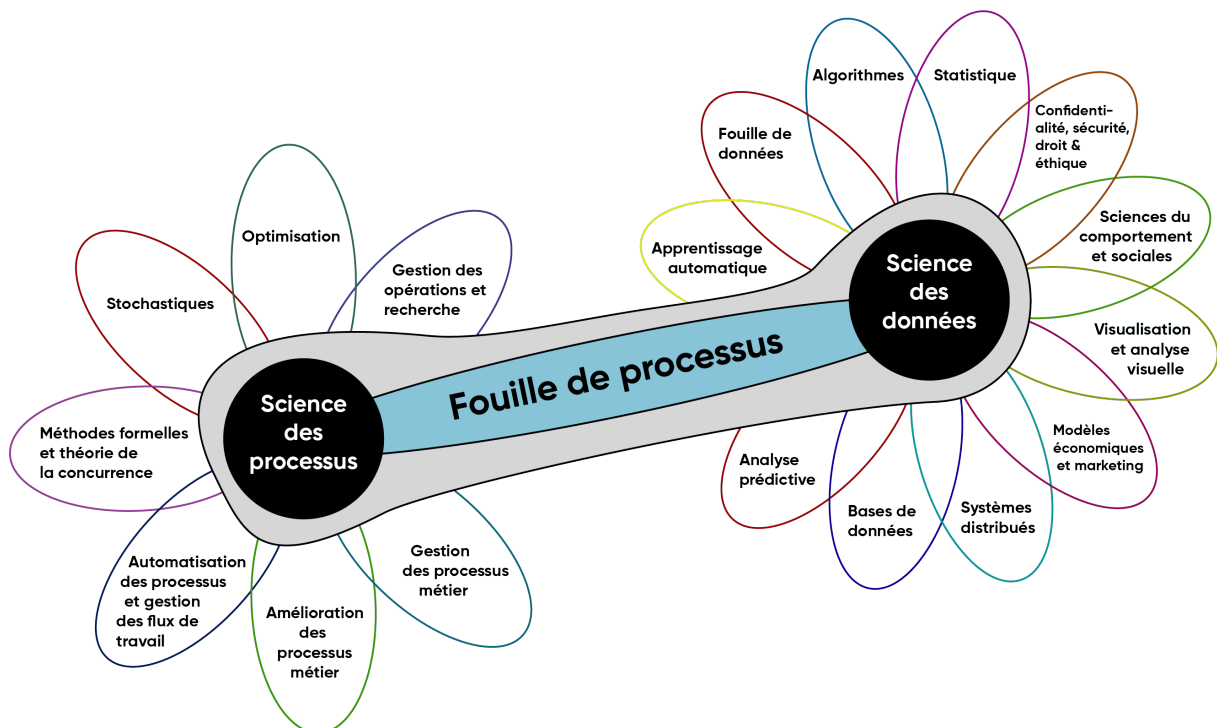


FIG. 1.2 : Le pont entre la science des processus et la science des données

La fouille de processus vise à reconstruire le déroulement réel des processus en analysant les traces d'exécution, établissant ainsi un lien crucial entre les processus réels

et les modèles théoriques. Cette comparaison permet de détecter d'éventuels écarts ou inefficacités entre la conception initiale et l'exécution réelle des processus.

Avant de commencer, on doit passer par quelques définitions nécessaires dans le contexte de fouille de processus, comme l'événement, le modèle de processus et le journal d'événement.

- **Processus métier** : Un processus métier est une séquence d'activités structurées et organisées visant à atteindre un objectif précis et mesurable au sein d'une organisation. Il s'agit d'une série d'étapes cohérentes et interdépendantes qui transforment une entrée en une sortie définie, apportant une valeur ajoutée à l'entreprise ou à ses clients.
- **Événement** : Une unité de base d'une trace d'exécution, représentant une action ou une décision spécifique dans le processus. Les événements sont caractérisés par un horodatage, un identifiant d'activité et d'autres attributs pertinents [31].
- **Modèle de processus** : Une représentation formelle du processus métier, décrivant les différentes étapes, les choix et les relations entre les activités. Les modèles de processus peuvent être représentés graphiquement ou sous forme de règles métier [1].
- **Journal d'événement (Event Log)** : Une collection structurée d'enregistrements représentant les événements qui se sont produits au sein d'un système ou d'un processus métier. Chaque enregistrement d'événement contient généralement des informations telles que l'horodatage, l'identifiant de l'activité, les attributs de l'activité et d'autres détails pertinents. Les journaux d'événements constituent la source de données principale pour le Process Mining [1].

Dans le cadre de la fouille de processus, l'analyse des logs est cruciale pour recueillir des données pertinentes. Ces logs, expliqués dans la section sur l'analyse des traces, représentent des traces brutes. Pour les analyser efficacement, il faut les transformer en traces modélisées, en identifiant les informations essentielles telles que les activités. Une fois le modèle de traces établi, le choix de l'algorithme de fouille de processus approprié est crucial pour obtenir le modèle final du processus.

Les traces modélisées suivant le modèle d'événement peuvent être représentées par le format XES. Après filtrage et transformation des traces brutes, les traces modélisées se présentent formellement comme un ensemble d'exécutions de processus (ensemble de traces d'événements). Chaque élément de cet ensemble est une séquence d'activités, aussi appelée "cas".

Par exemple, dans l'ensemble de traces d'événement $E = [\langle a, c, b \rangle^4, \langle c, a \rangle^2], \langle a, c, b \rangle$ et $\langle c, a \rangle$ sont des représentations d'exécutions sous forme de séquences d'activités effectuant successivement les activités a, c, b pour la première séquence et c, a pour la seconde. L'exposant indique le nombre d'exécutions d'une séquence dans les traces. Ainsi, dans l'ensemble de traces E , il y a quatre exécutions de la séquence d'activités $\langle a, c, b \rangle$ et deux exécutions de $\langle c, a \rangle$.

La fouille de processus vise à identifier les liens de causalité entre deux activités dans un ensemble de traces [6]. Considérons un ensemble de traces $E = \{e_1, e_2, e_3\}$ représentant l'exécution d'un processus de commande en ligne. Chaque trace représente une commande unique passée par un client. Les activités possibles dans ce processus sont :

- **Commande (C)** : Représente la création d'une nouvelle commande.
- **Paiement (P)** : Représente le paiement de la commande.
- **Expédition (E)** : Représente l'expédition de la commande au client.
- **Livraison (L)** : Représente la livraison de la commande au client.

Les traces de l'ensemble E sont les suivantes :

$$e_1 = \langle \text{Commande, Paiement, Expédition, Livraison} \rangle$$

$$e_2 = \langle \text{Commande, Paiement, Livraison, Expédition} \rangle$$

$$e_3 = \langle \text{Commande, Expédition, Paiement, Livraison} \rangle$$

Pour analyser les traces d'exécution, la fouille de processus se base sur la définition de quatre relations principales entre deux activités quelconques, notées a_1 et a_2 :

1. **Succession directe (direct succession)** : Notée $a_1 > a_2$, elle indique que l'activité a_1 est immédiatement suivie de l'activité a_2 dans au moins une séquence d'exécution. Par exemple, dans la séquence e_1 , on a Commande $>$ Paiement.
2. **Séquence (causality)** : Notée $a_1 \rightarrow a_2$, elle établit une relation de causalité entre les deux activités. Cela signifie que l'activité a_1 précède toujours l'activité a_2 dans toutes les séquences d'exécution. Par exemple, dans l'ensemble E , on a Commande \rightarrow Livraison, car la livraison ne peut avoir lieu qu'après la commande.
3. **Parallèle** : Notée $a_1 \parallel a_2$, elle indique que les activités a_1 et a_2 peuvent s'exécuter en parallèle, c'est-à-dire qu'il n'y a pas d'ordre strict entre elles. Dans l'exemple E , il n'y a pas de relation de parallélisme entre les activités.
4. **Choix (choice)** : Noté $a_1 \# a_2$, il signifie qu'il existe un choix exclusif entre les activités a_1 et a_2 . Cela implique que l'une ou l'autre des activités sera exécutée, mais pas les deux simultanément. Dans l'exemple E , il n'y a pas de relation de choix entre les activités.

Ainsi, la compréhension des relations entre les activités est essentielle pour analyser le déroulement d'un processus et identifier les points d'amélioration potentiels. La fouille de processus, en s'appuyant sur ces relations, permet d'extraire des connaissances précieuses des traces d'exécution et de contribuer à l'optimisation des processus métiers.

La figure 1.3 illustre les quatre relations principales qui peuvent s'établir entre deux activités :

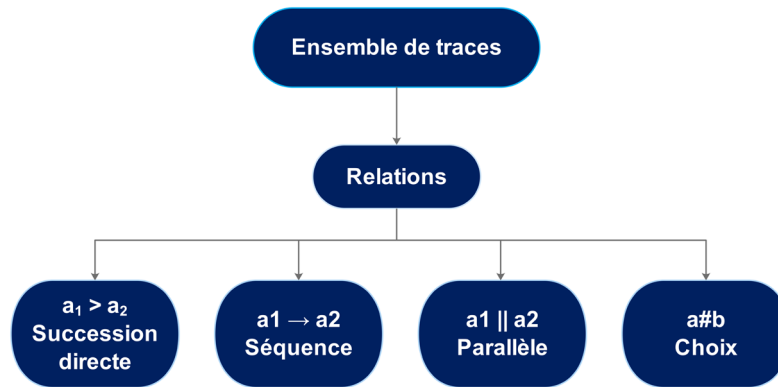


FIG. 1.3 : Les quatre relations principales entre deux activités

1.4 Avantages et défis du process mining

Le Process Mining offre de nombreux avantages aux organisations qui l'implémentent, notamment :

- **Visibilité accrue des processus** : La fouille de processus offre une vue d'ensemble des processus, de l'initiation à la terminaison, révélant des détails souvent invisibles. Cette visibilité aide à identifier les goulots d'étranglement, prioriser les actions d'amélioration et simplifier les workflows, optimisant ainsi l'efficacité opérationnelle.
- **Conformité et gestion des risques renforcées** : La fouille de processus renforce la conformité et la gestion des risques en assurant le respect des réglementations, réduisant ainsi les risques d'audits et de sanctions. Elle permet d'identifier les anomalies et écarts par rapport aux procédures, aidant à prévenir fraudes, erreurs et pertes financières.
- **Amélioration continue des processus** : L'amélioration continue des processus est un avantage majeur de la fouille de processus. Elle compare les performances réelles aux objectifs fixés, identifiant les domaines à améliorer. Les décisions d'optimisation, basées sur des données factuelles, garantissent des initiatives ciblées. En se concentrant sur les domaines les plus impactants, les organisations maximisent leur retour sur investissement.
- **Réduction des coûts et gain d'efficacité** : La fouille de processus aide à réduire les coûts et à améliorer l'efficacité. En identifiant les tâches redondantes et les temps d'attente inutiles, les organisations peuvent réduire les coûts substantiellement. La productivité augmente grâce à l'optimisation des workflows et à la réduction des erreurs. De plus, une meilleure allocation des ressources est possible en identifiant les goulots d'étranglement et en répartissant le travail plus efficacement.
- **Amélioration de la satisfaction client** : Enfin, la fouille de processus améliore la satisfaction client en optimisant les workflows et réduisant les erreurs, offrant

une expérience plus fluide. En résolvant rapidement les problèmes récurrents, les entreprises augmentent la fidélisation et développent une clientèle plus satisfaite.

Bien qu'offrant de nombreux avantages, la fouille de processus présente également divers défis qu'il convient de considérer pour en maximiser l'efficacité et la précision. Ces défis se manifestent à plusieurs niveaux, allant de la qualité des données à la complexité des modèles et à l'intégration avec les systèmes existants.

- **Qualité des données** : La qualité des données est cruciale pour la fouille de processus. Les journaux d'événements peuvent être incomplets ou erronés, affectant la précision des analyses [1]. Le manque de standardisation des formats de données entre systèmes complique l'extraction et l'intégration, nécessitant des efforts supplémentaires pour normaliser les informations [1].
- **Sources de données hétérogènes** : Une problématique majeure est la récupération d'informations de systèmes utilisant des bases de données hétérogènes. Fusionner ces données pour une analyse de qualité nécessite souvent un modèle pivot normalisé, ce qui peut être complexe et coûteux [1].
- **Bruit dans les données** : Les algorithmes de fouille de processus supposent que les traces sont correctes. Toutefois, le bruit – erreurs ou anomalies dans les données – peut fausser les modélisations. Les algorithmes doivent donc être robustes au bruit, distinguant chemins normaux et exceptions, surtout dans des contextes de sécurité où les événements rares sont cruciaux [1].
- **Traces incomplètes** : Dans des systèmes complexes, il est improbable que toutes les traces capturent tous les processus possibles. La complétude des traces est souvent limitée, rendant difficile l'obtention d'un modèle précis et généralisable [1].
- **Boucles et choix non-libres** : Les boucles et les choix non-libres posent des défis particuliers. Les boucles complexes peuvent créer des ambiguïtés entre activités dupliquées et boucles élémentaires [34]. Les choix non-libres nécessitent une modélisation sophistiquée car ils représentent des comportements non locaux et contrôlés, souvent difficiles à capturer avec précision [3]. Par exemple, dans la figure 1.4 qui représente un choix non libre très simple, il est possible de voir un choix au niveau de la place C mais t_3 ne peut être effectué que si t_1 a été effectué avant et pareillement avec t_3' et t_1' .

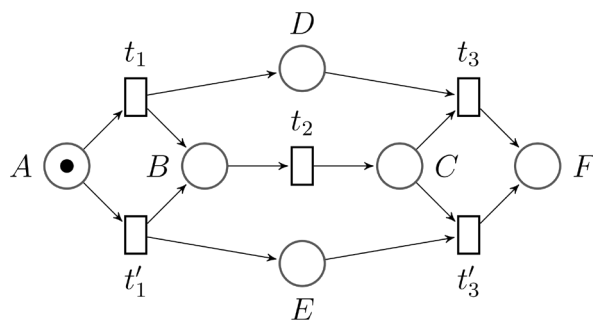


FIG. 1.4 : Choix non libre

- **Coût et investissement** : Les logiciels de fouille de processus, notamment les solutions d'entreprise, peuvent être coûteux. Leur utilisation efficace peut exiger une formation spécialisée ou le recours à des consultants experts, entraînant des coûts supplémentaires [1].
- **Facteurs humains et organisationnels** : L'adoption de la fouille de processus peut être freinée par la résistance au changement des employés et des managers. Son efficacité dépend fortement d'une culture organisationnelle favorable à l'utilisation des données et de l'implication active des parties prenantes concernées [1].

1.5 Motifs courants dans les processus

La modélisation et la conception de processus métier ont permis de définir des motifs de workflow, des structures récurrentes que l'on retrouve dans les processus. Ces motifs facilitent la compréhension et la communication des processus, et peuvent être utiles pour l'analyse et l'amélioration des processus métier.

Russell et al. (2006) [41] ont défini un cadre conceptuel pour la conception de processus, identifiant 20 motifs de workflow orientés vers le contrôle des processus. Parmi ces motifs, on retrouve les plus courants (illustrés dans la Figure 1.5) :

1. **Séquence** : Enchaînement de plusieurs transitions les unes à la suite des autres.
2. **Départ en parallèle (AND-split)** : Début de franchissement de transitions en parallèle.
3. **Synchronisation (AND-join)** : Fin des franchissements de transitions en parallèle.
4. **Choix exclusif (XOR-split)** : Obligation de choisir entre plusieurs activités possibles à franchir (une seule peut être choisie).
5. **Fusion (XOR-join)** : Fin d'un choix exclusif.

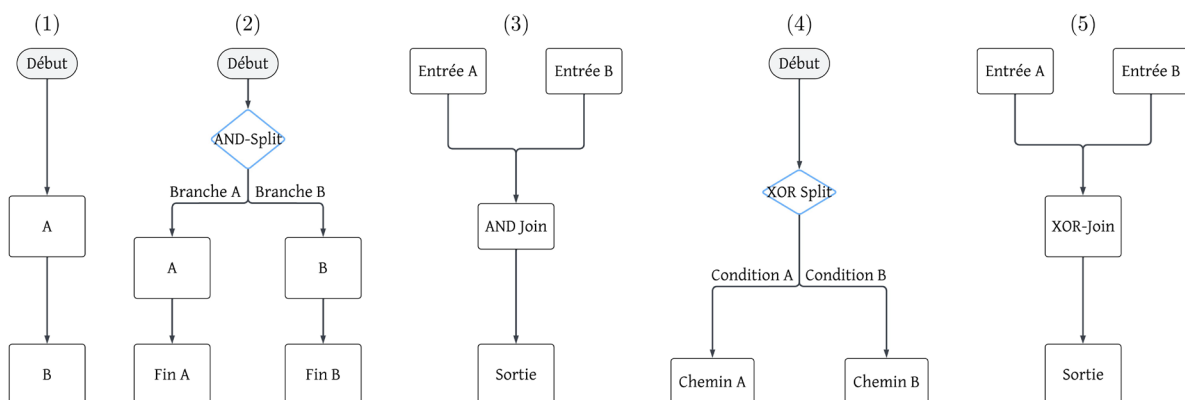


FIG. 1.5 : Motifs élémentaires de processus

En plus des motifs de base, il existe des structures plus complexes que ceux derniers comme les boucles (illustrés dans la Figure 1.6) :

6. **Boucles courtes (short loop)** : Une activité unique utilisable zéro ou plusieurs fois.
7. **Boucles mal formées (arbitrary cycles ou unstructured loop)** : Boucles apparaissant dans un cycle long.

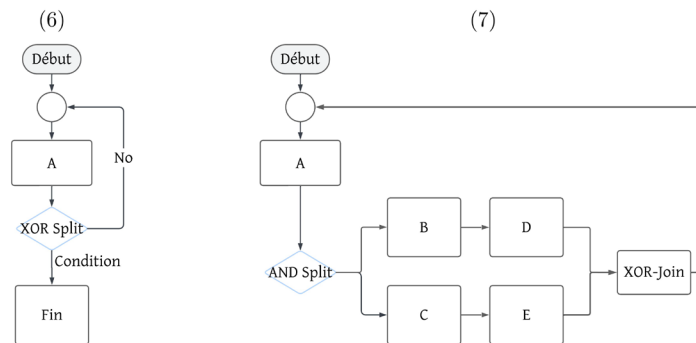


FIG. 1.6 : Exemple de boucles

La gestion des boucles, en particulier les boucles mal formées, pose un défi pour la modélisation et l'analyse des processus. La plupart des algorithmes de fouille de processus peuvent traiter les cas simples de boucles, mais des algorithmes plus avancés sont nécessaires pour découvrir les boucles plus complexes.

1.6 Modélisation des processus

La fouille de processus utilise différents modèles de processus pour représenter et analyser les exécutions de processus. Cette section présente les formalismes les plus courants pour modéliser un processus métier :

- **Réseaux de Petri** : Un modèle général permettant de représenter des processus sous forme de transitions et de places.
- **Workflow Nets et Process Tree** : Des restrictions des réseaux de Petri adaptées à la modélisation des processus métier.
- **Réseaux de causalité (Causal Nets)** : Un modèle utilisé pour décrire les relations de causalité entre les activités d'un processus.
- **BPMN (Business Process Model and Notation)** : Un modèle standardisé pour la représentation des processus métier.

Les algorithmes de fouille de processus n'utilisent pas tous le même modèle. Le choix du modèle dépend des objectifs de l'analyse et des caractéristiques du processus étudié. Ces modèles fournissent des outils pour représenter les processus et des outils mathématiques pour analyser leur comportement.

1.6.1 Réseaux de Petri (RdP)

Les réseaux de Petri, introduits par Carl Adam Petri en 1962, sont un outil mathématique et graphique puissant pour modéliser et analyser le comportement des systèmes discrets [36]. Ils sont particulièrement utiles pour représenter des systèmes parallèles, asynchrones et distribués, avec des ressources partagées et des synchronisations complexes. Ils sont utilisés dans de nombreux domaines : l'informatique, l'automatique, la logistique, la biologie, etc.

Définition :

Un réseau de Petri est un graphe biparti orienté composé de deux types de nœuds :

- **Places** : Elles représentent l'état du système et sont représentées par des cercles.
- **Transitions** : Elles représentent les événements ou actions qui modifient l'état du système et sont représentées par des rectangles.

Les places et les transitions sont reliées par des arcs orientés. Un arc reliant une place à une transition est dit "entrant", tandis qu'un arc reliant une transition à une place est dit "sortant".

Des jetons, ou marques, sont associés aux places pour indiquer leur état. Le nombre de jetons dans une place représente la quantité d'une entité abstraite, comme des pièces, des messages ou des ressources.

Réseau de Petri marqué : Un réseau de Petri marqué est le couple $N = (R, s)$ où :

- R est un réseau de Petri défini par le quadruplet $P, T, Pre, Post$ avec :
 - P un ensemble fini de places.
 - T un ensemble fini de transitions.
 - $P \cap T = \emptyset$.
 - $F \subseteq (PT) \cup (TP)$ est l'ensemble des arcs reliant les places et les transitions. Pre est la matrice d'incidence avant, décrivant l'ensemble des arcs entrants des transitions, et $Post$ est la matrice d'incidence arrière, définissant l'ensemble des arcs sortants des transitions.
- s est le marquage, une application associant à chaque place du réseau un ensemble de jetons. Ainsi, chaque place peut contenir un certain nombre de jetons, pouvant être nul. On note s_0 le marquage initial du réseau de Petri.

Les jetons dans un réseau de Petri se déplacent en respectant des règles d'évolution. Une transition est franchissable si chacune de ses places d'entrée contient au moins un jeton ($\forall p \in P, s(p) \geq Pre(p, t)$). Le franchissement d'une transition retire un jeton de chaque place d'entrée et en ajoute un à chaque place de sortie. Par exemple, le franchissement de la transition t_1 illustré dans la figure 1.7 consomme un jeton dans les places X et Q et en produit un dans la place Y. Ensuite, la transition t_2 consomme un jeton de Y et en produit un dans Z et Q. Le franchissement d'une transition correspond à l'occurrence d'un événement et une séquence de franchissements décrit l'exécution d'un processus.

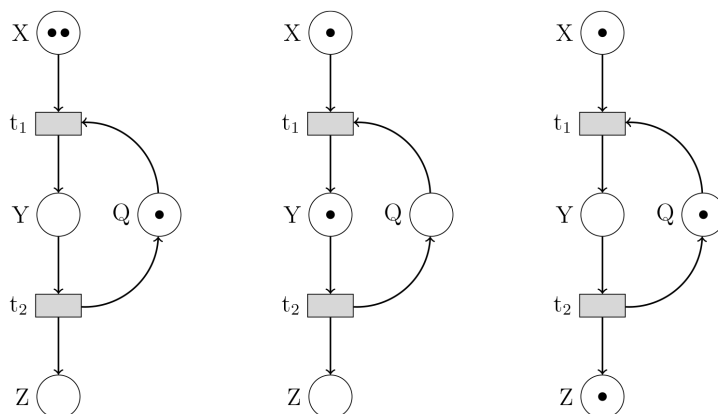


FIG. 1.7 : Évolution du marquage d'un RdP

Les réseaux de Petri décrivent un système de manière intentionnelle. Un graphe, où les sommets représentent tous les marquages possibles et les arcs les événements de transition, est appelé graphe des marquages accessibles. En traçant ce graphe, on peut déduire les propriétés du réseau de Petri, comme garantir qu'une place contient un nombre maximal de jetons ou qu'un état reste toujours accessible si le graphe est fortement connexe.

1.6.2 Workflow Net

Les Workflow Nets représentent un sous-ensemble spécifique des réseaux de Petri marqués. Ils sont caractérisés par une unique place marquée au début du processus et une unique place pour la terminaison du processus (voir Figure 1.8). En outre, tous les nœuds du réseau sont connectés par un chemin allant de la place initiale à la place finale [5].

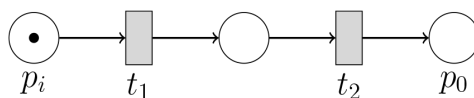


FIG. 1.8 : Exemple de Workflow-net

Formellement, un réseau de Petri R est un Workflow Net si et seulement si :

1. **Initialisation** : Il existe une et une seule place initiale, notée p_i , dans l'ensemble des places P : $\exists p_i \in P : \forall t \in T, \nexists Post(t, p_i)$.

2. **Terminaison** : Il existe une et une seule place de terminaison, notée p_0 , dans l'ensemble des places $P : \exists p_0 \in P : \forall t \in T, \nexists \text{Pre}(p_0, t)$.
3. **Connexité** : On ajoute une transition particulière \bar{t} reliant p_0 à p_i et on obtient ainsi $\bar{R} = (P, T \cup \{\bar{t}\}, F \cup \{(p_0, t), (t, p_i)\})$ fortement connecté.

Dans les travaux de l'Université de Technologie d'Eindhoven [7, 18], la notion de "soundness" est définie pour les Workflow Nets. Un Workflow Net est considéré comme "sound" si les conditions suivantes sont remplies :

1. **Terminaison garantie** : Pour tout marquage accessible depuis la place initiale, il existe une séquence de transitions qui permet d'atteindre le marquage final. Formellement, cela s'exprime par : $\forall s([i] \rightarrow s) \Rightarrow (s \rightarrow [o])$, où $[i]$ représente le marquage initial (contenant uniquement la place initiale p_i) et $[o]$ le marquage final (contenant uniquement la place finale p_0).
2. **Unicité du marquage final** : Le marquage final est le seul état accessible depuis l'état initial. Formellement, cela s'exprime par : $\forall M([i] \rightarrow s \Rightarrow s \geq [o]) \Rightarrow (s = [o])$, où M représente un marquage quelconque, s représente une séquence de transitions à partir de s et $s \geq [o]$ signifie que s peut mener à $[o]$.
3. **Vitalité du réseau** : Le réseau de Petri doit être vivant, ce qui signifie qu'il existe un marquage accessible depuis chaque place du réseau.

La propriété de "soundness" implique qu'il est possible d'atteindre le marquage final $[P_0]$ à partir de n'importe quel marquage initial. Cela garantit que le Workflow Net modélisé se termine toujours de manière correcte et atteint l'état final souhaité.

Les Workflow Nets peuvent être organisés en blocs, appelés Block-Structured Workflow Nets, permettant une représentation hiérarchique et récursive de workflows distincts [28]. Cette structure facilite la modélisation de processus complexes en les décomposant en sous-processus plus gérables.

1.6.3 Process Tree

Un process tree est une structure arborescente qui représente un processus de manière hiérarchique et complète [12, 28]. Il permet de visualiser l'ensemble des étapes et des relations qui composent le processus (voir figure 1.9).

- **Feuilles** : Les feuilles de l'arbre représentent les activités élémentaires du processus. Elles constituent les éléments de base de l'exécution du processus.
- **Nœuds** : Les nœuds de l'arbre représentent les relations entre les activités. Ils décrivent la façon dont les activités sont organisées et séquencées.

Les arbres de processus sont définis par :

- a , avec $a \in A$, est un process tree ;

- Soient M_1, \dots, M_n avec $n \geq 0$ des process tree et x un opérateur de process tree alors $x(M_1, \dots, M_n)$ est un process tree ; x pouvant être une composition séquentielle (\rightarrow), un choix exclusif (\times), une composition parallèle (\wedge) ou une boucle (\circlearrowleft) comme le montre la figure suivante.

Cette représentation est équivalente aux Block-Structured Workflow Net.

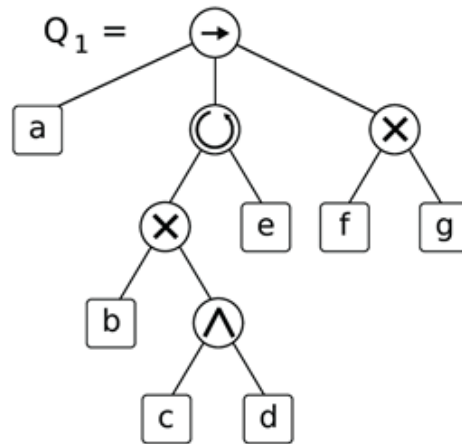


FIG. 1.9 : Exemple de Process Tree [10]

1.6.4 Business Process Model and Notation (BPMN)

Le Business Process Model and Notation (BPMN) est une norme de modélisation des processus métier. Il fournit une notation graphique basée sur les techniques traditionnelles de diagramme de flux pour représenter visuellement les processus d'entreprise dans un Diagramme de Processus Métier (DPM) [45].

Le Business Process Model and Notation (BPMN) a été créé par le Business Process Management Initiative (BPMI), qui a ensuite fusionné avec l'Object Management Group (OMG). La dernière version, BPMN 2.0, est largement supportée par les outils modernes, mais BPMN 1.2 reste populaire et de nombreux outils fonctionnent encore avec cette version. La figure 1.10 représente l'évolution du développement de BPMN avec le temps.

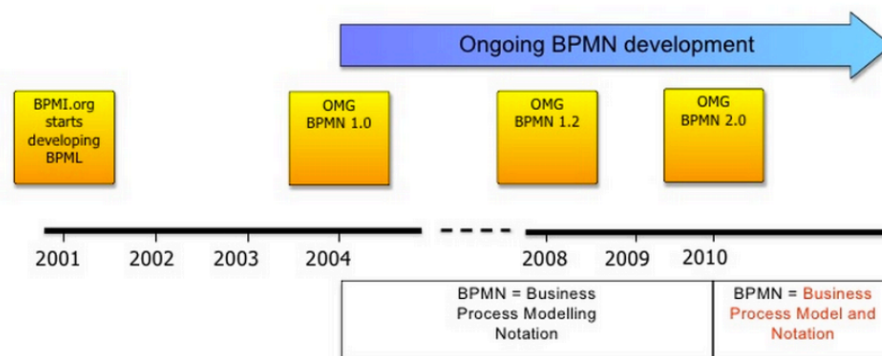


FIG. 1.10 : Évolution du développement de BPMN [11]

L'objectif du BPMN est de faciliter la modélisation des processus d'entreprise à la fois pour les utilisateurs techniques et les utilisateurs métiers. Il propose une notation intuitive pour les utilisateurs métiers tout en étant capable de représenter la sémantique complexe des processus. La spécification BPMN 2.0 fournit également une sémantique d'exécution ainsi qu'un mappage entre les éléments graphiques de la notation et d'autres langages d'exécution, en particulier le Business Process Execution Language (BPEL) [48].

Le BPMN utilise une variété d'éléments graphiques pour représenter les différentes composantes d'un processus métier. Les quatre catégories fondamentales du BPMN sont :

- **Objets de flux** : événements, activités, branchements.
- **Objets de connexion** : flux de séquence, flux de message, associations.
- **Couloirs** : pool ou voies.
- **Artefacts** : objets de données, groupes, annotations.

La figure 1.11 présente un modèle d'un processus métier en BPMN.

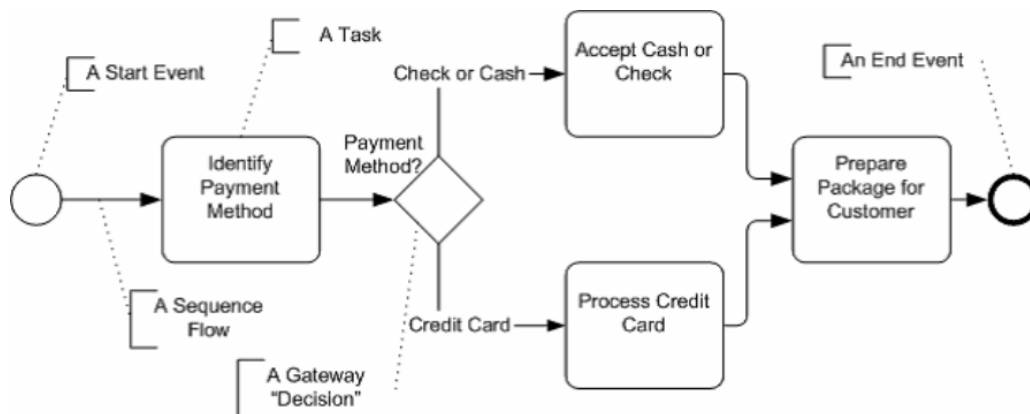


FIG. 1.11 : Exemple d'un modèle en BPMN [37]

1.7 Gestion des processus métier

La gestion des processus métier, ou Business Process Management (BPM) est une discipline qui vise à améliorer les performances d'une organisation en analysant, concevant, mettant en œuvre, contrôlant et optimisant ses processus métier [25, 21]. Son objectif principal est de rendre les flux de travail et les processus plus fluides et efficaces grâce à l'implantation d'outils informatiques dédiés.

En substance, la BPM s'attaque à deux problématiques majeures :

- **Réduction des erreurs humaines et des problèmes de communication** : Grâce à une modélisation claire et structurée des processus, la BPM permet d'identifier et d'éliminer les points de friction susceptibles de générer des erreurs ou des incompréhensions entre les acteurs impliqués.

- **Focalisation des intervenants sur leurs missions essentielles** : En automatisant les tâches répétitives et en simplifiant les procédures, la BPM libère le temps des collaborateurs, leur permettant de se concentrer sur des activités à plus forte valeur ajoutée, en adéquation avec leurs compétences et responsabilités.

Le cycle de vie du BPM est un cadre conceptuel qui décrit les différentes étapes d'une initiative BPM. Ces étapes, présentées dans la Figure 1.12, se répètent de manière itérative, permettant un raffinement continu des processus métier [19].

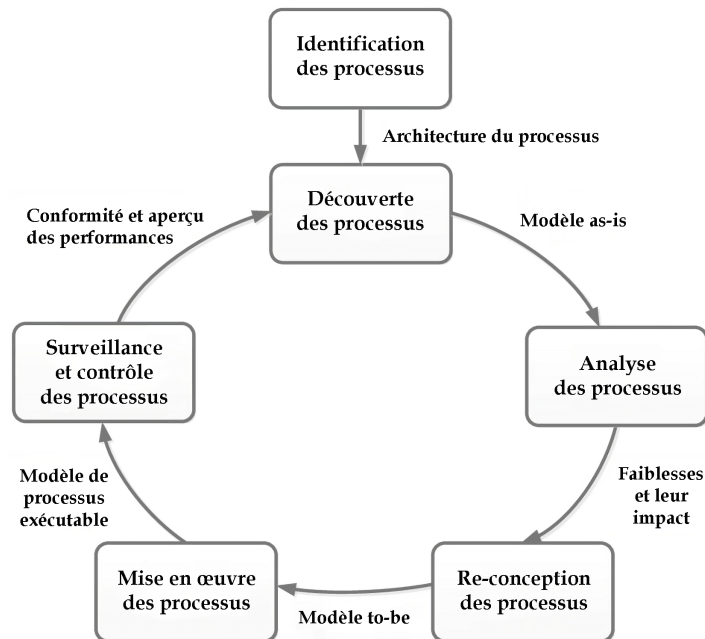


FIG. 1.12 : Cycle de vie du BPM

1. **Identification des processus** : Cette phase vise à définir les processus métier de l'organisation et à établir leurs contours. Cette démarche se traduit par la création d'une architecture de processus, mettant en lumière les relations entre ces processus clés. Des indicateurs de performance clés (KPI) sont également établis pour faciliter la hiérarchisation des processus et évaluer les interventions potentielles. Les recherches dans cette phase se concentrent principalement sur le développement de modèles, de méthodes et de systèmes pour les architectures de processus [20].
2. **Découverte des processus** : Cette étape consiste à modéliser un processus métier spécifique pour établir une compréhension commune, également appelée "modèle as-is". Les informations sur l'exécution du processus sont collectées à travers diverses méthodes telles que l'analyse de documents, les entretiens et les observations. La recherche se concentre sur l'adéquation des langages de modélisation des processus et leur impact sur la compréhension. Le "process mining" est une approche de recherche qui découvre automatiquement les processus métier à partir des journaux d'événements enregistrés dans les systèmes d'information opérationnels.
3. **Analyse des processus** : Cette étape implique l'examen du processus à l'aide du modèle "as-is" pour identifier les problèmes, leurs causes et leurs impacts. Les

techniques utilisées peuvent être qualitatives (analyse de la valeur ajoutée, analyse des gaspillages, analyse des causes premières) ou quantitatives (simulation, analyse des temps de cycle, théorie des files d'attente) [19].

4. **Re-conception des processus** : Cette phase vise à développer et à concevoir un processus métier "to-be" et son modèle de processus correspondant. Contrairement au modèle "as-is", le modèle "to-be" visualise comment le processus devrait être exécuté à l'avenir. Différentes méthodes de re-conception existent, regroupées dans ce que l'on appelle l'"orbite de re-conception" : transition vers la transformation, interne vs externe, technologique vs organisationnel [19]. En plus des mesures traditionnelles (coût, qualité, rapidité et flexibilité), de nouvelles approches se concentrent sur des objectifs différents. Par exemple, Wurm [50] proposent des "design patterns" pour adapter les processus métier à des besoins spécifiques.
5. **Mise en œuvre des processus** : Cette phase concrétise le modèle de processus "to-be" en se basant sur la gestion du changement et le développement d'applications informatiques pour l'exécution du processus. La recherche en BPM se concentre principalement sur ce dernier aspect, offrant des méthodes et outils pour traduire les modèles en logiciels. Les systèmes d'information sensibles aux processus, tels que les PGI et les systèmes de gestion des processus métier, sont utilisés pour soutenir cette mise en œuvre et l'exécution des processus.
6. **Surveillance et contrôle des processus** : La dernière phase du cycle de vie surveille l'exécution du processus et prend des mesures correctives si nécessaire. Les tableaux de bord de performance offrent une vue des indicateurs clés et des problèmes potentiels. Le process mining compare le comportement réel du processus avec le modèle "to-be" dans une étape de "conformance checking" (vérification de conformité) [5], fournissant ainsi des données pour la découverte et l'analyse des processus si le cycle de vie doit être relancé.

Gérer l'efficacité des processus métier est crucial pour la réussite de toute entreprise. C'est pourquoi le Business Process Intelligence (BPI) s'impose comme un outil indispensable pour optimiser les performances opérationnelles.

Le BPI regroupe un ensemble d'outils intégrés qui permettent de surveiller, analyser, contrôler et optimiser l'exécution des processus métier. Il s'agit essentiellement d'une application des principes du Business Intelligence (BI) aux processus métier, en exploitant les données générées par leur exécution pour en tirer des informations précieuses. Ses fonctionnalités clés sont les suivantes :

- **Analyse** : Comprendre le fonctionnement des processus en détail, identifier les goulots d'étranglement et les points d'amélioration.
- **Prévision** : Anticiper les tendances et les risques potentiels pour une meilleure prise de décision.
- **Surveillance** : Suivre en temps réel l'exécution des processus et détecter les anomalies.

- **Contrôle** : Ajuster et optimiser les processus en cours d'exécution pour maintenir une performance optimale.
- **Optimisation** : Identifier et mettre en œuvre des améliorations continues pour maximiser l'efficacité des processus.

Cette dernière fonctionnalité fait le corps de notre étude. L'optimisation des processus, également connue sous le nom de Business Process Improvement (BPI), est une discipline qui vise à identifier, analyser et améliorer les processus métier afin d'atteindre des objectifs spécifiques [22, 23].

1.7.1 Concepts clés de l'optimisation des processus

- **Efficacité** : La capacité à réaliser une tâche avec le minimum de ressources [22].
- **Performance** : La mesure de l'atteinte des objectifs d'un processus [22].
- **Amélioration continue** : Un processus itératif visant à identifier et à corriger les inefficacités des processus [22].
- **Parties prenantes** : Les individus ou groupes affectés par un processus [22].
- **Cartographie des processus** : Représenter visuellement les étapes d'un processus pour en identifier les inefficacités [22]. Cette cartographie permet de comprendre le déroulement du processus, d'identifier les points de départ et d'arrivée, les acteurs impliqués et les tâches réalisées.
- **Goulot d'étranglement** : Une étape du processus qui limite la capacité globale du système [23]. L'identification des goulots d'étranglement est cruciale pour améliorer l'efficacité du processus, car elle permet de se concentrer sur les points qui ont le plus d'impact sur la performance globale.
- **Métriques de performance** : Des indicateurs qui permettent de mesurer l'efficacité et la performance d'un processus [23]. L'utilisation de métriques de performance permet de suivre les progrès réalisés et de déterminer si les initiatives d'optimisation des processus sont efficaces.

Actuellement, les techniques les plus répandues pour l'optimisation des processus sont les suivantes :

- **Lean Management** : Focalisé sur l'élimination des gaspillages et l'amélioration continue.
- **Six Sigma** : Utilisé pour réduire la variabilité des processus et améliorer la qualité.
- **Théorie des Contraintes (TOC)** : Concentre sur l'identification et la gestion des contraintes qui limitent la performance des processus.

- **Business Process Reengineering (BPR)** : Implique une refonte radicale des processus pour atteindre des améliorations spectaculaires en termes de performance.

Le Process Mining joue un rôle crucial dans l'optimisation des processus en fournissant une vue basée sur les données des processus réels. Il permet de découvrir, surveiller et améliorer les processus en utilisant des données de journalisation des événements des systèmes d'information. Grâce au Process Mining, les entreprises peuvent identifier les inefficacités, comprendre les variations de processus et prendre des décisions basées sur des faits pour optimiser leurs opérations.

Parmi les exemples de réussites d'optimisation des processus :

- **Toyota** : Pionnière du Lean Management, a révolutionné la production automobile en réduisant les gaspillages et en améliorant l'efficacité [49].
- **General Electric (GE)** : A adopté Six Sigma pour améliorer la qualité et réduire les défauts, réalisant des économies substantielles et des gains de productivité [24].
- **Dell** : A utilisé le Business Process Reengineering pour réinventer sa chaîne d'approvisionnement, permettant une production juste-à-temps et une réduction significative des coûts [9].

Au cœur du BPI, le process mining joue un rôle crucial. Cette technique permet de découvrir et de comprendre les flux de travail réels des processus déployés en analysant les journaux d'événements générés par les systèmes d'information. Grâce au process mining, les entreprises peuvent identifier les inefficacités, les goulots d'étranglement et les opportunités d'amélioration au sein de leurs processus.

1.8 Propriétés du modèle miné

Le Process Mining génère automatiquement des modèles de processus basés sur les séquences d'événements enregistrés. Pour une compréhension optimale, il est crucial de produire des rapports complets incluant les modèles de processus, les indicateurs de conformité et une représentation des traces utilisées [39].

Quatre indicateurs principaux, largement reconnus par la communauté du Process Mining, permettent d'évaluer la qualité des modèles générés :

- **Justesse (Fitness)** : Capacité du modèle à reproduire toutes les traces disponibles.
- **Précision** : Limite les traces différentes pour éviter des séquences improbables.
- **Généralisation** : Capacité à s'adapter aux futures traces, y compris les nouvelles séquences.
- **Simplicité** : Complexité minimale du modèle tout en maintenant justesse, précision et généralisation.

L'utilisation conjointe des métriques d'évaluation et d'un graphe représentant la temporalité des événements permet d'obtenir des indications supplémentaires sur la construction du modèle et d'identifier d'éventuelles incohérences ou de valider certaines constructions. Chacune des quatre métriques attribue au modèle de processus une valeur comprise entre 0 et 1, 1 représentant la valeur optimale.

L'évaluation des indicateurs nécessite de mesurer la distance entre les traces d'exécution et le modèle de processus. Les algorithmes d'alignement calculent cette distance et déterminent la correspondance entre les traces et le modèle.

Les quatre indicateurs seront ensuite définis plus en détail dans les sections suivantes.

1.8.1 Alignement du modèle et des traces

Les traces d'exécution des processus peuvent être incomplètes ou comporter des erreurs, ce qui signifie que certaines traces ne correspondent pas parfaitement au processus défini par le modèle. Pour remédier à ce problème, on utilise le principe d'alignement, tel que défini par les algorithmes d'alignement, pour déterminer la distance entre une trace et le modèle [2].

Définitions clés pour l'alignement

- T_M : Ensemble des processus modélisés
- T_L : Ensemble des traces
- $A(T)$: Ensemble d'activités apparaissant dans T
- \perp : Activité nulle

Types d'activités dans l'alignement

- (x, y) est une activité dans le modèle : $x \in A(T_M)$ et $y = \perp$
- (x, y) est une activité dans les traces : $x = \perp$ et $y \in A(T_L)$
- (x, y) est une activité à la fois dans le modèle et dans les traces : $x \in A(T_M)$ et $y \in A(T_L)$
- (x, y) est une activité illégale : $x = \perp$ et $y = \perp$

Ensemble des activités légales

$$A_{LM} = \{(x, y) \in (A(T_M)A(T_L)) \mid x \in A(T_M) \vee y \in A(T_L)\} \quad (1.1)$$

Calcul de la distance entre le modèle et les traces

En associant un poids aux activités légales, on peut calculer la distance δ entre le modèle et les traces en faisant la somme des poids de chaque activité calculée à partir des alignements.

Distance entre une trace et un processus

Soit $\gamma(t_L, t_M)$ l'alignement entre la trace $t_L \in T_L$ et le processus $t_M \in T_M$, alors la distance entre t_M et t_L est égale à :

$$\delta(\gamma(t_L, t_M)) = \sum_{(x,y) \in \gamma} w(x, y) \quad (1.2)$$

Grâce à cette distance, nous pouvons chercher le meilleur alignement possible entre une trace et son modèle, noté $\lambda(t_L)$. Il nous suffit de chercher le processus pour lequel δ est minimale. Une fois cet alignement réalisé, nous pouvons réaliser les calculs des indicateurs introduits précédemment (justesse, précision, généralisation et simplicité) en utilisant les informations de correspondance entre la trace et le modèle.

1.8.2 Justesse

L'indicateur de justesse (fitness) mesure la capacité du modèle à reproduire fidèlement les traces d'exécution [39, 27]. Il s'agit de la proportion d'exécution qui correspond à un processus défini par le modèle dans sa totalité.

Calcul de la justesse

1. Alignement des traces sur le modèle : On calcule la distance entre chaque trace et les processus possibles du modèle en les alignant.
2. Distance totale : On somme les distances de toutes les traces.
3. Normalisation : On divise la distance totale par la distance maximale possible.

$$F(W, L) = \frac{\sum_{t \in L} \delta(\lambda(t))}{\delta_{max}} \quad (1.3)$$

1.8.3 Précision

L'indicateur de précision (precision) mesure la capacité du modèle à exclure les comportements impossibles [39, 27]. Il s'agit de vérifier que le modèle ne permet que des séquences d'événements observées dans les traces d'exécution.

Calcul de la précision

1. Activités possibles à chaque instant : On compare les activités possibles selon le modèle $enabled_M(e)$ et les activités observées dans les traces $enabled_L(e)$ pour chaque événement e dans les logs ϵ .
2. Indicateur de précision : On calcule $P(LM)$ en divisant le nombre d'événements pour lesquels les activités possibles correspondent aux activités observées par le nombre total d'événements.

$$P(L, M) = \frac{\sum_{e \in \epsilon} \frac{enabled_L(e)}{enabled_M(e)}}{|\epsilon|} \quad (1.4)$$

Il convient de noter qu'il est toujours possible de créer un modèle de processus capable de reproduire les traces observées en utilisant un modèle de processus en fleur (Figure 1.13). Ce type de modèle se compose d'un seul état (place) connecté à toutes les transitions par des boucles élémentaires. Ce modèle peut reproduire n'importe quelle séquence d'activités, à condition que ces activités soient incluses dans l'ensemble des activités observées. Cependant, un tel modèle manque de précision quant au comportement réellement observé.

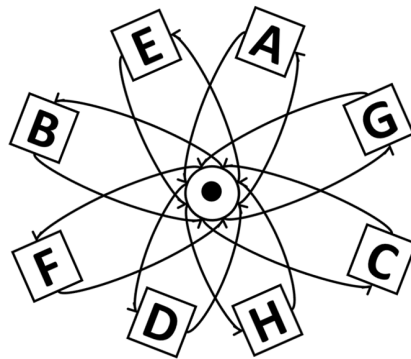


FIG. 1.13 : Modèle de processus en fleur ??

1.8.4 Généralisation

L'indicateur de généralisation (generalization) mesure la capacité du modèle à prédire des événements futurs [39, 27]. Il s'agit d'évaluer la mesure dans laquelle le modèle peut s'adapter à de nouvelles séquences d'événements non observées lors de sa construction.

Calcul de la généralisation :

1. **Fréquence des activités** : On calcule le nombre d'occurrences de chaque événement ($N(e)$) et le nombre d'activités candidates à chaque état possible du système ($O(e)$).
2. **Probabilité de transition** : On calcule la probabilité $P(e)$ que la prochaine visite à un état révèle une activité non vue précédemment.

3. **Indicateur de généralisation** : On calcule $G(L, M)$ en faisant la moyenne des probabilités $P(e)$ pour tous les événements et dans les logs (ϵ).

Formule :

$$P(e) = \begin{cases} \frac{O(e)(O(e)+1)}{N(e)(N(e)-1)} & \text{si } N(e) \geq O(e) + 2 \\ 1 & \text{sinon} \end{cases} \quad (1.5)$$

$$G(L, M) = 1 - \frac{\sum_{e \in \epsilon} P(e)}{|\epsilon|} \quad (1.6)$$

1.8.5 Simplicité

L'indicateur de simplicité (*simplicity*) mesure la complexité structurelle du modèle de processus [14]. Il s'agit d'évaluer la facilité d'interprétation et de compréhension du modèle par un utilisateur humain.

Calcul de la simplicité :

1. **Taille du modèle** : On calcule le nombre de places et de transitions présentes dans le réseau de Petri du modèle.
2. **Nombre d'activités** : On détermine le nombre d'activités présentes dans les traces d'exécution.
3. **Rapport taille/activités** : On compare la taille du modèle au nombre d'activités.

Formule :

$$S(L, M) = 1 - \frac{\#activits\ dupliques + \#activits\ manquantes}{\#noeuds\ du\ modle + \#classes\ d'vnements\ des\ logs} \quad (1.7)$$

Les quatre indicateurs (justesse, précision, généralisation et simplicité) permettent d'évaluer globalement la qualité du modèle découvert et ainsi la performance de l'algorithme utilisé vis-à-vis du jeu de données.

1.9 Techniques et algorithmes de fouille de processus

La fouille de processus (*process mining*) est un domaine en plein essor qui permet aux entreprises d'extraire des connaissances précieuses à partir des données d'exécution des processus métier. Ces connaissances peuvent ensuite être utilisées pour améliorer l'efficacité, la conformité et la performance globale des processus.

Dans ce contexte, trois techniques principales se distinguent :

1. **Découverte de processus métier automatisée** : L'ABPD (*Automated Business Process Discovery*) vise à découvrir automatiquement les processus métier réels à partir des journaux d'événements générés par les systèmes d'information. Cette technique permet de comprendre comment les utilisateurs exécutent réellement les processus, sans se fier aux modèles de processus existants qui peuvent être obsolètes ou inexacts [5].

2. **Vérification de la conformité des processus métier** : Le BPCC (*Business Process Conformance Checking*) compare le comportement réel des processus, tel qu'enregistré dans les journaux d'événements, au modèle de processus défini. Cette comparaison permet d'identifier les écarts entre le comportement attendu et le comportement réel, révélant ainsi des problèmes potentiels d'efficacité, de conformité ou de qualité des processus [5, 2].
3. **Analyse des performances** : La PA (*Performance Analysis*) va au-delà de la découverte et de la vérification de la conformité en évaluant l'efficacité et la performance des processus. Elle permet d'identifier les goulots d'étranglement, les gaspillages et les opportunités d'amélioration en analysant les indicateurs de performance clés tels que les temps de cycle, les taux d'erreur et d'autres mesures pertinentes.

Afin d'appliquer ces techniques, nous allons explorer trois principaux outils de fouille de processus :

1. **ProM**¹ : ProM est un environnement polyvalent et extensible pour l'exploration de processus. Il propose des plug-ins permettant d'extraire différents types de modèles à partir des journaux d'événements, tels que des modèles de processus et des modèles organisationnels [8]. En outre, ProM prend en charge la conversion et l'analyse des modèles. En utilisant des techniques de contrôle de conformité, il est possible de comparer les modèles à la réalité et d'améliorer les modèles existants avec des informations supplémentaires [5].

ProM offre des fonctionnalités essentielles pour la fouille de processus. Il permet l'extraction de modèles à partir de journaux d'événements, leur conversion entre différentes notations, une analyse approfondie pour identifier les goulots d'étranglement et les inefficacités, un contrôle de conformité et l'enrichissement des modèles avec des informations supplémentaires.

2. **Disco**² : Développé par Fluxicon, Disco est un outil de fouille de processus complet qui permet de réaliser des analyses de processus rapidement et facilement. Il est conçu pour rendre la fouille de processus accessible et agréable pour les utilisateurs. Disco offre des fonctionnalités avancées de visualisation et d'analyse des processus, facilitant ainsi la découverte et l'optimisation des processus métier [40].

Disco propose des fonctionnalités clés pour la fouille de processus, notamment la découverte automatique de processus à partir des journaux d'événements, l'analyse approfondie pour identifier les goulots d'étranglement et les inefficacités, la création de diagrammes clairs pour représenter les processus, la simulation de scénarios hypothétiques et les recommandations d'amélioration basées sur l'analyse des données.

3. **ARIS Express**³ : Initialement développé par IDS Scheer et acquis par Software AG en décembre 2010, ARIS Express est un outil gratuit de modélisation pour

¹<https://promtools.org/>

²<https://fluxicon.com/disco/>

³<https://aris.com/process-mining/>

l'analyse et la gestion des processus métier. Il prend en charge différentes notations de modélisation telles que BPMN 2 et les chaînes de processus événementielles (EPC). ARIS Express permet aux utilisateurs de modéliser, documenter et analyser les processus métier de manière intuitive [16].

ARIS Express propose des fonctionnalités clés pour la gestion des processus. Il permet la modélisation via des diagrammes BPMN et EPC, l'analyse des processus pour identifier les goulots d'étranglement et les inefficacités, ainsi que la simulation pour tester des scénarios hypothétiques. Il favorise la collaboration en permettant le partage des modèles et l'intégration avec d'autres outils de gestion des processus métier.

Au cœur de la fouille de processus se trouvent les algorithmes, ces outils mathématiques sophistiqués qui permettent d'extraire des connaissances précieuses des données brutes des journaux d'événements. Chaque algorithme possède ses propres forces et faiblesses, et le choix de l'algorithme le plus adapté dépend des objectifs spécifiques de l'analyse.

Algorithme alpha α

L'algorithme Alpha Miner est une méthode proposée par Van Der Aalst [6] pour reconstruire les liens de causalité dans les Workflow Nets à partir des relations existantes dans les logs d'exécution. Il permet de découvrir un modèle de processus à partir d'un ensemble de traces d'exécution. Cet algorithme suppose que les journaux d'événements sont complets et ne contiennent aucun bruit.

Avant de plonger dans l'algorithme Alpha Miner, il est crucial de comprendre les relations d'ordre entre les activités dans un journal de workflow. Ces relations permettent de capturer les enchaînements et les dépendances entre les tâches, éclairant ainsi le déroulement du processus.

Trois types de relations d'ordre existent :

- **Causalité** ($a \rightarrow b$) : b suit toujours a , mais jamais l'inverse.
 - Exemple : $(\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d})$; $(\mathbf{a},\mathbf{c},\mathbf{d})$; $(\mathbf{d},\mathbf{a},\mathbf{b})$
- **Simultanéité** ($a \parallel b$) : a et b peuvent se produire dans l'ordre ou dans l'ordre inverse.
 - Exemple : $(\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d})$; $(\mathbf{a},\mathbf{c},\mathbf{d})$; $(\mathbf{d},\mathbf{b},\mathbf{a})$
- **Exclusivité** ($a \# b$) : a et b ne peuvent jamais se produire dans la même séquence.
 - Exemple : $(\mathbf{a},\mathbf{c},\mathbf{d},\mathbf{b})$; $(\mathbf{a},\mathbf{c},\mathbf{d})$; $(\mathbf{b},\mathbf{d},\mathbf{a})$

L'algorithme Alpha Miner s'articule autour de huit étapes principales :

1. **Identification des activités présentes dans les logs** : On identifie toutes les activités distinctes présentes dans les traces d'exécution et on crée une transition pour chacune d'entre elles.

$$T_L = \{t \in T \mid \exists \sigma \in L \ t \in \sigma\} \quad (1.8)$$

2. **Identification de l'activité de départ** : On identifie l'activité qui apparaît en premier dans toutes les traces d'exécution et on la définit comme l'activité de départ du modèle.

$$T_I = \{t \in T \mid \forall \sigma \in L \ first(\sigma) = t\} \quad (1.9)$$

3. **Identification de l'activité de terminaison** : On identifie l'activité qui apparaît en dernier dans toutes les traces d'exécution et on la définit comme l'activité de terminaison du modèle.

$$T_0 = \{t \in T \mid \forall \sigma \in L \ last(\sigma) = t\} \quad (1.10)$$

4. **Détection des enchaînements d'activités** : On identifie les paires d'activités qui se suivent dans les traces d'exécution, en tenant compte des relations d'exclusion mutuelle (parallélisme).

$$X_L = \{(A, B) \mid A \subseteq T_L \wedge A \neq \emptyset \wedge B \subseteq T_L \wedge B \neq \emptyset \wedge \forall a \in A \ \forall b \in B \ a \rightarrow_L b \wedge \forall a_1, a_2 \in A \ a_1 \#_L a_2 \wedge \forall b_1, b_2 \in B \ b_1 \#_L b_2\} \quad (1.11)$$

5. **Réduction des paires d'activités** : On regroupe les paires d'activités identiques et on conserve uniquement les paires maximales, c'est-à-dire celles qui relient le plus grand nombre d'activités.

$$Y_L = \{(A, B) \in X_L \mid \forall A', B' \subseteq X_L \ (A' \subseteq A \wedge B' \subseteq B) \Rightarrow (A, B) = (A', B')\} \quad (1.12)$$

6. **Définition des places du modèle** : On définit une place pour chaque paire d'activités maximale, en considérant que cette place relie les activités de la paire.

$$P_L = \{p_{(A,B)} \mid (A, B) \in Y_L\} \cup \{i_L, o_L\} \quad (1.13)$$

7. **Identification des arcs du modèle** : On crée des arcs reliant les transitions et les places du modèle. Une transition est reliée à une place si l'activité correspondante appartient à la paire d'activités associée à la place.

$$F_L = \{((a, p_{(A,B)}) \mid (A, B) \in Y_L \wedge a \in A\} \cup \{(p_{(A,B)}, b) \mid (A, B) \in Y_L \wedge b \in B\} \cup \{(i_L, t_I) \mid t \in T_I\} \cup \{(t_0, o_L) \mid t \in T_0\} \quad (1.14)$$

8. **Construction du modèle de processus** : On regroupe les places, les transitions et les arcs définis dans les étapes précédentes pour obtenir le modèle de processus sous forme de réseau de Petri.

$$\alpha(L) = (P_L, T_L, F_L) \quad (1.15)$$

Les limites de l'algorithme Alpha Miner sont multiples. Il ne peut pas gérer les boucles dans les processus, ni identifier les choix non libres où plusieurs activités peuvent se dérouler simultanément. De plus, il peut être sensible au bruit dans les données d'exécution, ce qui peut entraîner des modèles inexacts. Ces défis mettent en évidence la nécessité d'autres approches pour améliorer la précision de la fouille de processus.

Algorithme alpha $\alpha+$

L'algorithme Alpha+, proposé par de Medeiros [35], est une extension de l'algorithme Alpha original qui permet de découvrir les boucles dans les modèles de processus.

L'algorithme Alpha+ introduit deux nouvelles relations pour identifier les boucles :

- **Boucle simple** ($a \Delta b$) : Une activité a se répète dans une boucle et b est l'activité suivante dans la séquence.

– Formule :

$$a\Delta b \iff \exists t = \langle e_1, e_2, \dots, e_n \rangle, t \in T \mid e_i = e_{i+2} = a \wedge e_{i+1} = b, 1 \leq i \leq n - 2 \quad (1.16)$$

- **Boucle double** ($a \diamond b$) : Deux boucles sont imbriquées, a étant la boucle externe et b la boucle interne.

– Formule :

$$a\diamond b \iff a \rightarrow b \wedge a \wedge b \rightarrow a \quad (1.17)$$

L'algorithme Alpha+ redéfinit les relations de parallélisme (\parallel) et de séquence (\rightarrow) pour tenir compte des boucles :

- **Séquence** ($a \rightarrow b$) : b suit a et il n'y a pas de boucle simple impliquant a et b.

– Formule :

$$a \rightarrow b \iff \neg(a\diamond b \wedge a\Delta b) \quad (1.18)$$

- **Parallélisme** ($a \parallel b$) : a et b peuvent se produire dans n'importe quel ordre, et il n'y a pas de boucle double les impliquant.

– Formule :

$$a\parallel b \iff a \rightarrow b \wedge a \wedge \neg(a\diamond b) \quad (1.19)$$

L'algorithme Alpha+ permet de découvrir les motifs de boucle de taille 2, c'est-à-dire les boucles simples et les boucles doubles.

Avant d'appliquer l'algorithme Alpha+, il est nécessaire d'effectuer un prétraitement des traces pour éliminer les boucles de taille 1, car elles ne peuvent pas être représentées dans le formalisme des réseaux de Petri utilisé par l'algorithme.

Algorithme alpha $\alpha++$

L'algorithme Alpha++, proposé par Van Der Aalst [1], est une extension de l'algorithme Alpha+ qui permet de traiter les choix non libres dans les modèles de processus.

Alpha++ introduit la notion de relations implicites, représentées par quatre nouvelles relations :

- **Précedence de choix** ($a \searrow b$) : a précède un choix entre a et b (XOR-split).
 - Formule : $a \triangleleft b \iff a \# b \wedge (\exists c \in T \mid c \rightarrow a \wedge c \rightarrow b)$
- **Implication de choix** ($a \nearrow b$) : a et b sont en choix exclusif et suivis d'une activité commune (XOR-join).
 - Formule : $a \triangleright b \iff a \# b \wedge (\exists c \in T \mid a \rightarrow c \wedge b \rightarrow c)$
- **Séquence indirecte** ($a \gg b$) : a et b sont en séquence mais ne se suivent pas directement.
 - Formule : $a \gg b \iff a \approx b \wedge (\exists e, t \in T \mid e \neq a \wedge e \neq b \wedge \neg(e \rightarrow a \vee e \leftarrow a))$
- **Implication directe ou indirecte** ($a \nabla b$) : a implique b de manière directe ou indirecte.
 - Formule : $a \nabla b \iff a \rightarrow b \vee a \gg b$

L'algorithme Alpha++ est décomposé en quatre étapes sont les suivantes :

1. Suppression des boucles de taille 1.
2. Application de l'algorithme Alpha sur les traces modifiées.
3. Calcul des dépendances implicites.
4. Calcul des arcs du modèle de processus.

L'algorithme Alpha++ (et ses dérivés) adopte une approche **bottom-up**. Il commence par identifier les relations au niveau le plus fin et les combine ensuite pour reconstruire le modèle global. En revanche, les techniques d'approche **top-down** infèrent les relations de granularité plus fine à partir de relations globales, en subdivisant les logs.

Heuristic Miner

L'Heuristic Miner est une technique de fouille de processus basée sur les règles de causalité de l'algorithme Alpha++. Elle fonctionne en calculant la fréquence d'apparition de ces règles pour identifier la règle la plus appropriée pour chaque relation entre les activités. Cette technique peut détecter les séquences hors contexte mais cela nécessite une phase de configuration complexe [47]. Son principe de fonctionnement est comme suit :

1. **Calcul des scores de règles** : Chaque règle possible reçoit un score en fonction du nombre d'apparitions de sa représentation dans les traces d'exécution.
2. **Détection de séquences hors contexte** : L'Heuristic Miner peut détecter des séquences hors contexte, mais cela nécessite une configuration complexe.

3. **Construction d'un graphe de dépendance** : Pour chaque activité, l'Heuristic Miner calcule la probabilité qu'elle implique une autre activité du processus.
4. **Sélection des implications** : Toutes les implications calculées au-dessus d'un seuil prédéfini sont retenues pour construire le réseau de Petri correspondant au modèle de processus.
5. **Réparation du modèle** : Une opération de réparation est généralement nécessaire pour corriger les imperfections du modèle généré.

Inductive Miner

L'algorithme Inductive Miner, développé par Leemans [28], permet de construire un modèle de processus sous forme d'arborescence (Process Tree). Cet algorithme fonctionne de la manière suivante :

1. **Approche globale** : L'Inductive Miner adopte une approche globale pour découvrir les modèles, en considérant les relations entre des ensembles d'activités plutôt qu'entre des activités individuelles.
2. **Découpage récursif des traces** : L'algorithme découpe récursivement les traces d'exécution en blocs, générant des règles qui lient les sous-blocs ainsi obtenus.
3. **Sélection de la découpe optimale** : Pour générer ces règles, l'Inductive Miner évalue la probabilité de chaque découpe possible et sélectionne la plus appropriée.
4. **Fonction de découpage variable** : La fonction de découpage utilisée dépend de la variante spécifique de l'algorithme Inductive Miner employée.

Inductive Miner Infrequent

L'algorithme Inductive Miner Infrequent, développé par Leemans [29], est une variante de l'algorithme Inductive Miner qui permet de filtrer les traces d'exécution en fonction des comportements peu fréquents.

L'objectif principal de l'Inductive Miner Infrequent est d'éliminer les bruits et les comportements aberrants des traces d'exécution afin d'obtenir des modèles de processus plus précis et robustes. Cet algorithme fonctionne de la manière suivante :

1. **Filtrage des traces rares** : Les traces ayant un nombre d'apparitions trop faible sont écartées.
2. **Pondération des arcs du graphe de causalité** : Les arcs du graphe de causalité sont pondérés en fonction de leur nombre d'apparitions dans les traces.
3. **Suppression des arcs peu utilisés** : Les arcs trop peu utilisés par rapport aux autres arcs sont supprimés.

1.10 Conclusion

Ce chapitre a exploré en profondeur les concepts fondamentaux des traces et de la fouille de processus, révélant leur potentiel immense pour l'analyse et l'optimisation des processus métier.

Les systèmes à base de traces constituent des outils précieux pour observer, collecter et analyser les interactions entre les utilisateurs et les systèmes numériques. En tirant parti des traces d'exécution, la fouille de processus permet de décrypter le déroulement réel des processus, d'identifier les inefficacités, de comprendre les comportements des utilisateurs et d'extraire des connaissances précieuses pour l'amélioration continue.

Divers algorithmes et techniques de fouille de processus ont été présentés, chacun offrant des approches distinctes pour analyser les relations entre les événements et construire des modèles de processus représentatifs. L'évaluation de la qualité des modèles minés s'appuie sur des propriétés telles que la justesse, la précision, la généralisation et la simplicité, permettant de comparer leur pertinence par rapport aux traces observées.

Malgré ces défis, la fouille de processus offre un potentiel immense pour les organisations. En identifiant les points d'amélioration, en optimisant l'efficacité des processus, en augmentant la qualité des produits et services et en renforçant la conformité réglementaire, la fouille de processus peut contribuer à propulser les entreprises vers de nouveaux sommets de réussite.

Le chapitre suivant examinera l'application pratique de ces techniques sur un cas réel, en utilisant les connaissances précédemment mentionnées.

Chapitre 2

Implémentation : Étude de Cas Hamoud Boualem

2.1 Introduction

Dans ce chapitre, nous mettons en pratique les concepts théoriques du process mining étudiés dans le chapitre précédent en les appliquant à un cas réel, celui de SPA Hamoud Boualem. Cette étude de cas a pour objectif d'illustrer l'efficacité du process mining dans l'optimisation des processus de chaîne logistique. Nous commencerons par une présentation détaillée de l'entreprise Hamoud Boualem, y compris son historique, sa structure organisationnelle, et ses principaux concurrents et fournisseurs.

Ensuite, nous procéderons à l'identification, la sélection et l'analyse d'un processus critique au sein de l'unité de Boufarik. À travers une série d'étapes méthodiques, nous analyserons les données disponibles, modéliserons et simulerons le processus choisi, et enfin, nous proposerons des recommandations pour son optimisation.

L'objectif final est de démontrer comment le process mining peut être utilisé pour améliorer l'efficacité et la performance des processus de supply chain, tout en offrant une feuille de route pour la mise en œuvre des recommandations proposées.

2.2 Présentation de l'organisme d'accueil

2.2.1 Historique de l'entreprise

L'établissement HAMOUD BOUALEM, société par actions (SPA), est l'une des plus anciennes entreprises de production de boissons gazeuses et de sirops, tous secteurs confondus, avec une histoire remontant à la première moitié du 19^{ème} siècle. À l'origine, Hamoud Boualem produisait des arômes, des boissons gazeuses et des blocs de glace. En 1878, l'entreprise fut établie dans le quartier de Belcourt, et au début du 20^{ème} siècle, elle déménagea à son emplacement actuel au 201 Rue HASSIBA BEN BOUALI, dans le quartier du RUISSEAU (EL ANNASSER).

Durant la première exposition universelle de Paris en 1889, les boissons HAMOUD BOUALEM furent classées hors concours et obtinrent la plus haute distinction, recevant vingt médailles d'or, qui font désormais partie de l'emblème et du logo présents sur les étiquettes des bouteilles de limonade blanche.

Aujourd'hui, HAMOUD BOUALEM se consacre à la gestion, l'exploitation et au développement de la production de boissons gazeuses, avec la marque "SELECTO" jouant un rôle prépondérant dans l'évolution de l'entreprise. Le nom "SELECTO" vient de l'idée de "Select", signifiant le choix et la sélection des meilleures matières pour son élaboration.

En tant qu'entreprise familiale, elle a toujours été dirigée par l'un des actionnaires. En plus de la société mère, le complexe industriel comprend plusieurs usines :

- Usine de BOUFARIK, à BLIDA (production de soda en bouteilles PET et en cannettes)
- Usine de OUED TLELAT, à ORAN (production de soda en bouteilles PET)
- Usine de HASSIBA, à Alger (production de bouteilles en verre)

De plus, HAMOUD BOUALEM collabore avec d'autres entreprises pour des prestations de services, telles que la production de soda en bouteilles en verre à BOUDOUAOU, dans la wilaya de Boumerdès.

L'annexe B présente une fiche d'identité de l'entreprise Hamoud Boualem.

2.2.2 Unité de Boufarik

Selon le secrétaire général et directeur des approvisionnements, Lyamine Lerari, « d'après les données de l'APAB, il existe près de 400 producteurs de boissons. La concurrence est rude, et si nous voulons encore être présents dans 10 ans, il est nécessaire de rester à la page et de nous adresser à un public jeune en dynamisant la marque. »

Pour relever ce défi, la SPA Hamoud Boualem a lancé et réalisé un méga projet : une nouvelle ligne de production située à Boufarik. Cette unité, qui a nécessité un investissement de près de 3 millions de dinars, est opérationnelle depuis le dernier trimestre de 2015 :

- Les trois derniers mois de 2015 ont été consacrés à tester le matériel et la recette.
- Le lancement officiel de la commercialisation a eu lieu en janvier 2016.

Cette unité produit des boissons gazeuses en bouteilles PET de différents formats (33 cl, 1 L, 2 L) et en canettes de 33 cl. Le PET représente environ 86 % du chiffre d'affaires de l'entreprise, raison pour laquelle notre étude se concentre sur ce type de produit.



FIG. 2.1 : Lignes de conditionnement de Hamoud Boualem - Boufarik

2.2.3 Structure organisationnelle

Pour comprendre le fonctionnement de l'entreprise Hamoud Boualem et l'organisation de ses différentes activités, il est essentiel de se pencher sur sa structure organisationnelle. La figure 2.2 illustre la structure hiérarchique de l'entreprise, montrant les différents départements et les interactions entre eux :

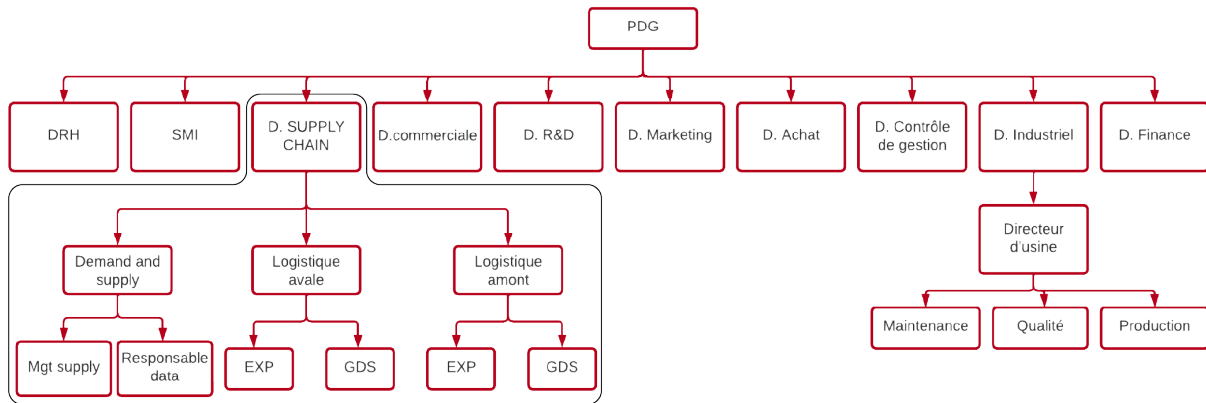


FIG. 2.2 : Organigramme de SPA Hamoud Boualem

Cette structure met en évidence les principales divisions de l'entreprise, ainsi que les sous-sections clés qui permettent une gestion fluide et coordonnée des opérations.

La direction encadrée en noir est celle qui fait l'objet de notre étude.

2.2.4 Concurrents et fournisseurs

Les fournisseurs jouent un rôle crucial dans la chaîne d'approvisionnement de Hamoud Boualem, garantissant la qualité et la disponibilité des matières premières nécessaires à la production de ses boissons gazeuses. La diversité des sources d'approvisionnement, qu'elles soient locales ou étrangères, permet à l'entreprise de maintenir une production constante tout en respectant les normes de qualité. Cependant, elle peut poser des problèmes majeurs, notamment avec les fournisseurs étrangers, en raison des conditions d'importation et des retards possibles.

Quant aux concurrents de Hamoud Boualem sur le marché algérien des boissons non alcoolisées, on trouve :

- Coca-Cola : une entreprise multinationale qui propose des boissons non alcoolisées telles que Coca-Cola, Fanta, Schweppes et Sprite, toutes commercialisées en Algérie.
- ABC Pepsi : une autre multinationale qui vend des sodas et autres boissons non alcoolisées en Algérie, notamment 7 UP, Pepsi et Mirinda.

Ces entreprises rivalisent pour attirer les consommateurs avec des stratégies de marketing agressives, une large gamme de produits et une forte présence sur le marché, rendant la concurrence particulièrement intense et exigeant une innovation constante de la part de Hamoud Boualem pour maintenir sa part de marché.

2.2.5 Direction Supply Chain

La direction Supply Chain de Hamoud Boualem joue un rôle crucial dans l'optimisation et la gestion des flux de matériaux et d'informations à travers l'ensemble de l'entreprise.

Sous la supervision du Directeur Supply Chain, cette direction est organisée en plusieurs sous-sections spécialisées pour garantir une efficacité maximale.

La gestion de la Supply Chain chez Hamoud Boualem repose sur trois volets principaux qui sont :

1. Logistique en amont : concerne l'approvisionnement en ressources ou en produits nécessaires à la production.
2. Logistique en aval : traite des processus de distribution des produits finis aux clients.
3. Planification de la demande et de l'approvisionnement (Demand & Supply Planning DSP) : implique la planification de la demande et de l'approvisionnement pour aligner la production avec les besoins du marché.

En outre, un quatrième volet a été adopté :

4. Outils d'Aide à la Vente (OAV) : ces outils sont conçus pour soutenir les activités de vente et optimiser les performances commerciales.

Dans cette section, on va concentrer sur le volet DSP, qui englobe les aspects suivants :

a. Planification de la Demande (Demand Planning) :

La planification de la demande chez SPA Hamoud Boualem est une étape essentielle en amont de la chaîne logistique. Elle permet à l'entreprise d'analyser les changements de son environnement, de gérer les déséquilibres en termes de délais de livraison, et d'augmenter ses parts de marché. L'objectif est de réaliser une planification opérationnelle à moyen terme, en étudiant le décalage entre les décisions d'achat et la satisfaction des clients, afin d'adopter une approche proactive pour maîtriser les ventes. Les étapes de cette planification sont résumées dans le schéma suivant :

La figure 2.3 résume les étapes de cette planification :

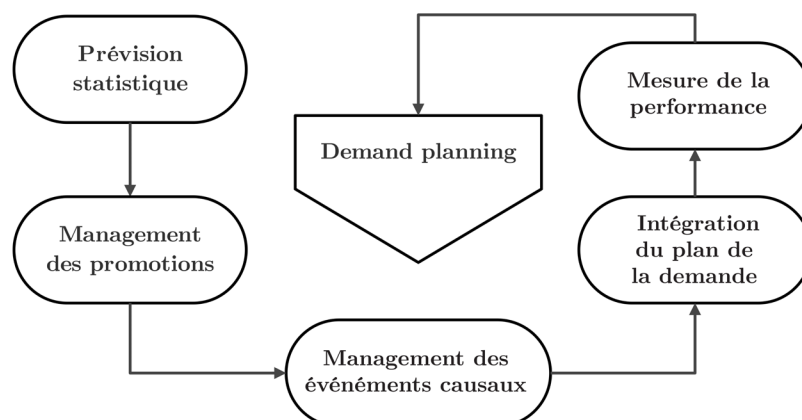


FIG. 2.3 : Étapes de planification de la Demande

b. Planification de l'approvisionnement (Supply Planning) :

La planification de l'approvisionnement chez SPA Hamoud Boualem aligne la capacité de production avec la demande prévue des clients, en établissant une planification de production détaillée pour garantir la disponibilité des produits finis en quantités suffisantes et au moment opportun. Cette planification inclut l'évaluation du TRS, la maintenance préventive et corrective, l'ordonnancement des produits, et la gestion des lots de production. Elle assure également la gestion optimale de la capacité de stockage et des niveaux de stock pour éviter les ruptures. Parallèlement, le plan d'approvisionnement des matières premières et des emballages couvre la gestion des stocks, la comparaison des consommations prévisionnelle et réelle, et la fixation des MOQ. Il prend en compte la DLUO, le lead time, et évalue le potentiel des fournisseurs pour garantir la disponibilité continue des matériaux nécessaires à la production.

La figure 2.4 résume les étapes de cette planification :

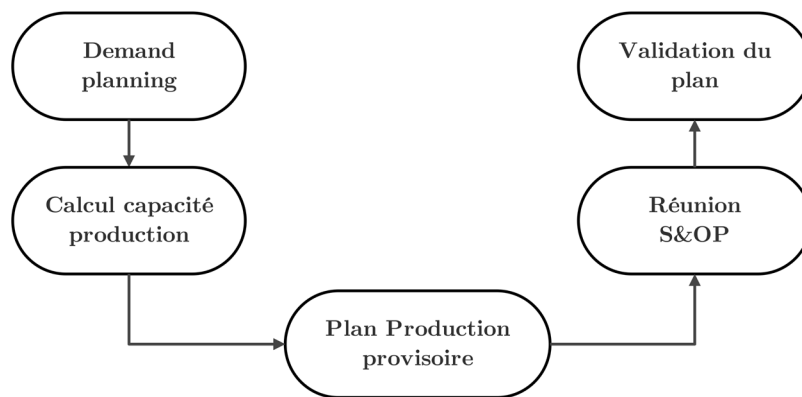


FIG. 2.4 : Étapes de planification de l'approvisionnement

2.2.6 Identification des processus

Dans cette section, les processus identifiés, à savoir le processus de production, le processus de commercialisation et celui de l'approvisionnement, ont été déterminés grâce à une analyse approfondie des activités clés de l'entreprise, des entretiens avec les responsables de la direction supply chain et de l'examen des documents internes de SPA Hamoud Boualem.

a. Processus de production :

L'entreprise fonctionne de manière à ce que la production soit son activité centrale. L'objectif est d'assurer la transformation efficace des matières premières en produits finis tout en maintenant un ordonnancement rigoureux, l'enregistrement systématique et la traçabilité de toutes les opérations. La direction de la production est active 24 heures sur 24 afin de répondre aux besoins et aux attentes des clients.

La direction supply chain fournit des prévisions de la demande (demand planning)

pour allouer les ressources nécessaires. Le processus de production commence par un ordre de fabrication, suivi de la préparation des matières premières.

Ensuite, l'eau est traitée pour répondre aux normes de qualité. Le sucre est ajouté et transformé en sirop simple. Ce sirop est mélangé avec de l'acide citrique pour obtenir les caractéristiques souhaitées, puis des arômes sont ajoutés et bien mélangés.

Le produit passe ensuite par les étapes de remplissage et d'emballage pour être conditionné. Enfin, le produit fini est prêt à être expédié aux clients, garantissant ainsi un produit de haute qualité conforme aux standards de l'entreprise.

• **Diagnostic du processus de production :** Le processus de production chez Hamoud Boualem présente plusieurs problèmes potentiels identifiés à travers le diagramme Ishikawa (figure 2.5). Au niveau de la main-d'œuvre, la mauvaise manipulation des équipements et le manque de formation du personnel peuvent entraîner des erreurs de production et affecter la qualité des produits. Les matériaux souffrent de problèmes liés à la qualité des matières premières et à des conditions de stockage inadéquates, pouvant détériorer les matières avant leur utilisation. Les méthodes non standardisées et un ordonnancement inefficace provoquent des variations de qualité et des inefficacités dans le flux de production. Les machines sont affectées par un entretien insuffisant et un manque d'automatisation, entraînant des pannes fréquentes et limitant l'efficacité. En termes de mesures, l'absence de KPIs et un suivi inefficace des performances rendent difficile l'évaluation et l'amélioration continue. Enfin, le milieu de travail est concerné par des mesures de sécurité insuffisantes et des risques environnementaux, ce qui peut entraîner des accidents et des perturbations de production. En adressant ces problèmes, Hamoud Boualem peut améliorer l'efficacité, la qualité et la sécurité de son processus de production.

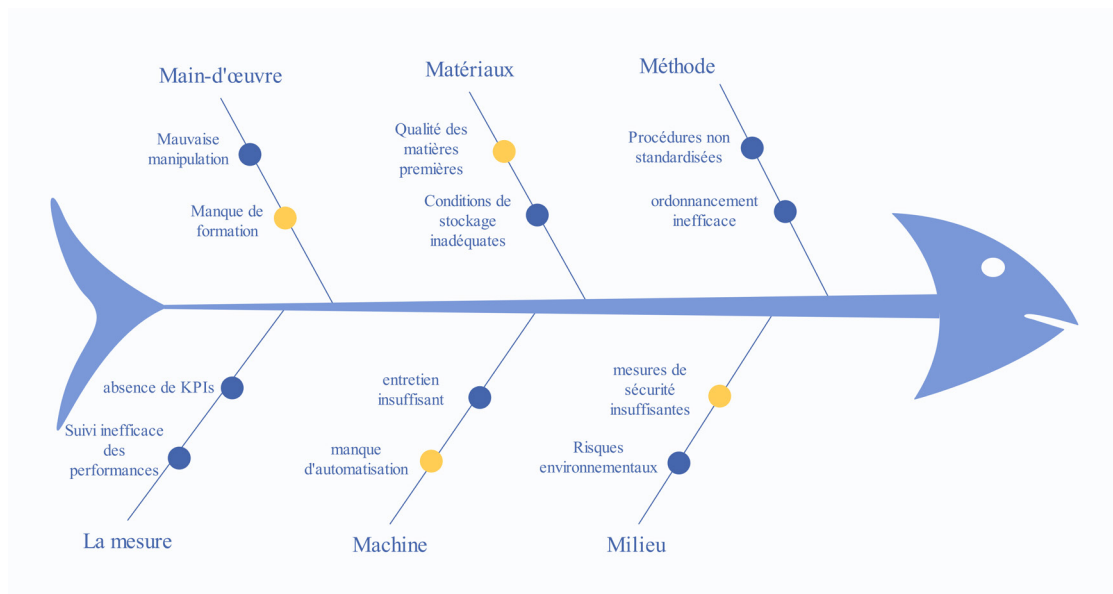


FIG. 2.5 : Diagramme d'Ishikawa du processus de production

b. Processus de commercialisation :

Le service commercial de SPA Hamoud Boualem est essentiel pour maintenir un

contact efficace avec plus de 800 clients. Supervisé par la direction générale, il gère les ventes locales et les exportations, assurant la satisfaction des clients.

La stratégie de ciblage est définie par plusieurs segments de marché. Les grands comptes achètent en grandes quantités, le commerce moderne vise les grandes surfaces comme UNO, Carrefour, et ARDIS, les dépositaires sont les grossistes, et la distribution directe concerne les clients utilisant leurs propres moyens de transport.

Le processus de commercialisation débute par la réception de la commande par divers moyens (appel, SMS, mail, déplacement). Ensuite, les stocks sont vérifiés pour garantir la disponibilité des produits. Un bon de commande est alors établi, détaillant la destination, la quantité, les délais et le type de transport. Un tableau de location est créé pour organiser la livraison.

Enfin, le produit est livré au client, complétant ainsi le processus de commercialisation. Ce système bien structuré permet à SPA Hamoud Boualem de répondre efficacement aux besoins des clients, assurant une satisfaction élevée.

• **Diagnostic du processus de commercialisation** : Parmi les problèmes de ce processus, les retards fréquents de livraison. En utilisant la méthode des 5 Pourquoi, nous pouvons identifier les causes profondes de ce problème et proposer des solutions ciblées pour réduire les retards de livraison.

On applique la méthode des 5 pourquoi comme suit :

Problème Identifié : Retards fréquents de livraison

Pourquoi? Parce que les commandes ne sont pas prêtes à temps pour l'expédition.

Pourquoi? Parce que le traitement des commandes prend plus de temps que prévu.

Pourquoi? Parce que les stocks ne sont pas toujours disponibles au moment de la préparation des commandes.

Pourquoi? Parce qu'il y a des inexactitudes dans les niveaux de stock enregistrés.

Pourquoi? Parce que les systèmes de gestion des stocks ne sont pas mis à jour en temps réel = **la cause racine**.

Les retards fréquents de livraison sont causés par des inexactitudes dans les niveaux de stock en raison de systèmes de gestion des stocks qui ne sont pas mis à jour en temps réel, entraînant des délais prolongés pour le traitement des commandes.

c. Processus de l'approvisionnement :

Le processus d'approvisionnement chez Hamoud Boualem comprend onze étapes clés pour garantir une gestion efficace des ressources et répondre aux besoins des clients.

La première étape consiste à déterminer la demande nationale pour chaque produit en utilisant des outils de prévision et des données historiques pour estimer la demande des clients pour chaque boisson gazeuse. Ensuite, il s'agit de déterminer la quantité de chaque produit à fabriquer sur chaque site, en tenant compte des prévisions nationales, des capacités de production, des stocks existants et des transferts inter-sites.

La troisième étape est de recevoir du fournisseur les données pour chaque matière

première. Cela implique d'obtenir des informations essentielles comme le lead time, la DLUO et la MOQ pour chaque matière première nécessaire à la production. Ensuite, il faut calculer la quantité de matière première nécessaire pour chaque site, en se basant sur le plan de production (PDP) et en tenant compte d'un taux de perte de 3

Une demande d'achat (DA) est créée pour chaque matière première identifiée, spécifiant les détails obtenus auprès des fournisseurs. Le service achat approuve ensuite la DA, vérifie sa conformité et négocie les conditions d'achat. Une fois approuvée, un ordre d'achat est émis aux fournisseurs sélectionnés, détaillant les quantités, les prix et les délais de livraison.

Parallèlement, il est crucial de déterminer la quantité d'emballage nécessaire pour chaque site en se basant sur le PDP, et de gérer l'approvisionnement de l'emballage par site pour optimiser les stocks et minimiser les coûts de stockage.

Un suivi précis des stocks des articles et des consommations est essentiel pour connaître les niveaux de stock en temps réel et prendre des mesures correctives si nécessaire. Enfin, le plan d'approvisionnement doit être ajusté régulièrement en fonction de l'évolution de la demande, des stocks et des performances des fournisseurs, garantissant ainsi que l'approvisionnement reste flexible et réactif aux changements du marché.

- **Diagnostic du processus d'approvisionnement** : En utilisant l'analyse fonctionnelle, nous avons identifié plusieurs problèmes potentiels dans le processus d'approvisionnement de Hamoud Boualem, comme présenté dans la table 2.1. Les principaux problèmes incluent des erreurs de prévision de la demande, une mauvaise gestion des stocks, des retards dans la création et l'approbation des demandes d'achat, et des inefficacités dans la communication avec les fournisseurs. Chaque problème a un impact direct sur l'efficacité, les coûts, et la satisfaction des clients. Les fonctions avec une criticité élevée sont cruciales pour la performance globale du processus d'approvisionnement et nécessitent une attention particulière pour éviter des perturbations majeures.

Fonction	Problèmes Potentiels	Impact	Criticité
Déterminer la demande nationale pour chaque produit	Erreurs de prévision de la demande	Surproduction ou sous-production, gaspillage de ressources, insatisfaction des clients	5
Déterminer la quantité de chaque produit à fabriquer sur chaque site	Mauvaise répartition de la production, capacités de production mal évaluées, manque de coordination	Inefficacité des ressources, coûts élevés, déséquilibres dans les niveaux de stock	5
Recevoir du fournisseur les données pour chaque matière première	Données incorrectes ou incomplètes, retards, manque de transparence	Problèmes de planification, délais de production, coûts supplémentaires	5

Calculer la quantité de matière première nécessaire pour chaque site	Estimations erronées, prise en compte insuffisante des pertes	Surstocks ou ruptures de stock, coûts de stockage élevés, inefficacité dans la production	3
Créer une demande d'achat (DA) pour chaque matière première	DA incorrectes ou incomplètes, délais dans la création, mauvaise communication	Retards dans la passation des commandes, problèmes de disponibilité des matières premières, coûts élevés	3
Approbation de la DA par le service achat	Délais d'approbation, manque de vérification	Retards dans l'approvisionnement, augmentation des coûts d'achat, risques de non-conformité	3
Passer un ordre d'achat après validation de la DA	Ordres incorrects ou retards, mauvaise communication	Retards dans la réception des matières premières, problèmes de production, coûts supplémentaires	3
Déterminer la quantité d'emballage nécessaire pour chaque site	Mauvaise estimation des besoins, prise en compte insuffisante des variations de production	Ruptures de stock d'emballage, problèmes de production, coûts supplémentaires	2
Gérer l'approvisionnement de l'emballage par site	Gestion inefficace des stocks d'emballage, surstocks ou ruptures de stock	Coûts de stockage élevés, inefficacité dans la production, perte de matières d'emballage	2
Suivre les stocks des articles par site et les inventaires	Inexactitudes dans les inventaires, manque de visibilité, suivi inefficace	Problèmes de production, coûts supplémentaires, inefficacité dans la gestion des stocks	2
Ajuster le plan d'approvisionnement si nécessaire	Manque de flexibilité, retards, mauvaise coordination	Surstocks ou ruptures de stock, problèmes de production, insatisfaction des clients	5

TAB. 2.1 : Analyse fonctionnelle du processus d'approvisionnement

2.2.7 Sélection du processus pour appliquer le Process mining

Pour justifier le choix d'un processus pour l'application de l'étude de process mining chez Hamoud Boualem, nous avons comparé ces trois processus clés (production, com-

mercialisation et approvisionnement) en utilisant la technique de BIAIS (Besoin, Impact, Accessibilité, Intégration, Support) ???. Cette comparaison évalue chaque processus selon ces critères pour déterminer lequel est le plus critique et bénéfique à optimiser.

La table 2.2 présente une comparaison entre ces processus :

Critère	Production	Commercialisation	Approvisionnement
Besoin	Critique pour la transformation des matières premières en produits finis.	Essentiel pour la satisfaction des clients et la gestion des commandes.	Critique pour assurer la disponibilité des matières premières nécessaires à la production.
	Criticité : 5	Criticité : 4	Criticité : 5
Impact	Impact direct sur la qualité et l'efficacité de la production.	Impact sur la satisfaction client et les performances commerciales.	Impact sur les coûts, la disponibilité des stocks et la production.
	Criticité : 5	Criticité : 4	Criticité : 5
Accessibilité	Données souvent bien documentées et accessibles.	Données disponibles mais parfois fragmentées.	Données bien documentées et accessibles.
	Criticité : 3	Criticité : 3	Criticité : 3
Intégration	Fortement intégrée avec les processus de maintenance et de qualité.	Lié aux processus de gestion des clients et des ventes.	Intégrée à la production, logistique et ventes.
	Criticité : 4	Criticité : 3	Criticité : 5
Support	Support fort de la direction pour maintenir la qualité de la production.	Support modéré pour améliorer la satisfaction client.	Support fort de la direction pour optimiser les coûts et assurer la disponibilité des stocks.
	Criticité : 5	Criticité : 3	Criticité : 5

TAB. 2.2 : Table comparative entre les processus

En comparant les trois processus selon la technique de BIAIS, le **processus d'approvisionnement** se distingue comme le choix le plus justifié pour l'application de l'étude de process mining chez Hamoud Boualem. Ce processus présente une criticité élevée (5) dans presque toutes les catégories, en particulier en termes d'intégration avec d'autres processus, et de support de la direction. L'optimisation du processus d'approvisionnement a le

potentiel d'améliorer significativement l'efficacité opérationnelle, de réduire les coûts, et d'assurer une meilleure gestion des stocks, impactant positivement l'ensemble de la chaîne de valeur.

2.3 Analyse du processus choisi

2.3.1 Etapes du processus

Le processus d'approvisionnement chez Hamoud Boualem se déroule en 11 étapes clés, qui sont les suivantes :

1. **Déterminer la demande nationale pour chaque produit** : Cette étape implique l'utilisation d'outils de prévision et de données historiques pour estimer précisément la demande des clients pour chaque boisson gazeuse sur une période donnée. Cela permet de planifier la production et l'approvisionnement en conséquence.
2. **Déterminer la quantité de chaque produit à fabriquer sur chaque site** : En tenant compte des prévisions nationales, des capacités de production de chaque site, des stocks existants et des transferts inter-sites, cette étape détermine la quantité de chaque boisson gazeuse à produire sur chaque site. Cela garantit une utilisation optimale des ressources de production et évite les surproductions ou les sous-productions.
3. **Recevoir du fournisseur les données pour chaque matière première** : Cette étape consiste à obtenir des informations cruciales auprès des fournisseurs concernant le lead time, la DLUO et la MOQ pour chaque matière première nécessaire à la production des boissons gazeuses. Ces données sont essentielles pour le calcul des besoins en approvisionnement et la planification des commandes.
4. **Calculer la quantité de matière première nécessaire pour chaque site** : En se basant sur le plan de production (PDP) et en prenant en compte un taux de perte de 3 % pour les rebuts ou le gaspillage pendant la production, cette étape calcule la quantité exacte de chaque matière première nécessaire pour chaque site. Cela permet d'optimiser les achats et de minimiser les stocks inutiles.
5. **Créer une demande d'achat (DA) pour chaque matière première** : Pour chaque matière première identifiée, une demande d'achat (DA) est créée, spécifiant le lead time, la DLUO et la MOQ obtenus auprès des fournisseurs. Cette DA sert de base à la passation de commandes ultérieures.
6. **Approbation de la DA par le service achat** : Le service achat examine la DA, vérifie sa conformité aux exigences et négocie les conditions d'achat avec les fournisseurs. Cette étape garantit que les achats sont effectués dans les meilleures conditions possibles pour l'entreprise.
7. **Passer un ordre d'achat après validation de la DA** : Une fois la DA approuvée par le service achat, un ordre d'achat officiel est émis aux fournisseurs sélectionnés.

L'ordre d'achat détaille les quantités de chaque matière première à commander, les prix convenus et les délais de livraison attendus.

8. **Déterminer la quantité d'emballage nécessaire pour chaque site :** En se basant sur le plan de production (PDP), cette étape détermine la quantité d'emballage nécessaire pour protéger et transporter les boissons gazeuses produites sur chaque site. Cela garantit que l'emballage est disponible en quantité suffisante pour répondre à la demande de production.
9. **Gérer l'approvisionnement de l'emballage par site :** Cette étape implique la gestion efficace des stocks d'emballage sur chaque site, en tenant compte des codes articles et de la nomenclature des produits. Cela permet d'optimiser les niveaux de stock, d'éviter les ruptures de stock et de minimiser les coûts de stockage.
10. **Suivre les stocks des articles par site et les consommations :** Un suivi précis des stocks d'articles, y compris les matières premières et l'emballage, est crucial pour chaque site. Cela permet de connaître les niveaux de stock en temps réel, d'identifier les écarts potentiels et de prendre des mesures correctives si nécessaire.
11. **Ajuster le plan d'approvisionnement si nécessaire :** Le plan d'approvisionnement doit être revu et ajusté régulièrement en fonction de l'évolution de la demande, des stocks, des performances des fournisseurs et d'autres facteurs pertinents. Cela garantit que l'approvisionnement reste flexible et réactif aux changements du marché.

En suivant rigoureusement ces 11 étapes, Hamoud Boualem peut garantir un processus d'approvisionnement fluide, efficace et rentable pour ses boissons gazeuses. Cela permet à l'entreprise de répondre à la demande des clients en temps et en heure, d'optimiser ses coûts, de minimiser les ruptures de stock.

2.3.2 Parties prenantes du processus

Ce processus d'approvisionnement implique plusieurs parties prenantes clés qui jouent un rôle crucial dans sa réussite. Chacune de ces parties prenantes a des responsabilités et des contributions spécifiques qui contribuent au bon déroulement du processus.

1. Service des ventes et supply chain :

- Responsable de la prévision de la demande des clients pour chaque produit sur.
- Collabore avec le service de production pour déterminer le plan de production (PDP) en fonction des prévisions de ventes.

2. Service de production :

- Élabore le plan de production (PDP) en tenant compte des capacités de production, des stocks existants et des transferts inter-sites.

- Communique le PDP au service supply chain pour qu'il puisse lancer le processus d'approvisionnement des matières premières et de l'emballage.

3. Service des achats :

- Envoie des demandes d'information (RFI) aux fournisseurs pour obtenir des données sur le lead time, la DLUO et la MOQ pour chaque matière première.
- Analyse les offres des fournisseurs et négocie les conditions d'achat.
- Émet des commandes d'achat aux fournisseurs sélectionnés.
- Suit les commandes d'achat et assure le respect des délais de livraison.

4. Fournisseurs :

- Fournissent les matières premières et l'emballage selon les spécifications et les délais convenus.
- Communiquent toute information pertinente concernant les retards de livraison, les changements de qualité ou les problèmes d'approvisionnement.

5. Direction générale :

- Supervise le processus d'approvisionnement et veille à son bon fonctionnement.
- Alloue les ressources nécessaires au bon déroulement du processus.
- Prend des décisions stratégiques pour optimiser la chaîne d'approvisionnement.

La collaboration efficace entre toutes ces parties prenantes est essentielle pour garantir un approvisionnement fluide, efficace et rentable des boissons gazeuses chez Hamoud Boualem. Une communication claire, un partage d'informations transparent et une résolution proactive des problèmes sont des éléments clés pour assurer le succès de ce processus collaboratif.

2.3.3 Documentation du processus

Avant de procéder à la modélisation, il est essentiel de bien comprendre le processus. La table 2.3 fournit une vue détaillée des principales informations et étapes du processus d'approvisionnement.

Nom de processus :	Processus d'approvisionnement		
Propriétaire du processus :	Ridha BRAHIM MAZARI		
Créé par :	Ridha BRAHIM MAZARI	Dernière mise à jour par :	Ridha BRAHIM MAZARI
Date de création :	11/05/2024	Date de la dernière révision :	21/06/2024
Objectif du processus :	Ce processus d'approvisionnement a pour objectif de garantir la disponibilité des produits finis en quantités suffisantes et au moment opportun pour répondre à la demande des clients, tout en optimisant les coûts et l'efficacité de la chaîne d'approvisionnement.		
Portée du processus :	Ce processus s'applique à l'ensemble des activités d'approvisionnement de l'entreprise, depuis la prévision des ventes jusqu'à la livraison des produits finis aux clients.		
Entrée de processus :	Prévisions des ventes ; Plan de production (PDP) par site de production ; Données des fournisseurs ; Stocks d'articles par site ; Consommations d'articles par site		

Limites du processus :	Les activités qui suivent immédiatement l'entrée du processus et qui précèdent immédiatement la sortie du processus définissent les limites du processus d'approvisionnement. Par conséquent, la limite initiale du processus est définie par les responsables d'approvisionnement. La limite finale du processus est définie par l'ajustement et le suivi du stock.
Flux de processus :	<ol style="list-style-type: none"> 1. Le processus commence une fois la date d'approvisionnement planifiée est arrivée. 2. On détermine d'abord la demande nationale pour chaque produit 3. Une fois déterminée, on détermine la quantité de chaque produit à fabriquer sur chaque site. 4. On fait ensuite recevoir les données des fournisseurs. 5. Sur la base de ces données, on calcule la quantité de matière première nécessaire pour chaque site. 6. Une demande d'achat doit être créée et envoyée au service achat pour l'approuver. 7. Après l'approbation, le service achat passe un ordre d'achat. 8. On fait ensuite déterminer la quantité d'emballage nécessaire pour chaque site et on gère l'approvisionnement de l'emballage par site. 9. On suit les stocks des articles par site et on effectue un ajustement sur le plan si nécessaire.
Sortie du processus :	La sortie du processus est le plan d'approvisionnement ajusté et optimisé.
Exceptions au déroulement normal du processus :	Des événements imprévus, tels que des catastrophes naturelles ou des pénuries de matières premières, peuvent perturber le déroulement normal du processus. Dans de tels cas, il est nécessaire de prendre des mesures correctives pour minimiser les impacts sur la chaîne d'approvisionnement.
Points de contrôle et mesures :	Précision des prévisions de ventes ; Taux de respect des délais de livraison ; Niveau des stocks d'articles ; Taux de perte ; Coût d'approvisionnement par produit ; Taux de rupture de stock ; Niveau de satisfaction

TAB. 2.3 : Documentation du processus d'approvisionnement

2.3.4 Modélisation du processus

Pour illustrer de manière claire et précise, le processus a été modélisé en utilisant le langage BPMN (Business Process Model and Notation). Cette modélisation permet de visualiser les différentes étapes et les interactions entre les services impliqués, notamment le service Supply Chain et le service achat, ainsi que les échanges avec les fournisseurs. La figure 2.6 présente le modèle détaillé du processus d'approvisionnement.

(Pour une meilleure visualisation, voir annexe C).

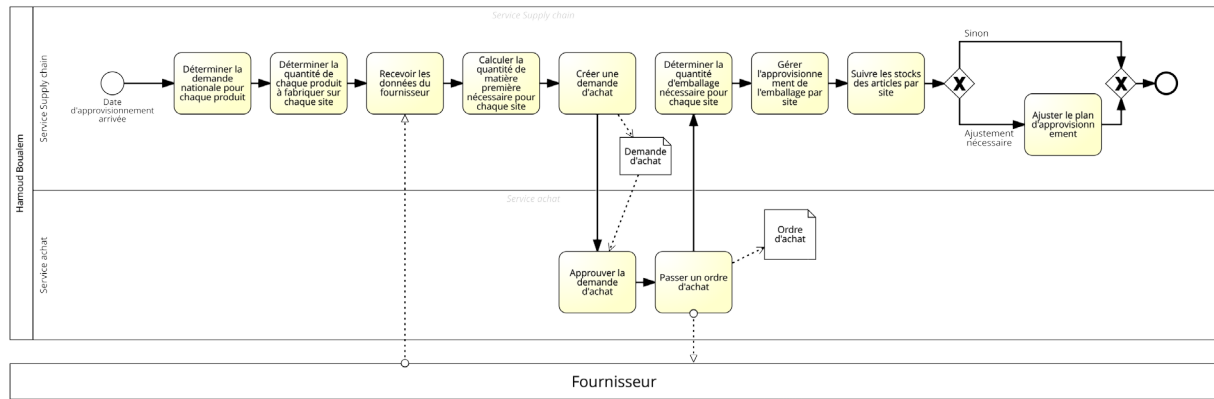


FIG. 2.6 : Modèle de processus d'approvisionnement en BPMN

2.3.5 Simulation du processus

La simulation de processus est un outil essentiel pour comprendre et optimiser les flux de travail dans les chaînes d'approvisionnement. Elle permet de modéliser les différentes étapes du processus, d'évaluer les performances et d'identifier les éventuelles déviations. Dans cette section, nous présentons la simulation du processus d'approvisionnement de notre étude, en utilisant l'outil Simul8¹.

Le tableau 2.4 présente les statistiques descriptives des durées de chaque étape sur 4 cycles de simulation (un mois) :

Étape	Moyenne (jours)	Écart-type (jours)	Temps total par cycle (jours)
Déterminer la demande nationale pour chaque produit	1.05	0.07	4.18
Recevoir les données du fournisseur	0.51	0.06	2.04
Déterminer la quantité de chaque produit	2.13	0.21	8.52
Calculer la quantité de matière première	1.44	0.18	5.78
Créer une demande d'achat	0.46	0.04	1.84
Approuver la demande d'achat	1.04	0.09	4.16
Passer un ordre d'achat	0.53	0.08	2.12
Déterminer la quantité d'emballage	0.98	0.07	3.92
Gérer l'approvisionnement de l'emballage	1.44	0.22	5.78
Suivre les stocks des articles par site	1.1	0.14	4.4
Ajuster le plan d'approvisionnement	2.06	0.23	8.24

TAB. 2.4 : Simulation du processus d'approvisionnement

¹<https://www.simul8.com/>

La simulation a permis de quantifier les durées pour chaque étape du processus et d'identifier les étapes critiques. La variabilité des durées, notamment pour les étapes telles que "Calculer la quantité de matière première nécessaire" et "Ajuster le plan d'approvisionnement", indique des opportunités d'amélioration pour réduire les écarts et optimiser le flux de travail.

2.3.6 Problèmes rencontrés dans l'approvisionnement

- **Complexité de la Chaîne d'approvisionnement** : Lorsque une entreprise produit 15 produits nécessitant chacun 15 matières premières, le responsable d'approvisionnement doit gérer 225 articles. Si chaque article provient d'un fournisseur différent, cela nécessite de négocier et d'évaluer avec 225 fournisseurs, ce qui peut entraîner des retards, des ruptures de stock, des erreurs de communication et des problèmes de qualité, et c'est le cas de Hamoud Boualem.
- **Fluctuations de la Demande** : Les prévisions de demande ne sont jamais totalement précises. Les variations soudaines de la demande nécessitent une augmentation rapide de la production, mais les fournisseurs peuvent ne pas être suffisamment flexibles pour répondre à ces changements.
- **Problèmes Logistiques** : Les retards de livraison et les ruptures de stock peuvent être causés par des problèmes de transport, des accidents ou des pannes.
- **Problèmes de Stockage** : Le manque d'espace de stockage, une mauvaise gestion des stocks et la détérioration des produits stockés peuvent perturber l'approvisionnement.
- **Problèmes de Qualité** : Les produits de mauvaise qualité nécessitent des remplacements ou des éliminations, entraînant des coûts supplémentaires, des ruptures de stock et des retards.
- **Problèmes de Financement** : Des retards de paiement peuvent entraîner la perte de fournisseurs, des ruptures de stock et une détérioration de la qualité en raison du changement de fournisseur.
- **Catastrophes Naturelles, Crises Sanitaires et Conflits Politiques** : Ces événements peuvent perturber l'approvisionnement en affectant l'importation et l'exportation, augmentant les prix, provoquant des problèmes de transport, des retards et la perte de fournisseurs (COVID-19 à titre d'exemple).

2.4 Préparation au process mining et analyse des données

2.4.1 Sources de données

L'acquisition des données nécessaires à l'analyse de process mining s'est avérée complexe en raison de l'hétérogénéité des systèmes d'information et des restrictions liées à la

confidentialité des données. En conséquence, un accès direct aux données réelles de l'entreprise n'a pas été possible. Pour surmonter cette limitation, une approche alternative basée sur la génération de données simulées a été adoptée.

Un script Python (voir Annexe G) a été développé pour générer des données synthétiques, en tenant compte de la structure et des caractéristiques typiques du processus d'approvisionnement de "Hamoud BOUALEM". Cet ensemble de données est constitué de traces d'exécution de différentes activités, en respectant les dépendances et les séquences logiques des étapes du processus. Chaque trace représente un cas unique avec des activités et des timestamps générés de manière aléatoire mais cohérente, permettant ainsi de modéliser des scénarios réalistes pour l'application des techniques de process mining. Les données simulées ont été structurées de manière à refléter fidèlement les opérations réelles de l'entreprise. Cette structuration minutieuse englobait l'ordre chronologique des activités accomplies, les indicateurs pertinents associés à chaque étape, ainsi que l'ensemble des attributs caractéristiques des opérations.

L'algorithme pour générer les données est le suivant :

1. Initialiser les listes pour stocker les données :

- case_id, timestamp, activité, site, demande_nationale, prévision, quantité_à_fabriquer, capacités_de_production, stocks_existants, lead_time, dl_uo, moq, quantite_necessaire, taux_de_perte, da_cree, da_approuvee, ordre_achat, quantite_emballage, gestion_stocks_emballage, consommations, ajustement_plan.

2. Définir les sites possibles :

- sites = ['SA', 'SB', 'SC'] (Les 3 sites mentionnés précédemment de l'entreprise : Boufarik, Hassiba et Oran)

3. Définir la date de départ :

- start_date = 1er janvier 2022

4. Définir les activités et leurs dépendances :

- activités = 'Activité A' : None, 'Activité B' : 'Activité A', ...

5. Définir les paramètres du dataset :

- num_cases = 2000 (Nombre total de cas)
- avg_activities_per_case = 6.5 (Nombre moyen d'activités par cas)
- week_delta = 1 semaine
- activity_prob = 0.8 (Probabilité d'inclure chaque activité dans une trace)

6. Pour chaque cas (de 0 à num_cases - 1) :

- Choisir un site aléatoirement parmi sites ;

- Définir la date de début du cas en ajoutant un multiple de `week_delta` à `start_date` ;
- Sélectionner aléatoirement un sous-ensemble d'activités avec une longueur aléatoire (entre 4 et `avg_activities_per_case`) ;
- Trier les activités sélectionnées selon l'ordre des dépendances.

7. Pour chaque activité dans le sous-ensemble sélectionné :

- Si l'activité a une dépendance qui n'a pas été incluse précédemment, sauter cette activité ;
- Générer un `case_id` unique pour le site et le cas ;
- Générer un timestamp aléatoire pendant les heures de travail (8-12 AM et 1-4:30 PM) ;
- Ajouter les valeurs générées pour chaque attribut dans les listes correspondantes.

8. Créer un DataFrame à partir des listes de données.

9. Enregistrer le DataFrame dans un fichier CSV.

L'algorithme génère un ensemble de données simulant des cas d'exécution de processus logistique en suivant une séquence d'activités avec des dépendances. Chaque cas est associé à un site et commence à une date de départ spécifique, incrémentée de manière hebdomadaire pour chaque nouveau cas. Les activités sont sélectionnées aléatoirement tout en respectant les dépendances définies. Pour chaque activité sélectionnée, un timestamp aléatoire est généré pendant les heures de travail, et des valeurs aléatoires sont attribuées aux attributs associés à cette activité. Ces données sont ensuite organisées dans un DataFrame et exportées en fichier CSV, fournissant ainsi un ensemble de données cohérent et prêt pour l'analyse de process mining.

L'utilisation de données simulées présente plusieurs avantages dans ce contexte. Elle permet de contourner les obstacles liés à l'accès aux données réelles et de garantir la confidentialité des informations sensibles. De plus, elle offre la possibilité de générer des scénarios hypothétiques et de tester l'impact de différentes configurations sur les processus logistiques.

Néanmoins, il est important de souligner que les données simulées ne constituent pas une réplique parfaite de la réalité. Une validation ultérieure avec des données réelles, lorsque cela sera possible, est recommandée pour confirmer la pertinence des résultats obtenus.

2.4.2 Nettoyage et prétraitement des données

Avant de procéder à l'analyse, une étape de nettoyage et de prétraitement des données simulées a été réalisée afin de garantir leur qualité et leur fiabilité. Cette étape visait à éliminer les anomalies, les incohérences et les informations superflues causées par les fonctions aléatoires du script utilisé, qui auraient pu altérer les résultats de l'analyse.

Les principales actions de nettoyage menées incluent :

- **Détection et suppression des doublons** : Les fonctions aléatoires peuvent générer des doublons et des enregistrements redondants qui doivent être identifiés et éliminés pour garantir l'unicité des données.
- **Traitement des valeurs manquantes** : Analyse et imputation des valeurs manquantes pour assurer la cohérence des ensembles de données.
- **Normalisation des timestamps** : Uniformisation du format des horodatages pour faciliter la comparaison et l'analyse temporelle.
- **Sélection des attributs pertinents** : Identification et conservation des attributs essentiels à l'analyse de process mining, comme l'identifiant, le nom de l'activité et le temps d'enregistrement.

Ces opérations de nettoyage et de prétraitement ont permis d'obtenir un ensemble de données simulées de qualité, prêt pour l'analyse de process mining. La rigueur appliquée à cette étape a contribué à fiabiliser les résultats obtenus et à renforcer la validité des conclusions tirées de l'étude.

2.4.3 Description du journal d'événements

Le journal d'événements, intitulé `process_data_extended.csv`, contient des enregistrements détaillés des activités liées au processus d'approvisionnement chez "Hamoud BOUALEM". Ce journal a été généré de manière simulée pour fournir une base de données représentative permettant l'analyse de process mining et couvre une période de deux ans. Chaque événement enregistré dans ce journal inclut un horodatage (timestamp) qui permet de situer précisément l'activité dans cette période. Le fichier contient un total de 5227 traces. Chaque ligne représente un événement unique dans le processus d'approvisionnement, incluant des informations cruciales telles que l'identifiant de cas, l'activité, le site, et d'autres attributs pertinents pour l'analyse.

Voici dans la table 2.5 un exemple du journal d'événements avec quelques activités extraites du fichier généré et traité avec quelques attributs pertinents :

case_id	timestamp	activité	site	demande nationale	prévision
SA0001	2022-01-01 08:15:30	Déterminer la demande nationale pour chaque produit	SA	1500	1200
SA0001	2022-01-02 09:30:45	Déterminer la quantité de chaque produit à fabriquer	SB	1500	1200
SA0001	2022-01-03 13:22:10	Recevoir du fournisseur les données concernant le lead time, la DLUO et la MOQ	SC	1500	1200
SA0002	2022-01-04 14:45:00	Calculer la quantité de matière première nécessaire	SA	1500	1200

TAB. 2.5 : Exemple du journal d'événements

2.4.4 Logiciels utilisé pour l'analyse

L'analyse a été réalisée en utilisant les outils ProM et Disco. ProM, grâce à ses capacités avancées et sa flexibilité, a permis d'explorer en profondeur les logs d'événements. Son large éventail de plug-ins pour la découverte de processus, la vérification de la conformité et l'amélioration des processus en a fait un choix idéal pour cette étude. Disco, connu pour sa rapidité et son ergonomie, a facilité le traitement et l'analyse des données. L'utilisation de ces deux outils puissants a permis une analyse complète et efficace des processus.

Après avoir importé le journal d'événements dans l'outil ProM, nous avons obtenu une vue d'ensemble des événements du journal et de leurs occurrences, comme le montre la figure 2.7.

Cette table identifie et classe les événements selon leurs noms et montre leur fréquence d'occurrence, et permet d'identifier les activités clés et les points potentiels d'optimisation pour améliorer l'efficacité globale du processus.

Event Name		
Event classes defined by Event Name		
All events		
Total number of classes: 11		
Class	Occurrences (absolute)	Occurrences (relative)
Déterminer la quantité d'emballage nécessaire pour chaque site en fonction du PDP	918	17,563%
Recevoir du fournisseur les données concernant le lead time, la DLUO et la MOQ pour chaque matière première	909	17,39%
Déterminer la quantité de chaque produit à fabriquer sur chaque site en fonction du forecast national et les capacités de production. Prendre en compte les stocks existants et les transferts inter-sites	906	17,333%
Déterminer la demande nationale pour chaque produit	898	17,18%
Calculer la quantité de matière première nécessaire pour chaque site en fonction du PDP et du taux de perte (3%)	538	10,293%
Gérer l'approvisionnement de l'emballage par site en tenant compte des stocks, des codes articles et de la nomenclature des produits	362	6,926%
Créer une demande d'achat (DA) pour chaque matière première en précisant le lead time, la DLUO et la MOQ	280	5,357%
Approuver de la DA par le service achat	147	2,812%
Suivre les stocks des articles par site et les consommations	141	2,698%
Passer un ordre d'achat après validation de la DA	81	1,55%
Ajuster le plan d'approvisionnement si nécessaire	47	0,899%

FIG. 2.7 : Résumé du journal d'événements

2.5 Découverte du processus

La découverte de processus est une étape cruciale dans l'analyse des processus d'affaires, permettant d'extraire et de comprendre les flux de travail à partir de données événementielles. Dans ce rapport, on va se concentrer sur le processus d'approvisionnement, tel que décrit par les étapes mentionnées dans le chapitre précédent.

Pour la découverte de processus, on a utilisé les outils ProM et Disco. Ces outils permettent d'extraire les modèles de processus à partir des logs d'événements et de visualiser les flux de travail de manière claire et intuitive. Les données événementielles ont été importées dans ces outils pour analyser et visualiser le processus d'approvisionnement.

La figure 2.8 montre les résultats de la découverte de processus en appliquant l'algorithme alpha++ (Pour une meilleure visualisation, voir Annexe E) :

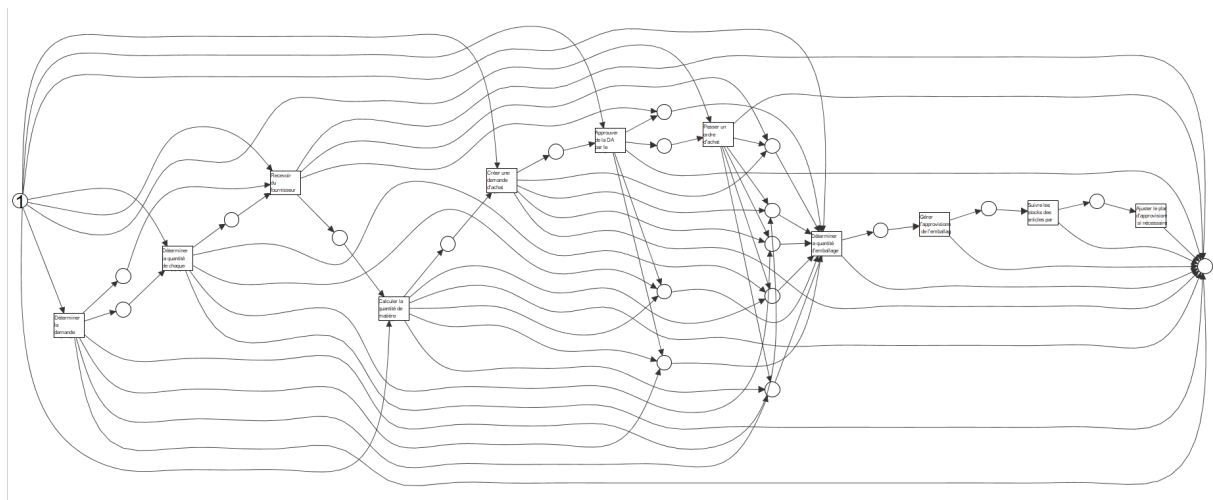


FIG. 2.8 : Modèle découvert avec alpha++

Cet algorithme a été appliqué en raison de sa capacité à découvrir des modèles de processus complexes et à gérer des structures comme les boucles imbriquées et la concurrence, courantes dans les environnements de chaîne d'approvisionnement. Alpha++ est également robuste face aux bruits et aux incohérences des données, ce qui garantit des résultats fiables.

L'algorithme Inductive Miner a également été utilisé, produisant la figure 2.9 :

FIG. 2.9 : Modèle découvert avec Inductive Miner

L'Inductive Miner a été également choisi en raison de sa capacité à produire des modèles de processus précis et compréhensibles, même à partir de logs d'événements complexes et bruyants. L'Inductive Miner excelle dans la gestion des structures de processus flexibles, telles que les boucles et la concurrence, tout en garantissant que les modèles générés sont toujours sons.

Une fois les données importées dans Disco, une représentation visuelle du processus a été automatiquement générée, comme le montre la figure 2.10 (Pour une meilleure

visualisation, voir Annexe F) :

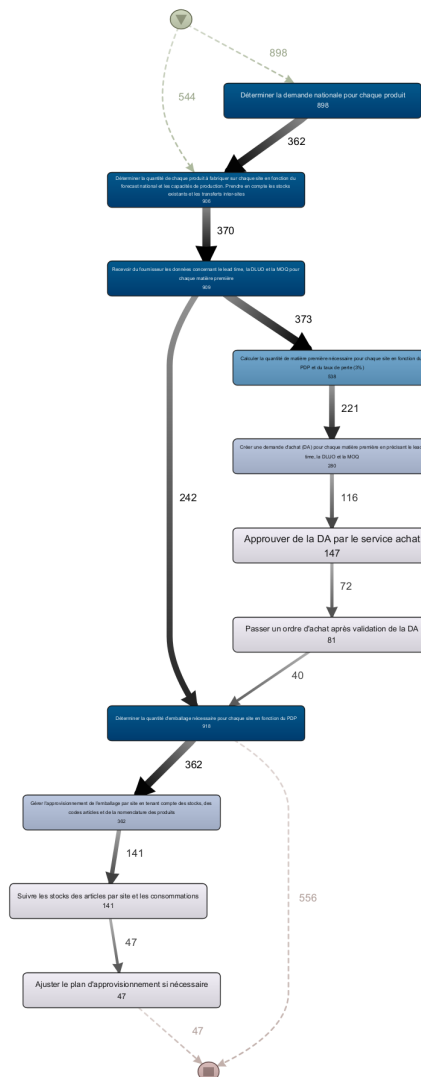


FIG. 2.10 : Modèle découvert avec Disco

Ce modèle montre une vue d'ensemble claire des activités et de leurs interconnexions, facilitant l'identification des séquences et des boucles dans le processus.

Les pointillés représentent les instances où les étapes prévues ont été contournées, menant à des chemins imprévus. Ces déviations indiquent des exceptions dans le processus, où les séquences d'actions attendues n'ont pas été suivies, révélant ainsi des anomalies ou des variations qui peuvent nécessiter une attention particulière pour améliorer la conformité et l'efficacité du processus.

Les différentes étapes du processus d'approvisionnement sont représentées par les places du diagramme de Petri (figure 2.10). Les places sont nommées par des étiquettes qui décrivent l'état du processus à cet instant. Les principales étapes du processus sont les suivantes :

1. Déterminer la demande nationale pour chaque produit (898 instances).
2. Déterminer la quantité de chaque produit à fabriquer sur chaque site en fonction du forecast national et les capacités de production. Prendre en compte les stocks existants et les transferts inter-sites (906 instances).
3. Recevoir du fournisseur les données concernant le lead time, la DL_UO et la MOQ pour chaque matière première (909 instances).
4. Calculer la quantité de matière première nécessaire pour chaque site en fonction du PDP et du taux de perte (3%) (538 instances).
5. Créer une demande d'achat (DA) pour chaque matière première en précisant le lead time, la DL_UO et la MOQ (280 instances).
6. Approuver de la DA par le service achat (147 instances).
7. Passer un ordre d'achat après validation de la DA (81 instances).
8. Déterminer la quantité d'emballage nécessaire pour chaque site en fonction du PDP (918 instances).
9. Gérer l'approvisionnement de l'emballage par site en tenant compte des stocks, des codes articles et de la nomenclature des produits (362 instances).
10. Suivre les stocks des articles par site et les consommations (141 instances).
11. Ajuster le plan d'approvisionnement si nécessaire (47 instances).

Les modèles de processus extraits révèlent plusieurs enseignements clés sur le flux de travail du processus d'approvisionnement. Voici une analyse détaillée des séquences d'activités, des boucles, et des points de décision critiques :

1. **Séquences d'Activités** : Les activités dans le processus d'approvisionnement sont principalement séquencées de manière linéaire, ce qui signifie que chaque étape du processus suit directement la précédente. Cette structure linéaire facilite la traçabilité et la gestion du flux de travail. Par exemple :
 - **Détermination de la Demande Nationale** : Cette étape initiale est suivie directement par la détermination de la quantité de chaque produit à fabriquer.
 - **Réception des Données Fournisseur** : Une fois les quantités déterminées, les informations du fournisseur sur le lead time, la DL_UO, et la MOQ sont directement intégrées pour calculer les besoins en matières premières.
 - **Création et Validation des DA** : La création des demandes d'achat (DA) suit le calcul des besoins, et chaque DA est ensuite soumise pour approbation.

Cette séquence linéaire assure une progression ordonnée du processus, réduisant les risques de retard ou d'erreur à chaque étape.

2. Boucles :

Certaines activités du processus présentent des boucles, indiquant des révisions fréquentes en fonction des consommations réelles et des conditions parfois changeantes. Par exemple :

- **Suivi des Stocks et Ajustement du Plan d'Approvisionnement** : Le suivi continu des stocks des articles par site et des consommations peut nécessiter des ajustements fréquents du plan d'approvisionnement, par exemple, un deuxième calcul de la quantité à fabriquer selon les nouvelles données du stock. Ces ajustements sont cruciaux pour maintenir des niveaux de stock optimaux et éviter les ruptures ou les surplus.

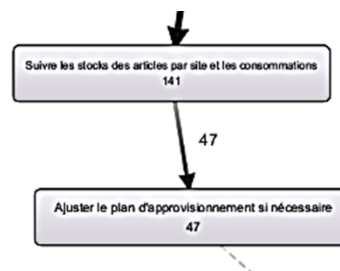


FIG. 2.11 : Boucle identifiée 1

- **Gestion de l'Emballage** : La gestion de l'approvisionnement en emballage peut également nécessiter des révisions en fonction des stocks disponibles et des besoins de production, on note 362 occurrences pour l'activité de gestion, suivies par seulement 141 occurrences pour l'activité suivante.

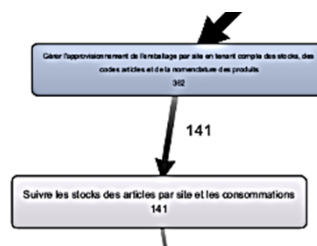


FIG. 2.12 : Boucle identifiée 2

Ces boucles sont essentielles pour la flexibilité et la réactivité du processus d'approvisionnement, permettant des ajustements en temps réel pour répondre aux variations de la demande et aux conditions de production.

3. Points de Décision :

Les modèles de processus révèlent plusieurs points de décision critiques qui influencent le déroulement du processus d'approvisionnement. Ces points de décision sont des étapes où des choix importants sont faits, affectant la direction et la progression du processus :

- **Réception des données des fournisseurs** : Lors de la réception des données des fournisseurs, il est possible de calculer les quantités nécessaires et de passer les commandes d'achat. En revanche, si les données concernant la matière première ne sont pas satisfaisantes, on passe directement à la détermination de la quantité d'emballage requise.



FIG. 2.13 : Point de décision identifié

- **Déterminer la quantité d'emballage nécessaire pour chaque site en fonction du PDP** : Les pointillés indiquent que le processus passe directement de la détermination de la quantité d'emballage nécessaire à la fin du processus, sans passer par les étapes intermédiaires de gestion de l'approvisionnement et de suivi des stocks. Cela pourrait indiquer des cas spéciaux ou des exceptions où les étapes intermédiaires ne sont pas nécessaires ou sont sautées pour des raisons spécifiques.

Ces points de décision sont des moments clés où les choix faits peuvent avoir un impact significatif sur l'efficacité globale et la performance du processus d'approvisionnement.

2.6 Analyse de Conformité

L'analyse de conformité consiste à comparer le modèle de processus découvert à un modèle de référence ou attendu, afin d'identifier les écarts et les conformités. Cette étape est essentielle pour garantir que les processus opérationnels se déroulent conformément aux attentes et pour identifier les opportunités d'amélioration.

Pour l'analyse de conformité, on a comparé le modèle de processus découvert avec le modèle de référence du processus d'approvisionnement déjà modélisé en BPMN. Les outils ProM et Disco ont été utilisés pour cette comparaison. Les écarts ont été identifiés et analysés pour comprendre les raisons sous-jacentes et les impacts potentiels sur les opérations.

L'analyse de conformité a révélé plusieurs conformités et écarts :

- **Analyse de Justesse** : Après avoir appliqué l'action « A* -ILP-based manifest replay » qui compare le modèle de référence avec le modèle découvert, l'algorithme a généré la figure 2.14.

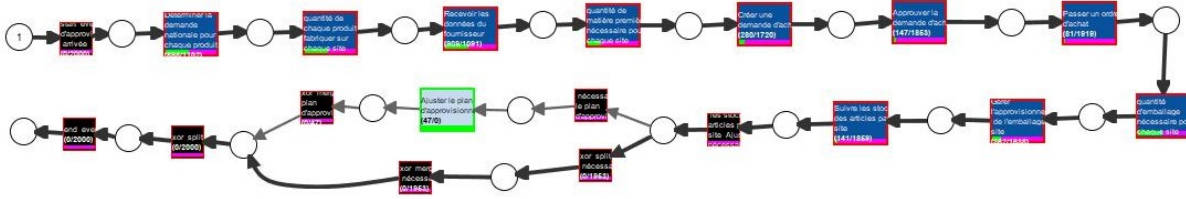


FIG. 2.14 : Analyse de justesse

La valeur de la justesse $F(W, L)$ a été déterminée :

$$F(W, L) = 0,39$$

On rappelle que W représente le Workflow Net et L une trace d'événements.

Cet indicateur mesure la conformité du modèle de processus aux traces d'événements observées. Avec une justesse de 0,39, cela signifie que le modèle reproduit seulement 39% des comportements observés dans le log d'événements généré. Cette faible valeur de justesse suggère que le modèle a des difficultés à capturer fidèlement les séquences d'activités et les transitions générées par le script Python. Les raisons possibles incluent des activités manquantes, des erreurs de séquençage, ou des divergences dans les règles de transition.

- **Analyse de Précision** : En appliquant l'action « Precision and Generalization » sur ProM, on a trouvé la valeur :

$$P(L, M) = 0,44$$

On rappelle que M représente le modèle de processus, et L une trace d'événements. La précision mesure la capacité du modèle à éviter les comportements non observés dans les logs d'événements. Une précision de 0,44 indique que le modèle inclut des transitions ou des séquences d'activités qui ne se produisent pas dans les données générées. Cette faible précision peut être due à un modèle trop permissif, autorisant des comportements qui ne sont pas reflétés dans le log d'événements simulé.

- **Analyse de Généralisation** : Après avoir appliqué l'action « Generalization Measure » sur ProM, on a trouvé :

$$G(L, M) = 0,50$$

La généralisation mesure la capacité du modèle à représenter des comportements possibles mais non observés dans le log d'événements. Une généralisation de 0,50 suggère que le modèle a une capacité modérée à généraliser les comportements. Cette valeur indique que le modèle inclut certains comportements non observés mais n'est pas excessivement général, ce qui peut aider à équilibrer la justesse et la précision.

- **Analyse de Simplicité** : En appliquant l'action « Measure of Structural Complexity » sur ProM, on a trouvé :

$$S(L, M) = 0,67$$

La simplicité mesure la complexité structurelle du modèle. Avec une valeur de 0,67, le modèle est relativement simple, ce qui est positif car un modèle simple est plus facile à comprendre et à gérer. Cependant, il est crucial de s'assurer que cette simplicité ne compromet pas la justesse et la précision du modèle.

Les résultats de l'analyse de conformité révèlent plusieurs points clés :

- **Justesse Faible** : Avec une justesse de 0,39, le modèle a des difficultés à correspondre aux comportements générés par le script Python.
- **Précision Faible** : Une précision de 0,44 indique que le modèle inclut de nombreux comportements non observés dans le log généré.
- **Généralisation Modérée** : Une généralisation de 0,50 montre que le modèle inclut certains comportements non observés mais reste modéré.
- **Simplicité Modérée** : Avec une simplicité de 0,67, le modèle maintient une complexité raisonnable.

Ainsi, l'analyse de conformité met en évidence les points forts et les faiblesses du modèle de processus actuel, fournissant des indications précieuses pour des améliorations futures. Étant donné que le log d'événements a été généré par Python, il est important de s'assurer que le script de génération reflète fidèlement les comportements attendus dans le processus pour obtenir des mesures de conformité plus précises.

• Causes possibles des écarts

Les problèmes d'approvisionnement peuvent avoir des origines diverses et complexes. Ils peuvent découler de facteurs internes, tels que des erreurs humaines ou des défaillances organisationnelles. Une compréhension approfondie des causes sous-jacentes de ces problèmes est essentielle pour mettre en place des solutions efficaces et durables. Voici une analyse détaillée des différentes raisons des problèmes d'approvisionnement.

- **Erreurs de Manipulation et Mauvaise Compréhension des Séquences** : La justesse faible (0,39) et la précision faible (0,44) indiquent que le personnel a des difficultés à exécuter correctement les étapes du processus. Ces faibles scores suggèrent une tendance aux erreurs de manipulation et à une mauvaise compréhension des séquences, ce qui entraîne des problèmes d'approvisionnement.
- **Résistance au Changement et Manque de Formation** : La faible justesse (0,39) et la précision (0,44) reflètent un manque de familiarité avec les nouvelles procédures. Cela est souvent dû à une résistance au changement et à un manque de formation adéquate, rendant le personnel incapable de suivre correctement les nouvelles méthodes d'approvisionnement.
- **Problèmes de Communication entre les Départements** : Le score de généralisation modéré (0,50) suggère que, bien que certaines pratiques soient correctement appliquées, des problèmes de communication persistent entre les départements. Ces malentendus peuvent causer des retards et des inefficacités dans l'approvisionnement.

- **Non-respect des Procédures** : La justesse faible (0,39) et la précision faible (0,44) révèlent que le personnel ne suit pas rigoureusement les procédures établies. Ce non-respect des procédures conduit à des erreurs et des retards, indépendamment des capacités des systèmes d'information.
- **Contraintes Réglementaires et Conditions Externes Imprévues** : Bien que la simplicité modérée (0,67) indique que les procédures sont relativement faciles à suivre, les contraintes réglementaires et les conditions externes imprévues contiennent de perturber les processus d'approvisionnement. Les changements dans les conditions réglementaires et les imprévus externes peuvent entraîner des retards et des coûts supplémentaires, affectant la régularité du flux d'approvisionnement.

2.7 Recommandations et optimisation

Pour optimiser le processus en tenant compte des écarts internes et externes, voici un plan d'optimisation détaillé avec des solutions spécifiques pour chaque cause identifiée :

- **Erreurs de Manipulation et Mauvaise Compréhension des Séquences**
Solution : Développer des guides détaillés et des procédures standardisées. Utiliser des vidéos de formation et des simulations pour améliorer la compréhension.
- **Résistance au Changement et Manque de Formation**
Solution : Mettre en place un programme de gestion du changement avec des sessions de formation régulières et des ateliers pour familiariser le personnel avec les nouvelles procédures.
- **Problèmes de Communication entre les Départements**
Solution : Utiliser des plateformes collaboratives et des outils de gestion de projet pour faciliter la communication et la coordination entre les départements. Organiser des réunions interdépartementales régulières.
- **Mauvaise Répartition des Responsabilités et Planification Inadéquate**
Solution : Clarifier les rôles et responsabilités avec des descriptions de poste détaillées. Utiliser des outils de planification avancée pour une meilleure allocation des ressources et planification des tâches.
- **Non-respect des Procédures**
Solution : Adopter une approche intégrée combinant formation, simplification des processus et suivi rigoureux.
- **Contraintes Réglementaires et Conditions Externes Imprévues**
Solution : Créer des plans de contingence pour répondre aux régulations changeantes et aux conditions imprévues. Mettre en place un système de veille réglementaire.

Le diagramme de la figure 2.15 présente une version optimisée du processus d’approvisionnement de l’entreprise ”Hamoud Boualem”, élaboré pour répondre aux défis identifiés dans l’analyse des écarts internes et externes. Cette optimisation vise à réduire les erreurs de manipulation, améliorer la communication entre les départements, et rendre le processus plus réactif aux fluctuations de la demande et aux problèmes logistiques. En intégrant des outils d’analyse prédictive, des systèmes de gestion des stocks en temps réel, et des flux de travail automatisés, ce modèle assure une gestion plus efficace et une meilleure coordination des activités d’approvisionnement. Les étapes clés du processus sont clairement définies pour garantir une exécution fluide et des ajustements proactifs en cas de besoin.

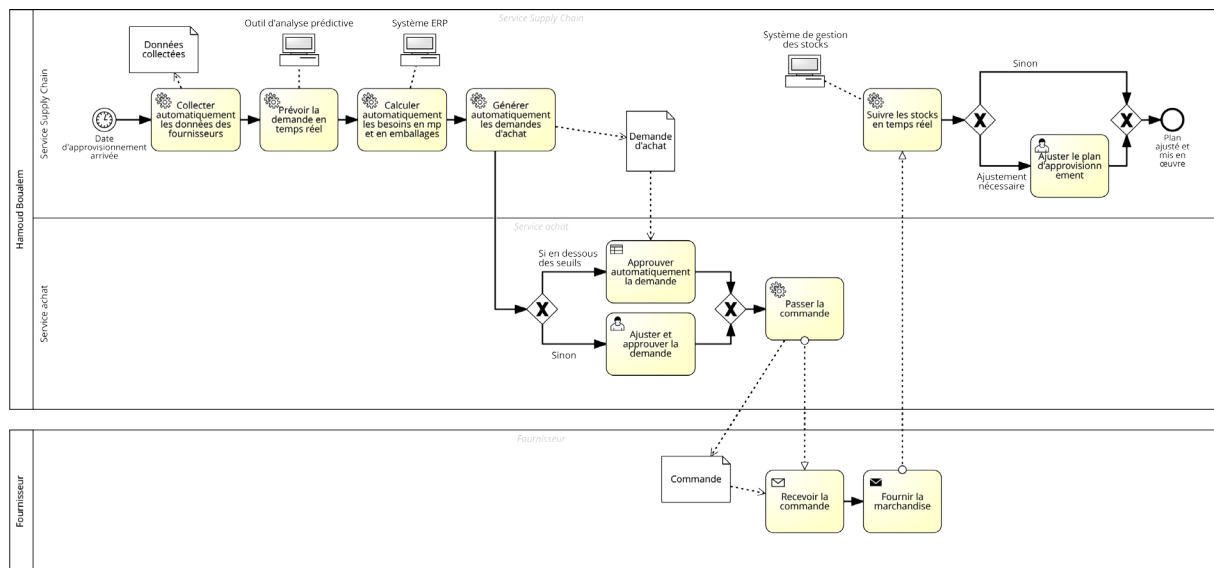


FIG. 2.15 : Modèle optimisé du processus d’approvisionnement

(Pour une meilleure visualisation, voir annexe D).

Le diagramme est structuré en trois principaux couloirs : Service Supply Chain, Service Achat, et Fournisseur, chacun étant responsable de tâches spécifiques qui contribuent à la fluidité et à l’efficacité globale du processus d’approvisionnement.

Service Supply Chain

- **Collecter automatiquement les données des fournisseurs** : Cette activité permet de collecter automatiquement les données des fournisseurs grâce à un système de gestion des fournisseurs (SGF). Cela garantit une collecte de données sans erreurs et en temps opportun, facilitant ainsi la planification et la gestion des approvisionnements.
- **Prévoir la demande en temps réel** : En utilisant des outils d’analyse prédictive, cette activité permet de prévoir la demande nationale en temps réel. Les prévisions précises aident à anticiper les besoins futurs et à éviter les ruptures de stock ou les surstocks.

- **Calculer automatiquement les besoins en matières premières et en emballages** : Cette activité calcule les besoins en matières premières et en emballages nécessaires pour chaque site de manière automatisée, basée sur les prévisions de demande. Cela assure une planification précise et réduit le risque de pénurie ou d'excédent.
- **Générer automatiquement les demandes d'achat** : Sur la base des calculs des besoins, cette activité génère automatiquement des demandes d'achat. Cela permet de lancer rapidement le processus d'approvisionnement sans intervention manuelle, augmentant ainsi l'efficacité.

Service Achat

- **Approuver automatiquement la demande** : Les demandes d'achat qui respectent les seuils prédéfinis sont automatiquement approuvées, réduisant ainsi les délais d'attente pour les approbations et accélérant le processus de commande.
- **Ajuster et approuver la demande** : Pour les demandes qui dépassent les seuils, une intervention humaine est nécessaire pour ajuster et approuver la demande. Cela permet une vérification supplémentaire et une flexibilité pour gérer des situations exceptionnelles.
- **Passer automatiquement les commandes** : Une fois les demandes approuvées, les commandes sont automatiquement passées aux fournisseurs, garantissant ainsi une transmission rapide et fiable des besoins d'approvisionnement.

Fournisseur

- **Recevoir la commande** : Les fournisseurs reçoivent les commandes et envoient un accusé de réception, confirmant ainsi la réception de la commande et permettant de suivre son statut.
- **Fournir la marchandise** : Les fournisseurs expédient les marchandises commandées à l'entreprise, assurant ainsi que les produits nécessaires sont livrés dans les délais.

Boucle de Suivi et Ajustement

- **Suivre les stocks en temps réel** : Cette activité permet de suivre les niveaux de stock en temps réel grâce à un système de gestion des stocks, ce qui facilite une gestion proactive et efficace des inventaires.
- **Ajuster le plan d'approvisionnement** : Basé sur les données de suivi des stocks, cette activité permet d'ajuster le plan d'approvisionnement pour éviter les ruptures ou les surstocks, assurant ainsi une continuité et une optimisation du processus d'approvisionnement.

Après l'optimisation, le processus d'approvisionnement a bénéficié de nombreuses améliorations technologiques. La collecte des données des fournisseurs est désormais automatique, ce qui réduit les erreurs et les délais. La demande est prédite en temps réel grâce à un outil d'analyse prédictive, permettant des ajustements rapides et précis aux changements de la demande. Les besoins en matières premières et en emballages sont calculés automatiquement via un système ERP, augmentant l'efficacité et la précision. Les demandes d'achat sont générées automatiquement, éliminant la nécessité de saisie manuelle et réduisant les délais. Le suivi des stocks est effectué en temps réel grâce à un système de gestion des stocks, améliorant la gestion des inventaires et réduisant les risques de ruptures ou de surstocks. L'approbation des demandes d'achat est automatisée pour les seuils définis, avec des ajustements manuels possibles si nécessaire, ce qui accélère le processus. Le processus de passation de commande est désormais détaillé et inclut la passation, la réception et la fourniture de la marchandise, assurant une meilleure traçabilité et coordination. Enfin, l'intégration accrue avec des outils d'analyse prédictive et des systèmes ERP améliore la synchronisation et l'optimisation du processus global.

• **Simulation du processus** Le tableau 2.6 présente les statistiques descriptives des durées de chaque étape sur les 4 cycles du processus optimisé :

Étape	Moyenne (jours)	Écart-type (jours)	Temps total par cycle (jours)
Collecter automatiquement les données des fournisseurs	0.48	0.03	1.92
Prévoir la demande en temps réel	0.96	0.07	3.84
Calculer automatiquement les besoins en matières premières et emballages	0.51	0.04	2.04
Générer automatiquement les demandes d'achat	0.54	0.05	2.16
Approuver automatiquement la demande	0.47	0.04	1.88
Ajuster et approuver la demande	1.04	0.1	4.16
Passer la commande	0.49	0.04	1.96
Suivre les stocks en temps réel	0.81	0.11	3.24
Ajuster le plan d'approvisionnement	1.48	0.1	5.92
Recevoir la commande	1.02	0.09	4.08
Fournir la marchandise	1.37	0.15	5.48

TAB. 2.6 : Simulation du processus optimisé

Les résultats montrent que le processus optimisé offre des durées plus courtes pour la plupart des étapes par rapport au processus original. Les étapes critiques telles que "Ajuster le plan d'approvisionnement" et "Fournir la marchandise" présentent toujours des variabilités, mais globalement, le processus optimisé est plus efficace.

2.8 Plan de la mise en œuvre des recommandations

2.8.1 Étapes

La mise en œuvre des recommandations s'articule autour de plusieurs étapes clés, chacune visant à améliorer un aspect spécifique du processus d'approvisionnement :

1. Automatisation de la collecte des données fournisseurs :

L'intégration des systèmes de gestion des fournisseurs (SGF) permet de collecter automatiquement les données nécessaires des fournisseurs. Cette étape vise à éliminer les erreurs manuelles et à accélérer le processus de collecte des données. La configuration des interfaces pour la collecte automatique des données garantit une mise à jour continue et précise des informations sur les fournisseurs.

2. Implémentation des outils d'analyse prédictive :

La sélection et le déploiement des logiciels d'analyse prédictive permettent de prévoir la demande en temps réel. Cette étape inclut également la formation du personnel à l'utilisation de ces outils pour s'assurer qu'ils peuvent interpréter et appliquer efficacement les prévisions. L'objectif est d'anticiper les besoins futurs et de planifier les approvisionnements en conséquence.

3. Calcul automatisé des besoins en matières premières et emballages :

En configurant des algorithmes de calcul basés sur les prévisions de demande, cette étape permet de déterminer automatiquement les quantités nécessaires de matières premières et d'emballages pour chaque site. L'intégration avec le système de planification des ressources de l'entreprise (ERP) garantit que les données sont cohérentes et facilement accessibles.

4. Génération automatique des demandes d'achat :

La mise en place des règles métier pour la génération des demandes d'achat automatisées permet de réduire les délais et d'assurer la cohérence des demandes. Cette étape inclut des tests rigoureux pour valider le processus automatisé et s'assurer qu'il fonctionne correctement avant de le déployer à grande échelle.

5. Automatisation du processus d'approbation :

En définissant des seuils pour l'approbation automatique, cette étape vise à accélérer le processus de validation des demandes d'achat. L'implémentation d'un système de gestion des approbations garantit que seules les demandes dépassant les seuils prédéfinis nécessitent une intervention humaine, réduisant ainsi les délais et les efforts administratifs.

6. Suivi et ajustement en temps réel des stocks :

Le déploiement d'un système de gestion des stocks en temps réel permet de surveiller en continu les niveaux de stock. La mise en place de tableaux de bord pour le suivi des stocks fournit des informations précises et actualisées, facilitant une gestion proactive des inventaires et des ajustements rapides en cas de besoin.

2.8.2 Ressources nécessaires

La mise en œuvre des recommandations nécessite des ressources variées :

Ressources humaines

- **Équipe de projet** : La mise en place des améliorations requiert une équipe de projet dédiée comprenant des chefs de projet, des analystes de processus, et des développeurs de logiciels. Ces professionnels sont essentiels pour planifier, exécuter et superviser chaque étape de la mise en œuvre.
- **Formation** : Des sessions de formation sont nécessaires pour préparer le personnel à utiliser les nouveaux systèmes et outils. Une formation adéquate garantit que les employés peuvent exploiter pleinement les capacités des technologies déployées.

Ressources matérielles

- **Infrastructures IT** : La mise à niveau des infrastructures IT, incluant des serveurs, des bases de données, et des équipements réseau, est cruciale pour supporter les nouveaux systèmes. Ces infrastructures doivent être robustes et sécurisées pour assurer une performance fiable et continue.
- **Logiciels** : L'acquisition des licences pour les outils d'analyse prédictive et les systèmes de gestion des fournisseurs est nécessaire. Ces logiciels jouent un rôle clé dans l'automatisation et l'optimisation des processus d'approvisionnement.

Ressources financières

- **Budget** : Un budget spécifique doit être alloué pour couvrir les coûts des matériels, des logiciels, et des formations. Une planification budgétaire rigoureuse est essentielle pour garantir que toutes les ressources nécessaires sont disponibles.
- **Investissement** : Des investissements financiers sont nécessaires pour la mise à niveau des infrastructures IT et la rémunération des consultants et des formateurs. Cet investissement est crucial pour assurer le succès de la mise en œuvre des améliorations.

2.8.3 Calendrier de mise en œuvre

Le calendrier de mise en œuvre est structuré en plusieurs phases pour assurer une intégration fluide et progressive des améliorations :

Phase 1 (0-3 mois) :

Préparation : Cette phase inclut la planification du projet, l'allocation des ressources, et la formation initiale du personnel. Les objectifs et les attentes sont clairement définis pour garantir une compréhension commune de la direction et des étapes à suivre.

Phase 2 (3-6 mois) :

Implémentation initiale : Durant cette phase, les systèmes de gestion des fournisseurs et les outils d'analyse prédictive sont déployés. Les configurations initiales sont testées et ajustées pour s'assurer qu'elles répondent aux besoins de l'entreprise.

Phase 3 (6-9 mois) :

Test et validation : Les processus automatisés pour le calcul des besoins et la génération des demandes d'achat sont mis en place. Des tests rigoureux sont effectués pour valider leur efficacité et identifier les ajustements nécessaires. Les résultats sont analysés pour garantir que les objectifs de performance sont atteints.

Phase 4 (9-12 mois) :

Suivi et ajustement : Le système de gestion des stocks en temps réel est déployé et les tableaux de bord pour le suivi sont mis en place. Des ajustements continus sont effectués en fonction des données collectées et des retours d'expérience pour optimiser le processus.

On peut résumer ce calendrier dans le diagramme de Gantt de la figure 2.16 :

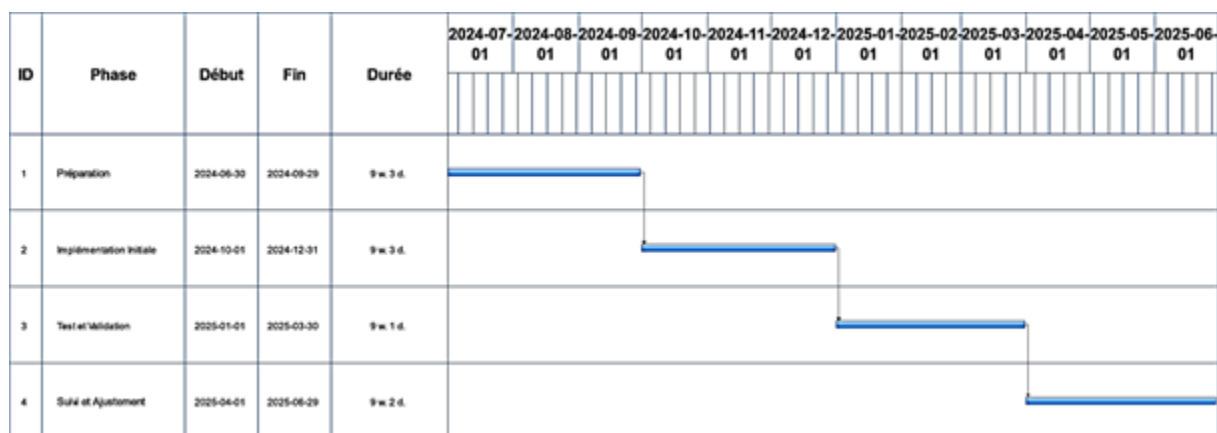


FIG. 2.16 : Diagramme de Gantt du plan de mise en oeuvre

2.8.4 Suivi et Évaluation

1. Comparaison entre le modèle d'origine et l'optimisé :

Le suivi et l'évaluation de la mise en œuvre sont essentiels pour mesurer l'efficacité

des améliorations apportées. La comparaison des proportions entre les durées moyennes des étapes des processus original et optimisé permet d'évaluer l'efficacité de l'optimisation. Voici une analyse des proportions et de leur impact sur la performance globale des processus.

Étape	Moyenne Original (jours)	Moyenne Optimisé (jours)	Proportion Moyenne (Optimisé/Original)
Déterminer la demande nationale pour chaque produit (E1)	1.05	0.48	0.46
Calculer la quantité de matière première (E2)	1.44	0.54	0.37
Approuver la demande d'achat (E3)	1.04	0.47	0.45
Passer un ordre d'achat (E4)	0.53	0.49	0.92
Suivre les stocks des articles par site (E5)	1.1	0.81	0.74
Ajuster le plan d'approvisionnement (E6)	2.06	1.48	0.72

TAB. 2.7 : Comparaison des durées moyennes des étapes entre le modèle original et le modèle optimisé

L'analyse des proportions entre les durées moyennes des étapes des processus original et optimisé permet d'évaluer l'efficacité de l'optimisation. La proportion de 0.46 pour "Déterminer la demande nationale" montre une réduction significative du temps nécessaire, presque de moitié, grâce à l'optimisation. Pour "Calculer la quantité de matière première", la proportion de 0.37 indique que le temps a été réduit de 63%, ce qui est un gain majeur.

L'automatisation a joué un rôle clé dans la réduction des temps d'exécution. Par exemple, l'étape "Approuver la demande d'achat" a une proportion de 0.45, montrant que le temps d'approbation a été réduit de 55 % grâce à l'automatisation. De plus, "Suivre les stocks des articles par site" et "Ajuster le plan d'approvisionnement" ont des proportions de 0.74 et 0.72 respectivement, indiquant une amélioration notable mais moins drastique par rapport aux autres étapes.

Certaines étapes montrent des gains moins significatifs. "Passer un ordre d'achat" présente une proportion de 0.92, ce qui signifie une réduction moins significative. Cela peut indiquer que cette étape était déjà assez optimisée ou qu'il y a moins de marge pour une amélioration significative.

Le graphique de la figure 2.17 illustre les changements dans la moyenne, l'écart-type et le temps total par cycle pour chaque étape du processus d'approvisionnement après l'optimisation. Chaque étape est représentée par une étiquette abrégée (E1, E2, E3, etc.) pour une meilleure lisibilité. Les lignes bleues montrent les mesures avant optimisation, tandis que les lignes vertes montrent les mesures après optimisation. Cette visualisation permet de comparer directement l'impact de l'optimisation sur chaque étape du processus.

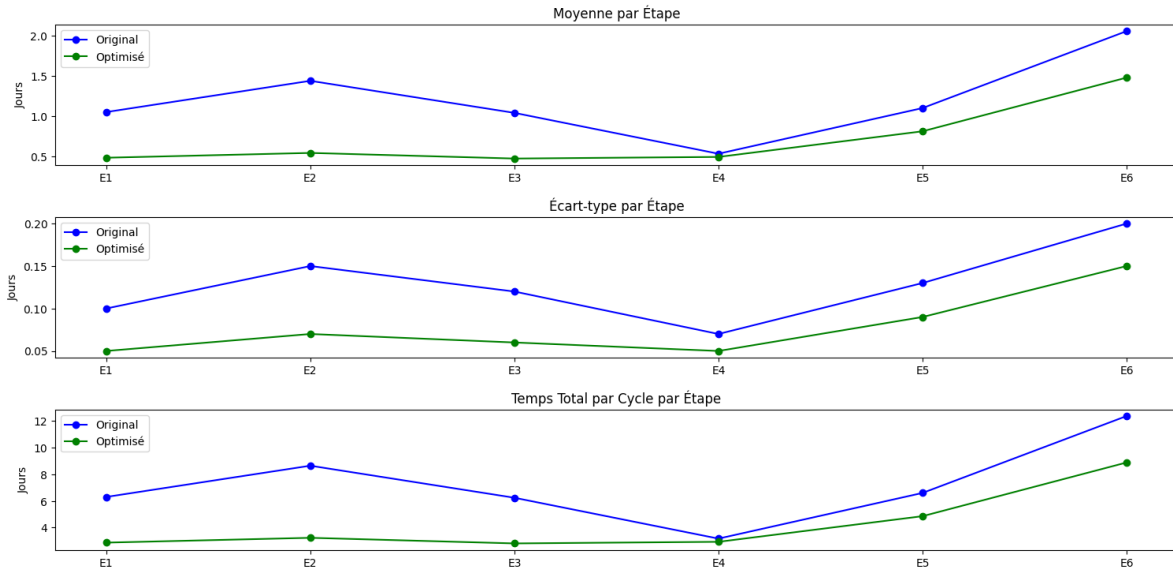


FIG. 2.17 : Les changements après l'optimisation

Ainsi, l'analyse des proportions montre que le processus optimisé est globalement plus efficace que le processus original. Les réductions de temps significatives dans les étapes clés et l'introduction d'automatisations ont permis d'améliorer la performance et de réduire les délais. L'optimisation a conduit à une meilleure gestion du temps et des ressources, rendant le processus d'approvisionnement plus réactif et efficient. Les étapes présentant des proportions élevées indiquent des gains importants en termes de réduction du temps, tandis que celles avec des proportions plus faibles montrent où des améliorations supplémentaires peuvent être explorées. Ces résultats confirment l'efficacité des stratégies d'optimisation mises en œuvre et soulignent les avantages de l'automatisation dans les processus de gestion de la chaîne d'approvisionnement.

2. Méthodes de suivi :

Tableaux de bord : L'utilisation de tableaux de bord pour suivre les KPI en temps réel permet une surveillance continue et proactive des performances.

Rapports mensuels : La génération de rapports mensuels offre une vue détaillée de la progression et des résultats, facilitant l'identification des domaines nécessitant des ajustements.

3. Évaluation périodique :

Revue trimestrielle : Des réunions trimestrielles permettent d'évaluer les résultats et d'ajuster les stratégies si nécessaire. Ces revues fournissent une opportunité de réflexion et de planification pour les phases suivantes.

Feedback du personnel : Le recueil des retours d'expérience des employés est crucial pour identifier les défis et les opportunités d'amélioration continue. Ces retours aident à adapter les formations et les processus pour mieux répondre aux besoins de l'équipe.

2.9 Conclusion

Ce chapitre a démontré l'application pratique du process mining dans le contexte de SPA Hamoud Boualem, mettant en lumière ses nombreux avantages et les défis associés. À travers une analyse détaillée et une modélisation rigoureuse du processus d'approvisionnement de l'unité de Boufarik, nous avons identifié plusieurs inefficacités et proposé des recommandations concrètes pour leur résolution. L'utilisation du process mining a permis de découvrir des motifs cachés dans les données, de vérifier la conformité des processus et de simuler des scénarios d'amélioration. Les recommandations formulées, si elles sont mises en œuvre efficacement, devraient conduire à une amélioration significative de la performance du processus étudié. Cette étude de cas souligne l'importance du process mining comme outil puissant pour l'optimisation des processus et fournit un cadre pratique pour son application dans d'autres contextes organisationnels.

Conclusion générale

Ce mémoire a exploré l'application du Process Mining pour l'optimisation des processus logistiques dans une entreprise réelle, SPA Hamoud Boualem. Les résultats obtenus démontrent l'efficacité de cette technique pour identifier les inefficacités, modéliser les flux de travail et proposer des recommandations d'amélioration concrètes. L'étude de cas de SPA Hamoud Boualem apporte plusieurs contributions majeures au domaine de l'optimisation des processus logistiques par le process mining :

- **Démonstration empirique de l'efficacité du process mining** : L'étude de cas fournit une preuve concrète de la capacité du process mining à identifier les inefficacités et à proposer des solutions d'amélioration.
- **Recommandations pratiques pour l'implantation du process mining** : Les recommandations formulées offrent un guide précieux aux entreprises souhaitant adopter cette technique et améliorer leurs processus logistiques.
- **Mise en lumière des limites et des perspectives de recherche** : L'analyse des limites et des perspectives de recherche ouvre la voie à de nouvelles investigations et à l'amélioration continue des pratiques d'optimisation des processus.

L'analyse des processus logistiques de SPA Hamoud Boualem a révélé plusieurs points d'amélioration potentiels, notamment :

- Délais de traitement des commandes excessifs.
- Manque de coordination entre les services.
- Utilisation insuffisante des outils technologiques.

Sur la base de ces constats, des recommandations détaillées ont été formulées, visant à :

- Simplifier les procédures de la chaîne logistique.
- Améliorer la communication et la collaboration entre les services.
- Investir dans des solutions technologiques intégrées.

L'implantation de ces recommandations devrait conduire à des améliorations significatives en termes de réduction des coûts opérationnels, d'amélioration des délais de livraison et d'augmentation de la satisfaction client.

Malgré les résultats encourageants, il est important de reconnaître les limites de cette étude. La qualité et la complétude des données générées ont pu influencer la précision des modèles générés. De plus, l'implantation des recommandations nécessite un engagement fort de la part de l'entreprise et des ressources financières adéquates.

Des pistes de recherche futures pourraient inclure :

- Intégration de techniques d'apprentissage automatique pour prédire les performances des processus et proposer des améliorations en temps réel.
- Réalisation d'une étude comparative avec d'autres entreprises du secteur pour identifier les meilleures pratiques et les facteurs de succès.
- Élaboration d'un cadre méthodologique pour l'implantation du process mining.

Le process mining s'affirme comme un outil puissant pour l'optimisation des processus logistiques, offrant des bénéfices tangibles en termes d'efficacité, de rentabilité et de satisfaction client. L'étude de cas de SPA Hamoud Boualem illustre le potentiel du process mining à améliorer la performance opérationnelle et à renforcer la compétitivité des entreprises. La poursuite de l'effort d'amélioration continue, guidée par une approche data-driven et une implication forte des parties prenantes, permettra aux entreprises de tirer pleinement parti des avantages offerts par le process mining.

En conclusion, ce mémoire souligne le rôle crucial du process mining dans la transformation digitale des entreprises et leur quête d'excellence opérationnelle.

Bibliographie

- [1] W. M. P. Van Der AALST. *Data Science in Action*. Berlin, Heidelberg : Springer, 2016.
- [2] W. M. P. Van Der AALST, A. ADRIANSYAH et B. Van DONGEN. “Replaying history on process models for conformance checking and performance analysis”. In : *Wiley Interdisciplinary Reviews : Data Mining and Knowledge Discovery* 2.2 (2012), p. 182-192.
- [3] W. M. P. Van Der AALST, J. C. A. M. BUIJS et B. F. van DONGEN. “Towards Improving the Representational Bias of Process Mining”. In : *International Symposium on Data-Driven Process Discovery and Analysis*. Berlin, Heidelberg : Springer, 2011, p. 39-54.
- [4] W. M. P. Van Der AALST et A. WEIJTERS. “Process Mining : A Research Agenda”. In : *Computers in Industry* 53.3 (2004), p. 231-244.
- [5] W. M. P. Van Der AALST et al. “Soundness of Workflow Nets : Classification, Decidability, and Analysis”. In : *Formal Aspects of Computing* 23.3 (2011), p. 333-363.
- [6] W. M. P. Van Der AALST et al. “Workflow Mining : A Survey of Issues and Approaches”. In : *Data & Knowledge Engineering* 47.2 (2003), p. 237-267.
- [7] W. M. P. Van der AALST. “Structural characterizations of sound workflow nets”. In : (1996).
- [8] W. M. P. van der AALST. *Process Mining : Discovery, Conformance and Enhancement of Business Processes*. Berlin, Heidelberg : Springer, 2011.
- [10] Mauricio ARRIAGADA et al. “Strategies to Automatically Derive a Process Model from a Configurable Process Model Based on Event Data”. In : *Applied Sciences* 7 (oct. 2017).
- [12] R. P. J. C. BOSE et W. M. P. Van Der AALST. “Context Aware Trace Clustering : Towards Improving Process Mining Results”. In : *2009 SIAM International Conference on Data Mining*. Sparks, NV, USA : SIAM, 2009, p. 401-412.
- [13] M. BOZKAYA, J. GABRIELS et J. M. van der WERF. “Process diagnostics : A method based on process mining”. In : *International Conference on Information, Process, and Knowledge Management (eKNOW'09)*. Cancun, Mexico : IEEE, 2009, p. 22-27.

- [14] J. C. A. M. BUIJS, B. F. van DONGEN et W. M. P. Van Der AALST. “Quality dimensions in process discovery : The importance of fitness, precision, generalization and simplicity”. In : *International Journal of Cooperative Information Systems* 23.1 (2014), p. 1-40.
- [15] A. CORDIER et al. “Trace-Based Reasoning : Modeling Interaction Traces for Reasoning on Experiences”. In : *Proceedings of the Florida Artificial Intelligence Research Society Conference (FLAIRS)*. Florida, USA, 2013, p. 363-368.
- [16] R. DAVIS. *ARIS Design Platform : Getting Started with BPM*. Berlin, Heidelberg : Springer, 2012.
- [17] A. DIX. “Human-Computer Interaction”. In : *Encyclopedia of Database Systems*. Sous la dir. de L. LIU et M. T. ÖZSU. Boston, MA : Springer US, 2009, p. 1327-1331.
- [18] P. DIXIT. “Interactive Process Mining”. Proefschrift. PhD Thesis. Department of Mathematics et Computer Science, 2019.
- [19] M. DUMAS et al. *Fundamentals of Business Process Management*. 2nd. Springer, 2018.
- [20] F. GONZALEZ-LOPEZ et G. BUSTOS. “Business Process Architecture Design Methodologies – A Literature Review”. In : *Business Process Management Journal* 25.6 (2019), p. 1317-1334.
- [21] V. GROVER et al. “The Influence of Information Technology Diffusion and Business Process Change on Perceived Productivity : The IS Executive’s Perspective”. In : *Information & Management* 34.3 (1998), p. 141-159.
- [22] M. HAMMER et J. CHAMPY. *Reengineering the Corporation : A Manifesto for Business Revolution*. HarperBusiness, 2007.
- [23] H. J. HARRINGTON. *Business Process Improvement : The Breakthrough Strategy for Total Quality, Productivity, and Competitiveness*. McGraw-Hill, 2010.
- [24] M. J. HARRY et R. SCHROEDER. *Six Sigma : The Breakthrough Management Strategy Revolutionizing the World’s Top Corporations*. New York : Currency, 2000.
- [25] W. J. KETTINGER, J. T. C. TENG et S. GUHA. “Business Process Change : A Study of Methodologies, Techniques, and Tools”. In : *MIS Quarterly* 21.1 (1997), p. 55-80.
- [26] R. KIMBALL et J. CASERTA. *The Data Warehouse ETL Toolkit : Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*. Hoboken, NJ, USA : John Wiley & Sons, 2011.
- [27] Joffrey LEBLAY. “Vers une nouvelle forme d’accompagnement des processus dans les systèmes interactifs : apport de la fouille de processus et de la recommandation”. Français. NNT : 2019LAROS021, tel-02522177. Thèse de doctorat. Université de La Rochelle, 2019.
- [28] S. J. J. LEEMANS, D. FAHLAND et W. M. P. Van Der AALST. “Discovering Block-Structured Process Models from Event Logs - A Constructive Approach”. In : *International Conference on Applications and Theory of Petri Nets and Concurrency*. Milan, Italy : Springer, 2013, p. 311-329.

- [29] S. J. J. LEEMANS, D. FAHLAND et W. M. P. Van Der AALST. “Discovering block-structured process models from event logs containing infrequent behaviour”. In : *International conference on business process management*. Beijing, China : Springer, 2013, p. 66-78.
- [30] G.-C. LOGHIN. “Observer un Environnement Numérique de Travail pour réguler les activités qui s’y déroulent”. Thèse de Doctorat. France : Université de Savoie et Université Technique de Cluj-Napoca, 2008.
- [31] G. LOMBARDI, M. SAUER et G. Di CESARE. “How ”Vitality Forms” Influence Our Mood. In : Atmosphere and Mood. Two Sides of the Same Phenomenon”. In : *Art Style* 11.3 (mar. 2023). Sous la dir. de M. SAUER et Z. WANG, p. 127-139.
- [32] K. LUND et A. MILLE. *Traces, Traces d’interactions, Traces D’apprentissages Définitions, Modèles Informatiques, Structurations, Traitements et Usages*. Juin 2009.
- [33] Kristine LUND et Alain MILLE. *Traces, Traces d’interactions, Traces D’apprentissages Définitions, Modèles Informatiques, Structurations, Traitements et Usages*. Juin 2009.
- [34] A. K. A. Alves De MEDEIROS et al. *Process mining : extending the alpha-algorithm to mine short loops*. BETA publicatie : working papers. Technische Universiteit Eindhoven, 2004.
- [35] A. K. A. de MEDEIROS et al. *Process mining : Extending the alpha-algorithm to mine short loops*. Technical report. Eindhoven : Eindhoven University of Technology, 2004.
- [36] Tadao MURATA. “Petri nets : Properties, analysis and applications”. In : *Proceedings of the IEEE* 77.4 (1989), p. 541-580.
- [37] Maryam RADGUI, Rajaa SAIDI et Salma MOULINE. “Extracting reusable fragments from business process using BPMN”. In : sept. 2012.
- [38] Á. REBUGE et D. R. FERREIRA. “Business process analysis in healthcare environments : A methodology based on process mining”. In : *Information Systems* 37.2 (2012), p. 99-116.
- [39] A. ROZINAT et W. M. P. Van Der AALST. “Conformance checking of processes based on monitoring real behavior”. In : *Information Systems* 33.1 (2008), p. 64-95.
- [40] A. ROZINAT et C. W. GÜNTHER. “Disco : Discover Your Processes”. In : *BPM 2014: Business Process Management Workshops* (2014), p. 40-44.
- [41] N. RUSSELL et al. “Workflow Control-Flow Patterns : A Revised View”. In : *BPM Reports* 6.22 (2006), p. 1-134.
- [42] L. S. SETTOUTI. “Systèmes à Base de Traces Modélisées : Modèles et Langages pour l’exploitation des traces d’Interactions”. Thèse de Doctorat. France : Université Claude Bernard Lyon 1, 2011.
- [43] L. SETTOUTI et al. “A Trace-Based Systems Framework : Models, Languages and Semantics”. In : (jan. 2009).
- [44] L. SETTOUTI et al. “Système à base de trace pour l’apprentissage humain”. In : jan. 2006, p. 25-27.

- [45] S. SIMPSON. *An XML Representation for Crew Procedures, Final Report NASA Faculty Fellowship Program*. Rapp. tech. Johnson Space Center, 2004.
- [46] H. M. W. VERBEEK et al. “XES, XESame, and ProM 6”. In : *Information Systems Evolution*. Sous la dir. d’O. PASTOR, A. FLORY et J.-L. CAVARERO. T. 72. Lecture Notes in Business Information Processing. Berlin, Heidelberg : Springer, 2010, p. 60-75.
- [47] A. WEIJTERS, W. M. P. Van Der AALST et A. K. A. de MEDEIROS. *Process mining with the heuristics miner-algorithm*. Tech. Rep. WP 166. Technische Universiteit Eindhoven, 2006, p. 1-34.
- [48] S. WHITE. *Business Process Modeling Notation v1.0*. Rapp. tech. Business Process Management Initiative (BPMI), mai 2004.
- [49] J. P. WOMACK, D. T. JONES et D. ROOS. *The Machine That Changed the World*. New York : Rawson Associates, 1990.
- [50] B. WURM et al. “Design Patterns for Business Process Individualization”. In : *17th International Conference on Business Process Management (BPM 2019)*. Springer, 2019, p. 370-385.
- [51] R. ZARKA et al. “TStore : A Trace-Base Management System using Finite-State Transducer Approach for Trace Transformation”. In : *MODELSWARD 2013 - Proceedings of the 1st International Conference on Model-Driven Engineering and Software Development*. Fév. 2013.

Annexes

Annexe A

Table comparative des systèmes à base de traces existants :

Système	Collecter les données	Transformer les traces	Adapté aux algorithmes	Domaine appliqué
Streaming Application Trace Miner (SATM)	Oui	Non	Non	Débogage des applications de streaming via les traces d'exécutions
Collaboration Analysis Tool (ColAT)	Oui (multiples sources)	Non	Non	Recherche ethnographique
Aplusix	Oui	Non	Non	Éducation (aide à l'apprentissage de l'algèbre)
Modeling USEs and Tasks for Tracing Experience (MUSSETTE)	Oui	Oui	Non	Approche générale, pouvoir développer et diversifier
Trace Analysis Tool for Interaction ANalysis (TATIANA)	Oui	Non	Non	Chercheurs dans une séance de travail collaboratif
VISU	Oui	Non	Non	Éducation (répondre aux besoins de l'enseignement apprentissage des langues en ligne)
Système à base de traces pour le calcul d'indicateur dans Moodle (SBT-IM)	Oui	Non	Non	Moodle
Projet d'Intégration de l'eXpérience pour l'Enseignement à Distance	Oui	Non	Non	Analyse d'apprentissage en ligne

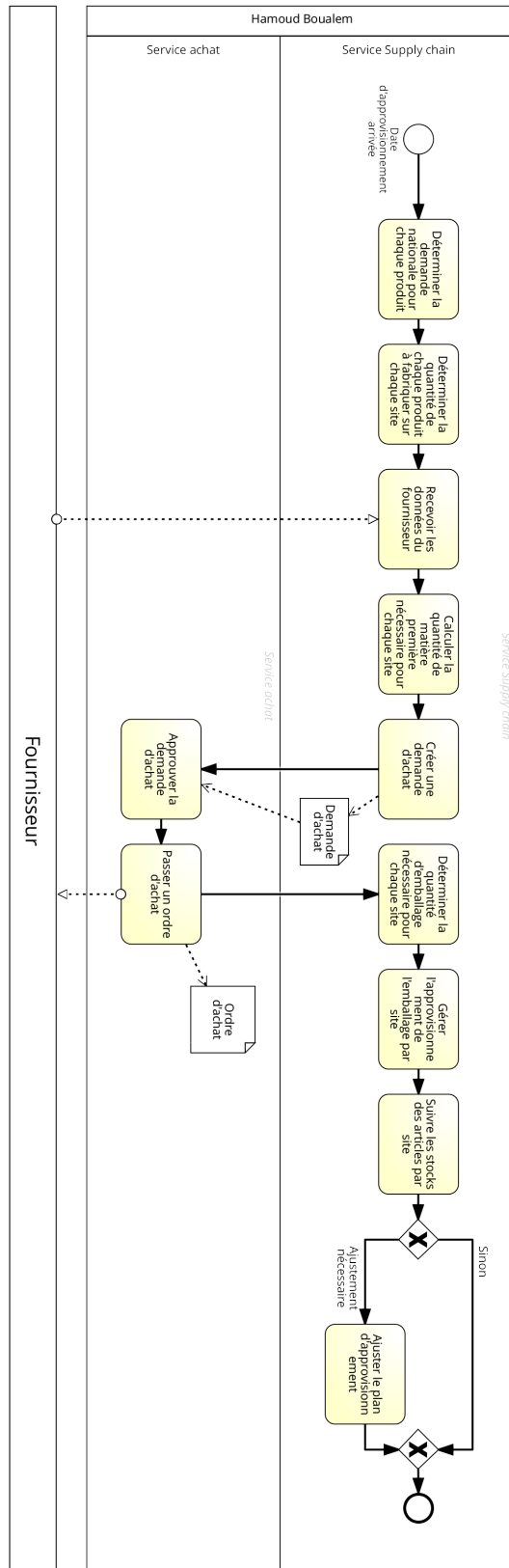
Annexe B

Fiche d'identité de l'entreprise Hamoud Boualem :

Nom commercial	Hamoud Boualem
Adresse du siège social	Hydra, Alger
Site Internet	https://www.hamoud-boualem.com
N° RCS	/
N° SIRET	/
Code APE	/
Date de création	1876
Forme juridique	SPA
Capital social	
Nationalité	Algérienne
Type d'organisation	<ul style="list-style-type: none"><input type="radio"/> Société commerciale<input checked="" type="radio"/> Entreprise individuelle<input type="radio"/> Association<input type="radio"/> Coopérative de type associatif<input type="radio"/> Établissement public<input type="radio"/> Collectivité locale
Nom des dirigeants	Boualem Hamoud
Activité de l'entreprise	Boissons Gazeuse
But de l'activité	<ul style="list-style-type: none"><input checked="" type="radio"/> Lucratif<input type="radio"/> Non lucratif
Nature de l'activité	<ul style="list-style-type: none"><input checked="" type="radio"/> Marchande<input type="radio"/> Non marchande
Secteur d'activité	<ul style="list-style-type: none"><input checked="" type="radio"/> Primaire<input type="radio"/> Secondaire<input type="radio"/> Tertiaire
Taille	Chiffre d'affaires : / Effectif : 900
Marché	<ul style="list-style-type: none"><input type="radio"/> Local<input type="radio"/> Régional<input checked="" type="radio"/> National<input type="radio"/> International

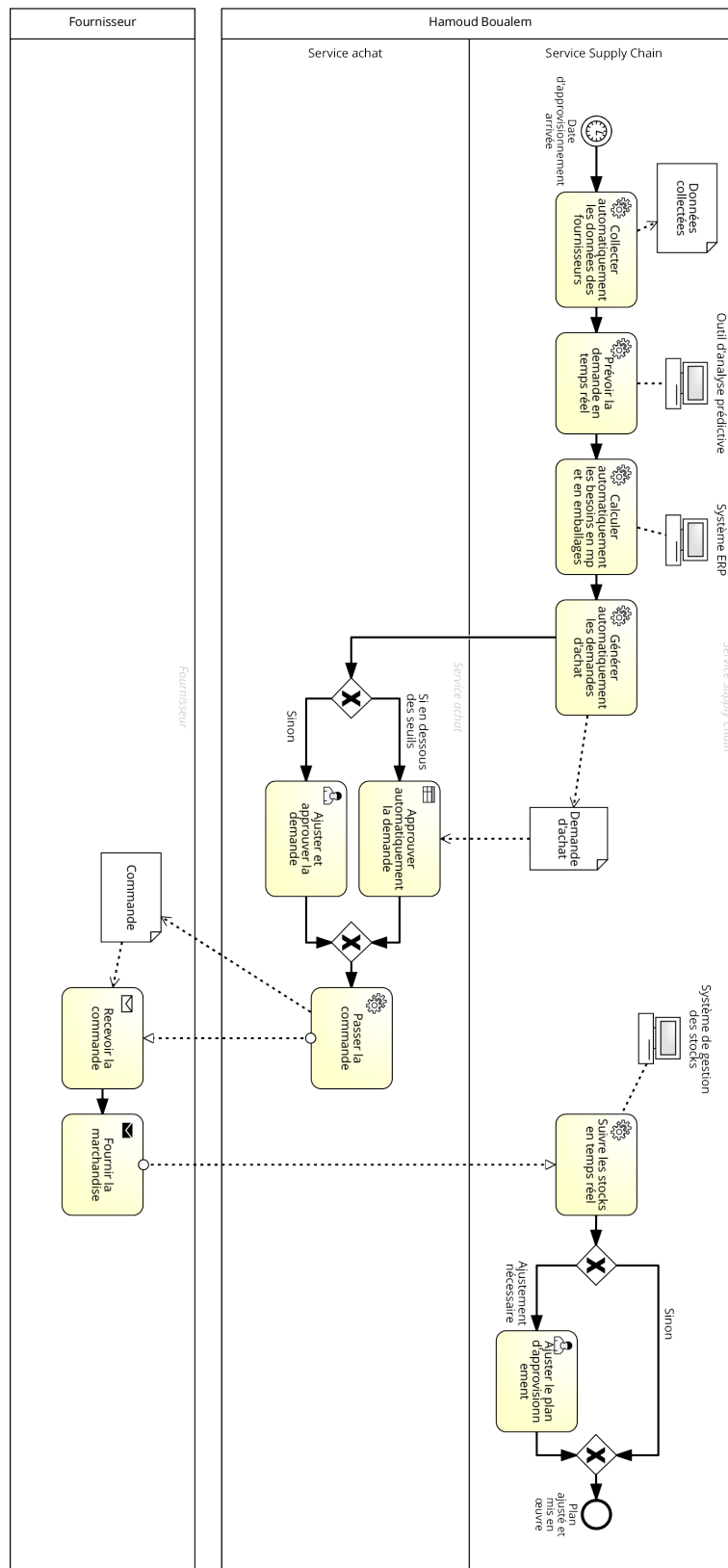
Annexe C

Modèle de processus d'approvisionnement de l'entreprise Hamoud Boualem :



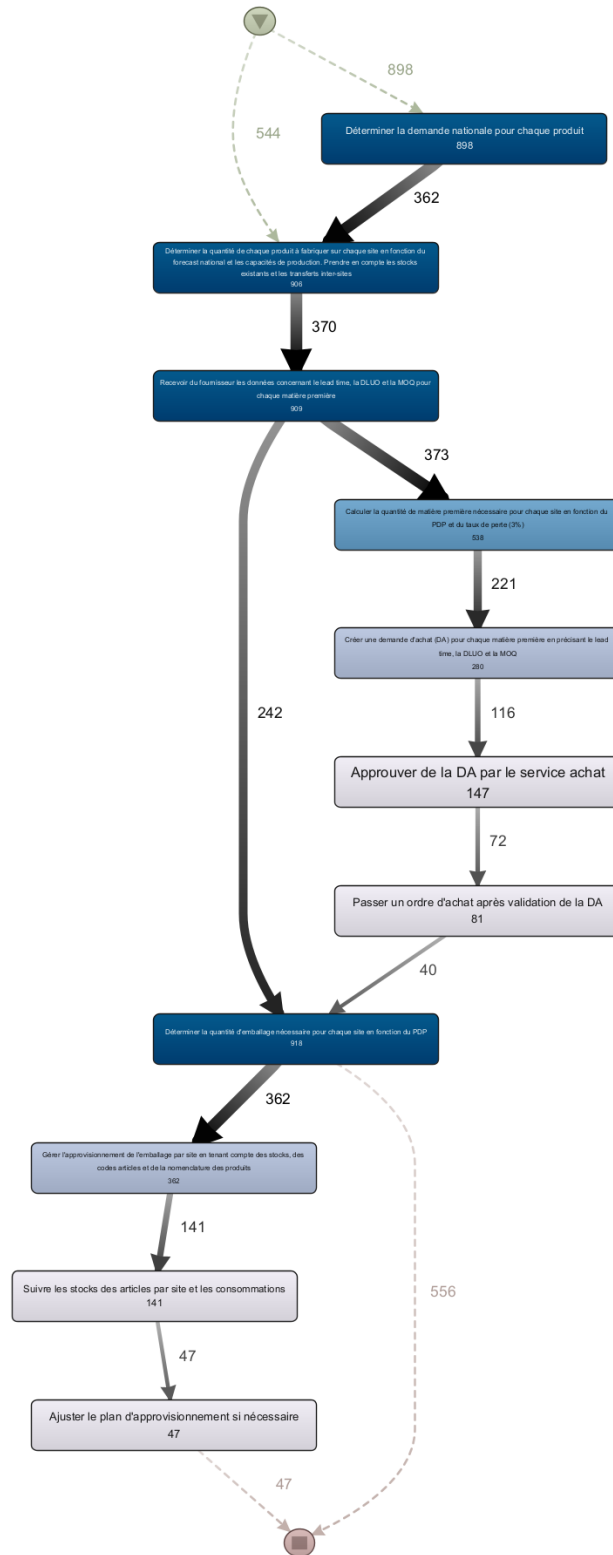
Annexe D

Modèle de processus d'approvisionnement optimisé de l'entreprise Hamoud Boualem :



Annexe F

Modèle découvert par Disco :



Annexe G

Code python de la génération des données :

```
1 import pandas as pd
2 import random
3 from datetime import datetime, timedelta
4
5 # Initialize lists for storing data
6 case_id = []
7 timestamp = []
8 activity = []
9 site = []
10 demande_nationale = []
11 prevision = []
12 quantite_a_fabriquer = []
13 capacites_de_production = []
14 stocks_existants = []
15 lead_time = []
16 dluo = []
17 moq = []
18 quantite_necessaire = []
19 taux_de_perte = []
20 da_cree = []
21 da_approuvee = []
22 ordre_achat = []
23 quantite_emballage = []
24 gestion_stocks_emballage = []
25 suivi_stocks = []
26 consommations = []
27 ajustement_plan = []
28
29 # Define possible sites
30 sites = ['SA', 'SB', 'SC']
31
32 # Define start date
33 start_date = datetime(2022, 1, 1)
34
35 # Define the set of activities with dependencies
36 activities = {
37     'Déterminer la demande nationale pour chaque produit' : None,
38     'Déterminer la quantité de chaque produit à fabriquer sur chaque site en
        fonction du forecast national et les capacités de production. Prendre
        en compte les stocks existants et les transferts inter-sites' : '
        Déterminer la demande nationale pour chaque produit',
39     'Recevoir du fournisseur les données concernant le lead time, la DLUO et
        la MOQ pour chaque matière première' : None,
40     'Calculer la quantité de matière première nécessaire pour chaque site en
        fonction du PDP et du taux de perte (3%)' : 'Recevoir du fournisseur
```

```

    les données concernant le lead time, la DLUO et la MOQ pour chaque
    matière première',
41 'Créer une demande d\'achat (DA) pour chaque matière première en
    précisant le lead time, la DLUO et la MOQ' : 'Calculer la quantité de
    matière première nécessaire pour chaque site en fonction du PDP et
    du taux de perte (3%)',
42 'Approuver de la DA par le service achat' : 'Créer une demande d\'achat (
    DA) pour chaque matière première en précisant le lead time, la DLUO
    et la MOQ',
43 'Passer un ordre d\'achat après validation de la DA' : 'Approuver de la
    DA par le service achat',
44 'Déterminer la quantité d\'emballage nécessaire pour chaque site en
    fonction du PDP' : None,
45 'Gérer l\'approvisionnement de l\'emballage par site en tenant compte des
    stocks, des codes articles et de la nomenclature des produits' : '
    Déterminer la quantité d\'emballage nécessaire pour chaque site en
    fonction du PDP',
46 'Suivre les stocks des articles par site et les consommations' : 'Gérer l
    \'approvisionnement de l\'emballage par site en tenant compte des
    stocks, des codes articles et de la nomenclature des produits',
47 'Ajuster le plan d\'approvisionnement si nécessaire' : 'Suivre les stocks
    des articles par site et les consommations'
48 }
49
50 # Convert activity keys to a list
51 activity_keys = list(activities.keys())
52
53 # Define the number of cases and average number of activities per case to
    target around 13,000 rows
54 num_cases = 2000 # Total number of cases
55 avg_activities_per_case = 6.5 # Average number of activities per case
56
57 # Generate data ensuring random but logical sequences
58 week_delta = timedelta(weeks=1)
59 activity_prob = 0.8 # Probability to include each activity in a trace
60
61 for i in range(num_cases) :
62     current_site = random.choice(sites)
63     case_start_date = start_date + (i // len(activity_keys)) * week_delta
64     trace_activities = random.sample(activity_keys, k=random.randint(4, int(
        avg_activities_per_case))) # Randomly select activities
65     trace_activities = sorted(trace_activities, key=lambda x : activity_keys.
        index(x)) # Maintain order based on dependencies
66     last_activity = None # Track the last valid activity
67
68     for j, act in enumerate(trace_activities) :
69         if last_activity and activities[act] and activities[act] !=
            last_activity :
70             continue # Skip this activity if its dependency was skipped

```

```

71
72     current_case_id = f"{current_site}-{i+1}"
73
74     # Generate random hour, minute, and second within working hours (8-12
75     # AM and 1-4:30 PM)
76     hour = random.choice([random.randint(8, 11), random.randint(13, 16)])
77     minute = random.randint(0, 59)
78     second = random.randint(0, 59)
79
80     # Add timestamp with hours, minutes, and seconds
81     timestamp.append(case_start_date + timedelta(days=j, hours=hour,
82     # minutes=minute, seconds=second))
83     case_id.append(current_case_id)
84     activity.append(act)
85     site.append(current_site)
86     demande_nationale.append(random.randint(1000, 5000))
87     prevision.append(random.randint(800, 4800))
88     quantite_a_fabriquer.append(random.randint(500, 3000))
89     capacites_de_production.append(random.randint(600, 3200))
90     stocks_existants.append(random.randint(200, 1500))
91     lead_time.append(random.randint(1, 10))
92     dluo.append(random.randint(30, 365))
93     moq.append(random.randint(50, 500))
94     quantite_necessaire.append(random.randint(100, 1000))
95     taux_de_perte.append(round(random.uniform(0.01, 0.05), 2))
96     da_cree.append(random.randint(0, 1))
97     da_approuvee.append(random.randint(0, 1))
98     ordre_achat.append(random.randint(0, 1))
99     quantite_emballage.append(random.randint(100, 1000))
100     gestion_stocks_emballage.append(random.randint(0, 1))
101     suivi_stocks.append(random.randint(0, 1))
102     consommations.append(random.randint(100, 800))
103     ajustement_plan.append(random.randint(0, 1))
104
105     last_activity = act # Update the last valid activity
106
107 # Create a DataFrame
108 data = pd.DataFrame({
109     'case_id' : case_id,
110     'timestamp' : timestamp,
111     'activity' : activity,
112     'site' : site,
113     'demande_nationale' : demande_nationale,
114     'prevision' : prevision,
115     'quantite_a_fabriquer' : quantite_a_fabriquer,
116     'capacites_de_production' : capacites_de_production,
117     'stocks_existants' : stocks_existants,
118     'lead_time' : lead_time,
119     'dluo' : dluo,

```

```
118     'moq' : moq,
119     'quantite_necessaire' : quantite_necessaire,
120     'taux_de_perte' : taux_de_perte,
121     'da_cree' : da_cree,
122     'da_approuvee' : da_approuvee,
123     'ordre_achat' : ordre_achat,
124     'quantite_emballage' : quantite_emballage,
125     'gestion_stocks_emballage' : gestion_stocks_emballage,
126     'suivi_stocks' : suivi_stocks,
127     'consommations' : consommations,
128     'ajustement_plan' : ajustement_plan
129 })
130
131 # Save to CSV
132 data.to_csv('process_mining_data.csv', index=False)
```

Résumé

Dans un environnement économique en constante évolution, les entreprises modernes font face à des défis croissants pour maintenir leur compétitivité. La gestion de la chaîne logistique (SCM) joue un rôle central dans cet effort, impactant directement la performance opérationnelle et la satisfaction des clients. Ce mémoire explore le potentiel du Process Mining (PM) pour optimiser les processus logistiques. Le Process Mining combine des techniques de data science et de gestion des processus métier pour analyser les données des systèmes d'information, permettant de visualiser, comprendre et améliorer les flux de travail existants. Une étude de cas sur l'entreprise SPA Hamoud Boualem, leader algérien dans le secteur des boissons gazeuses, démontre l'applicabilité du process mining pour améliorer l'efficacité et réduire les inefficacités dans une chaîne logistique réelle.

Mots clés : Process Mining, chaîne logistique, optimisation, efficacité, SPA Hamoud Boualem.

Abstract

In a constantly evolving economic environment, modern companies face growing challenges to maintain their competitiveness. Supply Chain Management (SCM) plays a central role in this effort, directly impacting operational performance and customer satisfaction. This thesis explores the potential of Process Mining (PM) to optimize logistics processes. Process mining combines data science techniques and business process management to analyze information system data, enabling visualization, understanding, and improvement of existing workflows. A case study on SPA Hamoud Boualem, an Algerian leader in the soft drink sector, demonstrates the applicability of PM to enhance efficiency and reduce inefficiencies in a real logistics chain.

Keywords : Process Mining, supply chain, optimization, efficiency, SPA Hamoud Boualem.

ملخص

في بيئة اقتصادية متغيرة باستمرار، تواجه الشركات الحديثة تحديات متزايدة للحفاظ على تنافسيتها. تلعب إدارة سلسلة التوريد دورًا مركزيًا في هذا الجهد، مما يؤثر بشكل مباشر على الأداء التشغيلي ورضا العملاء. يستكشف هذا البحث إمكانات تقنية استخراج العمليات لتحسين العمليات اللوجستية. تجمع تقنية استخراج العمليات بين تقنيات علم البيانات وإدارة العمليات التجارية لتحليل بيانات نظم المعلومات، مما يمكن من تصور وفهم وتحسين تدفقات العمل الحالية. دراسة حالة لشركة حمود بوعلام، الرائدة الجزائرية في قطاع المشروبات الغازية، توضح قابلية تطبيق تقنية استخراج العمليات لتحسين الكفاءة وتقليل العيوب في سلسلة توريد حقيقية.

كلمات مفتاحية : الكلمات المفتاحية: استخراج العمليات، سلسلة التوريد، تحسين، كفاءة، حمود بوعلام.