



HAL
open science

GESTION PROACTIVE DU CHANGEMENT DANS LES PROJETS DE REINGENIERIE DES PROCESSUS METIERS

Mamadou Samba Camara

► **To cite this version:**

Mamadou Samba Camara. GESTION PROACTIVE DU CHANGEMENT DANS LES PROJETS DE REINGENIERIE DES PROCESSUS METIERS. Sciences de l'ingénieur [physics]. Université Paris VIII Vincennes-Saint Denis, 2009. Français. NNT: . tel-00638199

HAL Id: tel-00638199

<https://theses.hal.science/tel-00638199>

Submitted on 4 Nov 2011

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

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

ECOLE DOCTORALE COGNITION, LANGAGE, INTERACTION

THESE

Présentée pour obtenir le grade de

Docteur de l'Université de Paris 8 – Vincennes-Saint-Denis

Discipline : Productique-Génie Industriel

Par CAMARA Mamadou Samba

**GESTION PROACTIVE DU CHANGEMENT DANS LES
PROJETS DE REINGENIERIE DES PROCESSUS METIERS**

Thèse dirigée par le Professeur EL MHAMEDI Abderrahman

Soutenue le 10 Novembre 2009

Devant le jury composé de

M. PINGAUD Hervé

Rapporteur

Professeur à l'Ecole des Mines d'Albi Carmaux Campus

M. BOUCHER Xavier

Rapporteur

Maître de conférences et HDR à l'Ecole Nationale Supérieure des Mines de Saint-Etienne

M. EL MHAMEDI Abderrahman

Directeur de thèse

Professeur à l'Université de Paris 8

M. KERMAD Lyes

Co-directeur de thèse

Maître de conférences à l'Université de Paris 8

M. EL MOUDNI Abdellah

Examineur

Professeur à l'Université de Technologie de Belfort-Montbéliard

Thèse préparée au sein de l'équipe MGSI (Modélisation et Génie des Systèmes Industriels) à l'UIT de Montreuil-**Université de Paris 8 – Vincennes-Saint-Denis**

Résumé :

La problématique de cette étude est la gestion du changement dans les projets d'implémentation de progiciel. L'approche proposée est la gestion proactive du changement. Elle est basée sur des actions à priori pour réduire l'impact négatif du changement. Ces actions sont optimisées et ciblées sur certains processus de l'entreprise identifiés par prédiction. Notre objectif est de combiner la gestion quantitative et non-quantitative du changement dans les projets ERP afin de lever leurs limites. En effet, la littérature montre que les méthodes non-quantitatives proposent des actions de gestion du changement, sans pour autant accompagner ces actions d'une optimisation et d'un ciblage formalisé. Les techniques quantitatives, quant à elles, proposent des modèles mais ne définissent pas des actions claires.

La démarche proposée repose sur l'application du processus de datamining. Nous nous sommes intéressés à trois concepts principaux qui sont l'ampleur, l'effort de gestion et l'impact du changement. L'idée est de tirer parti des informations collectées dans des projets d'implémentation de progiciel antérieurs, pour établir des relations quantitatives et objectives entre ces trois concepts.

La première composante de l'application du datamining est l'utilisation d'un modèle pour faire des prédictions sur l'impact potentiel du changement. Cet impact est mesuré de manière prédictive par rapport à la variation des critères de performance du processus après sa restructuration et son implémentation dans le progiciel. L'unité d'analyse dans l'étude est le processus métier. La seconde composante de l'application du processus de datamining est la prise de décision basée sur les prédictions du modèle. L'ampleur et l'effort de gestion du changement sont considérés comme des explications pour l'impact du changement. Ils sont utilisés aussi comme levier dans les actions à définir pour la gestion de cet impact.

Notre approche intervient principalement dans la phase « d'adéquation et de configuration » du projet ERP (Tomas, 2000) dans laquelle les processus sont restructurés. Elle vient compléter la simulation, le prototypage (qui inclut l'analyse des écarts) pour l'acceptation d'un processus « en-devenir ».

Mots clés : Réingénierie Organisationnelle, Changement Organisationnel, Progiciel de Gestion Intégré, Statistique Bayésienne, Gestion des Risques

Abstract:

The issue of this work is change management in enterprise resources planning projects. The proposed approach is called proactive change management. It is based on prior actions to reduce negative impact of change. These actions are optimized and directed on some business processes identified by prediction. Our goal is to combine quantitative and non-quantitative change management in ERP projects in order to overcome their limits. Indeed, the literature shows that non-quantitative methods have proposed change management actions, but do not give means for optimizing and targeting these actions. The quantitative techniques used in ERP projects have proposed models but do not define clear change management actions.

The proposed approach is based on the application of data mining process. We are interested by three main concepts : magnitude, management effort and impact of change. The idea is to capitalize on information from previous ERP project to establish quantitative and objective relations between the three concepts.

The first component of the data mining process application is the use of the model for predictions on potential impact of change. The impact was evaluated predictive way against the variation on key performance indicators of business process, after its reengineering and its implementation of ERP. The unit of analysis of our study is business process. The second component of the data mining process application is decision making based on predictions given by the model. Magnitude and management effort of change are considered as explications of change impact. They are also used as lever in actions to define the management of this impact.

Our approach intervenes principally in the business process reengineering step of ERP project where processes are reconfigured. It comes to complete simulation and prototyping (including gap analysis) for the acceptance of a “to-be” process.

Key words: Business Process Reengineering, Organizational Change, Enterprise Resource Planning, Bayesian Statistic, Risk Management

A Absa CAMARA

A toute ma famille, pour vos encouragements.

A Kaw Amadou CAMARA et toute sa famille pour leur soutien

Remerciements

Je tiens tout d'abord à exprimer ma profonde reconnaissance à Monsieur Abderrahman EL MHAMED, Professeur à l'Université de Paris 8, pour avoir encadré ces travaux de thèse. Ses qualités humaines, ses remarques pertinentes, ses conseils m'ont été d'un grand soutien et ont largement contribué à la réalisation de ce mémoire.

Je souhaite exprimer également ma profonde reconnaissance à Monsieur Lyes KERMAD, Maître de conférences à l'Université de Paris 8 co-directeur de ces travaux, pour sa disponibilité et l'attention qu'il a portée à ce travail.

Je remercie Monsieur M. Hervé PINGAUD, Professeur à l'Ecole des Mines d'Albi Carmaux et Monsieur Xavier BOUCHER, Maître de conférences et HDR à l'Ecole Nationale Supérieure des Mines de Saint-Etienne, pour avoir accepté d'être rapporteurs de mon mémoire de thèse, et pour les conseils et les remarques constructives qu'ils m'ont apportés.

Je remercie également Monsieur Abdellah EL MOUDNI, Professeur à l'Université de Technologie de Belfort-Montbéliard, pour sa participation, dans mon jury de thèse, en tant qu'examineur.

Ces travaux ont été menés au sein de l'équipe MGSI à l'IUT de Montreuil. A ce titre, je tiens à remercier Monsieur Bernard BETANT, Directeur de l'IUT, ainsi que tout le personnel de l'IUT. Je remercie également tous les enseignants et les collègues doctorants membres de l'équipe.

A tous ce qui m'on aidé de près (Yasmine, j'en oublie) ou de loin dans la réalisation de ce travail, je dis un grand merci.

Merci aux amis de BNF-Bastille (Manu, Manga, Mingou, Rob, Marcel) et à tous les amis que je connais depuis Sanar (Fall, Dem, Rose)

Partie 1	ETAT DE L'ART ET DEMARCHE.....	20
Chapitre 1.1	Introduction de la première partie.....	21
Chapitre 1.2	Implémentation de progiciel ERP.....	22
1.2.1.	Introduction.....	22
1.2.2.	Déroulement de l'implémentation de progiciel.....	22
1.2.3.	Problématiques de l'implémentation de progiciel.....	24
1.2.4.	Le BPR dans le projet ERP.....	26
1.2.5.	Dimensions de l'ampleur, de l'effort de gestion et de l'impact.....	32
1.2.6.	Conclusion.....	35
Chapitre 1.3	Gestion non-quantitative du changement.....	36
1.3.1.	Introduction.....	36
1.3.2.	Facteurs de risque et facteurs de succès en projet ERP.....	36
1.3.3.	Guides à n étapes de gestion du changement.....	37
1.3.4.	Intégration organisationnelle et gestion du changement.....	42
1.3.5.	Le modèle de maturité pour les processus métier.....	46
1.3.6.	Conclusion.....	49
Chapitre 1.4	Gestion quantitative du changement.....	51
1.4.1.	Introduction.....	51
1.4.2.	Mesure du changement.....	51
1.4.3.	Analyse de la criticité et simulation.....	53
1.4.4.	La clusterisation des processus.....	55
1.4.5.	Raisonnement à base de cas.....	57
1.4.6.	Méthodes d'explication.....	58
1.4.6.1.	Régression linéaire multiple.....	59
1.4.6.2.	Analyse de corrélation.....	60
1.4.6.3.	Prédiction par régression linéaire.....	61
1.4.6.4.	Méthodes de classement.....	63
1.4.7.	Les approches quantitatives très utilisées.....	67
1.4.8.	Conclusion.....	69
Chapitre 1.5	Gestion proactive du changement.....	71

1.5.1. Introduction.....	71
1.5.2. Intégration des approches quantitatives et qualitatives.....	71
1.5.3. Processus de datamining.....	73
1.5.4. Datamining pour la proactivité dans la gestion du changement	75
1.5.5. Le processus comme unité d'analyse.....	79
1.5.5.1. Les motivations de l'analyse au niveau processus.....	79
1.5.5.2. Les implications de l'analyse au niveau processus.....	82
Chapitre 1.6 Conclusion de la première partie	86
Partie 2 DEFINITION DES VARIABLES DU PROBLEME.....	87
Chapitre 2.1 Introduction de la deuxième partie.....	88
Chapitre 2.2 Le processus métier dans notre démarche.....	90
2.2.1. Introduction.....	90
2.2.2. Ingénierie des processus métier	90
2.2.2.1. Modélisation des processus	93
2.2.2.2. Business Process Modeling Notation (BPMN)	94
2.2.3. Cycle de vie du processus dans notre démarche.....	100
2.2.4. Relation d'influence entre les dimensions	104
2.2.5. Quelles variables définir sur les processus ?	107
2.2.6. Conclusion	111
Chapitre 2.3 Définition des variables à expliquer.....	112
2.3.1. Introduction.....	112
2.3.2. Métriques opérationnelles.....	112
2.3.3. Quelle est la variation dans le délai ?	114
2.3.4. Quelle est la variation dans le coût ?	117
2.3.5. Quelle est la variation dans la qualité ?	120
2.3.6. Conclusion	122
Chapitre 2.4 Définition des variables explicatives	123
2.4.1. Introduction.....	123
2.4.2. L'analyse structurelle des processus (ASP).....	124
2.4.3. Procédé de calcul des métriques de changement en GL.....	125

2.4.4.	Définition des métriques structurelles de changement	127
2.4.5.	Définition des métriques structurelles de variation	136
2.4.5.1.	Variation dans la complexité du flux de contrôle	137
2.4.5.2.	Variation dans la profondeur d'imbrication maximale	138
2.4.5.3.	Variation dans la simplicité	139
2.4.5.4.	Variation dans l'intégration	141
2.4.5.5.	Variation dans la flexibilité	143
2.4.6.	Mesure de l'effort de gestion du changement	144
2.4.1.	Conclusion	145
Chapitre 2.5	Conclusion de la deuxième partie	148
Partie 3	APPRENTISAGE ET UTILISATION DU MODELE	149
Chapitre 3.1	Introduction de la troisième partie	150
Chapitre 3.2	Choix de la méthode de modélisation et de validation	151
3.2.1.	Introduction	151
3.2.2.	Modélisation dans la gestion du changement dans les projets ERP	151
3.2.3.	Choix de la technique de sélection des variables	153
3.2.1.	Apprentissage des réseaux bayésiens par recherche gloutonne	155
3.2.1.1.	Distribution normale	156
3.2.1.2.	Distributions de probabilités	157
3.2.1.3.	Détermination des paramètres joints	158
3.2.1.4.	Calcul des paramètres locaux à priori	160
3.2.1.5.	Calcul du score local	162
3.2.1.6.	Détermination des paramètres locaux à postériori	163
3.2.1.7.	Détermination de l'équation de prédiction	163
3.2.2.	Validation croisée avec k sous-échantillons	164
3.2.3.	Détermination de la taille de la base	167
3.2.4.	Conclusion	168
Chapitre 3.3	Collecte des données et apprentissage	169
3.3.1.	Introduction	169
3.3.2.	Création de la base d'apprentissage	169

3.3.3. La validation croisée avec dix sous-échantillons.....	172
3.3.3.1. Déroulement de l'expérience 1	173
3.3.3.2. Déroulement de l'expérience 10.....	184
3.3.4. Calcul de l'erreur globale et choix du modèle.....	188
3.3.5. Conclusion	189
Chapitre 3.4 Utilisation du modèle sur les processus	190
3.4.1. Introduction.....	190
3.4.2. Restructuration.....	191
3.4.3. Extraction des métriques explicatives.....	194
3.4.3.1. Extraction des métriques structurelles de changement.....	194
3.4.3.2. Extraction des métriques structurelles de variation	198
3.4.3.3. Initialisation des métriques d'effort de gestion du changement	200
3.4.4. Prédiction.....	201
3.4.5. Validation du processus « en-devenir ».....	203
3.4.6. Conclusion	207
Chapitre 3.5 Conclusion de la troisième partie.....	208
Annexe 1 : Méthodologie de mise en place d'un ERP	227
Annexe 2 : Exemple sur les réseaux bayésiens	230
Paramètres joints	233
Paramètres locaux à priori.....	235
Calcul du score et des paramètres à postériori	238
Equation de prédiction.....	242
Annexe 3 : Programme banlist	243

Liste des tableaux

Tableau 1 Facteurs clés de succès en projet ERP (Nah, et al., 2003).....	24
Tableau 2 Stratégies de BPR en implémentation de SAP (Cooke, et al., 1998)	29
Tableau 3 Hypothèses relatives au changement (Saint-Leger, 2005)	39
Tableau 4 Métriques de changement en projet ERP (Esteves, 2002B).....	52
Tableau 5 Variables projet ERP (Barki, et al., 2004) et (Parr, et al., 2000).....	65
Tableau 6 Relation entre le périmètre et les sorties d'un projet (Barki, et al., 2004)...	65
Tableau 7 Noyau BPMN (OMG, 2007B p. 18).....	95
Tableau 8 Extension du noyau BPMN (OMG, 2007B p. 24)	96
Tableau 9 Variation opérationnelle durée processus traitement commandes	116
Tableau 10 Variation opérationnelle coût processus traitement de commandes.....	120
Tableau 11 Variation opérationnelle non-qualité processus traitement commandes .	122
Tableau 12 Synthèse des variables à expliquer	122
Tableau 13 Métriques de base : objets de flux	124
Tableau 14 Métriques de base : objets de connexion.....	125
Tableau 15 Métriques de base : Swimlane	125
Tableau 16 Métriques de base : Artéfacts	125
Tableau 17 Métriques de complexité logiciel	126
Tableau 18 Métriques de changement logiciel.....	127
Tableau 19 Détails simplicité PTC.....	140
Tableau 20 Synthèse variables explicatives	146
Tableau 21 Détails simplicité PGR	199

Liste des figures

Figure 1 Objectif proactivité	21
Figure 2 Flux des dérivables entre les douze phases (Tomas, 2000 p. 145)	23
Figure 3 Phase d'adéquation et de configuration.....	32
Figure 4 Relations entre l'ampleur, l'effort de gestion et l'impact.....	32
Figure 5 Régression linéaire simple	60
Figure 6 CRISP-DM.....	74
Figure 7 Datamining pour la gestion proactive du changement.....	76
Figure 8 Définition des variables du problème	89
Figure 9 Méta-modèle BPMN.....	99
Figure 10 Cycle de vie du processus	101
Figure 11 Processus traitement de commandes en-cours (Scheer, 1998 p. 16)	102
Figure 12 Processus traitement de commandes «implémenté» (Scheer, 1998 p. 17)	103
Figure 13 Relation d'influence entre les dimensions.....	106
Figure 14 Définition « top-down » des métriques par la GQM	108
Figure 15 Coût processus	119
Figure 16 Relations "responsabilités" supprimées dans PTC.....	129
Figure 17 Relations "responsabilités" rajoutées dans PTC	130
Figure 18 Associations objets de donnée vers activités supprimées dans PTC	131
Figure 19 Associations des objets de donnée vers activités rajoutées dans PTC.....	131
Figure 20 Associations activités vers objets de donnée supprimées dans PTC	132
Figure 21 Associations des activités vers objet de donnée rajoutées dans PTC	133
Figure 22 Flux de séquence événements vers activités supprimées dans PTC	134
Figure 23 Flux séquence des événements vers les activités rajoutées dans PTC.....	135
Figure 24 Flux séquence des activités vers les événements supprimées dans PTC...	135
Figure 25 Flux de séquence des activités vers les événements rajoutées dans PTC..	136
Figure 26 Apprentissage, validation et utilisation du modèle	150
Figure 27 La distribution normale.....	157
Figure 28 Cinq premières observations de l'historique	169

Figure 29 Cinq dernières observations de l'historique	170
Figure 30 Création des sous-échantillons à partir de la base de départ.....	170
Figure 31 Création des bases d'apprentissage et de test pour les 10 expériences	171
Figure 32 Cinq premiers éléments base de test expérience 1	171
Figure 33 Cinq derniers éléments base de test expérience 1	172
Figure 34 Cinq premiers éléments base d'apprentissage expérience 1	172
Figure 35 Cinq derniers éléments base d'apprentissage expérience 1	172
Figure 36 Recherche gloutonne de l'expérience 1	174
Figure 37 Réseau bayésien expérience 1	175
Figure 38 Détails du réseau bayésien expérience 1	175
Figure 39 Paramètres nœud MOVC réseau expérience 1	177
Figure 40 Paramètres nœud MOVD réseau expérience 1	178
Figure 41 Paramètres nœud MOVNQ réseau expérience 1	178
Figure 42 Commandes R pour la prédiction sur l'échantillon de test.....	180
Figure 43 Résultats des prédictions sur l'échantillon de test	181
Figure 44 Valeurs réelles dans l'échantillon de test	181
Figure 45 Enregistrement des erreurs nœuds de réponse et expérience 1	183
Figure 46 Affichage des erreurs nœuds de réponse et expérience 1	183
Figure 47 Recherche gloutonne expérience 10	184
Figure 48 Réseau expérience 10.....	185
Figure 49 Relations réseau expérience 10.....	185
Figure 50 Paramètres réseau expérience 10	186
Figure 51 Echantillon de test expérience 10.....	187
Figure 52 Résultat prédiction expérience 10	187
Figure 53 Erreurs expérience 10.....	187
Figure 54 Erreurs des dix expériences.....	188
Figure 55 Erreur en généralisation validation croisée	189
Figure 56 Prise en compte du changement dans l'acceptation du processus.....	190
Figure 57 Processus d'assurance décès « en-cours » (Küster, et al., 2005)	192
Figure 58 Processus d'assurance décès « en-devenir » (Küster, et al., 2005).....	193

Figure 59 Relations "responsabilités" supprimées dans PGR	194
Figure 60 Relations "responsabilités" rajoutées dans PGR	194
Figure 61 Associations objets de donnée vers activités supprimées dans PGR	195
Figure 62 Association objets de donnée vers activités » rajoutées dans PGR	195
Figure 63 Associations activités vers objets de donnée supprimées dans PGR	196
Figure 64 Associations activité vers objets de donnée rajoutées dans PGR	196
Figure 65 Flux de séquence événements vers activités supprimées dans PGR	197
Figure 66 Flux séquence événements vers activités rajoutées dans PGR	197
Figure 67 Flux de séquence activités vers événements supprimées dans PGR	198
Figure 68 Flux de séquence activités vers événements rajoutées dans PGR	198
Figure 69 Variables explicatives PGR	200
Figure 70 Résultats première prédiction PGR	201
Figure 71 Commandes de prédiction PGR	202
Figure 72 Résultats seconde prédiction PGR	204
Figure 73 Equilibre entre changement, écart et simulation	206
Figure 74 Réseau de départ exemple	230
Figure 75 Greedy search exemple	231
Figure 76 Réseau final exemple	232
Figure 77 paramètres joints pour la configuration [1] des parents discrets	234
Figure 78 Entrées des paramètres à priori	236
Figure 79 paramètres à priori	237
Figure 80 Entrées calcul du score et des paramètres à postériori	239
Figure 81 Calcul du score et des paramètres à postériori	240
Figure 82 Somme des log des scores des trois nœuds	241
Figure 83 Prédiction avec la distribution de probabilités mise à jour	242

Liste des équations

Équation 1 Régression linéaire multiple.....	59
Équation 2 Somme des carrées des erreurs	59
Équation 3 Estimation des paramètres	60
Équation 4 Régression linéaire simple	60
Équation 5 Métrique opérationnelle de variation dans le délai	115
Équation 6 Métrique opérationnelle de variation dans le coût.....	120
Équation 7 Métrique opérationnelle de variation de la non-qualité	121
Équation 8 MSCAES.....	129
Équation 9 MSCAER	130
Équation 10 MSCITS	130
Équation 11 MSCITR.....	131
Équation 12 MSCIPS	132
Équation 13 MSCIPR.....	133
Équation 14 MSCESS	134
Équation 15 MSCESSR.....	134
Équation 16 MSCEGS.....	135
Équation 17 MSCEGR	136
Équation 18 Métrique structurelle de variation de la complexité flux de contrôle	138
Équation 19 Métrique structurelle de variation de la profondeur maximale.....	139
Équation 20 Simplicité (Tjaden, 2001)	140
Équation 21 Métrique structurelle de variation de la simplicité.....	140
Équation 22 Intégration Processus (Tjaden, 2001).....	142
Équation 23 Intégration processus (Aubert, et al., 2003).....	143
Équation 24 Flexibilité (Tjaden, 2001)	144
Équation 25 Métrique effort de formation.....	144
Équation 26 Métrique effort d'accompagnement	145
Équation 27 Vraisemblances à postériori et facteur de bayes	155
Équation 28 Score réseau	156

Équation 29 Décomposition du score.....	156
Équation 30 Paramètres nœud continu.....	157
Équation 31 Calcul de la variance.....	158
Équation 32 Distribution de la moyenne jointe.....	158
Équation 33 Distribution de la covariance jointe.....	158
Équation 34 Nujoint, Rhojoint et Alphajoint.....	159
Équation 35 Mujoint et sa décomposition.....	159
Équation 36 Matrice des coefficients de propagation des changements.....	159
Équation 37 Vecteur des coefficients de régression.....	159
Équation 38 Sigmajoint.....	160
Équation 39 Eléments diagonaux de la covariance.....	160
Équation 40 Eléments non-diagonaux de la covariance.....	160
Équation 41 Phijoint.....	160
Équation 42 Décomposition du Phijoint.....	160
Équation 43 Distribution des paramètres de régression.....	161
Équation 44 Mu a priori du nœud.....	161
Équation 45 Décomposition de mu a priori du nœud.....	161
Équation 46 Phi à priori du nœud.....	161
Équation 47 Rho à priori du nœud.....	161
Équation 48 Tau à priori du nœud.....	162
Équation 49 Score nœud continu.....	162
Équation 50 Tau a posteriori du nœud.....	163
Équation 51 Mu à posteriori du nœud.....	163
Équation 52 Rho à posteriori du nœud.....	163
Équation 53 Phi à posteriori du nœud.....	163
Équation 54 Equation de Prédiction.....	163
Équation 55 Erreur expérience t.....	165
Équation 56 Fonction de perte cas qualitatif.....	165
Équation 57 Fonction de perte cas quantitatif.....	165
Équation 58 Erreur globale validation croisée.....	165

INTRODUCTION GENERALE

Depuis les années 1975-1980, l'environnement technico-économique des entreprises est caractérisé par une forte instabilité et des mutations brusques. Ceci est principalement dû au renversement du rapport entre l'offre et la demande, à la mondialisation et la globalisation de l'économie et aux mutations techniques et technologiques. Cette instabilité de l'environnement a obligé, et oblige encore actuellement, les entreprises à adopter de nouveaux schémas de comportement et à modifier en profondeur leurs organisations afin d'être capables de produire mieux, plus rapidement et moins cher. La survie de l'entreprise dans un marché compétitif global requiert l'amélioration de l'intégration interne et de la collaboration externe (Al-Mashari, et al., 2006). Les ERP (Enterprise Resource Planning) sont un des fondements cruciaux pour satisfaire ces exigences et déterminer l'entreprise du futur (Al-Mashari, et al., 2006). Un ERP, appelé aussi PGI¹, peut être défini comme un produit logiciel destiné à la gestion intégrale de l'entreprise. C'est un progiciel composé de modules intégrés (achat, vente, production, comptabilité...) qui communiquent, autour d'une base de données unique (Kermad L., 2003). A travers les formes d'interdépendances variées qu'il introduit, l'ERP encourage une approche transversale de l'organisation des entreprises ; ce qui permet aux utilisateurs de sortir des îlots fonctionnels (El Amrani, et al., 2004 p. 64).

Le changement est considéré dans la littérature comme l'un des facteurs les plus critiques pour le succès de l'implémentation d'ERP et de sa mise en œuvre (Nah, et al., 2003 ; Esteves, 2000). En effet, les projets ERP peuvent être synonymes de remise en cause de la manière de fonctionner de l'entreprise. La raison est la nécessité d'adapter les processus de l'entreprise en vue de la mise en correspondance du modèle de référence du progiciel et celui de l'entreprise. L'ERP agit alors comme un déclencheur clé de la réorganisation de l'entreprise à travers ses modèles de processus métiers de références, représentant les « bonnes pratiques » qui sont structurées par la connaissance accumulée au fil du temps par l'éditeur (Al-Mashari, et al., 2006).

Le changement est généralement introduit en amont du projet ERP lors de la ré-ingénierie des processus métier. La ré-ingénierie des processus est identifiée dans la littérature comme une phase du projet d'implémentation d'un ERP. Cette phase est appelée BPR dans (Kermad, 2003), Business Blueprint dans ASAP (Tatsiopoulos, et al., 2003) ou phase d'adéquation et de configuration dans (Tomas, 2000). C'est donc à cette partie de l'implémentation que nous allons nous intéresser principalement dans notre analyse.

Dans la littérature de la gestion du changement dans les projets ERP, nous trouvons des méthodes non-quantitatives et quantitatives. Alors que les méthodes non-quantitatives ne

¹ Progiciel de Gestion Intégré

permettent pas de cibler et d'orienter l'effort de gestion du changement dont elles préconisent l'utilisation, les méthodes quantitatives ne proposent pas d'actions de gestion du changement claires. Nous pensons que la difficulté de proposer des actions dans la gestion quantitative est liée à la non-définition d'une unité d'analyse unique. De plus, les méthodes de gestion quantitative qui évaluent l'impact du changement ne le font pas par rapport à la performance des processus qui sont au cœur de l'architecture d'un ERP.

Notre proposition est une combinaison de méthodes quantitatives et non-quantitatives déjà utilisées dans la gestion du changement en projet ERP. Les actions de gestion du changement que nous allons proposer sont issues de la gestion non-quantitative. Le ciblage des processus les plus touchés est réalisé avec l'aide de la prédiction utilisée par plusieurs auteurs dans des approches quantitatives de gestion du changement dans les projets ERP. Les prédictions sont réalisées sur l'impact opérationnel du changement sur les processus. L'unité d'analyse de notre étude est fixée au niveau processus. L'ampleur et l'effort de gestion du changement sont considérés comme les deux sources d'influences de l'impact du changement. La prédiction implique donc de définir des variables mesurant l'ampleur, l'effort de gestion et l'impact du changement sur les processus.

Nous allons présenter une méthode de gestion proactive du changement dans les projets ERP qui consiste à proposer des actions de gestion du changement en amont du projet lors de la phase de réorganisation des processus métier. Ces actions de gestion du changement cibleront en priorité les processus les plus impactés par le changement. Elles seront organisées selon le processus de datamining (Tufféry, 2007 p. 9), (Gokhale, et al., 1997).

Notre travail présente, dans une première grande partie, l'état de l'art de la gestion du changement dans les projets ERP. Cette partie commence par un chapitre de présentation du processus d'implémentation des ERP. Nous y précisons les dimensions, dans un projet ERP, de trois concepts que notre étude prend en compte : l'ampleur, l'effort de gestion et l'impact du changement. Nous poursuivrons cette première partie par la description de la gestion non-quantitative du changement dans les projets ERP et de leurs guides à n étapes de gestion du changement dans un deuxième chapitre. Enfin dans le dernier chapitre, nous présenterons les principaux travaux concernant la gestion quantitative pour finir par la présentation de l'utilisation du datamining pour assurer la gestion proactive du changement.

La deuxième partie de cette thèse a pour objectif de définir les variables de notre problème. Son premier chapitre précise la place du processus dans notre démarche. Ce chapitre commence par une étude de l'ingénierie des processus métier pour identifier les sources de données relatives aux processus. Nous présenterons ensuite les relations d'influence entre les dimensions de l'ampleur, de la gestion et de l'impact du changement au niveau processus. Le chapitre se termine par une détermination des variables à définir à l'aide de l'approche *Goal Question Metric*.

Le second chapitre concerne la définition de variables à expliquer. Ces variables sont les métriques opérationnelles de variation dans les indicateurs de performance des processus. Enfin, le dernier chapitre de cette seconde partie nous permettra de définir les variables explicatives à travers la définition des métriques liées à l'ampleur du changement. Nous définirons ensuite les variables qui mesurent l'effort de gestion du changement.

La troisième partie de cette thèse présente l'apprentissage et l'utilisation du modèle prédictif. Son premier chapitre est relatif à la sélection de la méthode de prédiction. Ce choix se base sur la nature des variables du problème, l'état de l'art de la prédiction en datamining, et l'état de l'art de la quantification du changement dans les projets ERP. Nous déterminons dans cette partie la taille de la base qui sera utilisée pour la construction du modèle.

Le second chapitre est relatif à l'apprentissage et à la validation du modèle. Il décrit dans un premier temps le processus de création de la base d'apprentissage. Les techniques choisies au chapitre précédant (validation croisée et un apprentissage de réseaux bayésiens) sont ensuite utilisées pour créer le modèle prédictif. Dans le troisième chapitre, nous présentons l'utilisation du modèle pour la validation d'un processus à restructurer.

PARTIE 1 ETAT DE L'ART ET DEMARCHE

Chapitre 1.1 Introduction de la première partie

Dans cette partie nous commencerons par présenter les différentes questions traitées dans le cadre de l'implémentation de progiciel. Nous allons montrer l'intérêt de la problématique de gestion du changement dans ce type de projet. Pour y répondre, nous avons commencé par analyser le changement dans ces projets. Cela nous a amenés à nous intéresser à trois concepts que sont l'ampleur, la gestion et l'impact du changement ; et à déterminer les dimensions de ces derniers. Nous avons ensuite fait un état de l'art de la gestion quantitative et non-quantitative du changement dont nous avons étudié les limites par rapport à la nature du changement dans les projets ERP. Les analyses de la gestion quantitative, de la gestion non-quantitative et des dimensions des trois concepts sont les éléments qui nous ont permis de définir la proactivité comme objectif dans notre gestion du changement². L'objectif de proactivité repose sur la nécessité de combiner la gestion quantitative et non-quantitative en tenant compte de l'ampleur, de l'effort de gestion et de l'impact du changement (Figure 1).

Nous allons ensuite présenter la démarche à travers les éléments qui permettent d'atteindre cet objectif de proactivité³. Le premier élément pour atteindre la proactivité est l'application du processus de datamining. Le second point est le choix du processus métier comme unité d'analyse⁴. Le troisième élément est l'évaluation de l'impact du changement par rapport à la performance des processus.

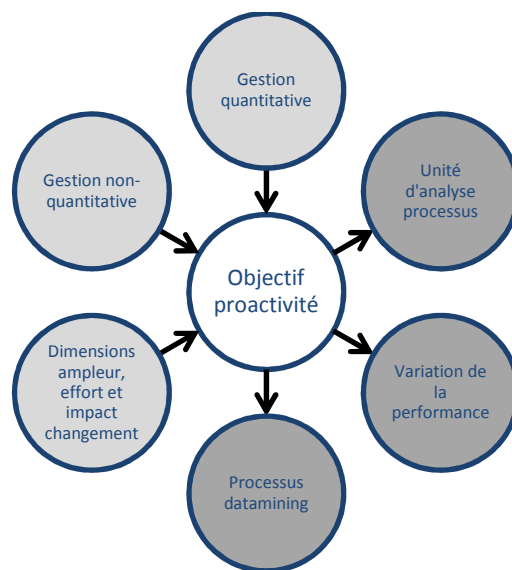


Figure 1 Objectif proactivité

² Flèches dirigées vers l'objectif Figure 1

³ Flèches partant de l'objectif Figure 1

⁴ Niveau de définition des variables

Chapitre 1.2 Implémentation de progiciel ERP

1.2.1. Introduction

L'objectif de ce chapitre est de comprendre le déroulement du projet ERP et les problématiques qui y sont liées. Nous allons mettre l'accent sur la question de la gestion du changement dans ces projets. L'étude de la phase de réingénierie nous permet ensuite de déterminer la manière dont le changement est introduit par le projet ERP. Les concepts d'ampleur, d'effort de gestion et d'impact du changement seront analysés.

1.2.2. Déroulement de l'implémentation de progiciel

Selon Lawrence, et al. (1967), les organisations sont soumises en permanence à des forces contraires qui visent soit à spécialiser les différents services (différentiation) soit à assurer l'unité de l'organisation (intégration). L'organisation est divisée en sous-systèmes qui tendent chacun à développer des compétences et des structures spécifiques pour répondre au mieux aux demandes de leurs environnements proches (fournisseurs, clients, autres services) (Perotin, 2004 p. 35). Si beaucoup d'entreprises sont attirées par les PGI pour leurs promesses d'un changement organisationnel en profondeur, c'est que ce dernier est censé apporter des gains de productivité et d'efficacité, notamment par la capacité des PGI à rendre l'organisation plus « intégrée » (Perotin, 2004 p. 34). Des caractéristiques spécifiques des PGI, modules interconnectés et référentiel unique de données notamment, semblent en effet faciliter les échanges d'informations entre services, et donner ainsi une plus grande cohérence à l'organisation, par le partage de représentations communes (Perotin, 2004). L'ERP permet à une organisation de gagner des avantages en compétitivité en économisant des ressources et en répondant à l'environnement constamment changeant (Forcht, et al., 2007).

Certains fournisseurs de progiciel proposent leur propre méthodologie d'implémentation comme SAP avec ASAP « AcceleratedSAP » (Esteves, et al., 2002A). D'autres méthodologies sont proposées par les entreprises de conseil comme "The Total Solution" d'Ernst & Young et le « Fast Track Workplan » de Deloitte & Touche (Bruges, 2002). Il y'a aussi des travaux académiques qui proposent des modèles de projet ERP (Tomas, 2000 p. 145) (Bernard, et al., 2002) (Kermad L., 2003) (Saint-leger G., 2002) (Deixonne, 2001 p. 74). Parmi ces méthodologies de mise en place d'un ERP, nous avons choisi de nous référer à celle proposée par Tomas (2000 p. 145) à cause de la clarté de la description de la

phase dans laquelle les processus sont restructurés⁵. Cette méthodologie s'articule (Figure 2) en douze phases principales⁶

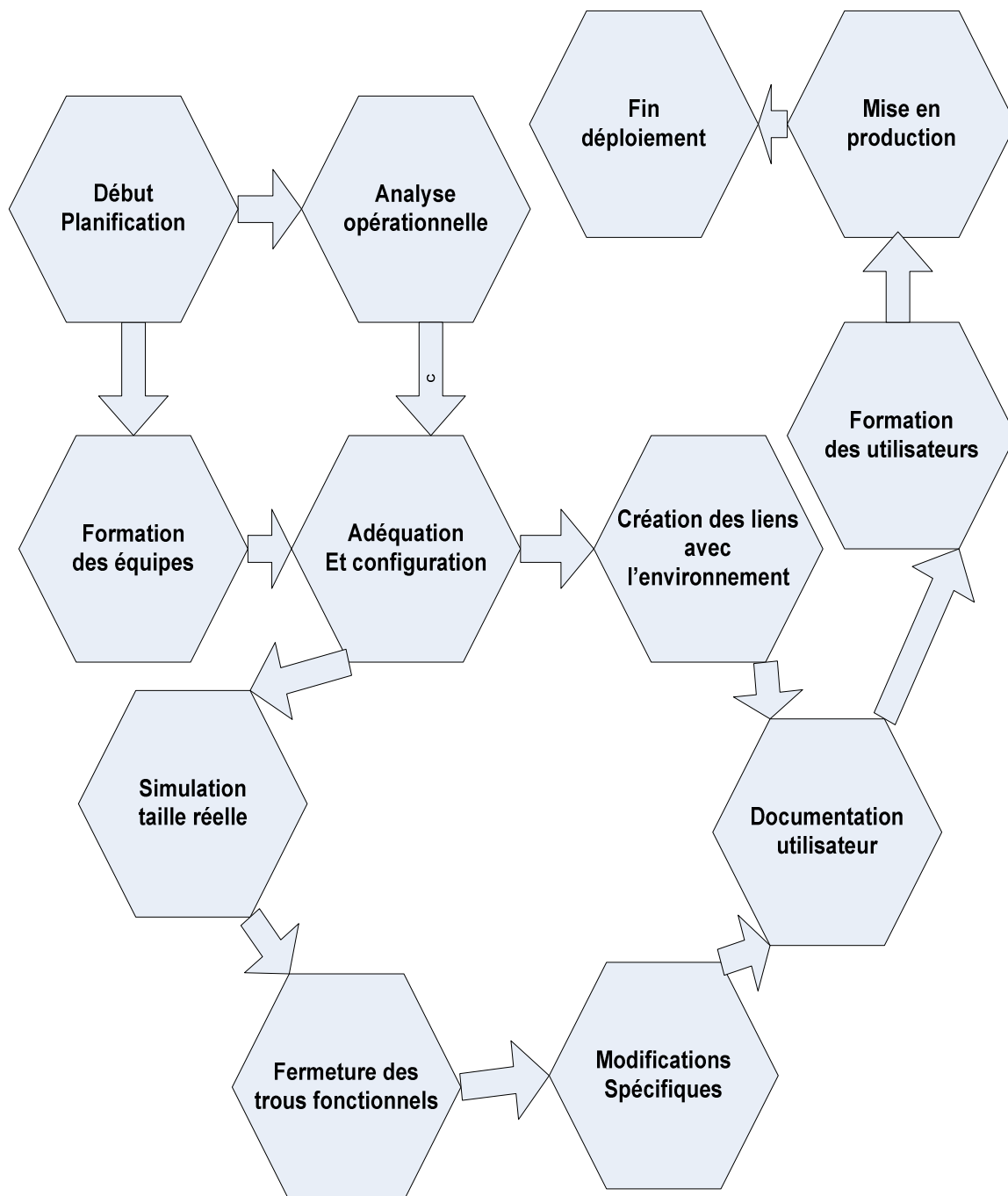


Figure 2 Flux des dérivables entre les douze phases (Tomas, 2000 p. 145)

⁵ Phase d'adéquation et de configuration (Tomas, 2000 p. 177)

⁶ Les dérivables de ces phases sont présentés en Annexe (Annexe 1 : Méthodologie de mise en place d'un ERP)

1.2.3. Problématiques de l'implémentation de progiciel

Les critères de réussite sont une manière de présenter, par ordre d'importance éventuellement, les questions importantes en implémentation de progiciel. Nah, et al. (2003) ont réalisé une revue des facteurs de succès critiques pour l'implémentation de progiciel à partir de plusieurs études empiriques (Tableau 1). Onze facteurs sont identifiés et classés en termes de nombre de citations. Le facteur « culture et programme de gestion du changement » arrive en première place ex-æquo dans Nah, et al. (2003). Ce facteur a été placé en deuxième position dans (Nah, et al., 2001) et dans (Esteves, 2000). Cela nous a convaincus de l'intérêt de la gestion du changement en implémentation de progiciel et nous a amenés à nous intéresser à ce point particulier.

Tableau 1 Facteurs clés de succès en projet ERP (Nah, et al., 2003)

Culture et programme de gestion du changement	9
Composition et travail d'équipe du groupe chargé du projet	9
BPR et personnalisation faible	8
Soutien du top management	8
Vision et business plan	7
Gestion de projet	7
Communication	6
Suivi et évaluation de la performance	6
Existence d'un champion de projet	6
Méthodologies de développement, de test et de mise au point	6
Adaptation des systèmes hérités et congruence ERP avec le business	2

Nous présentons quelques études menées dans le champ de l'implémentation pour identifier leur positionnement, leurs apports et leurs limites par rapport à la question du changement. Nous commençons par Darras (2004 p. 173) dont le travail a consisté dans la proposition d'un cadre de référence pour la gestion des projets de conception (côté éditeur) et d'exploitation d'ERP (côté client de l'éditeur) dans des PME/PMI. Ce travail s'est finalisé par la définition d'un cadre de référence permettant de placer le modèle de référence de la solution logicielle et le modèle de l'entreprise dans le cadre de la gestion de projet. L'auteur propose, dans son cadre de référence, des mécanismes adaptés afin de mettre en correspondance ces deux modèles. Le cadre produit couvre toute la gestion du projet, de la conception au déploiement, et est basé sur les points forts suivants (Darras, 2004 p. 174):

- la définition de l’architecture dans le développement d’un ERP,
- la place de l’ingénierie des Systèmes d’Information et l’importance d’un modèle de référence de la solution logicielle,
- le rôle de la modélisation d’entreprise pour une meilleure compréhension de l’organisation de l’entreprise,
- la mise en correspondance du modèle de l’entreprise et du modèle de référence de la solution pour la juste adéquation entre organisation et Système d’Information.

Darras (2004) travaille sur la mise en correspondance du modèle de l’entreprise et du modèle de référence de la solution mais ne s’intéresse pas à l’impact du changement introduit par cette opération.

Heili (2004 p. 9) part de l’hypothèse selon laquelle, la contribution des ERP au pilotage organisationnel est fortement liée à la compréhension de la dynamique de transformation des activités résultant de l’adoption de ces progiciels. Il propose un modèle conceptuel permettant l’appréhension de la dynamique d’approbation de l’ERP dans ses dimensions cognitives, sociales et politiques (Heili, 2004 p. 10). Ces contributions sont articulées autour des points ci-dessous :

- une grille d’analyse des relations entre l’informatisation et les constituants du contrôle organisationnel (Heili, 2004 p. 11).
- une analyse de la relation organisation-technologie s’intéressant à l’importance du phénomène d’approbation (Heili, 2004 p. 10).
- il pose enfin les bases d’une théorie du traitement de la question de l’approbation en mobilisant conjointement deux champs : la théorie de l’activité appliquée à l’outillage de gestion et la compréhension de l’organisation sous forme d’une communauté de communautés (Heili, 2004 p. 17).

Nous considérons l’approche de Heili (2004) comme sur-socialisée parce que, d’une part, il a une compréhension de l’organisation sous forme d’une communauté de communautés. D’autre part, il analyse la dynamique d’approbation de l’ERP dans des dimensions cognitives, sociales et politiques. Les approches dites sur-socialisées sont celles qui intègrent la reconnaissance du changement en tant qu’activité sociale impliquant des gens de différents groupes sociaux qui auront tendance à interpréter les problèmes et les situations de manières différentes et même souvent contradictoires (Collins, 1998 p. 82).

Dans cette étude, nous avons fait le choix de nous orienter vers les approches plutôt rationalistes ou sous-socialisées. En effet, dans les projets ERP, le changement poursuit des objectifs mesurables et doit donc être évalué par rapport à ces objectifs. L’approche que nous voulons proposer repose sur les notions de quantification (du changement) et de prédiction (de son impact) qui sont des caractéristiques des approches sous-socialisées. Les approches

rationalistes de l'analyse, de la planification et de gestion du changement donnent une grande importance dans la contrôlabilité du changement (Collins, 1998 p. 84).

Nous nous sommes ensuite intéressés aux travaux de Millet (2008), Saint-Leger (2005) et Perotin (2004). Ces travaux seront présentés dans la section relative aux méthodes non-quantitatives de gestion du changement car ils intègrent des « guides à n-étapes » de gestion du changement dans les projets ERP. Les « guides à n-étapes » sont des approches rationalistes parce qu'elles prétendent que le changement a des sorties prédictibles et qu'elles se prêtent à une planification détaillée de sa gestion (Burnes, 1996). Le modèle de changement des « guides à n-étapes » inclut les caractéristiques clés suivantes (Collins, 1998 p. 84) :

- Une analyse rationnelle du changement organisationnel
- Une approche séquentielle de la planification et de la gestion du changement
- Un ton et une perspective généralement optimistes

1.2.4. Le BPR dans le projet ERP

Les technologies de l'information sont des déclencheurs critiques du changement (García, et al., 2004). La reconfiguration des processus d'entreprise s'effectue souvent dans le cadre de projets d'évolution du système d'information (Morley, et al., 2007 p. 75). Dans le projet ERP, nous nous intéresserons particulièrement d'adéquation et de configuration qui introduit le changement par la restructuration des processus ou BPR (Tomas, 2000 p. 177). Contrairement aux réorganisations structurelles qui modifient le regroupement des activités mais souvent sans toucher les tâches opérationnelles, le BPR introduit des bouleversements au niveau des postes de travail (Morley, et al., 2007). Hammer définit le processus métier comme suit : c'est une séquence organisée d'activités connexes qui, ensemble, créent un résultat ayant de la valeur pour le client (Hammer, 2002 p. 80). En s'inspirant de la norme ISO (ISO, 2000), la plupart des entreprises déclinent leurs processus en (Debauche, et al., 2004 p. 17):

- processus métier (ou opérationnel, de réalisation, business process) : vendre, concevoir, fabriquer un bien ou un service ;
- processus de support : gérer les ressources humaines, les ressources matérielles, maintenir et gérer le système d'information ;
- processus de management : définir une stratégie, organiser, planifier, communiquer, gérer la qualité (enquêtes, audits, documents du système qualité, analyse et traitement des données, indicateurs ...).

La démarche du BPR commence par une identification et une description des processus existants à optimiser. Une analyse des processus est ensuite menée en recherchant des dysfonctionnements et des solutions d'optimisation. Elle se poursuit ensuite par une conception, une formalisation et une validation des processus cibles. Sa dernière étape est la mise en place des nouveaux processus. Dans le BPR, les améliorations recherchées pour un

processus s'inscrivent dans le cadre d'un objectif général de management. On peut distinguer les catégories d'améliorations suivantes (Morley, et al., 2007 p. 76) :

- efficacité : on attend avant tout un meilleur fonctionnement du processus, en particulier par la réduction de la durée du cycle d'exécution et par la qualité des décisions;
- relation client : il s'agit principalement d'améliorer la qualité perçue par le client pour améliorer sa satisfaction ;
- efficience : l'objectif majeur est la réduction des coûts ;
- flexibilité : on recherche surtout à obtenir un système flexible pouvant être modifié rapidement en cas d'évolution des contraintes et/ou de la stratégie.

Selon Hammer, et al., (2003 p. 62), l'exigence de simplicité de la réingénierie a d'énormes conséquences sur la façon de concevoir les processus et de profiler les organisations, et cela à divers niveaux (Hammer, et al., 2003 p. 62):

- regroupement de plusieurs postes en un seul,
- le travail est réalisé là où c'est le plus logique,
- les vérifications et contrôles sont réduits : ils ne sont gardés que s'ils sont économiquement justifiés ;
- les unités de travail évoluent : de service fonctionnel en équipe responsable d'un processus ;
- les postes de travail évoluent : de tâche simple à un travail multidimensionnel,
- les rôles évoluent : de postes contrôlés, vers des postes à responsabilité et autonomes ;
- la préparation à l'exercice d'un métier évolue : de la formation vers l'éducation,
- les organigrammes évoluent : de hiérarchiques à plats.

Plusieurs stratégies de BPR peuvent être utilisées en implémentation de progiciel. Ces stratégies vont de l'approche « déclenché par la technologie » à l'approche « tableau vierge » (O'Leary, 2000 p. 73). Dans une réingénierie de type « tableau vierge », le design du système commence par un tableau vierge (O'Leary, 2000 p. 75). Les processus sont reconfigurés pour répondre aux besoins et aux exigences de l'entreprise. Dans cette approche, théoriquement, il n'y a pas d'artefacts ou de processus restrictifs prédéfinis. Idéalement, les développeurs peuvent générer un design de système optimal pour une organisation particulière. Dans le cas d'un ERP, cette approche implique l'ordonnancement séquentiel suivant : restructurer d'abord puis choisir le progiciel ensuite (O'Leary, 2000 p. 75). Puisque, le progiciel est paramétré pour répondre aux besoins du design, le coût d'implémentation d'une réingénierie de type « tableau vierge » dépasse celui d'une réingénierie de type « déclenchée par la technologie ». Les désavantages de l'approche « tableau vierge » sont (O'Leary, 2000 p. 80):

- l'absence de structure pour aider le design,
- il n'y pas de raison pour la réingénierie,

- les designs peuvent être sous-optimaux,
- ils n'y a pas de limites pour le design initial,
- un design choisi peut s'avérer infaisable,
- un design peut ne pas fonctionner avec l'ERP choisi,
- il peut être coûteux en temps ou en argent de développer et d'implémenter le design choisi,
- il peut y avoir un grand nombre de consultants et de rapports intermédiaires résumant leurs travaux,
- il peut n'y avoir aucun progiciel disponible.

Dans la réingénierie « déclenchée par la technologie », la technologie de l'information est choisie pour réaliser la réingénierie avant qu'aucun système, artefact ou processus ne soit reconfiguré (O'Leary, 2000 p. 74). Cela entraîne le fait que les choix de réingénierie sont dirigés en fonction d'une technologie particulière choisie pour réaliser la réingénierie (Al-Mashari, et al., 2006). Par exemple, si la technologie est un ERP, le système ERP est choisi d'abord ; ensuite, les processus de l'entreprise sont choisis parmi ceux que le progiciel peut recevoir. Ainsi, l'ERP conduit la réingénierie des processus (O'Leary, 2000 p. 74). Les désavantages du « déclenchée par la technologie » sont en grande partie les inverses des avantages du « tableau vierge » (O'Leary, 2000 p. 80):

- la réingénierie est limitée par un outil particulier utilisé pour l'implémentation ;
- la réingénierie est limitée par la connaissance sur les artefacts et les processus qui sont inclus dans l'outil ;
- l'évolution du système peut être limitée par la technologie ;
- le design choisi est un design auquel d'autres entreprises peuvent avoir accès ;
- il peut y avoir une confusion pour savoir s'il s'agit d'une implémentation de technologie ou d'une activité de réingénierie ;
- il peut ne pas y avoir de meilleures pratiques disponibles dans l'ERP pour certaines fonctionnalités, limitant son usage.

Dans la plupart des projets de réingénierie, le choix du « tableau vierge » ou « déclenchée par la technologie » n'est pas discret ; il y a plutôt un continuum entre les deux (O'Leary, 2000). Typiquement, la réingénierie se produit quelque part sur ce continuum. Soit plus vers « tableau vierge » ou plus vers « déclenchée par la technologie », utilisant plus ou moins de portions de chaque (O'Leary, 2000). Ce continuum apparaît sous d'autres formes dans les travaux de Cooke, et al. (1998). Cooke, et al. (1998) ont mené une enquête sur les stratégies de BPR dans l'implémentation de l'ERP SAP (Tableau 2). Dans cette enquête la stratégie « BPR avant la sélection de SAP » correspond à la réingénierie par « tableau vierge » et est seconde pour la fréquence d'utilisation mais dernière pour l'efficacité. La stratégie « BPR vers le modèle de référence de SAP » correspond à la réingénierie « déclenchée par la technologie » et est la plus fréquemment utilisée mais aussi la plus

efficace. Les deux autres stratégies ne correspondent exactement ni à l'une ou à l'autre des types de réingénierie, elles sont donc à l'intérieur du continuum.

Tableau 2 Stratégies de BPR en implémentation de SAP (Cooke, et al., 1998)

Stratégie	Efficacité (1 à 5)	Fréquence d'utilisation (rang)
BPR vers le modèle de référence de SAP	4.0	1
BPR tous les processus, même si SAP ne les affecte pas.	3.8	4
BPR des processus affectés par SAP seulement	3.7	3
BPR avant la sélection de SAP	3.5	2

O'Leary (2000) a identifié l'étendue du changement dans l'organisation et dans le progiciel en projet ERP.

Changements minimaux dans l'organisation et dans le progiciel (O'Leary, 2000 p. 120). Le progiciel doit être assez robuste pour s'adapter aux processus de l'entreprise ; et l'organisation doit pouvoir subir des changements quand le progiciel ne correspond pas exactement aux processus de l'entreprise. L'étendue du changement décrit se situe dans le continuum.

Changements importants dans l'organisation et minimaux dans le progiciel (O'Leary, 2000 p. 123). Les utilisateurs changent leurs processus métiers pour utiliser le progiciel. Ce dernier permet à l'entreprise de fonctionner de manière plus efficace. Cette étendue du changement correspond à la réingénierie « déclenchée par la technologie ».

Changements minimaux dans l'organisation et étendus dans le progiciel (O'Leary, 2000 p. 124). Certaines firmes peuvent changer leur progiciel pour le faire correspondre à leur processus existant ou des meilleures pratiques non disponibles dans l'ERP. Changer le progiciel rend difficile et coûteux son maintien et sa mise à jour. Cette étendue du changement aussi se situe dans le continuum.

Changements étendus dans l'organisation et dans le progiciel (O'Leary, 2000 p. 126): Demande des ressources et du temps substantiel. L'avantage est qu'en cas de succès, l'entreprise obtient le progiciel et les processus qu'il avait identifiés comme désirables. Cet étendu de changement correspond à la réingénierie de type « tableau vierge ».

Le cadre principal de notre étude est le projet ERP. Nous devons faire une restriction de ce cadre par rapport à une stratégie de BPR spécifique. Pour les raisons ci-dessous, nous estimons, en effet, qu'il est difficile de proposer une démarche de gestion du changement applicable quelque soit le type de réingénierie utilisé :

- l'ampleur du changement dans l'organisation dépend du type de réingénierie utilisé dans le projet ERP ;
- nous devons identifier le moment où est introduit le changement dans le projet ERP et y situer l'utilisation du modèle prédictif. Cet instant d'introduction du changement dépend de la stratégie de BPR utilisée.

Nous avons choisi d'analyser le changement dans les projets ERP utilisant la stratégie BPR « déclenché par la technologie » car elle est plus réaliste et moins complexe (O'Leary, 2000), plus fréquemment utilisée (Cooke, et al., 1998) et moins coûteuse (Gendron, 1996). Les inconvénients de la réingénierie de type « tableau vierge » sont en fait très critiques et sont tous dus au fait que la réingénierie est faite alors qu'aucun outil n'est choisi. Il faut noter cependant, qu'il n'est pas possible de positionner notre étude sur une borne fixe de l'intervalle du fait de la dispersion même des projets ERP sur le continuum. Le cadre de notre étude est fixé aux projets ERP, avec une stratégie de réingénierie se situant dans demi-portion du continuum la plus proche de la borne de réingénierie « déclenchée par la technologie ».

Ce que nous venons de présenter correspond à l'identification des stratégies de BPR au niveau entreprise. Nous pouvons essayer d'avoir une vision plus détaillée du déroulement de la restructuration au niveau de chaque processus. Nous verrons en effet que les choix sont réalisés sur chaque processus. Selon Tomas (2000 p. 165), quatre domaines de processus interagissent dans la définition des processus qui feront partie de la mise en production dans un projet ERP :

- le domaine des processus utilisés aujourd'hui dans l'entreprise : processus « en-cours »
- le domaine des processus que l'entreprise souhaite utiliser demain : processus « en-devenir »
- le domaine des processus configurables que l'ERP autorise par son architecture, ses fonctionnalités et ses possibilités : processus « possibles »
- le domaine des processus qui seront effectivement mis en production et qui résultent des trois domaines ci-dessus : processus « implémentés ».

C'est dans la phase d'adéquation et de configuration, qui est en amont du projet, que les changements sur les processus sont proposés et validés. La phase d'adéquation et de configuration se décompose en plusieurs sous-phases (Tomas, 2000 p. 177):

- a) l'apprentissage : c'est une phase de familiarisation et de recherche qui permet aux EMO⁷ de connaître les principales caractéristiques de l'ERP (Tomas, 2000 p. 179) ;
- b) l'adéquation : parmi les possibilités et les limites de l'ERP, l'adéquation recherche les concepts, les fonctions et les caractéristiques qui se rapprochent le plus près du processus considéré (Tomas, 2000 p. 179).

⁷ Équipes de Mise en Œuvre

- c) la configuration : dans cette étape les EMO saisissent les données de paramétrisation de l'ERP selon les décisions prises dans l'adéquation (Tomas, 2000 p. 186);
- d) le prototypage : le but de cette étape est de vérifier que les processus qui ont été configurés dans l'étape précédente fonctionnent de manière satisfaisante (Tomas, 2000 p. 188). En fonction des résultats obtenus, quatre situations peuvent apparaître :
 1. le processus donne toute satisfaction. Il y'a lieu alors de s'intéresser à l'étape suivante d'acceptation du processus ;
 2. le processus ne fonctionne pas de façon satisfaisante. Il y a lieu alors de retourner à l'étape précédente de la configuration, de modifier le script ERP en conséquence, de le prototyper à nouveau et ainsi de suite. Cette boucle Adéquation/ Configuration/ Prototypage s'arrête soit à l'obtention du résultat désiré (cas numéro 1) ; soit, lorsque toute les configurations possibles ayant été essayées, le résultat escompté n'a pas pu être finalement atteint (cas numéro 3 et 4) ;
 3. le processus ne fonctionne pas de façon satisfaisante, mais l'entreprise se satisfait du résultat atteint par la configuration de l'ERP et accepte de modifier son processus initial en conséquence. L'acceptation peut alors débiter ;
 4. le processus ne fonctionne pas de façon satisfaisante et l'entreprise confirme et conserve son processus initial. Elle identifie un trou fonctionnel en pratiquant une analyse des écarts ou « Gap Analysis » ;
- e) l'acceptation : l'issue de la boucle Adéquation/ Configuration/ Prototypage est documentée dans l'encyclopédie du projet, les livrables sont transférés sur la machine de production et les scripts ERP associés y sont implémentés (Tomas, 2000 p. 189).

Nous nous intéressons dans cette étude à la phase d'adéquation et de configuration (Figure 3). Cette phase est déclenchée par la restructuration d'un processus « en-cours » en un processus « en-devenir ». Elle consiste à exécuter la boucle Adéquation/ Configuration/ Prototypage sur le processus « en-devenir ». Cette boucle se termine lorsque l'acceptation de ce processus est effective. C'est exactement dans cette phase que le changement opéré sur les processus est décidé puis validé. L'aide à la décision doit être apportée dans cette phase avant de passer à la validation (Camara, et al., 2005).

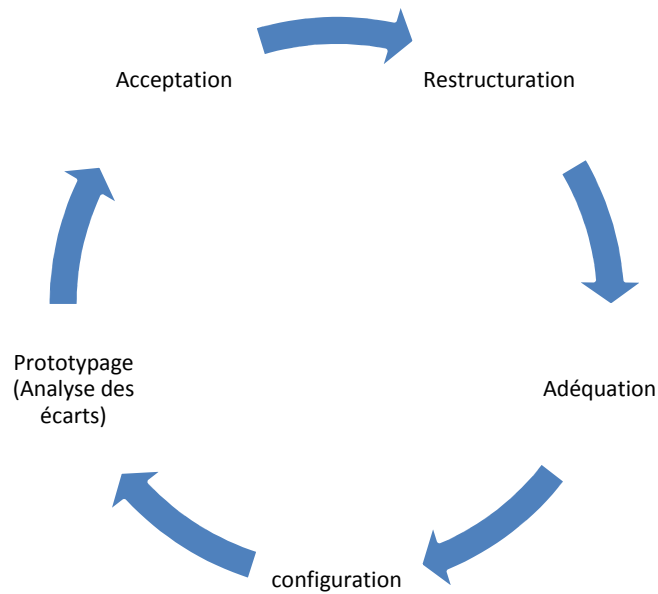


Figure 3 Phase d'adéquation et de configuration

1.2.5. Dimensions de l'ampleur, de l'effort de gestion et de l'impact

L'objectif de ce travail est l'analyse de l'impact du changement apporté par la réingénierie des processus métier. Les concepts majeurs de cette problématique sont l'ampleur, l'effort de gestion et l'impact du changement. Nous considérons que l'impact du changement est le résultat de l'ampleur et de l'effort fourni pour gérer le changement.

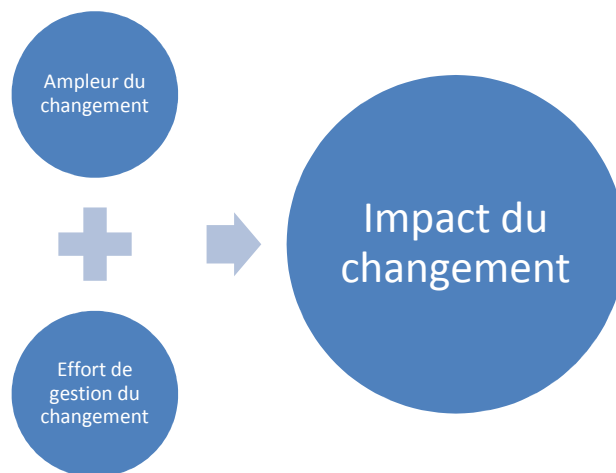


Figure 4 Relations entre l'ampleur, l'effort de gestion et l'impact

Pour atteindre notre objectif, il nous a semblé nécessaire de déterminer les différentes dimensions de chacun de ces trois concepts. Nous faisons référence à une dimension du changement, de son impact ou de sa gestion, comme un aspect au travers duquel se manifeste celui-ci. Selon Bernard, et al. (2002), le changement au sein d'une organisation peut s'opérer à plusieurs dimensions qui peuvent être concrètes ou abstraites. Nous pensons que les métriques définies dans la littérature autour de ces trois concepts peuvent constituer une troisième catégorie : celle des dimensions quantifiées. Le quantifié, le concret et l'abstrait représentent, en effet, des degrés d'abstraction croissants. Les dimensions d'un même type peuvent aussi être classées selon le niveau physique dans lequel ils se manifestent dans l'organisation. Pour le changement, nous avons identifié dans la littérature les dimensions suivantes :

- **Dimensions abstraites de l'ampleur du changement**
 - Niveau organisation
 - la culture (Bernard, et al., 2002) (García, et al., 2004),
 - les valeurs (Bernard, et al., 2002),
 - la philosophie (García, et al., 2004)
- **Dimensions concrètes de l'ampleur du changement :**
 - Niveau individuel et poste de travail ou activité :
 - les rôles (ROL)
 - les responsabilités (RSP) (Hammer, et al., 2003)
 - les compétences requises (COR)
 - Niveau processus
 - Changement radical des processus métiers (CR) (Hammer, et al., 2003)
 - l'écart entre les processus actuels et les processus cibles (EPAC) (Bernard, et al., 2002)
 - les activités composant le processus (ACT) ; suppression des activités sans valeurs ajoutés (Hammer, et al., 2003)
 - le degré d'intégration des processus (DIP) (Barki, et al., 2003)
 - complexité des nouveaux processus (CPX) (Bernard, et al., 2002)
 - le système d'évaluation des performances (Bernard, et al., 2002),
 - Niveau organisation
 - le changement apporté par la réingénierie dans l'entreprise (Kernad L., 2003) (Bernier, et al., 2003) (Bernard, et al., 2002)
 - les responsabilités
- **Dimensions quantifiées de l'ampleur du changement**
 - Niveau organisation
 - la variable magnitude BPR (Barki, et al., 2004)
 - la variable étendue BPR (Barki, et al., 2004)
 - la variable profondeur BPR (Barki, et al., 2004)

- la variable de transformation organisationnelle (Wischnevsky, et al., 2001)
- la variable de changement dans la stratégie (Wischnevsky, et al., 2001)
- les variables de changement de la structure (Wischnevsky, et al., 2001)
- la variable de changement dans le pouvoir (Wischnevsky, et al., 2001)

Les conséquences ou l'impact de ce changement peuvent aussi être classées dans les catégories abstraites, concrètes et quantifiés.

- **Dimensions abstraites de l'impact du changement**

- Niveau utilisateurs et groupe
 - la non-satisfaction des utilisateurs (NSU) (Bernard, et al., 2002),
 - la résistance au changement (RST) (Morley, et al., 2007 p. 103)
 - la sous-utilisation ou le rejet du nouveau système (Holland, 1998).
 - l'incapacité de répondre à un tel changement (Tatsiopoulos, et al., 2003) ; les employés ne seront pas facilement à même d'exécuter le processus (CEP) (Bernard, et al., 2002)

- **Dimensions concrètes de l'impact du changement**

- individu
 - qualité du travail
 - productivité employés (PROD) (Bernard, et al., 2002)
- processus :
 - amélioration spectaculaire de la performance des processus (ASP) (Hammer, 1990) ; efficacité des processus métiers
 - efficacité (réduction de la durée du cycle d'exécution) (RCE) (Morley, et al., 2007 p. 76)
 - efficience (réduction des coûts) (Morley, et al., 2007 p. 76)
 - relation client : (améliorer la qualité perçue) (Morley, et al., 2007 p. 76)
- organisation
 - avantages compétitifs,
 - résultats métiers
- inter-organisationnel
 - impact sur la chaîne logistique,
 - structure de l'industrie

- **Dimensions quantifiées de l'impact**

- Niveau projet
 - la variable durée projet ERP (Barki, et al., 2004)
 - la variable effort Projet ERP (Barki, et al., 2004)
 - la variable budget projet ERP (Barki, et al., 2004)
- Niveau organisation
 - Les variables de performance de l'organisation (Wischnevsky, et al., 2001)

Les dimensions de l'effort de gestion du changement représentent les leviers disponibles pour accompagner le changement et atténuer son impact.

- **Dimensions abstraites de l'effort de gestion du changement**
 - la vision stratégique (Nah, et al., 2003)
 - l'engagement de la direction (Aladwani, 2001)
 - la variation du temps (Perotin, 2004)
 - la minimisation des coûts (Aladwani, 2001)
 - la différenciation (Aladwani, 2001)
- **Dimensions concrètes de l'effort de gestion du changement**
 - la formation (Al-Mashari, et al., 2006) (Perotin, 2004)
 - la communication (Al-Mashari, et al., 2006) (Aladwani, 2001)
 - l'implication (Aladwani, 2001)
 - l'éducation (Aladwani, 2001)
 - le management personnel (Al-Mashari, et al., 2006)
 - les mécanismes d'intégration (Barki, et al., 2003)
 - les dispositifs favorisant l'intégration (Perotin, 2004)

1.2.6. Conclusion

Ce chapitre nous a permis d'identifier le changement comme problématique à traiter dans cette étude. Il a permis aussi, de positionner l'aide à la décision que nous voulons apporter dans la phase d'adéquation et la configuration du projet ERP. Nous avons analysé trois concepts importants pour notre étude : l'ampleur, l'effort de gestion et l'impact du changement. La relation générale qui lie ces concepts a été déterminée. Les dimensions de ces trois concepts ont été présentées selon le niveau d'abstraction et le niveau physique dans l'entreprise.

Chapitre 1.3 Gestion non-quantitative du changement

1.3.1. Introduction

Dans la gestion du changement en projet ERP, la plupart des approches proposées sont des approches non-quantitatives. Nous allons présenter une première catégorie d'approches qui sont basées sur les facteurs de risque et les facteurs de succès. Les approches de type guide à n étapes de gestion du changement seront présentées. Nous présenterons ensuite la gestion du changement qui analyse l'intégration organisationnelle dans le projet ERP. L'utilisation des modèles de maturité pour la gestion du changement sera étudiée.

1.3.2. Facteurs de risque et facteurs de succès en projet ERP

Certaines approches de gestion du changement étudient les réussites et les échecs pour fournir des préconisations afin d'éviter les échecs dans les futurs projets ERP (Heili, 2004). Ces auteurs proposent une identification des facteurs de risque et des facteurs de succès du projet ERP. Dans les facteurs de risque, ils citent des éléments relatifs au changement comme :

- le Manque de gestion du changement (Deloitte, 1999), une faible gestion du changement (Society, 2000) ;
- la sous-estimation de la difficulté de gérer de changement (Markus, et al., 2000), sous-estimation de l'effort requis pour une gestion du changement réussi (Stefanou, 2001) (Appleton, 1997) ;
- la minimisation de la gestion du changement et la formation par rapport aux problèmes techniques comme la gestion de projet, la configuration et l'architecture de la technologie de l'information (Deloitte&Touche, 1998) ;
- la concentration sur les éléments mesurables et faible attention sur les éléments non-mesurables (attitudes, valeurs, feelings) (Al-Mashari, et al., 2006) ;
- la résistance au changement (Malhotra, 1998) (Deloitte, et al., 1993) ;
- la volonté de saper les efforts de restructuration par les personnes qui sont intéressées par le maintien du statu quo (Boar, 1993) ;
- l'ampleur du changement apporté sur les processus et l'organisation (Bernard, et al., 2002).

Dans les facteurs de succès, ils citent l'effort de gestion du changement qui doit être consenti pour prévenir l'impact négatif du changement :

- la gestion du changement, gestion effective du changement organisationnel (Esteves, et al., 2002A)

- s'intéresser à la manière dont les gens seront formés et motivés pour utiliser la technologie afin de changer la façon dont ils font leur travail (Shields, 2001 p. 20) ;
- appliquer le concept d'apprentissage organisationnel pour le succès de la gestion du changement (Simon, 1994) ;
- s'occuper des questions organisationnelles, comportementales et culturelles pour le succès de l'implémentation d'ERP (Stefanou, 1999) ;
- considérer l'attitude des employés à l'égard du changement, (Aladwani, 2001) ;
- la formation et l'éducation (Al-Mashari, et al., 2006) ;
- identifier les sources potentielles de résistance et développer un plan pour les réduire (Aladwani, 2001) ;
- le leadership (Zairi, et al., 1995) ;
- l'engagement de la direction (Esteves, 2000) ;
- la vision stratégique (Esteves, 2000).

1.3.3. Guides à n étapes de gestion du changement

D'autres auteurs proposent des démarches incluant des guides à n-étapes de gestion du changement spécifiques aux projets ERP avec des actions définies et clairement positionnées dans le projet. Millet (2008 p. 138), par exemple, considère l'implémentation d'un progiciel comme la contribution des SI à la performance des organisations par la rencontre de trois points de vue : celui des exigences managériales de l'intégration flexible, celui de l'ingénierie d'entreprise dirigée par les modèles, celui de la maîtrise de la maturité des systèmes d'information. Il propose ainsi un cadre d'étude pour une ingénierie d'entreprise tenant compte des contraintes d'intégration dont sont porteurs les modèles des différents points de vue. Ce cadre d'étude de l'intégration repose sur (Millet, 2008 p. 138) :

- un cadre de modélisation d'entreprise et un méta-modèle issu des normes récentes que l'auteur tente d'interpréter et de compléter pour prendre en compte des référentiels métiers issus de son expérience des ERP;
- des graphes de dépendances permettant l'analyse des problématiques d'alignement des processus, de l'organisation, des systèmes d'information, de l'infrastructure, vues comme la maîtrise des dépendances entre les acteurs, composants et objets de l'entreprise,
- une caractérisation dans le méta-modèle et les graphes de dépendances de la maturité des usages pour prendre en compte l'acteur humain dans l'ingénierie afin d'accompagner les projets et les actions contribuant à la performance de l'entreprise à travers des systèmes d'information.

L'auteur cherche à caractériser les conséquences du projet sur un modèle organisationnel de l'entreprise. Il envisage de dimensionner les tâches de conduite du

changement nécessaires pour assurer la faisabilité d'une solution (Millet, 2008 p. 228). Pour faire ce dimensionnement, il faut, selon lui, relier le système existant au modèle en cours d'ingénierie. Cela permet d'« affecter » les objets du système aux acteurs de l'organisation pour éclairer le « qui fait quoi », autant dans le système cible que pour le projet. Il faut pouvoir ensuite lister les connaissances nécessaires aux différents acteurs, et vérifier comment le futur système produit les indicateurs existants ou redéfinis (Millet, 2008). Nous considérons que cette identification des connaissances nécessaires aux acteurs est un pas important dans la gestion du changement car elle permet de planifier et d'optimiser la formation à proposer à ces derniers.

Dans Saint-Leger (2005 p. 397), l'hypothèse centrale est formulée de la manière suivante : « La mise en œuvre et l'utilisation des systèmes d'information intégrés de type ERP dans les entreprises de production de biens et de services, de taille moyenne, deviendront efficaces et efficientes quand les dirigeants mettront en place dès l'amont du projet, et de manière pédagogique, un processus global de conduite du changement dans lequel seront pris en compte les particularités de fonctionnement des organisations et l'implication des acteurs ».

L'auteur commence par créer une arborescence des thèmes (Système, Impacts multidimensionnels d'un système ERP, Processus de changement) et des sous-thèmes. Il définit un corps d'hypothèses en associant à chaque sous-thème une hypothèse descriptive, une hypothèse explicative et une hypothèse prescriptive (Saint-Leger, 2005 p. 30). La dimension descriptive correspond à une information d'observation, de constat alimentée et démontrée à l'aide de matériaux expérimentaux et bibliographiques. La dimension explicative fournit une information qui explique ou donne une ou des explications possibles au constat formulé dans l'hypothèse. La dimension prescriptive représente ce que le chercheur pense qu'il est nécessaire de faire pour éviter les situations dysfonctionnelles observées. Une recherche-action est menée principalement dans deux entreprises et a permis l'évaluation⁸ de chaque triplet d'hypothèses avec les acteurs. Parmi les hypothèses de Saint-Leger (2005 p. 459), nous ne présentons que celles se rapportant au changement (Tableau 3). Ces dernières ont été toutes validées totalement dans un projet et validées partiellement dans un autre au minimum.

⁸ Hypothèse validée, partiellement validée, non validée ou non testée

Tableau 3 Hypothèses relatives au changement (Saint-Leger, 2005)

Hypothèses Descriptives	Hypothèses Explicatives	Hypothèses Prescriptives
HD2 Les directions d'entreprises se déchargent de la responsabilité de la conduite du changement sur les prestataires de services.	HE2 Les directions d'entreprise interprètent le changement comme un package standard livré avec le progiciel.	HP2 Le changement peut être accompagné, mais il doit avant tout venir de l'intérieur de l'entreprise.
HD13 Quand le projet est réussi, l'entreprise est en position de faire évoluer les utilisateurs vers l'acquisition de nouvelles fonctionnalités du progiciel.	HE13 La phase post-ERP révèle les dysfonctionnements fossilisés au moment du projet et/ou pendant la phase pré-projet.	HP13 Le changement et son plan de communication doivent être anticipés dès le début du projet et non au moment où les premières difficultés apparaissent.
HD16 Les DRH sont rarement impliquées dans les projets ERP et la conduite du changement.	HE16 Un lien existe parfois au moment de la formation sinon le projet est considéré comme étant du ressort des experts techniques et /ou aux mains des Directions opérationnelles. Les Directions Générales ne voient pas l'intérêt de la participation des DRH alors que des compétences vont être modifiées au sein de l'entreprise.	HP16 La présence de la DRH est nécessaire dans les projets ERP en particulier lors de la conception du programme de conduite du changement et le suivi de la formation intégrée.
HD23 Les pratiques internes façonnées à l'échelle des années ne peuvent s'estomper le temps d'un projet ERP.	HE23 Les programmes de conduite du changement comprimés l'espace d'un projet et tournés le plus souvent vers l'unique objet ERP ne sont pas suffisamment étoffés pour provoquer une métamorphose des organisations.	HP23 L'identification des particularités de fonctionnement des entreprises permet de confirmer, ou pas, si les meilleures conditions sont réunies pour s'engager dans un projet d'intégration de type ERP.

Al-Mashari (2006) définit quatre catégories des facteurs critiques de succès d'une implémentation de progiciel : gestion du système d'information, gestion du BPR, gestion du projet et gestion du changement. Dans la catégorie gestion du changement, il identifie les facteurs suivants (Al-Mashari, et al., 2006) :

- **communication** : la communication est la manière la plus efficace pour régler le problème du gap entre les supporteurs et les détracteurs du changement ;
- **implication des supérieurs** : créer une politique de changement compréhensible et attractive et ouvrir un environnement de discussion pour définir le meilleur chemin du changement organisationnel sont les rôles que la direction doit jouer ;
- **moyens et formation** : les moyens -qui signifient le nombre de membres de l'équipe projet, et les capacités et compétences et ces membres- sont un élément clé de la gestion du changement. Une formation, un recrutement et une allocation des ressources appropriés augmentent les capacités organisationnelles et atténuent la résistance au changement;
- **management personnel** : le management personnel est un système de support pour les membres individuels pour s'adapter en douceur et participer positivement au projet de changement organisationnel.

Aladwani (2001) suggère un cadre conceptuel intégré pour la gestion des problèmes de la résistance des employés à l'implémentation d'un ERP. Ce cadre dédié au *top management* est orienté processus et est composé de trois phases : formulation de la connaissance, implémentation de la stratégie et évaluation du statut.

La phase de formulation de la connaissance. Cette première étape vise à gérer effectivement le changement introduit par la TI en identifiant et en évaluant les attitudes des utilisateurs individuels et des groupes d'influence. Cette analyse doit s'intéresser aux questions telles que :

- Quels sont les individus et/ou groupes résistants ?
- Quels sont leurs besoins ?
- Quelles croyances et quelles valeurs est-ce qu'ils ont ?
- Quels sont leurs intérêts ?

La phase d'implémentation de la stratégie. L'équipe peut utiliser la connaissance concernant les utilisateurs potentiellement résistants issue de la phase précédente pour mettre en place des stratégies qui peuvent surmonter la résistance face à l'ERP. Dans une tentative de changer l'attitude des utilisateurs potentiels, le management peut, en premier, essayer d'affecter la composante cognitive de l'attitude des utilisateurs. Une stratégie majeure pour atteindre cet objectif est la communication. Une stratégie de communication efficace est d'informer les utilisateurs potentiels des bénéfices de l'ERP. Une autre stratégie de

communication est de donner une description générale de la manière de fonctionner du système ERP implémenté.

La seconde étape dans la phase d'implémentation de la stratégie est d'influencer la composante affective de l'attitude des utilisateurs. La première stratégie qui peut être utilisée par les managers est la minimisation du coût⁹. Par exemple, si les utilisateurs réalisent que l'ERP est une opportunité pour améliorer leur travail, et par conséquent le rendre plus attirant avec un coût additionnel minimal, alors ils vont probablement développer un intérêt pour le système. De la même façon, les groupes d'influence dans l'organisation prêtent aussi attention à l'aspect coût de l'effort d'implémentation.

Une autre stratégie qui pourrait affecter l'attitude d'adoption de l'utilisateur potentiel est la différenciation. Dans un contexte de projet ERP, la perception des utilisateurs de la haute qualité du système va sûrement avoir un impact positif sur leur attitude envers le système. Certains ERP ont une interface utilisateur difficile à manier, ce qui est un problème. Généralement les utilisateurs du système ne mesurent pas scientifiquement les attributs-qualité du système, chaque utilisateur construit plutôt sa perception du système à partir de son expérience réelle (ou socialement construite). De plus, la formation est un autre levier important de la réussite de l'implémentation d'un ERP. La formation offre une bonne opportunité pour aider les utilisateurs à s'ajuster au changement qui a été introduit par l'ERP, et aide à construire des attitudes positives à l'encontre du système.

Une autre partie de la phase d'implémentation de la stratégie est d'obtenir l'adhésion et le support des personnes connues et des leaders d'opinion. Convaincre les leaders de groupe de participer au processus d'implémentation et faire qu'ils aient l'impression d'être des acteurs clés va assurer leur engagement. Une autre stratégie est de choisir prudemment le moment de l'introduction du nouveau système. Par exemple, n'introduisez pas un ERP quand une masse critique de vos employés se sentent menacés par le système ou ont l'impression d'être forcés de l'accepter. Enfin, l'engagement et le support de la direction est la stratégie ultime qui va garantir les conditions nécessaires à l'introduction réussie du changement apporté par l'ERP dans l'organisation.

La phase d'évaluation du statut. Il faut d'abord un système de mesure de la performance pour assurer que les résultats sont atteints sur le plan métier (business). A côté de cela, c'est important d'avoir un système de performance pour suivre la progression des efforts de gestion du changement dans le projet ERP. Si le feedback est négatif, le management pourra voir qu'il y a toujours une forte résistance du personnel au changement opérationnel résultant de l'implémentation de l'ERP. Dans ce cas, le top management peut ré-identifier les besoins des utilisateurs et réévaluer l'exécution des stratégies de gestion du changement adoptées pour trouver une correspondance acceptable entre les deux.

⁹ Cost minimization

L'analyse des écarts¹⁰ (Tomas, 2000 p. 188) permet de déterminer les écarts au niveau structurel entre les processus « en devenir » et les « possibles ». C'est une évaluation des fonctions fournies par le système ERP par comparaison aux processus opérationnels nécessaires à supporter le métier (Inter nec, 1999). Le changement auquel nous nous intéressons ici est l'écart entre le processus « en-cours » et le processus « en-devenir ». Le changement opéré sur les processus, et les trous fonctionnels, étudiés par l'analyse des écarts, évoluent en sens inverse. En effet, un paramétrage faible de l'ERP, c'est-à-dire un petit nombre de trous fonctionnels, voudrait dire des changements significatifs pour faire correspondre l'organisation à l'ERP (Barki, et al., 2004). Par contre, un paramétrage poussé de l'ERP diminuerait les trous fonctionnels mais n'impliquerait qu'une faible reconfiguration des processus (Barki, et al., 2004). L'analyse des écarts, par la mise en relief quelle donne sur l'écart, permet d'avoir une idée sur l'ampleur du changement sur les processus. De plus, elle ne s'intéresse pas au niveau opérationnel des processus.

1.3.4. Intégration organisationnelle et gestion du changement

Le concept d'intégration est la pierre angulaire des systèmes ERP car leur implémentation offre une structure intégrative complète et la capacité d'intégrer tous les aspects métiers (Scott, et al., 2001). Le degré d'intégration de l'organisation évolue en fonction des changements amorcés ou subis par l'organisation, comme notamment dans le cas de la mise en place d'un ERP (Perotin, 2004 p. 36). Barki, et al (2003) sont allés dans le même sens en affirmant que : « puisque, l'objectif clé des ERP est de rendre les organisations plus intégrées, le changement dans l'intégration organisationnelle entre la post et la pré-implémentation peut être considéré comme positif ». L'enseignement que nous pouvons tirer des ces affirmations est le suivant : « dans le cadre de l'implémentation d'un ERP, l'intégration organisationnelle peut être considérée comme une dimension du changement sur les processus ou sur l'entreprise dans son ensemble ». Comme toutes les dimensions du changement, l'intégration peut apporter des effets positifs dans certaines conditions, ou négatifs dans d'autres conditions. Dans la suite, nous considérerons donc que tout effort de facilitation, de gestion ou d'accompagnement de l'intégration comme un effort de gestion du changement.

Perotin (2004 p. 115) s'est attaqué à la problématique suivante : « comment piloter le processus de mise en place d'un PGI pour obtenir un degré d'intégration élevé ? ». L'auteur a utilisé une analyse processuelle de cas de mise en place auxquelles il a participé. L'approche processuelle consiste à reconstituer le déroulement du processus à partir des événements, des étapes et des opérations qui le constituent. Il s'est attelé à décrire les différentes caractéristiques de l'intégration et la manière dont ces dernières ont été gérées dans les projets

¹⁰ Gap Analysis

ERP qu'il a suivis. Selon lui, le concept d'intégration se matérialise à travers les problèmes transverses. L'auteur identifie les tests d'intégration, une phase des projets étudiés, comme une occasion de matérialiser les problèmes transverses parce qu'ils doivent permettre de réaliser la simulation de processus de gestion transversaux à plusieurs services ou groupes d'acteurs (Perotin, 2004 p. 151). Ils consistent en l'exécution des scénarios globaux représentatifs de l'activité du site. Ses discussions avec les acteurs autour de thèmes de la réalisation de l'intégration des informations et des processus organisationnels lui ont permis d'identifier (Perotin, 2004 p. 58) :

- les origines des problèmes « transverses »
 - différences de niveaux de détail souhaités entre plusieurs acteurs d'un même flux d'informations,
 - dépendances chronologiques ou hiérarchiques dans la mise à jour des données ;
- les dispositifs favorisant l'intégration
 - initiatives individuelles de coordination,
 - gestion optimale et participative du travail en groupe, usage large de la messagerie ;
- les conséquences de l'intégration
 - manque de flexibilité dans la gestion des erreurs (propagation des erreurs potentielles, phénomène accru par l'interdépendance des fonctions et des données),
 - responsabilisation des acteurs chargés de la saisie (liés au manque de flexibilité : erreur difficilement rectifiable),
 - complexité et opacité des calculs ou des règles de gestion,
 - cohérence du système d'information (administration et maintenance) et faible évolutivité.

L'auteur a aussi analysé le changement et son accompagnement dans l'étude de cas et les décrit comme suit (Perotin, 2004 p. 59) :

- les modifications perçues
 - ajustement des effectifs et des compétences,
 - analyse plus fine de l'activité;
- accompagnement du changement
 - une perturbation, perçue comme forte, nécessite de favoriser la formation, la communication et l'implication des personnels.

L'auteur identifie les tests de validation du nouveau système par les utilisateurs, une autre étape du planning, comme importants pour la résistance au changement. Dans la phase de validation des utilisateurs, il s'agit du test par un échantillon d'utilisateurs finaux, sur site, avec des données et la configuration technique locales (micro-ordinateurs et imprimantes définitifs), des processus conçus et approuvés précédemment. Ceux-ci doivent avoir au préalable reçu la formation au progiciel et aux transactions utilisées.

Dans son étude de cas, l'auteur a observé aussi l'utilisation du temps et ses variations (accélérer, ralentir) comme faisant partie des modes de management du processus de changement (Perotin, 2004 p. 213). En effet, la période dévolue à la mise en œuvre d'un ERP voit des transformations du rythme de travail et donc des perceptions de l'écoulement du temps par les acteurs. Ces décalages dans le temps peuvent être mis au service du changement organisationnel (Perotin, 2004 p. 209). Ils peuvent être vus comme des déclencheurs, aidant les individus soit à prendre du recul pour réfléchir sur leurs activités, soit au contraire, à les plonger, par un sens accru de l'urgence, dans l'action (Staudenmayer, et al., 2002).

L'approche retenue par la direction d'un des projets étudiés dans la phase de déploiement est un mélange d'urgence et de pragmatisme (Perotin, 2004 p. 210). L'accent est mis sur la réalisation des fonctionnalités indispensables, il ne s'agit pas de créer un système idéal. Le rythme du projet est volontairement rapide, le planning est donc très resserré et le travail doit donc être mené avec un sentiment d'urgence permanent, surtout dans la phase de déploiement.

L'auteur note aussi, dans un des projets étudiés, la mise en place d'une cellule de gestion du changement constitué de trois personnes possédant une expérience des grands projets de refonte organisationnelle. Cette cellule favorise l'emploi des méthodes de gestion du changement : animation de réunions, élaboration de supports et contenus de formation, accompagnement des équipes projet responsables de site dans le processus du changement (Perotin, 2004 p. 181). Elle doit également s'assurer que là où les métiers et les structures évoluent, une aide substantielle et effective est apportée. Cette cellule recueille les besoins des utilisateurs en termes d'assistance et de formation. La nature de ses interventions peut aller de la simple mise en relation entre deux interlocuteurs (demandeur et possesseur d'informations), jusqu'à l'élaboration d'un programme de formation sur des points particuliers (fiscalité, gestion de projet, etc.). Nous pouvons dire que l'auteur est arrivé à observer comment se matérialisaient et comment ont été gérés le changement et l'intégration (une dimension du changement) dans les projets suivis.

Barki, et al. (2003), ont d'abord affiné le concept d'intégration organisationnel (IO) en proposant un cadre conceptuel qui relie cinq concepts (interdépendance¹¹, barrières à l'IO¹², mécanismes pour réaliser l'IO¹³, les efforts IO¹⁴, les bénéfices de l'IO). Ils ont séparé les processus opérationnels (activités principales à valeur ajoutée) des processus fonctionnels qui sont les activités de support à la chaîne de valeur. Une application dans laquelle des

¹¹ Séquentielle, réciproque et cloisonnée (polled)

¹² Spécialisation fonctionnelle, cadres de référence et considération politiques

¹³ Ajustement mutuel, supervision directe, standardisation du travail, standardisation des sorties, standardisation des compétences et des connaissances et standardisation des normes

¹⁴ Faible, modéré, élevé et très élevé

hypothèses de recherche sont développées pour l'implémentation d'ERP, est donnée à titre d'exemple :

- H1A : les implémentations des modules PP (Production Planning) et MM (Materials Management) de SAP R/3 seront généralement caractérisées par une interdépendance séquentielle et/ou réciproque, alors que l'implémentation du module RH (Ressources Humaines) sera généralement caractérisée par une interdépendance cloisonnée.
- H2A : l'implémentation des modules opérationnels d'un ERP requiert plus d'efforts que celui de ses modules fonctionnels.
- H3A : implémenter des modules opérationnels externes de l'ERP va requérir plus d'efforts que d'implémenter ses modules opérationnels internes.
- H4A : implémenter des modules fonctionnels externes de l'ERP va requérir plus d'efforts que d'implémenter ses modules fonctionnels internes.
- H5A : la spécialisation fonctionnelle sera une barrière plus importante quand on implémente des modules fonctionnels (internes ou externes) d'un ERP, que quand on implémente ses modules opérationnels (internes ou externes).
- H6A : la présence de structures de référence distinctes sera une barrière plus importante quand on implémente des modules fonctionnels (internes ou externes) d'un ERP, que quand on implémente ses modules opérationnels (internes ou externes).
- H7A : les barrières politiques seront plus importantes lorsqu'on implémente des modules (fonctionnels ou opérationnels) orientés interne que quand on implémente des modules (fonctionnels ou opérationnels) orientés externe.
- H8A : les efforts requis, pour implémenter un ERP, seront plus considérables quand les barrières à l'IO sont présentes.
- H9A : en général, implémenter les modules opérationnels d'un ERP nécessitera une utilisation de mécanismes d'intégration plus complexes que ceux nécessaires pour implémenter les modules fonctionnels.
- H10A : utiliser les mécanismes d'intégration demandera des efforts plus importants pour l'implémentation des modules opérationnels d'un ERP que pour implémenter ses modules fonctionnels.
- H11A : utiliser un mécanisme d'intégration approprié pour implémenter les modules opérationnels d'un ERP (i.e. supervision directe, standardisation des sorties, standardisation du travail, planification, ajustement mutuel) requiert moins d'efforts qu'utiliser des mécanismes d'intégration inappropriés (i.e. standardisation des normes, standardisation des compétences et des connaissances).
- H12A : utiliser des mécanismes d'intégration appropriés pour implémenter les modules fonctionnels d'un ERP (i.e. standardisation des normes, standardisation des compétences et des connaissances) requiert moins d'efforts que d'utiliser des

mécanismes d'intégration inappropriés (i.e. supervision directe, standardisation des sorties, standardisation du travail, planification, ajustement mutuel).

- H13A : en général, l'implémentation d'un ERP améliorera la performance organisationnelle.
- H14A : sous des conditions de choc ou de turbulence environnementale, l'implémentation d'un ERP sera négativement liée à la performance.
- H15A : sous des conditions de choc ou de turbulence environnementale, l'utilisation des mécanismes de réduction de la complexité va modérer la relation négative entre l'implémentation de l'ERP et la performance.

L'auteur a fait des hypothèses sur l'effort nécessaire pour utiliser les mécanismes d'intégration afin de lever les barrières à l'IO, selon le type d'interdépendance ou le type d'IO, et les bénéfices éventuels. Ces hypothèses définissent de manière précise les conditions d'utilisation des mécanismes d'intégration. Nous considérons ainsi l'ensemble de ces hypothèses comme un moyen de gestion de l'intégration, donc de gestion du changement. Le travail n'est pas au niveau processus et ne comporte pas de mesure de l'intégration. En effet, les auteurs parlent de modules ERP qui sont des éléments très larges ne permettant qu'une vision globale des choses.

1.3.5. Le modèle de maturité pour les processus métier

Les modèles de maturité sont apparues à la fin des années 1980 avec la création du *Process Maturity Framework* par Watts Humphrey (1989). Le *Process Maturity Framework* a ensuite évolué en *Capability Maturity Model for Software CMM* (Paulk, et al., 1995) et ensuite en *Capability Maturity Model Integration CMMI* (Chrissis, et al., 2003 p. 7). Il est ainsi devenu le standard prééminent pour l'évaluation de capacité des organisations qui développent des systèmes logiciels intensifs. Le Business Process Maturity Model (BPMM) est une version des modèles de maturité spécifique aux processus métiers. Le BPMM a été rédigé pour guider l'amélioration des processus. Il a été érigé en standard par l'OMG afin d'aider les organisations à être mieux préparées pour le déploiement de technologie telles que les progiciels (OMG, 2007A). Ils considèrent que la plupart des problèmes rencontrés se manifestent par des faiblesses dans les processus métiers ciblés par l'application d'entreprise. Le BPMM est divisé en cinq niveaux de maturité qui représentent les états au travers desquels l'organisation est transformée lorsque ses processus et sa capacité sont améliorés. Ces niveaux comprennent (OMG, 2007A p. 73):

1. **Initial** (management de type « lutte contre incendie ») : il n'y a pas d'objectif spécifique. Le succès de ces organisations dépend des compétences d'héroïques personnes et non de l'utilisation de processus testés ;
2. **Géré** (management d'unité de travail) : l'objectif est de créer le fondement d'un management dans chaque unité de travail ou projet ;

3. **Standardisé** (management des processus) : l'objectif est d'établir et d'utiliser une infrastructure commune de processus organisationnels et des avantages associés aux processus pour assurer une cohérence dans la manière dont le travail est réalisé pour produire les biens et services dans l'organisation ;
4. **Prévisible** (management de la capacité) : l'objectif est de manager et d'exploiter la capacité de l'infrastructure de processus organisationnels et des avantages associés aux processus pour assurer des résultats prévisibles avec une variation contrôlée ;
5. **Innovant** (gestion du changement) : l'objectif est d'améliorer de manière continue les processus de l'organisation et les biens et services résultants à travers la prévention des défauts et des problèmes, la capacité continue, et des améliorations innovantes planifiées.

Au niveau 4, les changements sont réalisés sur les processus pour réduire la variation, produire une performance stable ; et obtenir des résultats prévisibles. Au niveau 5, les améliorations du processus sont réalisées pour combler l'écart entre les performances et les résultats actuels, et les performances et les résultats désirés. Le changement décrit dans le niveau 5 a le même objectif que celui introduit par la restructuration lors de l'implémentation de progiciel ERP : améliorer les performances des processus. La différence réside dans le fait que le changement au niveau de maturité 5 est continu et incrémental alors que celui du projet ERP est discontinu et radical. A chaque niveau de maturité, le modèle fait correspondre des processus de domaine qui collectivement permettent d'atteindre la maturité du niveau en question. Chaque processus de domaine consiste en une collection de pratiques ou sous-pratiques qui indiquent ce qui doit être réalisé. Pour atteindre les différents niveaux de maturité, la spécification propose des directives regroupées dans les catégories suivantes :

- les directives pour la mesure et l'analyse,
- les directives pour l'inspection du produit du travail,
- les directives pour la résolution des problèmes et des décisions,
- les directives pour la gestion des risques,
- et les directives pour la gestion du changement organisationnel.

Chaque pratique ou sous-pratique d'un processus de domaine fait appel à une ou plusieurs directives des catégories présentées ci-dessus. Nous nous sommes principalement intéressés aux directives de la catégorie gestion du changement organisationnel. Les directives de cette catégorie sont regroupées en activités comme suit (OMG, 2007A p. 436) :

Activité 1 : maintenir le parrainage du programme de changement. Le parrainage du programme de changement par la direction est établi et maintenu.

Activité 2: maintenir la vision, les stratégies et les objectifs du changement. La description des objectifs et des stratégies du changement organisationnel est établie et maintenue par la direction.

Activité 3: maintenir la justification du changement. Les raisons métier qui justifient le changement organisationnel sont établies et maintenues par la direction.

Activité 4: maintenir la description de la solution. La description de la solution pour le changement organisationnel est établie et maintenue.

Activité 5: maintenir le plan de changement. Les plans pour développer et déployer le changement organisationnel sont établis et maintenus.

Activité 6: développer la solution de changement. Les plans pour développer et déployer le changement sont implémentés.

Activité 7: déployer la solution de changement. La solution de changement organisationnel est déployée conformément aux plans.

Activité 8: superviser le programme de changement. Les activités et les résultats du changement organisationnel sont supervisés par rapport aux stratégies, aux objectifs et aux plans, et les actions correctives appropriées sont réalisées.

Activité 9: incorporer les leçons d'amélioration apprises. Les leçons apprises dans la planification, le développement et le déploiement du changement organisationnel sont enregistrées et incorporées aux activités de changement organisationnel de manière appropriée.

Activité 10: communiquer les informations d'amélioration. Les informations, les statuts, les mesures et les autres résultats du changement organisationnel sont fournis à ceux qui sont affectés et communiqués à travers l'organisation.

Les directives de gestion du changement interviennent dans les pratiques ou les sous-pratiques (SP) suivantes selon le niveau de maturité considéré :

- Niveau 2
 - SP : maintien de la justification de l'amélioration des processus,
 - SP : définir les stratégies d'amélioration de processus les mieux appropriées à l'organisation et à ses produits et services,
- Niveau 3
 - SP : maintenir les plans d'action processus,
 - SP : déployer les avantages des processus et leurs changements,
 - SP : identifier les barrières potentielles du déploiement des améliorations,
- Niveau 5
 - SP : déterminer et documenter les stratégies pour traiter les barrières potentielles de chaque déploiement d'amélioration,
 - SP : gérer le déploiement des améliorations.

Les directives de gestion du changement constituent une sorte de « guide à n étapes » pour la gestion du changement. Le BPMM intègre une partie gestion quantitative du changement. Le niveau de maturité 5 du BPMM est celui où le changement est le mieux pris en compte. Le niveau de maturité 5 est composé des processus de domaine suivants (OMG, 2007A p. 84) :

- planification des améliorations organisationnelles,
- alignement de la performance organisationnelle,
- amélioration continue de la capacité,
- amélioration innovatrice de l'organisation,
- déploiement des améliorations organisationnelles.

Nous nous sommes intéressés au processus de déploiement des améliorations organisationnelles car il contient des pratiques qui portent sur la gestion prédictive et quantitative du changement. Le processus de déploiement des améliorations organisationnelles améliore continuellement et quantitativement la performance et la qualité dans l'organisation en transférant les améliorations dans une utilisation systématique (OMG, 2007A). Le déploiement des améliorations organisationnelles implique (OMG, 2007A p. 424) :

- sélectionner les améliorations à déployer et planifier leur déploiement,
- prédire statistiquement les coûts et bénéfices des améliorations sélectionnées,
- gérer le déploiement des améliorations selon les plans,
- mesurer les coûts et les bénéfices et les comparer aux prédictions.

Le modèle préconise de mesurer l'impact des améliorations en termes de coûts et de bénéfices sans préciser comment et avec quelles variables. La quantification du changement n'est donc pas entièrement intégrée. La proactivité et la prédiction sont recommandées dans le modèle pour gérer le changement, mais ne sont pas déterminées de manière précise.

1.3.6. Conclusion

Les premières approches de gestion du changement dans les projets ERP auxquelles nous nous sommes intéressées sont celles qui définissent les facteurs de risque et les facteurs de succès. La limite principale de ces approches est le fait qu'elles ne définissent pas la manière d'éviter les accueils ou comment mettre en place les facteurs de succès. D'autres approches que nous pourrions qualifier de guides à n étapes permettent de dépasser cette limite en proposant des actions claires et ordonnées ; ou des stratégies. Ces méthodes reposent pour la plupart sur des moyens de gestion du changement tels que la communication, les réunions, la formation, l'accompagnement, etc. (Al-Mashari, et al., 2006) (Aladwani, 2001). D'autres auteurs ont proposé des approches qui s'intéressent à la gestion de l'intégration en tant que dimension du changement dans les projets ERP (Perotin, 2004) (Barki, et al., 2003). Ces approches peuvent être considérées comme des guides à n étapes de la gestion du

changement car ils proposent des dispositifs ou des mécanismes pour faciliter l'intégration. Les différents niveaux des modèles de maturité comportent aussi des guides à étapes de gestion du changement (OMG, 2007A).

La limite commune à tous ces guides à n étapes est qu'ils ne permettent pas d'identifier de manière claire les parties de l'organisation où doit être orienté en priorité l'effort de gestion du changement. L'origine de ce problème est la déconnexion totale de ces méthodes par rapport aux techniques quantitatives qui pourraient les aider à cibler leurs actions et ainsi être plus efficaces. Nous pensons que les actions et stratégies de la gestion non-quantitative restent les leviers principaux pour la gestion de l'impact du changement. Toute quantification du changement ne doit donc être considérée que comme complément pour orienter ces stratégies. La seconde limite de ces techniques non-quantitatives est qu'elles ne proposent que des actions de réduction directe de l'impact du changement par l'effort de gestion. Les actions liées à l'ajustement de l'ampleur du changement ne sont pas explorées. Cette limite, comme la première, a pour cause la non-quantification. La quantification pourrait permettre de mesurer et d'ajuster l'ampleur du changement.

Chapitre 1.4 Gestion quantitative du changement

1.4.1. Introduction

Nous présentons, dans ce chapitre, les démarches qui intègrent la formalisation quantitative de la gestion du changement. Il s'agit pour nous de voir si ces démarches présentent ou non des limites par rapport à la gestion du changement dans les projets ERP. Nous vérifierons notamment si les trois concepts que sont le changement, son accompagnement et son impact sont pris en compte dans la quantification proposée. Nous vérifierons également si l'impact du changement est mesuré par rapport à l'amélioration de la performance du processus. Nous déterminerons aussi le niveau auquel se situe l'analyse quantitative appliquée, et la pertinence de ce niveau par rapport à la gestion du changement. Nous avons un premier groupe d'auteurs qui se sont limités à la définition de métriques liées au changement. D'autres auteurs utilisent des techniques issues du datamining ou de la statistique. Nous avons regroupé les travaux de ces auteurs selon la technique de modélisation utilisée.

1.4.2. Mesure du changement

Eriksson (2001) fait une comparaison entre un processus de vente avant et après sa restructuration. Il définit des critères de comparaison, équivalents à des métriques, qu'il calcule pour les deux versions du processus. Il calcule par exemple le nombre d'acteurs et d'activités du processus avant et après la reconfiguration. Nous pensons que cette manière de déterminer les critères de comparaison ne montre pas toute la transformation opérée sur le processus (Camara, et al., 2005). Sur le processus qu'il a pris comme exemple, il y a douze activités avant et six activités après. En suivant son raisonnement, cela nous amènerait à dire que le changement est de six activités. Supposons que huit activités ont été complètement supprimées sur ce processus, et que deux nouvelles ont été rajoutées aux quatre restantes pour faire un total de six activités pour le nouveau processus. Nous pouvons voir qu'en réalité il y a eu huit suppressions d'activités et deux rajouts, ce qui fait dix activités changées. Dans la plupart des cas, les mesures définies ne rendent pas compte de l'ampleur du changement. La mesure de l'impact et de l'effort de gestion du changement n'est pas traitée. Seule la mesure du changement est abordée et l'utilisation des métriques définies n'est pas présentée.

Esteves (2002B) quant à lui propose une approche qui est liée à la fois aux processus, au projet et à l'entreprise. Il définit dans son article dix métriques dans les dimensions suivantes : la gestion du changement, les utilisateurs, le processus ou le produit(ERP) (Tableau 4). L'intérêt de cet article réside dans le fait que certaines des métriques proposées (2, 5, 6 et 8) sont des mesures du changement sur les processus ou sur l'entreprise dans le cadre de projets ERP. Les métriques définies ne sont pas toutes des caractéristiques d'un seul

et même élément. En effet cinq métriques sont définies au niveau processus (1, 2, 3, 7 et 9), quatre au niveau projet/entreprise (4, 5, 6 et 8) et une au niveau produit (10). L'unité d'analyse présente donc des variables définies à des niveaux différents. Nous verrons, dans le chapitre suivant, les difficultés qui peuvent découler de cela dans le cas où les métriques définies sont utilisées comme variables dans la construction d'un modèle prédictif par exemple. La quantification du changement est assurée par des mesures qui rendent compte de l'ampleur du changement. Les mesures d'impacts ne sont pas définies. La démarche proposée s'arrête à la définition de métriques et ne dit pas comment ces dernières seront analysées et dans quel type d'étude quantitative elles seront intégrées.

Tableau 4 Métriques de changement en projet ERP (Esteves, 2002B)

Idx	Métrique	Description
1	Ampleur de la refonte	Ampleur de la refonte qui est nécessaire pour chaque processus métier
2	Nombre d'activités touchées	Nombre d'activités qui sont touchées pour chaque processus métier repensé
3	Nombre de départements touchés	Nombre de départements qui sont liés à chaque processus métier repensé
4	Utilisateurs concernés	Utilisateurs impliqués dans le processus de BPR
5	Nombre de processus à repenser	Nombre de processus métiers qui doivent être repensés
6	Nombre de processus touchés	Nombre de processus qui ont besoin de modifications en raison de la refonte d'un autre processus métier.
7	complexité du processus métier	« Somme du nombre d'activités, du nombre de personnes dans chaque activité, du nombre des flux de matières en provenance et à destination du processus, et du nombre de flux de matières entre les activités du processus »
8	Effort BPR	Cette métrique est liée à la part du métier impliqué dans la refonte. Elle est composée du nombre total de départements, du nombre de processus métiers redessinés et du nombre de personnes impliquées dans chaque phase.
9	Durée refonte processus	Estimation du temps nécessaire à la refonte de chaque processus métier
10	Nombre des processus ERP touchés	Nombre de processus d'ERP qui ont besoin de modifications en raison du processus de BPR.

1.4.3. Analyse de la criticité et simulation

Sienou, et al. (2006), proposent par exemple, une approche qui intègre la gestion des risques à la gestion des processus. Ils ont traité la question du changement sur les processus. L'idée générale dans cette étude est d'analyser l'incertitude sur un processus pour répondre aux questions suivantes :

- l'entreprise doit-elle reconnaître quelle n'est plus en mesure d'assurer la mise en œuvre du processus avec satisfaction ?
- avec quel niveau de confiance le processus conduira aux résultats escomptés ?

La démarche proposée consiste à appliquer à chaque phase du processus de BPM¹⁵ (planning, implémentation, contrôle et amélioration) la boucle de gestion des risques (Sienou, et al., 2006). Cette boucle est composée de l'identification, de l'évaluation, du traitement et du suivi des risques. L'application de la boucle de gestion des risques est présentée à travers un processus de fabrication et de vente d'ordinateurs sur commande (Sienou, et al., 2006) :

- identification des risques :
 - venant de l'environnement extérieur
 - R1 (Marché) : changement dans les exigences du marché,
 - R2 (Politique) : changement dans la réglementation sur les composants informatiques,
 - venant de l'entreprise
 - R3 (Processus) : capacité de fabrication,
 - R4 (Ressources) : changement dans les contrats fournisseurs ;
- évaluation : chaque risque est analysé par rapport à sa probabilité d'occurrence et son impact. L'impact est mesuré sur l'objectif qui est, dans l'exemple, une augmentation de la marge de 15%. Les résultats sont présentés dans un modèle d'évaluation sur lequel est affiché le seuil d'acceptabilité. Les risques qui dépassent ce dernier feront l'objet d'un traitement ;
- traitement : l'entreprise entreprend des actions pour réduire les causes du risque (donc sa probabilité) ou son impact.

Comme nous avons pu le constater à travers cet exemple, l'approche proposée permet de prendre en compte de manière formalisée la question du risque lié au changement. Elle comporte une évaluation quantitative de la probabilité du changement (R1, R2 et R4) et de son impact. Les auteurs n'ont, cependant, pas présenté en détails l'évaluation de ces deux caractéristiques. L'étude n'étant pas dédiée aux projets ERP, ils ne s'intéressent pas aux mêmes axes du changement que nous. Dans les projets ERP, il existe un premier axe de changement dans l'environnement de l'entreprise et donc externe à l'entreprise. Le second axe de changement est celui interne à l'entreprise et qui touche sa structure, ses processus et

¹⁵ Business Process Management

son organisation. Ce second axe de changement structurel est impacté par le premier axe de changement externe. Le troisième axe est l'axe opérationnel de l'entreprise qui est à son tour impacté par les changements sur l'axe structurel interne. Les auteurs travaillent entre l'axe externe (R1, R2 et R4) et l'axe opérationnel interne (R3). C'est-à-dire qu'ils relient directement les changements dans l'environnement externe de l'entreprise à ces performances opérationnelles. Dans les projets ERP, les événements extérieurs à l'entreprise conduisent à la décision de restructurer les processus. Ces événements externes, n'impactant pas directement les performances, ont une importance moindre dans le projet ERP. C'est pour cela que nous ne tenons pas compte de l'axe environnement externe utilisé dans Sienou, et al. (2006) .

Tatsiopoulos, et al. (2003) proposent pour les projets ERP, une approche de gestion des risques. Cette proposition est structurée autour des phases d'Identification, d'Analyse, de Contrôle et de Feedback. Les risques identifiés dans son étude de cas sont orientés projet et sont catégorisés de la manière suivante (Tatsiopoulos, et al., 2003):

- questions industrielles
 - les utilisateurs ne pourront pas répondre à un tel changement,
 - les problèmes logiciels peuvent causer des retards dans l'implémentation ;
- questions concernant l'étendue du projet
 - le périmètre du projet est trop large pour une implémentation dans les délais,
 - la complexité fonctionnelle va empêcher l'implémentation ;
- questions concernant le management du projet
 - l'indisponibilité du staff va affecter le délai de livraison du projet,
 - les délais dus aux tiers vont causer un retard général du projet ;
- questions concernant l'approche,
- questions concernant la formation, le support.

Ces risques sont ensuite hiérarchisés en utilisant l'exposition du projet face à chaque risque qui est le produit entre sa probabilité d'occurrence et son impact. Dans cette hiérarchie, le risque « les utilisateurs ne pourront pas répondre à un tel changement » est le plus important après celui de « l'indisponibilité du staff va affecter la fin du projet ». Cette hiérarchisation est suivie d'une analyse quantitative des risques les plus importants ; ces deux étapes composent la phase d'analyse. Ainsi, une analyse par la simulation de Monte Carlo est présentée pour le risque « l'indisponibilité du staff va affecter la fin du projet ». L'auteur atteste qu'une approche similaire a été proposée pour les autres risques importants, sans pour autant le présenter. Nous considérons que la transposition de cette méthode ne peut pas se faire d'une manière automatique car son application varie selon ce qui est analysé. Par exemple analyser le délai (Giard, 1991 p. 218) et analyser le coût (Vose, 1996 p. 128) en gestion de projet avec la simulation de Monte Carlo sont deux tâches comportant beaucoup de différences. Nous pensons qu'il serait plus intéressant de présenter la simulation de Monte Carlo, ou toute autre méthode quantitative, appliquée au risque lié au changement. En effet, la simulation de Monte Carlo sur les risques liés au délai ou au coût est une problématique traitée dans le domaine de la gestion de projet (Giard, 1991) (Vose, 1996). La quantification du changement et de son impact est énoncée mais n'est pas démontrée.

1.4.4. La clusterisation des processus

L'urbanisme organisationnel a pour objectif de s'appuyer sur une structuration des processus pour faciliter la gestion de l'évolution dans l'entreprise (Chapron, 2006). Elle repose sur une représentation du système d'information qui permet d'identifier des regroupements découplés selon le principe d'un couplage externe faible et d'un couplage interne fort. Le but de la méthodologie est de délimiter des zones de processus cohérents vis à vis de la gestion de l'évolution, c'est à dire des zones que l'on fera évoluer de manière conjointe au sein de l'entreprise dans le but de faciliter la gestion globale des changements. Les blocs constituent des unités de gestion du changement et le diagnostic des blocs permet de mettre en évidence l'alignement d'un changement vis à vis de la stratégie de l'entreprise et du système existant (Chapron, 2006).

Le critère utilisé pour évaluer la notion de « cohérence au regard de l'évolution » est le degré de dépendance entre processus. Les types de dépendances pris en compte dans cette étude sont : dépendance par les ressources, dépendance par les informations et dépendance par les acteurs. Ces dépendances élémentaires sont agrégées pour obtenir la dépendance principale entre deux processus. Une fois le graphe d'interrelation entre les processus défini, il est possible de lui appliquer les méthodes de clusterisation afin de définir des regroupements¹⁶ entre ces processus. Après la clusterisation, l'auteur propose deux niveaux de diagnostic.

Diagnostic des cartes de cluster. Ce niveau de diagnostic est lié aux différentes cartes de dépendances et à la constitution des clusters. Le premier diagnostic s'intéresse à la qualité des cartes générées et aux informations structurelles qu'elles peuvent apporter sur l'évolution du système. Ce diagnostic lié à la qualité des cartes de clusters se base sur la définition de trois indicateurs : le degré de découplage des clusters, le degré de cohésion interne et le nombre de clusters. Le degré de découplage représente la distance (dépendance) moyenne entre les clusters de la carte. Le degré de cohésion interne représente la dépendance moyenne entre les processus d'un même cluster de la carte. Le diagnostic est basé sur la comparaison de cartes de deux systèmes, afin de faire apparaître l'évolution des indicateurs de qualité des cartes de clusters. L'augmentation du degré de découplage entraîne (Chapron, 2006 p. 162):

- des informations de plus en plus découplées entre clusters,
- une orientation vers une organisation plus décentralisée et
- des applications davantage dédiées à chaque cluster.

L'augmentation de la dépendance moyenne entre les processus d'un même cluster implique (Chapron, 2006 p. 163)

- des processus de plus en plus dépendants pour le partage d'informations ;
- des acteurs de plus en plus organisés par cluster,

¹⁶ clusters

- des applications de plus en plus structurées autour des processus d'un même cluster.

Le second diagnostic s'intéresse au calcul de l'effort à fournir pour transformer l'état d'un système vers un autre état. Cet effort dépend des différences existantes entre les deux configurations du système et des couplages qui existent au travers de ces dépendances. L'effort de transformation se basera sur l'identification des écarts entre les deux cartes et à la quantification de ces écarts. L'étendue est un premier indicateur sur la quantité de liens constituants des clusters modifiés entre les deux cartes (Chapron, 2006 p. 167). La quantification des écarts s'intéresse à l'évaluation de l'amplitude de l'évolution. L'amplitude de l'évolution correspond à l'importance des modifications sur les processus touchés. Pour chaque écart de la configuration entre deux processus, l'auteur propose de calculer la différence en valeur absolue entre les dépendances initiales et finales entre ces deux processus. Ces valeurs sont ensuite utilisées pour définir l'indicateur d'amplitude pour l'évaluation des écarts entre deux cartes.

Diagnostic intra-cluster. Cette partie du diagnostic est orientée sur les caractéristiques des clusters. Le premier point étudié dans ce diagnostic est la maturité de ces derniers. La mesure des niveaux de maturité des différents processus d'un même cluster permet d'évaluer les faiblesses de certains processus du cluster qui, fortement liés à d'autres, peuvent représenter un risque (Chapron, 2006 p. 177). La vision globale des maturités des différents clusters d'une carte de processus fournit une information supplémentaire dans le choix de l'évolution à apporter au système d'information. Pour définir l'agrégation de la maturité des processus au sein d'un cluster, l'auteur suggère deux approches alternatives (Chapron, 2006 p. 178) :

- l'utilisation de la moyenne des maturités ainsi que sa variance
- le cluster est évalué à partir de la maturité la plus faible de ses processus

Le second point du diagnostic intra-cluster est relié à l'apport stratégique des clusters. La stratégie correspond à l'ensemble des choix d'objectifs et de moyens qui orientent à moyen et à long terme les activités d'une entreprise. L'évolution de l'ensemble des composantes de l'entreprise doit s'appuyer sur la stratégie choisie par l'entreprise pour orienter ses choix et ses modes de fonctionnements. A chaque processus, l'auteur associe un niveau d'apport à la stratégie. Ce niveau d'apport stratégique représente l'impact de la performance du processus à la stratégie actuelle de l'entreprise. La mise en évidence de l'apport des différents processus d'un groupement permet d'évaluer l'importance de ce cluster dans l'ensemble de la carte des processus. Elle permet aussi d'identifier au sein d'un cluster les dépendances qui, au cours d'une évolution, peuvent impacter des processus définis comme critiques pour la stratégie de l'entreprise. Pour définir l'agrégation de l'apport stratégique au sein d'un cluster, l'auteur suggère les deux mêmes approches alternatives que pour la maturité (Chapron, 2006 p. 180) :

- l'utilisation de la moyenne des apports ainsi que l'écart type
- le cluster est évalué à partir de l'apport le plus élevé au sein de ses processus

Dans cette étude, l'ampleur du changement est mesuré au niveau de l'entreprise avec le degré de découplage et sur l'effort de transformation. Un diagnostic est réalisé en utilisant ces informations. Au niveau du regroupement de processus, le diagnostic est réalisé à l'aide des quantifications

- de l'ampleur du changement avec le degré de couplage interne et la moyenne des maturités ;
- et de l'impact du changement avec la moyenne des apports stratégiques.

Au niveau de l'entreprise et du regroupement de processus, nous pensons que le travail réalisé peut être complété par une mise en relation des mesures de l'impact (apport) avec les mesures de l'ampleur du changement (couplage, maturité). Au niveau du processus, les informations sont collectées sur les dépendances, sur la maturité et sur l'apport stratégique. Aucun diagnostic n'est mené cependant à ce niveau. Dans cette étude, les informations sur les dépendances sont collectées au niveau des processus. Des métriques sont créées ensuite par agrégation aux niveaux des regroupements de processus et des cartes de clusters. Les résultats de cette étude, aux niveaux des entreprises et des regroupements de processus, auront un apport certain en cas d'application en implémentation de progiciel. La seule limite de l'application de cette approche dans les projets ERP est l'absence de diagnostic au niveau du processus.

1.4.5. Raisonnement à base de cas

Raphaeli, et al. (2004), ont proposé d'appliquer l'approche de raisonnement à base de cas (RBC) pour gérer le changement dans les projets d'implémentation de systèmes d'entreprise¹⁷. Ils considèrent le changement organisationnel dans ces types de projets comme des processus complexes caractérisés par des événements et des problèmes imprévus. Le RBC fonctionne sur un ensemble de cas dérivés de l'expérience et enregistrés dans une base de données de type Cause-Evénement-Action-Résultat(CEAR). Quand un nouveau problème est rencontré, le RBC retrouve un cas similaire dans la base CEAR et, si nécessaire, l'adapte pour fournir la solution désirée (Aamodt, et al., 1994). Dans la démarche proposée, les cas contiennent les informations suivantes (Raphaeli, et al., 2004):

- description du problème
 - les causes de l'événement : les causes des événements pouvant se produire durant l'implémentation,
 - données de contexte : fournissent une vue de l'entreprise et du système, telle que la taille de l'entreprise, le type de système d'entreprise, la culture régionale, etc.;

¹⁷ ERP, CRM (Customer Requirement Management) et PDM (Product Data Management)

- profil de l'entreprise : indicateur de risques dans le statut de la compagnie sur six aspects : stratégie&objectif, management, structure, processus, connaissance&savoir-faire et dynamique sociale.
- solution du problème
 - action appropriée pour traiter l'origine du problème.

L'auteur a choisi de s'attaquer directement aux causes des événements et non aux événements eux mêmes, comme c'est le cas d'habitude en RBC. Ce choix est basé sur la supposition qu'un événement peut découler de causes totalement différentes. La recherche de cas similaires commence par la sélection des cas ayant le même contexte que le cas cible. L'étape suivante est une recherche d'équivalence basée sur la similarité entre les causes des événements et la similarité entre les profils des entreprises. Cette méthode augmente l'efficacité de la recherche car un seul sous-ensemble de cas est examiné.

L'auteur s'est donné comme objectif de proposer un support pour le processus de changement organisationnel lié à l'implémentation des systèmes d'entreprise. Cependant, les événements pris en compte dans son modèle ne permettent pas de faire ressortir le changement organisationnel. Sa description des événements et de leurs causes semble plus s'orienter vers les événements qui modifieraient le processus d'implémentation lui-même plutôt que des événements liés au changement organisationnel¹⁸.

Les variables relatives aux données de contexte et au profil de l'entreprise sont définies au niveau entreprise et système d'entreprise qui sont deux niveaux équivalents. Les causes des événements et les solutions n'étant pas clairement identifiées et positionnées, leurs niveaux d'analyse ne peuvent pas être déterminés. Le niveau d'analyse des cas qui est le niveau d'analyse de l'ensemble de l'étude ne peut pas être ainsi identifié. L'unicité du niveau d'analyse qui est une chose très importante dans les méthodes qui collectent des observations (section 1.5.5.1), n'est donc pas respectée.

1.4.6. Méthodes d'explication

Certains auteurs ont proposé des méthodes de gestion du changement basées sur la prédiction, le classement ou l'analyse de corrélation. Ces sont des techniques d'explication qui consistent à prédire les valeurs d'un attribut endogène (dépendante, de réponse, à expliquer) à partir d'autres attributs exogènes (indépendante, prédicteur, variable explicatives) (Han, et al., 2001 p. 24) (Rakotomalala, 200B). Nous commencerons par présenter la régression linéaire multiple qui est l'une des techniques de prédiction souvent utilisées par ces auteurs.

¹⁸ *An ES implementation (ESI) process, as other system development processes, is a complex and dynamic process that cannot be fixed from the start. The process is characterized by the occurrences of unplanned problems and events. These situations may lead to major restructuring of the process with severe implications to the whole company*

1.4.6.1. Régression linéaire multiple

La régression linéaire multiple permet de représenter la relation de dépendance entre une variable cible continu¹⁹ et des prédicteurs continus (Han, et al., 2001 p. 320). Cette relation est une connaissance extraite d'une base d'exemples composée de n observations supposées être indépendantes. Le modèle linéaire ou équation linéaire peut s'écrire sous forme matricielle comme suit (Draper, et al., 1998 p. 7) (Cohen, et al., 2003 p. 34):

Équation 1 Régression linéaire multiple

$$Y = X\beta + \varepsilon$$

Y : le vecteur colonne de dimension n contenant les observations de la variable cible, n représentant la taille de la base d'exemples.

X : la matrice de taille n*(p+1) composée d'une colonne de 1 suivies de p colonnes des vecteurs observations des variables explicatives. p représente le nombre des variables explicatives

β : est le vecteur de dimension (p+1) des paramètres à estimer. $\beta_0 = m$ est l'ordonnée à l'origine qui représente la valeur de la variable cible Y quand les variables explicatives sont nulles. Les β_j sont les coefficients de régression pour estimer Y à partir de X. Chaque β_j représente le taux de changement en unité de Y par unité de changement dans la variable explicative X j.

ε : vecteur de dimension n des erreurs (erreur standard).

La méthode la plus utilisée pour résoudre cette équation est la méthode des moindres carrés. Le principe des moindres carrés consiste à rechercher les valeurs des paramètres qui minimisent la somme des carrés des erreurs, à savoir :

Équation 2 Somme des carrés des erreurs

$$S = \sum_{i=1}^n \varepsilon_i^2$$

En général, il n'y a pas de solution pour le système $Y = X\beta$ car il est surdéterminé, il comporte plus d'équations que d'inconnues (Lea07). Un système d'équations équivalent appelé « les équations normales » peut être utilisé pour trouver la solution (Cohen, et al., 2003 p. 498). Si $X'X$ est inversible alors ces équations normales admettent une solution (Draper, et al., 1998) (Cohen, et al., 2003 p. 631) et :

¹⁹ On parle de prédiction (Han, et al., 2001 p. 24) (Rakotomalala, 200B)

Équation 3 Estimation des paramètres

$$\beta = (X'X)^{-1} (X'Y)$$

Lorsque le nombre des prédicteurs est limité à 1, nous avons une régression linéaire simple (Équation 4) :

Équation 4 Régression linéaire simple

$$Y = X\beta + \varepsilon = \beta_0 + \beta_1 X + \varepsilon$$

L'estimation de la régression linéaire est représentée par une droite linéaire (\hat{Y}). Pour chaque point de données du diagramme, une erreur (ε) est associée à la distance entre le point (Y) et la droite de régression (\hat{Y}).

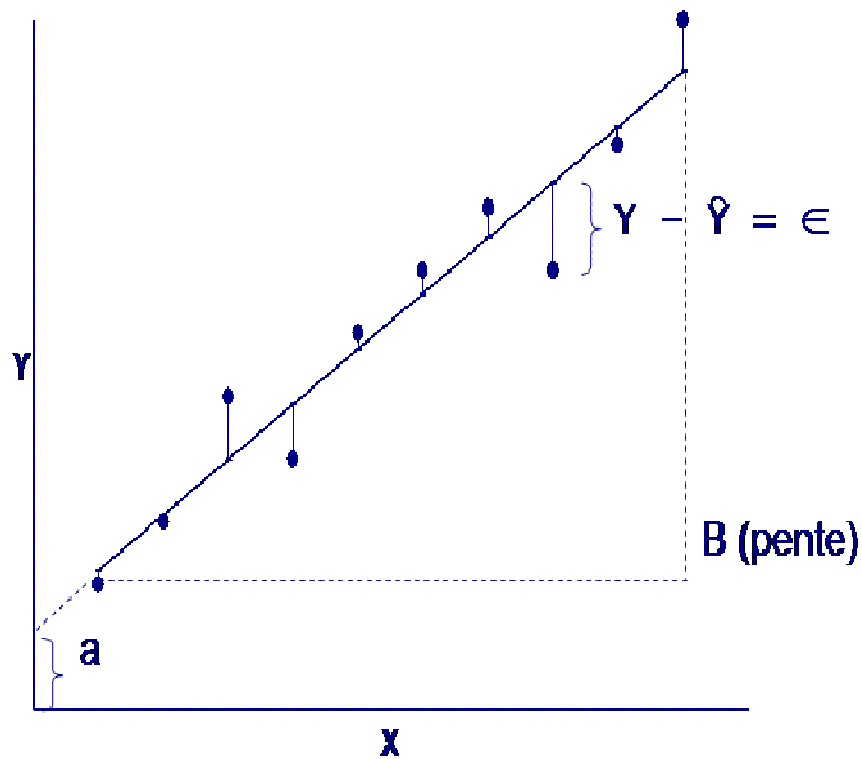


Figure 5 Régression linéaire simple

1.4.6.2. Analyse de corrélation

Maaloul (2003) s'est intéressé au changement en travaillant avec des corrélations de Pearson. Le coefficient de corrélation de Pearson est une mesure de la qualité d'une régression linéaire. Maaloul (2003) propose un modèle conceptuel de mesure de la satisfaction des utilisateurs d'un système ERP. L'auteur commence par définir les variables suivantes :

- la qualité du système implanté,
- la qualité de l'information qu'il fournit,
- l'utilité perçue par les utilisateurs,
- l'ingénierie du changement qui comprend les variables suivantes :
 - l'implication de la direction générale,
 - l'implication des utilisateurs,
 - la communication,
 - la formation,
 - la stratégie d'implémentation,
- la satisfaction des utilisateurs : l'objectif du modèle est de déterminer l'influence des variables définies ci-dessus sur cette variable.

Des coefficients de corrélation de Pearson sont calculés entre chacune de ces variables et celle de « satisfaction des utilisateurs ». Ces coefficients représentent pour chaque variable le pourcentage que la variation de celle-ci explique dans la variation de la variable « satisfaction des utilisateurs ». L'auteur réalise une mesure de l'effort de gestion du changement à travers les variables d'ingénierie du changement. L'impact du changement est évalué par rapport à la satisfaction des utilisateurs. La satisfaction des utilisateurs se situant selon nous au niveau abstrait, est difficilement projetable au niveau quantitatif. L'ampleur du changement n'est pas quantifiée. L'analyse n'est pas menée au niveau processus mais sur le projet.

1.4.6.3. Prédiction par régression linéaire

Wischnevsky, et al. (2001), proposent une recherche empirique afin d'explorer les effets de la transformation organisationnelle sur les performances des entreprises de manière systémique. Selon eux, les études menées jusque-là n'ont examiné que les effets de changements stratégiques isolés sur les résultats des organisations (Rajagopalan, et al., 1997). Wischnevsky, et al. (2001), ont examiné cinquante banques américaines entre 1975 et 1995. Pendant cette période, ces banques ont entrepris des changements stratégiques et structurels majeurs, devenant le moteur principal de l'expansion spectaculaire des marchés de banques commerciales et des services fournis. Les sources de données de cette étude comprennent des rapports annuels aux actionnaires, des données historisées de statistiques bancaires. La définition des variables est fondée sur la supposition selon laquelle, la transformation organisationnelle est un processus révolutionnaire qui exige des changements rapides, discontinus et concomitants dans la stratégie, la structure, la distribution du pouvoir et le système de contrôle (Romanelli, et al., 1994) :

Les variables dépendantes. Les variables à expliquer concernent la performance de l'organisation qui est représentée par les ratios de profitabilité et sont mesurées par le retour sur les actifs (RSA) et le retour sur les capitaux (RSC).

Les variables indépendantes. Les variables explicatives sont relatives à la transformation organisationnelle et ses conditions d'interaction avec la modification de l'environnement.

- la variable transformation Organisationnelle est une variable discrète égale à 1 quand la banque affiche des changements substantiels dans la stratégie, la structure et la force de distribution dans une période de 2 ans et 0 sinon (Wischnevsky, et al., 2001) (Romanelli, et al., 1994).
- la variable de changement dans la stratégie, définie comme une liste des suppressions ou des modifications majeures dans l'importance donnée au marché d'un produit ou d'un service, incluant les changements dans : (1) la diversification géographique ; (2) les activités non-bancaires ; et (3) les caractéristiques des prêts (commercial-Client et étranger).
- les variables de changement de la structure se concentrent sur la centralisation-décentralisation, incluant : (1) la centralisation à l'échelle nationale ; (2) la centralisation à l'échelle de l'état ; et (3) le changement dans la forme organisationnelle.
- la variable de changement dans le pouvoir est mesurée par : (1) la rotation des cadres (2), la concentration du pouvoir, le degré selon lequel les cadres sont concentrés ou distribués ; et (3) le pouvoir à la tête de la banque²⁰.

Parce qu'il n'existe pas de critère objectif pour déterminer quand un changement dans un intervalle de stratégie, de structure ou de pouvoir constitue un changement substantiel, l'auteur a appliqué de façon séparée quatre niveaux de critère à chaque variable, à partir des valeurs de changement relativement large de 30, 40, 50 et 60 pour cent (Romanelli, et al., 1994). Quand la valeur de la variable excède le niveau du critère défini, le changement est considéré comme substantiel et est codé comme 1, sinon, il est codé comme 0.

- La variable de modification de l'environnement est mesurée comme une variable discrète, basée sur l'occurrence de changements législatifs majeurs qui impactent la capacité des banques de se développer géographiquement.

Pour chaque variable dépendante (RSA et RSC) et pour chaque niveau de critère utilisé pour déterminer les occurrences de transformations organisationnelles (30, 40, 50 et 60 pour-cent), quatre modèles sont estimés²¹. Les modèles 1 et 3 testent l'effet de la transformation organisationnelle sur la performance, respectivement, un an et deux ans après la transformation. Les modèles 2 et 4 testent les effets d'un terme d'interaction sur la performance pour justifier la transformation organisationnelle en réponse à une modification de l'environnement. La méthode de régression utilisée est la méthode des moindres carrés.

²⁰ Le degré selon lequel le pouvoir à la tête de la banque égale le pouvoir dans la banque

²¹ Ce qui fait un total de 32 régressions

Dans Wischnevsky, et al. (2001), seuls les résultats du modèle 1 pour RSA et RSC au critère de niveau 30% affichent une significativité statistique à 0.5. Le manque de coefficient significatif dans les 14 autres des 16 modèles 1 et 3 indique que la transformation organisationnelle n'a pas d'effets bénéfiques ou délétères sur la performance. De plus, aucun des termes d'interaction (transformation organisationnelle*modification de l'environnement) des 16 modèles 2 et 4 n'était significatif à 0.5. Cela ne donne donc aucune base pour conclure qu'une transformation organisationnelle entreprise en réponse à une modification de l'environnement est soit bénéfique ou nuisible pour la performance des entreprises (Wischnevsky, et al., 2001).

Dans cette étude, la quantification du changement et celle de son impact est réalisée. Toutes les variables sont définies au niveau de l'entreprise, ce qui assure une définition claire de l'unité d'analyse. La régression est utilisée pour créer des modèles mathématiques et réaliser des prédictions sur l'impact du changement sur la performance. Cependant, les prédictions obtenues permettent seulement de comprendre la relation d'influence entre le changement organisationnel et la performance. Des actions proactives ne sont pas proposées à partir des résultats obtenus.

1.4.6.4. Méthodes de classement

Le classement fait référence à une méthode d'explication dont la variable à expliquer est discrète (Han, et al., 2001 p. 24) (Rakotomalala, 200B). Parr, et al. (2000), proposent un classement de l'implémentation d'ERP avec une taxonomie de trois types : « vanille », « compréhensive » et « mi-chemin ». Les projets « vanille » (6 à 12 mois et 1 à 2 M \$ US) sont les moins ambitieux, ont un faible risque dans l'implémentation et affectent un petit nombre d'utilisateurs potentiels. Ils sont concentrés sur l'implémentation des fonctionnalités clés de l'ERP avec un BPR minimal. A l'opposé, les projets « compréhensive » (> 4 ans et > 10 M \$ US) sont des implémentations ambitieuses qui affectent un nombre large d'utilisateurs dans de multiples sites. Ils entraînent un effort majeur de BPR où l'ensemble complet des fonctionnalités de l'ERP sont installées. Enfin, les implémentations « mi-chemin » (>12 mois et > 3 M \$ US) affectent un nombre relativement large d'utilisateurs dans des sites multiples, mais n'implémentent que les fonctionnalités majeures du cœur de l'ERP avec un peu de BPR.

La détermination du type d'implémentation résulte de la combinaison de caractéristiques de l'implémentation appartenant à cinq catégories : allocation de ressources, étendue technique, étendue physique, étendue BPR et stratégie d'implémentation des modules²² (Tableau 5). Dans cette étude, les variables explicatives mesurent l'étendue du changement au niveau de l'entreprise ou du projet.

La prédiction de l'impact du changement est réalisée par la mise en relation des caractéristiques du projet des catégories explicatives (périmètre BPR, périmètre physique,

²² Quels sont les modules à implémenter et comment seront-ils intégrés au système existant.

périmètre technique et stratégie d'implémentation des modules) avec les caractéristiques de la catégorie ressource à l'aide d'un tableau. Le changement peut ainsi être géré de façon proactive dans la mesure où il est possible de fixer les caractéristiques des catégories explicatives pour impacter la durée du projet et son budget. L'impact du changement sur l'entreprise n'est pas pris en compte. Cela empêche d'analyser le changement dans les projets ERP par rapport à son objectif principal qui est l'amélioration de la performance.

Barki, et al. (2004), se sont aussi intéressés à l'identification des caractéristiques des projets ERP en procédant de la manière suivante : identifier d'abord les concepts théoriques essentiels et identifier les caractéristiques importantes qui les reflètent pour développer une taxonomie et une mesure de ces concepts. Ils ont travaillé principalement avec le concept de périmètre du projet en adaptant certaines variables de Parr, et al. (2000) liées à ce concept (Tableau 5). Ils ont aussi repris les variables de catégorie « ressources du projet » de Parr, et al. (2000) en y ajoutant la variable « effort projet » en les considérant cette fois comme variables de réponse. En effet, ils ont supposé que les caractéristiques de catégorie « ressources projet », ne pouvant pas être exactement connues avant la fin du projet, sont des conséquences du processus d'implémentation et à ce titre, elles doivent être séparées des caractéristiques antécédentes. L'étude est réalisée à partir de données contenant les caractéristiques pour 31 projets d'implémentation d'ERP. Une analyse factorielle est d'abord menée pour déterminer les corrélations entre les variables explicatives. Les variables « étendue ERP » et « étendue BPR » étant très corrélées, donnent naissance à un premier facteur appelé « étendue ». L'analyse a montré que les variables « Augmentation de l'automatisation des processus métiers », « magnitude du BPR » sont corrélées positivement entre elles. Elles sont corrélées négativement avec la variable « Customisation » car une customisation²³ étendue de l'ERP correspond à un BPR faible. A l'inverse, une faible customisation voudra dire un BPR significatif pour faire correspondre l'organisation à l'ERP. Un second facteur appelé « Magnitude » est construit sur la base de la corrélation entre ces trois variables. Enfin, un troisième facteur, nommé « profondeur », est construit à partir de la corrélation entre les variables « profondeur ERP » et « profondeur BPR ».

Des valeurs discrètes B (bas) ou E (élevé) sont affectées aux projets de l'échantillon pour chacune des trois dimensions du périmètre selon qu'elle est située au dessus ou en dessous de la moyenne de l'échantillon. Chacune des huit catégories ainsi créées est représentée par au moins un projet dans l'échantillon (Tableau 6 colonne N). Des tests de régression sont ensuite menés pour explorer les relations entre les dimensions du périmètre de l'implémentation d'ERP et les sorties du projet. Les résultats indiquent que la « profondeur » est positivement associée aux trois sorties du projet (durée, effort et budget). Il a été aussi montré que « l'étendue » du projet affecte positivement le « budget » du projet. Ensemble,

²³ Paramétrage

l'étendue, la profondeur et la magnitude expliquent 43%, 36% et 49% de la variance dans la « durée » du projet, « l'effort » du projet et le « budget » du projet, respectivement.

Tableau 5 Variables projet ERP (Barki, et al., 2004) et (Parr, et al., 2000)

[Parr, et al., 2000]		[Barki, et al., 2004]	
Variable	Mesure	Variable	Mesure
Allocation de ressource	Temps	Durée du projet	nombre de mois
		Effort projet	nombre d'homme par mois (000)
Allocation de ressource	Budget	Budget du projet	Millions de \$ US
Etendue technique	1= pas de modification de l'ERP ; 2= modifications mineures ; 3=modifications majeures	Customisation	Ampleur de la modification à l'ERP pour le customiser (1-10)
Etendue physique	1= un seul site ; 2= sites multiples, régional ; 3= sites multiples, international	Etendue ERP	1= un seul site ; 2= sites multiples, un état ; 3= sites multiples, plusieurs états, 4= sites multiples, international
Etendue physique-nombre d'utilisateurs	1= petit, <100 ; 2= moyen, <200 ; 3=large, >200)	Profondeur ERP	Nombre d'utilisateurs
		Augmentation de l'automatisation processus Métiers	(% des processus automatisés après ERP) – (% des processus automatisés avant ERP)
		Magnitude BPR	(% d'activité dans les processus restructurés qui ont été modifiées) * (ampleur de la modification 1-10)
Etendue BPR	1=alignement à l'ERP; BPR Global ; 3 BPR local	Etendue B PR	1= un petit nombre de personnes dans un département, 2= un département, 3=plus d'un département, 4 une région, 5=plus qu'une région
		Profondeur BPR	Nombre d'employés dont les activités ont changé

Tableau 6 Relation entre le périmètre et les sorties d'un projet (Barki, et al., 2004)

Périmètre du projet			N	Moyennes sorties du projet		
Etendue	Profondeur	Magnitude		Durée (mois)	Efforts (hommes mois)	Budget (M \$ US)
B	B	B	5	12.4	939	3.58
B	B	E	5	15.5	1166	23.74
B	E	B	2			
B	E	E	1			
E	B	B	8			
E	B	E	4			
E	E	B	1			

E	E	E	4	38.5	6920	
---	---	---	---	------	------	--

L'approche utilise des variables liées au périmètre du projet qui mesurent le changement au niveau du projet (Tableau 6). L'impact du changement est mesuré sur la durée, le budget et l'effort du projet (Tableau 6). Nous considérons que ces caractéristiques du projet sont liées à des objectifs secondaires par rapport à l'objectif d'amélioration des performances de l'entreprise. Les variables explicatives sont discrétisées afin de créer un modèle de classement sous forme de tableau. La proactivité est gérée de la même manière que dans Parr, et al. (2000), par la modification des paramètres liés au périmètre du projet afin d'impacter les sorties du celui-ci (durée, budget et effort).

Dans Anderson (2001), l'objectif est de prédire l'impact de changements potentiels dans les variables de structure organisationnelle, de senior leadership et de vision stratégique sur la variable prise de décision. L'approche proposée utilise les réseaux bayésiens pour prédire l'impact de ces changements potentiels. Les variables utilisées sont les suivantes :

- structure de l'organisation
 - structure : la structure de l'organisation est appropriée,
- sénior leadership
 - intégrité : les cadres supérieurs présentent les plus hautes normes d'intégrité,
 - confiance : les employés font confiance aux membres de la direction,
 - inspiration : les dirigeants sont source d'inspiration,
- vision stratégique :
 - connaissance : le comité de direction comprend les menaces de la concurrence et les opportunités,
 - vision réaliste : la vision stratégique est réaliste et atteignable,
 - vision convaincante : la vision stratégique est convaincante,
- prise de décision :
 - focalisation : le comité de direction concentre les efforts et les ressources sur les questions critiques,
 - faits : les décisions sont fondées sur des faits, pas sur des politiques d'entreprise,
 - problèmes : les problèmes sont définis et résolus rapidement,
 - apprentissage : les processus de prise de décision favorisent l'apprentissage organisationnel.

Nous pouvons dire que cette proposition intègre la gestion du changement et la mesure de l'impact de celui-ci. La quantification du changement est satisfaite avec l'utilisation d'un certain nombre de variables. L'impact est mesuré sur les variables de prise de décision qui

vont impacter la performance. L'impact n'est donc pas directement mesuré sur la performance de l'entreprise. Nous ne considérons pas cela comme une limite car l'auteur ne se positionne pas dans les projets ERP. La prédiction est formalisée à travers les réseaux bayésiens. L'unité d'analyse est fixée au niveau des sites ou de l'entreprise.

1.4.7. Les approches quantitatives très utilisées

Après avoir analysé des travaux isolés, nous présentons brièvement le rapport à la gestion du changement de certaines approches quantitatives très utilisées qui peuvent intervenir dans les projets ERP. Nous parlerons essentiellement de l'analyse structurelle des processus, de la simulation et de la supervision des processus. L'analyse structurelle des processus définit deux types de métriques : les métriques opérationnelles des processus et les métriques structurelles des processus (Aguilar, et al., 2006A) (Tjaden, 2001). Les métriques structurelles sont relatives aux propriétés statiques du processus (Aguilar, et al., 2006A) (Tjaden, 2001). Les métriques opérationnelles sont celles qui mesurent les performances du processus à travers des occurrences d'exécution de ce dernier (Tjaden, 2001). L'Analyse structurelle de processus peut être utilisée dans la gestion du changement dans la mesure où elle permet de comparer les processus « en-cours » et ceux « en devenir » selon plusieurs critères liés à leur structure, comme la complexité, la simplicité, la flexibilité ou l'intégration.

Les auteurs de l'analyse structurelle des processus se sont plus concentrés sur les métriques structurelles. Seul Tjaden (2001) définit sommairement des métriques opérationnelles et émet l'hypothèse de l'influence des métriques structurelles sur celles opérationnelles. Le modèle qui relie ces deux catégories de métriques n'est pas défini. Cependant, les métriques telles qu'elles sont définies en l'analyse structurelle ne permettent pas de mesurer directement le changement et son impact sans faire de comparaison ou de transformation. C'est-à-dire que les métriques sont définies à partir d'un seul modèle de processus alors que la quantification directe du changement nécessite de travailler avec deux modèles au moins²⁴.

L'approche de simulation des processus montre le déroulement de l'exécution d'un processus dans un environnement de simulation avec des ressources définies. Une simulation est basée sur la logique contenue dans le modèle du processus et peut utiliser des données d'exécution réelles (Bea, 200X). Elle permet une analyse dynamique des performances du processus « en-cours » et de celui « en-devenir » (Greasley, 2000). La simulation s'intéresse au niveau opérationnel du processus. Les résultats de la simulation dépendent des probabilités sur les différents chemins du processus.

La simulation intervient dans la gestion du changement car elle permet de comprendre, de mesurer et d'évaluer l'impact de l'introduction d'un nouveau processus par modification

²⁴ Nous faisons référence à la démarche de calcul des métriques de changement utilisée par (Eriksson, 2001)

d'un processus existant ou par une réallocation des ressources (Jin, et al., 2001). Il y a une prédiction de la performance des processus « en devenir » qui est reliée au changement sur leurs structures. La quantification du changement et de son impact par la simulation n'est pas complète pour les raisons suivantes :

- les dimensions du changement prises en compte sont des éléments liés au flux du processus, aux durées des activités, à la disponibilité des ressources etc. Les dimensions du changement, tels que les rôles, les compétences requises, les responsabilités des employés ne sont pas analysées ;
- aucune métrique n'est définie au niveau des structures des processus. La quantification du changement sur les structures n'est pas directe, car les modifications sur les structures ne sont visibles que dans le calcul des mesures opérationnelles ;
- le changement et son impact sur la performance sont étudiés de manière séparée sur le processus « en-cours » et celui « en-devenir ». Les structures des deux processus ne sont pas comparées, ce qui fait que l'ampleur du changement n'est pas visible directement.

Le pilotage de la performance opérationnelle (Business Activity Monitoring) est une approche qui permet d'accéder en temps-réel aux indicateurs de performance critiques des opérations métiers (Gartner Group, 2002). Le BAM intervient au niveau opérationnel des processus « en-cours » ou « implémentés ». Cette approche peut permettre, par comparaison des mesures de performance des processus « en-cours » et « implémentés », de vérifier les résultats des changements opérés sur les processus. Elle ne s'inscrit pas dans une gestion proactive du changement dans les projets ERP, car ne permettant qu'une mesure à posteriori de l'impact du changement. Aussi, cette approche ne prend pas en compte le niveau structurel.

Les métriques structurelles et opérationnelles définies dans les approches que nous venons de présenter ne sont pas destinées à l'analyse du changement ou de son impact. C'est pour cette raison qu'elles ne sont mesurées que sur un seul modèle de processus. Une seule valeur de ces métriques ne suffit pas pour constater l'évolution des propriétés mesurées sur le processus. La simulation et le pilotage opérationnel des processus définissent les mêmes métriques calculés à priori et à posteriori respectivement. Ces métriques dans leur état actuel ne permettent pas de mesurer l'impact du changement. Dans cette étude, nous émettons l'hypothèse selon laquelle, l'impact du changement doit être mesuré en termes de variation sur la performance des processus. Les métriques opérationnelles doivent être étudiées pour analyser leur adaptabilité pour la mesure de l'impact du changement. Pour mesurer l'ampleur du changement, nous pensons que les modèles de deux versions du processus doivent être utilisés.

1.4.8. Conclusion

Les approches proposées par Anderson (2001) et Wischnevsky, et al. (2001), ne sont pas dédiées à la restructuration des processus ou de l'entreprise. Elles ne définissent pas leurs unités d'analyse au niveau processus. Nous pouvons néanmoins nous inspirer de ces approches qui sont des démarches de prédiction complètes comprenant la définition des variables, la construction des modèles prédictifs et leur utilisation. Dans Eriksson (2001), les métriques définies sur la structure ne permettent pas de prendre en compte tout le changement opéré sur le processus restructuré. Cependant, l'idée de quantifier le changement par comparaison des structures des deux versions du processus est une bonne solution pour mesurer la profondeur du changement sur le processus.

Les approches de Parr, et al. (2000) et Barki, et al. (2004) sont quant à elles destinées à la restructuration des processus métiers dans le cadre de projets ERP. Elles comprennent aussi la démarche de prédiction complète. L'ampleur du changement est mesuré au niveau de l'entreprise et du projet mais son impact n'est pas mesuré par rapport à l'amélioration de la performance. Parr, et al. (2000), Barki, et al. (2004) et Anderson (2001) proposent des actions proactives exclusivement dirigés vers la réduction des causes. La réduction de l'impact qu'ils considèrent passe par la réduction des causes. Les actions directes sur les conséquences du changement à travers l'effort de gestion ne sont pas proposées. Dans Esteves (2002B), la démarche est aussi dédiée aux projets ERP. Elle se limite à la définition de métriques qui mesurent le changement sur plusieurs niveaux.

Chapron (2006) propose des mesures de l'impact (apport) avec les mesures de l'ampleur du changement (couplage, maturité). La relation entre ces deux catégories de mesures n'est cependant pas analysée. Les diagnostics de l'auteur sont plus orientés vers les regroupements de processus ou l'entreprise dans son ensemble. Cela peut constituer une limite si la méthode est appliquée pour la gestion du changement dans les projets ERP.

La simulation peut être utilisée pour comparer de manière prédictive les performances de deux versions d'un même processus suite à sa restructuration. Ces prédictions dépendent des flux et de l'allocation des ressources dans les deux structures. La mesure du changement n'est pas effectuée directement par comparaison des deux modèles comme dans Eriksson (2001). L'approche de pilotage opérationnel des processus (BAM) peut aider à comparer à posteriori les performances des processus restructurés. Le changement sur les structures est indirectement pris en compte dans cette mesure de la performance. Nous allons étudier l'adaptabilité des métriques définies en simulation et en pilotage opérationnel des processus pour mesurer l'impact du changement. L'analyse structurelle des processus définit des métriques structurelles sur les modèles de processus. Ces métriques ne sont pas utilisables dans leur état actuel pour mesurer le changement mais sont un bon point de départ pour cela.

Chapitre 1.5 Gestion proactive du changement

1.5.1. Introduction

Dans ce chapitre, nous expliquerons d'abord la nécessité d'intégrer la gestion quantitative et non-qualitative pour obtenir une gestion proactive du changement. Nous montrerons ensuite l'intérêt de l'utilisation du datamining pour notre étude, par rapport à son application dans d'autres domaines. Nous allons présenter le processus de datamining. Nous terminerons par la présentation des grandes lignes de l'application du datamining pour la gestion du changement dans les projets ERP.

1.5.2. Intégration des approches quantitatives et qualitatives

Après avoir étudié la gestion quantitative et non-quantitative du changement dans les projets ERP, nous pensons que la limite principale de chacune d'elles est de ne pas assez utiliser l'autre. En effet, la gestion non-quantitative se concentre sur la réduction de l'impact du changement. Elle aurait besoin des méthodes quantitatives pour l'aider à cibler le déploiement de ses moyens de gestion du changement. Ce ciblage passerait par une identification des éléments de l'entreprise les plus affectés par la restructuration à l'aide de modèles prédictifs. La gestion quantitative du changement définit des mesures du changement et de son impact et développe parfois des modèles prédictifs avec celles-ci. La plupart des méthodes quantitatives n'abordent pas la question des actions de gestion du changement et se limitent à une modélisation. La difficulté de proposer des actions de gestion du changement est liée selon nous à une mauvaise définition de l'unité d'analyse ou des variables.

Nous proposons de combiner la gestion quantitative et non-quantitative dans un processus de datamining pour lever ces limites. Le datamining est, en essence, un ensemble de techniques qui permettent d'accéder aux informations appropriées, précises et utiles qui sont cachées dans les bases de données (Adriaans, et al., 1996 p. 127). Dans les projets ERP, la gestion du changement se fait plutôt en aval avec l'implication, la formation et l'accompagnement des utilisateurs finaux par rapport à la nouvelle configuration des processus ou de leur travail. Si ces efforts de gestion du changement ne sont pas ciblés, ils représenteront un poids pour l'organisation du projet et son coût. Cela entraînerait aussi un manque d'efficacité dans la gestion du changement.

La première raison pour laquelle nous nous sommes intéressés au datamining est qu'elle propose des solutions à ce besoin de ciblage dans un certain nombre de domaines. Dans le domaine de la gestion de la relation client, par exemple, un des gains attendus du

datamining est l'identification des clients les plus rentables, et la concentration sur eux des efforts des commerciaux (Tufféry, 2007 p. 9). Le datamining s'appuie sur des bases de données, ou de plus en plus sur des entrepôts de données qui permettent de stocker le profil de chaque client. Dans le domaine de la prévention de la non-qualité en génie logiciel, le datamining est utilisé pour l'identification des modules susceptibles d'être non fiables avant la phase de test. Cela permet de concentrer les efforts de test sur ces modules pour assurer un gain de temps et de productivité (Gokhale, et al., 1997). En effet, une grande partie du travail réalisé dans le développement d'un logiciel concerne la phase de test (Gokhale, et al., 1997). Selon lui, cela est dû au fait que l'on consacre un temps de test et de vérification égale à tous les modules alors que 80% des erreurs se trouvent dans 20% du code (Gokhale, et al., 1997). Des techniques statistiques sont ainsi mises en œuvre pour construire des modèles prédictifs permettant d'évaluer au mieux une caractéristique qualité donnée à partir de mesures descriptives de l'objet à évaluer. L'objectif est d'assurer la qualité des logiciels par la définition de mesures pertinentes et l'analyse des retours d'expérience.

La seconde raison de notre intérêt pour le datamining est notre volonté de tirer parti des informations historisées disponibles autour de notre problématique de recherche. Selon Adriaans, et al. (1996 p. 2), la plupart des organisations produisent en une semaine plus de données que ne peuvent lire beaucoup de personnes durant toute leur vie. Elles sont confrontées ainsi au paradoxe de l'augmentation des données car plus de données signifie moins d'information. Les organisations qui filtrent, sélectionnent et interprètent le mieux les données vont avoir plus de chances pour survivre ; et à cause de cela, l'information elle-même est devenue un facteur de production (Adriaans, et al., 1996 p. 2). Nous considérons que, sur plusieurs projets ERP, les outils de modélisation, d'analyse, d'exécution et de supervision des processus métiers produisent une quantité importante de données qu'il faut exploiter. Nous pensons que la gestion du changement gagnerait à explorer ces données.

Selon Tufféry (2007 p. 1), le datamining permet de limiter la subjectivité humaine dans le processus de décision, et aussi, grâce à la puissance grandissante des outils informatiques, de traiter de plus en plus rapidement de grands nombres de dossiers. Cette qualité est la troisième raison qui nous a poussés à adopter le datamining. En effet, la littérature de la gestion du changement dans les projets ERP propose un certain nombre d'actions pour accompagner le changement. Ces actions sont basées sur des hypothèses émises sur les relations entre les dimensions de l'ampleur, de l'effort de gestion et de l'impact du changement. Ces hypothèses sont construites de manière empirique par l'expérience de chercheurs ou de praticiens mais leur validité réelle n'est pas prouvée. Nous constatons là une certaine forme de subjectivité dans les décisions prises, pour les actions de gestion du changement, sur la base de ces hypothèses. L'idée de cette étude est de rassembler ces hypothèses et de les soumettre à la validation du datamining. Le rejet ou l'acceptation de chaque hypothèse repose sur sa vérification quantitative par les données historisées.

Aussi, selon Boucher, et al. (2006), l'utilisation de l'aide à la décision pour la gestion du changement est fortement recommandée par les sciences de l'ingénieur. C'est une rationalisation du processus de décision, lié à la gestion du changement, qui repose sur une définition de mesures quantitatives et objectives (indicateurs), une phase de prise de décision plus ou moins structurée, et une phase de déploiement utilisant les outils de gestion de la performance (Boucher, et al., 2006). Le datamining permet de réaliser toutes ces actions dans un processus formalisé. Nous présenterons une description du processus de datamining dans la section suivante. Nous présenterons ensuite les grandes lignes de l'application de ce processus pour la proactivité dans la gestion du changement.

1.5.3. Processus de datamining

Le datamining, ou fouille de données, est l'ensemble des méthodes et techniques destinées à l'exploration et l'analyse de (souvent grandes) bases de données informatiques, de façon automatique ou semi-automatique, en vue de détecter dans ces données des règles, des associations, des tendances inconnues ou cachées, des structures particulières restituant l'essentiel de l'information utile tout en réduisant la quantité de données (Tufféry, 2007 p. 4). Le spectre d'application du datamining et de la statistique est très large (Tufféry, 2007 p. 1):

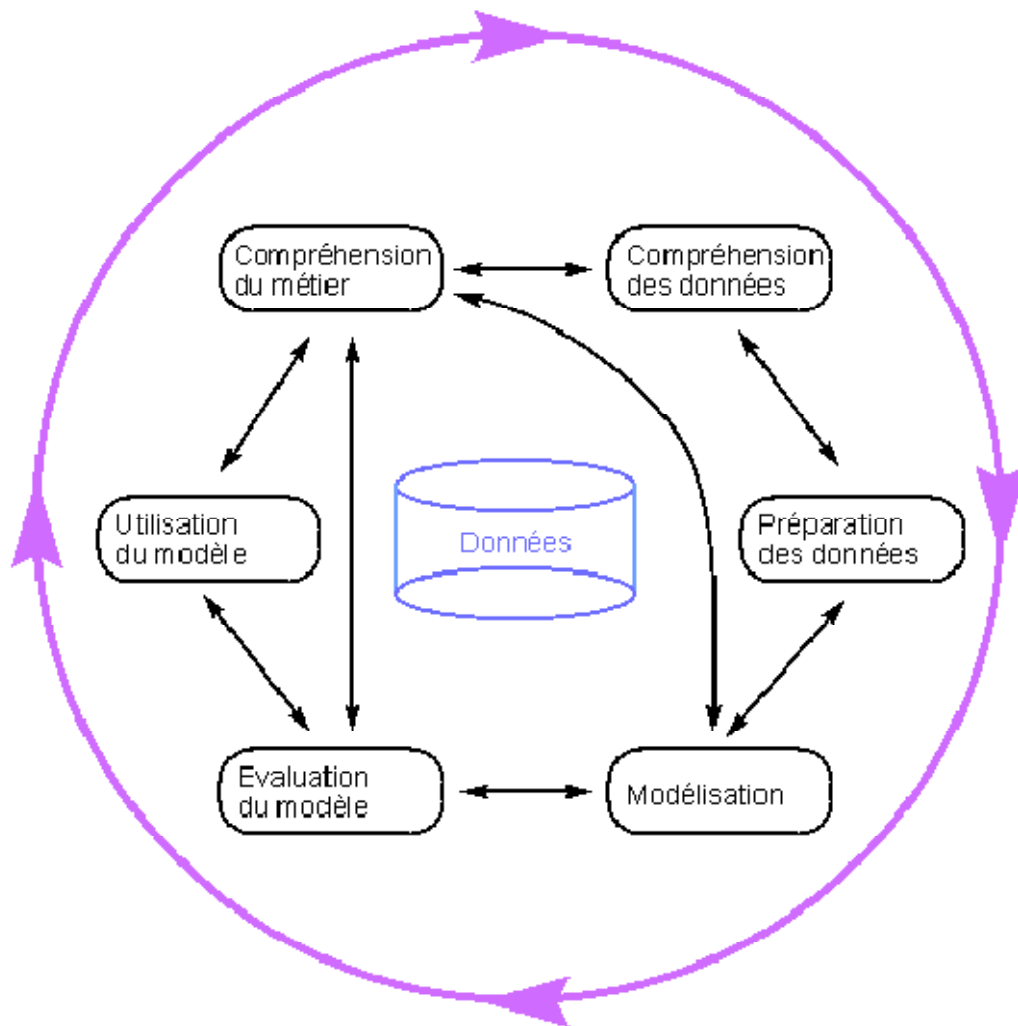
- génomique, Astrophysique,
- gestion de la relation client,
- aide au pilotage aéronautique,
- e-commerce,
- prévention du terrorisme, détection automatique de la fraude dans la téléphonie mobile ou l'utilisation des cartes bancaires,
- contrôle qualité, pilotage de la production,
- enquêtes en sciences humaines, études biologiques, médicales et pharmaceutiques ;
- études agronomiques et agro-alimentaires,
- prédiction d'audience TV.

Le CRISP-DM ²⁵ est considéré comme la méthodologie la plus utilisée pour le datamining (Adderley, et al., 2006 p. 22). Il a été conçu par un consortium d'entreprises pour être utilisé avec n'importe quel outil de datamining dans n'importe quel domaine (Adderley, et al., 2006 p. 22). Le CRISP-DM identifie six phases dans le processus de datamining (

Figure 6) (Milutinovic, et al., 2002 p. 76).

²⁵ Cross Industry Standard Process for Data Mining

Figure 6 CRISP-DM



Compréhension du métier. Cette première phase doit permettre de comprendre les objectifs et les besoins d'un point de vue métier, et ainsi convertir cette connaissance en une définition de problème de datamining, et un plan permettant d'atteindre ces objectifs.

Compréhension des données. La phase de compréhension des données commence avec une collecte des données et se poursuit avec des activités qui ont pour objectif de se familiariser avec les données, d'identifier les problèmes de qualité des données, découvrir les premières connaissances dans les données, ou détecter les sous-ensembles intéressants pour former des hypothèses sur les informations cachées.

Préparation des données. Cette phase couvre toutes les activités permettant de construire l'échantillon final (le jeu de données qui sera fourni à l'outil de modélisation). Les tâches de préparation des données sont souvent exécutées plusieurs fois, et non dans un ordre

prescrit. Ces tâches incluent la sélection de tables, d'enregistrements et d'attributs, ainsi que la transformation et le nettoyage des données pour l'outil de modélisation.

Modélisation. Dans cette phase, diverses techniques de modélisation sont sélectionnées et appliquées, et leurs paramètres ajustés aux valeurs optimales. Généralement, il existe plusieurs techniques pour le même type de problème de datamining. Certaines techniques ont des exigences spécifiques sur la forme des données. Par conséquent, un pas en arrière vers la préparation des données est souvent nécessaire.

Évaluation du modèle. A cette étape du projet, un (ou des) modèle qui semble avoir une haute qualité, du point de vue de l'analyse de données, a été construit. Avant de procéder au déploiement final du modèle, il est important de bien évaluer le modèle, et de revoir les étapes exécutées pour construire le modèle, pour être certain qu'il satisfait correctement les objectifs métiers. A la fin de cette phase, une décision sur l'utilisation ou non des résultats du datamining doit être prise.

Utilisation du modèle. La mise au point du modèle n'est pas la fin du processus de datamining. Une fois les connaissances extraites des données, elles doivent encore être organisées et présentées de façon à les rendre utilisables par les destinataires du modèle. Selon les besoins, le déploiement peut être aussi simple que de fournir une synthèse descriptive des données ou aussi complexe que de mettre en œuvre un processus de fouille de données répétable. Dans tous les cas, c'est l'utilisateur, et non l'analyste de données, qui va mettre en œuvre la phase de déploiement. Il est néanmoins toujours important que l'utilisateur comprenne d'emblée quelles actions devront être menées afin de véritablement faire usage du modèle.

1.5.4. Datamining pour la proactivité dans la gestion du changement

Dans notre application du datamining (Figure 7), la phase compréhension du métier repose essentiellement sur la description de l'objectif de proactivité dans la gestion du changement qui est l'objectif principal de cette étude. L'objectif de proactivité dans la gestion du changement est basé sur les limites des méthodes non-quantitatives de gestion du changement. Ces dernières proposent des actions, ou parfois même des stratégies, pour accompagner le changement et gérer son impact. Cependant, elles n'intègrent pas un moyen d'identifier de manière précise les processus, les employés, ou les services les plus touchés dans l'entreprise pour y orienter les ressources de gestion du changement. Il faut noter que ces ressources constituées d'efforts de communication, de réunions, de formation, d'accompagnement, d'application d'autres stratégies ne sont pas illimitées. La finalité de l'application du datamining dans cette étude est d'isoler les éléments les plus impactés par la restructuration pour y concentrer l'action de gestion du changement.

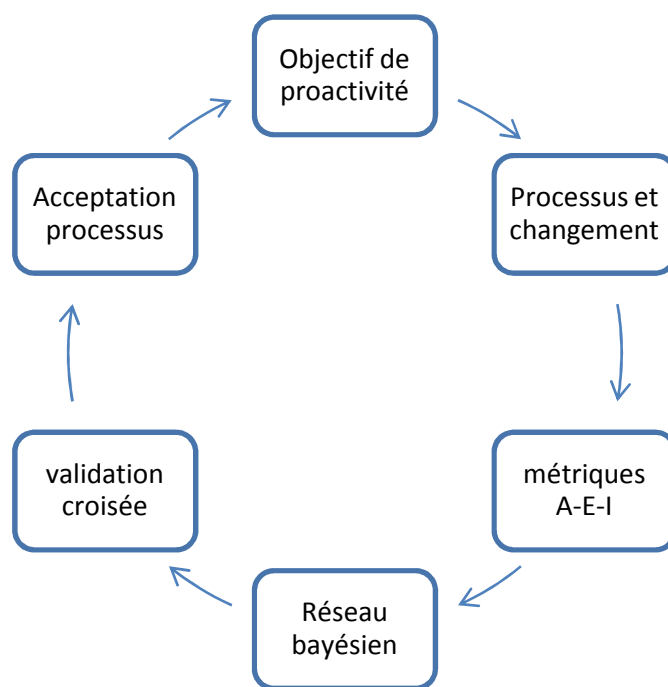


Figure 7 Datamining pour la gestion proactive du changement

Nous avons vu que, dans les projets ERP, le changement est introduit à travers les processus et qu'il doit être accepté, validé avant d'être implémenté. Il nous semble que toute analyse de l'impact du changement doit se dérouler à cette étape d'acceptation, pour éviter de valider et d'implémenter un changement dont les conséquences potentielles sont inconnues. A l'acceptation, le changement n'est qu'à l'étape de proposition et donc son impact ne s'est pas encore réalisé effectivement. Mesurer l'impact du changement à ce moment là, ne peut se faire donc que de manière prédictive. La prédiction est la construction d'un modèle prédictif dont les variables à expliquer sont les mesures de l'impact du changement. Les mesures du changement font partie des variables prédictives du modèle. Le changement se réalisant sur la structure de l'entreprise et de ses processus doit être mesuré au niveau structurel. Les autres variables explicatives sont les mesures de l'effort de gestion du changement.

La définition d'une unité d'analyse permet aussi de comprendre comment le processus de datamining sera mis en place. La construction d'un modèle prédictif exige de définir d'abord une unité d'analyse. L'unité d'analyse est l'entité majeure qui est analysée dans l'étude. Dans cette étude, nous avons fixé l'unité d'analyse au niveau des processus métier. L'une des raisons principales de ce choix est que le changement est introduit sur les processus pour améliorer leurs performances. Les autres motivations de ce choix et ses conséquences sont détaillées dans la section 1.5.5.

Processus et changement. Dans la compréhension des données, nous parlerons des sources de données liées au processus que sont la modélisation et la supervision des processus

métier. L'objectif principal de cette phase est de repérer les données manipulées en ingénierie ou en réingénierie du processus qui peuvent être intégrées au processus de datamining. Nous parlerons aussi des relations entre les dimensions de l'ampleur, de l'effort de gestion et de l'impact du changement. Nous présenterons ensuite les variables à définir sur les processus.

Métriques A-E-I (Ampleur-Effort-Impact). La phase de préparation des données comprend la définition des variables qui seront utilisées dans la modélisation. Ces variables sont des mesures de l'ampleur, de l'effort de gestion et de l'impact du changement.

Mesure de l'impact du changement. La mesure de l'impact du changement définit les variables à prédire sur l'unité d'analyse (Camara, et al., 2007b). La mesure de l'impact du changement se fait sur la variation de la performance opérationnelle du processus (Camara, et al., 2007b). L'impact sur les indicateurs de performance du processus est le critère de jugement du changement dans notre étude. En effet, les résultats de la réingénierie sont attendus sur l'axe opérationnel. Dans l'axe opérationnel, le processus est considéré de manière dynamique. L'intérêt est porté sur l'exécution des instances du processus et ses performances.

Mesure du changement. La mesure de l'ampleur du changement est une partie importante de la prédiction car elle permet de définir, sur l'unité d'analyse, une partie des variables explicatives du modèle (Camara, et al., 2007b). La mesure du changement se fait sur l'axe structurel car, c'est la partie du processus qui est manipulée lors de sa conception ou de sa réingénierie. Cet axe permet de s'intéresser à des caractéristiques statiques du processus telles que la complexité, la flexibilité, l'intégration, etc.

Mesure de l'effort de gestion. La prédiction ne peut être complète que si elle intègre des variables explicatives représentant l'effort de gestion du changement. La prise en compte de l'effort de gestion du changement dans la prédiction se justifie par les raisons suivantes :

- l'effort déployé pour gérer le changement influe, au même titre que le changement lui-même, sur l'impact du changement,
- les variables relatives à l'effort de gestion sont utilisées comme des variables pour ajuster, optimiser les moyens à mettre en œuvre sur chaque processus,
- elle offre une vision globale de tous les moyens mis en œuvre, ce qui permet de vérifier et d'anticiper leur disponibilité

Réseau bayésien. La phase de modélisation repose sur l'apprentissage d'un réseau bayésien par recherche gloutonne. Les réseaux bayésiens sont une technique statistique qui permet de modéliser des relations d'influence entre variables. La recherche gloutonne est une approche utilisée pour sélectionner un modèle bayésien parmi un grand nombre de possibilités. Dans notre étude, nous désirons faire des prédictions de l'impact du changement à partir de mesures de l'ampleur et de l'effort de gestion. Cela correspond à la fonction de prévision ou d'apprentissage supervisé dans la typologie des techniques de datamining (Han, et al., 2001 p. 24) (Rakotomalala, 200B). Nous avons choisi les réseaux bayésiens parce qu'ils

permettent de réaliser cette fonction. La seconde motivation de ce choix est l'adéquation entre les types de nos variables et ceux supportés par les réseaux bayésiens.

La troisième raison du choix de l'apprentissage des réseaux bayésiens par recherche gloutonne est la possibilité d'apprendre la structure du réseau à partir des données (Camara, et al., 2007a). En effet, partir d'un réseau bayésien donné implique d'accepter des hypothèses issues de la littérature ou d'en émettre d'autres sur les relations d'influence entre les dimensions de l'ampleur, de l'effort de gestion et de l'impact du changement. Cela reviendrait à introduire de la subjectivité dans la modélisation en plus de celle liée au choix des variables de départ. La quatrième raison de ce choix réside dans la capacité de la recherche gloutonne à sélectionner les variables pour avoir une meilleure exactitude prédictive du modèle.

Validation croisée. La phase d'évaluation du modèle bayésien repose sur la validation croisée avec k sous-échantillons. Comme la validation croisée avec k sous-échantillons inclus l'apprentissage, les phases d'évaluation et de modélisation ne forment en réalité qu'une seule phase dans notre processus de datamining. La validation croisée avec k sous-échantillons assure la précision du modèle qui représente sa capacité intrinsèque à reconnaître la variable à prédire dans la population (Rakotomalala, 1997). Nous avons choisi cette méthode de validation car elle assure un biais et une variance minimale en évitant de calculer le taux d'erreurs directement sur l'échantillon ou sur des fichiers de validation qui se recouvrent²⁶ (Rakotomalala, 1997). Elle permet aussi de prendre en compte toutes les sources de variabilité en ne réalisant pas l'estimation de l'erreur exclusivement sur l'échantillon de validation (Rakotomalala, 1997).

Acceptation processus. La phase d'utilisation du modèle intervient dans la phase d'adéquation et de configuration de l'implémentation de progiciel. Cette phase comprend dans sa description initiale une boucle adéquation-configuration-prototypage qui conduit vers l'acceptation du processus en cours de restructuration (Figure 3). Notre démarche de gestion proactive vise à intégrer l'analyse du changement dans cette prise de décision. Elle propose d'exécuter une boucle restructuration-extraction-prédiction-acceptation dans cette phase d'adéquation et de configuration, en parallèle à celle que cette dernière contient déjà.

L'extraction est le calcul, pour le processus à restructurer, des variables explicatives qui sont les mesures de l'ampleur et des moyens de gestion du changement. La prédiction est l'application du modèle pour obtenir des estimations de variables de réponse du processus. Les prédictions portent sur l'impact opérationnel du changement et sont obtenues en introduisant dans le modèle les variables explicatives. Cette prédiction permet de cibler les processus les plus critiques et constitue une première partie de la proactivité (Camara, et al., 2006).

²⁶ Qui ont des observations en commun.

L'acceptation combine les prédictions avec les informations issues du prototypage et de la simulation pour valider ou non le processus « en-devenir » proposé. La partie de la prise de décision basée sur la gestion proactive vise la réduction de l'impact du changement sur la performance du processus. Elle propose des actions de réduction directe et indirecte de cet impact. Les actions proactives de réduction directe sont celles qui agissent pour diminuer l'impact du changement (Camara, et al., 2006). Ces actions sont basées sur l'activation du levier « effort de gestion du changement ». Les actions proactives de réduction indirecte sont celles qui agissent sur l'impact du changement en passant par la réduction de la cause c'est-à-dire l'ampleur du changement. Ces actions utilisent ce qu'on peut appeler le levier « ampleur du changement » qui consiste à proposer une nouvelle structure pour le processus « en-devenir » (Camara, et al., 2006). Notre démarche, à travers ces deux alternatives, vise à apporter de la flexibilité dans la gestion du changement. Cette flexibilité est importante car, dans la réingénierie des processus, il existe d'autres contraintes en dehors de la gestion du changement, comme l'alignement à l'ERP.

Notre proposition est très proche de l'aide à la décision basée sur les scénarios de transformation (Boucher, et al., 2006) (Chapron, 2006). Le concept de scénario de transformation est une adaptation par l'ingénierie de la notion de trajectoires d'organisation dans le domaine du management (Boucher, et al., 2006). Leur objectif commun est de supporter la phase de diagnostic lors du changement d'une ancienne configuration vers une nouvelle (Boucher, et al., 2006). Dans notre proposition, il s'agit d'aider à choisir parmi plusieurs configurations d'un processus « en-devenir » en intégrant le changement par rapport à la version « en-cours » et les moyens de gestion du changement à mettre en place.

1.5.5. Le processus comme unité d'analyse

Dans cette étude, les processus métier représentent les objets sur lesquels les variables seront définies²⁷, les données collectées²⁸ et les conclusions tirées²⁹. Nous présentons ci-dessous les motivations et les conséquences de ce choix.

1.5.5.1. Les motivations de l'analyse au niveau processus

Le choix du processus métier comme unité d'analyse permet d'avoir une correspondance avec l'unité d'analyse du projet l'ERP et à notre problématique, un niveau de granularité adéquat, et une unicité du niveau d'analyse.

²⁷ Unité d'analyse

²⁸ Unité d'observation

²⁹ Unité de généralisation

Correspondance avec l'unité d'analyse du projet l'ERP et notre problématique.

L'unité d'analyse doit être choisie de manière à correspondre avec l'objectif de l'étude. L'objectif de notre étude est la gestion proactive du changement dans les projets ERP. Cela nécessite de réaliser des prédictions sur la performance des processus. Cet objectif est centré sur les processus. Notre approche est dédiée aux projets ERP dont l'unité d'analyse est le processus. Le choix de l'unité d'analyse au niveau processus correspond donc à l'objectif de notre étude et à celui du projet ERP (Camara, et al., 2005). En effet, selon García, et al. (2004), l'unité d'analyse en BPR est le processus en opposition aux départements ou aux domaines fonctionnels. Le BPR étant une pierre angulaire dans l'implémentation des progiciels, l'unité d'analyse dans un projet ERP est aussi le processus. Dans les projets ERP, les changements sont introduits par le BPR, processus par processus. Nous pensons que l'unité d'analyse de la gestion du changement, ou de la mesure ou de la prédiction de l'impact du changement doit aussi être le processus.

Niveau de granularité. Dans cette étude, nous voulons identifier les parties de l'entreprise les plus impactées par le changement pour y concentrer l'effort de gestion du changement. Le niveau de granularité doit donc être pris en compte dans la détermination de l'unité d'analyse. Cela signifie que l'unité d'analyse ne doit être ni trop générale ni trop détaillée. Nous avons vu que les dimensions du changement, de l'effort de gestion et de son impact se situent à différents niveaux comme acteur-poste de travail, processus, entreprise ou inter-entreprise. L'inter-entreprise ne peut pas être choisie comme unité d'analyse car le périmètre du projet ERP lui-même et les moyens de gestion du changement se limitent à l'entreprise. De même, la question de savoir dans quelle entreprise les moyens de gestion du changement doivent être concentrés n'est pas pertinente. En effet, il n'y a qu'une seule entreprise concernée par le projet.

L'unité d'analyse doit être cherchée parmi les sous-systèmes de l'entreprise comme le processus ou les postes de travail-acteurs. Nous pensons que le niveau poste de travail-acteur est trop détaillé. En effet, à cause du grand nombre d'acteurs dans le périmètre du projet ERP, ce choix entraînerait une distribution des moyens de gestion du changement plus complexe. Le niveau processus quant à lui n'est ni trop global ni trop détaillé. En effet, il y'a moins de processus que de postes de travail-acteurs, ce qui rendra la distribution des moyens de gestion du changement moins complexe.

Unicité du niveau d'analyse. Le fait de fixer dès le départ l'unité d'analyse au niveau processus nous permet d'assurer l'unicité du niveau d'analyse. Toutes les variables sont définies ainsi à ce niveau. Cela nous permet d'éviter les problèmes liés à la nature hiérarchique de la population. Quand la population étudiée a une structure hiérarchique, la définition de l'unité d'analyse rencontre les problèmes suivants (Arrègle, 2003) (Hox, 1995):

- Echantillonnage à plusieurs niveaux
- Définition des variables à plusieurs niveaux

Echantillonnage à plusieurs niveaux. La recherche multi-niveaux désigne l'étude simultanée d'au moins deux niveaux d'analyse (Arrègle, 2003). Dans une recherche pour l'éducation par exemple, la population est composée d'écoles et d'élèves dans ces écoles. L'échantillonnage procède en deux étapes : un certain nombre d'écoles sont sélectionnées, et un échantillon d'élèves est choisi dans chacune de ces écoles. Dans un tel échantillonnage, les observations individuelles ne sont pas complètement indépendantes à cause du processus de sélection dans l'école³⁰, d'une part ; et de l'histoire, de l'environnement ou des conditions qu'ils partagent, d'autre part (Hox, 1995). L'hypothèse d'indépendance des observations est violée. Cela entraîne une estimation très faible de l'erreur standard par les tests statistiques conventionnels et produit beaucoup de résultats faussement significatifs (Hox, 1995). La procédure de correction habituelle est de calculer l'erreur standard par la méthode des moindres carrés, d'estimer ensuite la corrélation entre les individus de chaque groupe et d'employer une formule de correction à l'erreur standard (Hox, 1995).

Définition des variables à plusieurs niveaux. Arrègle (2003) donne l'exemple de la recherche sur les performances des employés qui amène à considérer et mesurer des variables relevant de différents niveaux d'analyse. Des variables caractérisant la formation, les compétences et la motivation peuvent selon lui être collectées au niveau de l'individu pour expliquer sa performance. D'autres variables telles que la taille, le mode de management, l'ambiance, le pourcentage annuel de promotion, le soutien accordé par les autres membres de l'équipe, peuvent avoir un impact sur son activité. Ces variables sont mesurées au niveau du service et ont la même valeur pour tous les individus se trouvant dans un service donné.

Cette hétérogénéité des niveaux d'analyse peut se traiter de trois façons différentes. Les deux premières approches permettent de placer toutes les variables à un niveau unique par agrégation ou désagrégation pour appliquer ensuite une méthode d'analyse standard comme les moindres carrés. L'agrégation signifie que les variables du niveau inférieur sont déplacées au niveau supérieur en utilisant par exemple la moyenne des valeurs des variables de niveau inférieur (Hox, 1995 p. 2). Si les données sont agrégées, cela veut dire que différentes valeurs provenant de plusieurs sous-unités sont combinées en un nombre de valeurs plus petit du niveau supérieur. Cela occasionne une perte d'information et une perte de force de l'analyse statistique (Arrègle, 2003) (Hox, 1995 p. 4). La désagrégation consiste quant à elle à déplacer les variables à un niveau plus bas, en assignant par exemple à chaque unité de niveau inférieur un indicateur représentant l'unité de niveau supérieur dans laquelle elle se trouve (Hox, 1995 p. 2). Peu de données issues d'un petit nombre d'unités sont transformées en des valeurs pour un nombre plus large de sous-unités. La désagrégation amène à utiliser le nombre de cas désagrégés comme la taille d'échantillon (Hox, 1995 p. 4).

³⁰ Le fait, par exemple, que les élèves de même classe socio-économique soient sélectionnés dans les mêmes écoles diminue l'indépendance des observations

Nous avons une significativité fautive car ces cas désagrégés ne sont pas indépendants (Arrègle, 2003) (Hox, 1995 p. 5).

La troisième possibilité est d'utiliser les modèles linéaires hiérarchiques ou modèle de régression multi-niveaux. Le modèle de régression multi-niveaux assume qu'il y a un ensemble de données hiérarchiques, avec une variable dépendante mesurée au niveau le plus bas et des variables explicatives à tous les niveaux (Hox, 1995) (Osborne, 2000A). Dans une régression à deux-niveaux par exemple, une régression linéaire est réalisée séparément dans chaque unité (groupe) pour prédire la variable dépendante du bas niveau à partir des variables explicatives de ce même bas niveau (Hox, 1995 p. 10). L'étape suivante dans la régression à deux-niveaux est de prédire la variation des coefficients de régression en introduisant des variables explicatives de niveau supérieur (Hox, 1995 p. 11).

1.5.5.2. Les implications de l'analyse au niveau processus

Le choix est donc fait de prendre le processus métier comme unité d'analyse pour notre étude. Ce choix pose plusieurs problèmes : difficulté de collecter les données, degré de similitude entre processus, et hétérogénéité des méthodes et des outils de modélisation et des outils de supervision.

Difficulté de collecter les données. La contrainte principale par rapport à l'application de cette démarche est la difficulté de collecter les données nécessaires à l'apprentissage du modèle. En effet ces données sont composées des modèles et d'informations d'exécution des versions « en-cours » et « implémenté » de processus restructurés dans des projets ERP antérieurs. La collecte de ces données est rendue difficile par le fait qu'elles doivent provenir de projets différents donc d'entreprises différentes. De plus, le modèle construit avec ces données ne peut être utilisé que pour des projets différents de ceux dans lesquels elles ont été collectées. En plus de cela, ces données doivent nécessairement être recueillies à des moments précis. En effet, les enregistrements des données opérationnelles pour tous les processus de l'historique doivent être réalisés pendant deux intervalles de temps définis autour de l'implémentation du progiciel, un avant et un autre après³¹. Les mesures structurelles du changement doivent être calculées sur des modèles de processus correspondant à ces deux intervalles. Pour contourner cette difficulté, nous avons fait le choix d'utiliser des données générées aléatoirement (Camara, et al., 2006). L'objectif de cette étude n'est donc pas d'appliquer entièrement la démarche proposée, mais de démontrer sa pertinence, son déroulement et sa faisabilité technique.

Similitude entre les processus. La question qui est posée est de savoir si tous les processus métiers sont assez similaires pour permettre de tirer des conclusions générales sur

³¹ Ces intervalles peuvent être fixés par exemple au dernier mois d'exécution avant l'implémentation et au premier mois après celle-ci

leur restructuration. Pour le moment, le seul critère de similarité pour les processus pouvant être utilisé dans notre démarche est relatif au champ d'application. En effet, les processus que nous considérons dans cette étude doivent être implémentés³² ou en cours de restructuration³³ dans un projet ERP où le progiciel est choisi avant la restructuration³⁴. La question de la similarité des processus se pose donc en termes de rajout ou non de critères de similarité supplémentaires. Deux hypothèses existent selon qu'on répond à cette question par l'affirmative ou pas.

1. Si nous supposons que tous les processus ne sont pas similaires pour faire une généralisation, nous devons intégrer des données de contexte. Ces données de contexte peuvent être liées au secteur d'activité, à la taille, à l'ambiance de résistance face au changement de l'entreprise qui contient le processus ; au type de progiciel ; ou à la couverture fonctionnelle du projet. Deux solutions peuvent nous permettre d'intégrer ces données de contexte. La première solution consiste à réduire le périmètre de l'étude en appliquant des restrictions par rapport à un contexte particulier. Cela reviendrait à fixer par exemple un domaine d'activité et/ou à une ambiance et/ou une couverture fonctionnelle pour la modélisation et l'utilisation du modèle. Dans ce cas, la difficulté de collecter les données serait accentuée parce que seul les processus de même contexte pourront être utilisés lors de la collecte des données. L'utilisation du modèle sera elle aussi restreinte au sous-ensemble de processus de même contexte. La seconde solution est d'ajouter aux observations des variables qui représentent le contexte. Dans le cas où ces variables sont d'un niveau supérieur au processus, ce choix entraînerait une structure hiérarchique dans les données. Pour traiter cela, l'agrégation ne peut pas être envisagée car le niveau d'analyse principale est déjà fixé au processus. La solution doit être trouvée entre la désagrégation et la modélisation multi-niveaux. La désagrégation pose le problème de l'indépendance des observations qui conduit à une significativité fautive (Arrègle, 2003) (Hox, 1995). La modélisation multi-niveaux augmente la complexité par rapport à la modélisation sur un seul niveau.
2. Si nous considérons que tous les processus métiers sont assez similaires pour permettre une généralisation, nous pouvons modéliser et utiliser le modèle sans prendre en compte le contexte. Cette solution permet

³² Condition valable pour collecter un processus dans l'échantillon d'apprentissage

³³ Condition valable si on veut appliquer le modèle obtenu sur le processus

³⁴ Restriction relative à la stratégie de BPR utilisée dans le projet

d'éviter les contraintes liées à la prise en compte du contexte. De plus, la notion de contexte n'apparaît ni explicitement ni implicitement dans la plupart des relations établies dans la littérature entre l'ampleur, l'effort de gestion et l'impact du changement. Devant la difficulté de trancher la question de la prise en compte du contexte, nous optons pour la solution la plus simple, c'est-à-dire ne pas prendre en compte le contexte. Le choix d'intégrer ou non le contexte rentre dans le cadre de la définition de l'unité d'analyse et de ses variables. Ces définitions ne sont pas définitives car elles peuvent être modifiées sur la base d'informations issues de l'apprentissage ou de l'utilisation du modèle. Dans le cas où le contexte serait important en réalité, cela conduirait à un taux d'erreurs lors de l'apprentissage du modèle ou à des retours d'expériences insatisfaisantes dans les actions proactives lors de son utilisation. Une redéfinition de l'unité d'analyse et des variables serait à envisager dans ces cas de figures.

En raisonnement à base de cas, la similarité des observations est assurée par un test entre tous les cas et le cas cible. Le contexte peut être intégré à travers les variables de description du cas sans que cela pose le problème de la définition des variables à plusieurs niveaux ; ou n'accroît la difficulté de collecter les données. Le raisonnement à base de cas permet donc de résoudre le problème de la similarité des observations de manière plus simple. Si nous avons choisi de travailler avec l'apprentissage des réseaux bayésiens, c'est pour d'autres raisons que nous présenterons au chapitre trois.

Hétérogénéité des méthodes et outils de modélisation et de supervision. Plusieurs outils et méthodes sont utilisés actuellement pour la modélisation et la supervision des processus. Dans la démarche proposée, nous devons collecter des métriques structurelles et opérationnelles sur des processus issus de différents projets ERP donc modélisés ou supervisés avec des outils, ou des méthodes, différents. Cela pose le problème de l'interopérabilité des métriques. L'interopérabilité des métriques structurelles de processus signifie que les métriques sont indépendantes du langage de spécification et ont la même valeur pour un processus donné, quelque soit le langage de modélisation (Cardoso, 2005) (Gruhn, et al., 2005). Pour les métriques opérationnelles, il s'agit d'assurer l'indépendance par rapport aux outils de supervision. Dans le cas où les métriques ne peuvent être calculées que dans un langage, un outil donné, la non-interopérabilité conduit à la restriction du périmètre de collecte des observations et participe à l'aggravation de la difficulté de collecter les données. Dans le cas où la non-interopérabilité se manifeste par des valeurs de métriques différentes pour des processus modélisés avec des langages différents, alors cela induit des erreurs dans les données donc dans le modèle et ses prédictions. Les méta-modèles de processus (Darras, 2004) (Roque, 2005) ou les ontologies de processus (Camara, et al., 2008) sont des solutions qui peuvent être utilisées pour atteindre cette interopérabilité. Cependant,

nous avons fait le choix de ne pas traiter l'interopérabilité des variables définies sur les processus, dans cette étude. Pour ce faire, nous avons choisi la notation BPMN³⁵, un langage de modélisation standard, à partir duquel toutes nos métriques seront définies.

³⁵ Business Process Management Notation

Chapitre 1.6 Conclusion de la première partie

Cette partie nous a permis de préciser notre problématique et de présenter la démarche que nous désirons suivre. La démarche proposée est une application du processus de datamining. Notre démarche utilise certaines des approches de gestion non-quantitative pour les actions qu'elles mettent en place. La phase de modélisation de notre démarche s'inspirera des techniques utilisées déjà pour la gestion quantitative du changement. La première phase de compréhension du métier a été présentée dans cette partie. Elle a permis de définir l'impact du changement comme ce qui doit être prédit. L'unité d'analyse y a été définie au niveau processus. Le processus de datamining est utilisé pour valider l'acceptation d'un processus en restructuration et définir l'effort de gestion du changement à prévoir.

Les phases de compréhension du métier et de préparation des données sont présentées dans la partie suivante. Les phases de modélisation, d'évaluation et d'utilisation du modèle sont regroupées dans la dernière partie.

PARTIE 2 DEFINITION DES VARIABLES DU PROBLEME

Chapitre 2.1 Introduction de la deuxième partie

Cette partie correspond à la « définition des variables » dans notre démarche. Nous posons ici les bases nécessaires pour l'apprentissage et l'utilisation du modèle qui sera présenté dans la partie suivante. Cette partie est guidée par l'objectif général de l'étude qui est la gestion proactive de l'impact du changement sur les processus. Nous avons un premier chapitre qui s'intéresse à la place du processus dans notre démarche. Ce chapitre correspond à la compréhension des données dans le processus de datamining. Il traite de l'ingénierie des processus pour identifier les sources de données opérationnelles et structurelles du processus. Les relations entre les dimensions de l'ampleur, de l'effort de gestion et de l'impact du changement sont identifiées au niveau processus. Pour définir les variables du problème, nous appliquons la méthode Goal Question Metric (GQM) (Figure 8) qui nous permet d'identifier les métriques à définir à partir des objectifs de l'étude et des questions qui leurs sont associées. L'application de cette méthode tient compte de l'unité d'analyse pour définir les objectifs, les questions et les métriques à son niveau. La méthode GQM prend en entrée l'objectif de gestion proactive et ses sous-objectifs pour définir sa partie Goal (Objectif). Les questions s'inspirent des dimensions de l'ampleur, de l'effort de gestion et de l'impact du changement identifiées au niveau processus

La définition effective des métriques est présentée dans les chapitres « définition des variables à expliquer » et « définition des variables explicatives ». Ces deux chapitres correspondent à la partie Metric de l'approche GQM et à la phase de préparation des données dans le processus de datamining. La définition des variables à expliquer se fait par la mesure de la variation opérationnelle des critères de performance. La définition des variables à expliquer s'inspire des métriques opérationnelles définies en supervision des processus. La définition des variables explicatives repose, quant à elle, sur la mesure de l'ampleur et de l'effort de gestion du changement. La définition des variables explicatives s'appuie essentiellement sur le méta-modèle de description de processus BPMN. Elle s'inspire en partie des métriques structurelles définies en analyse structurelle des processus.

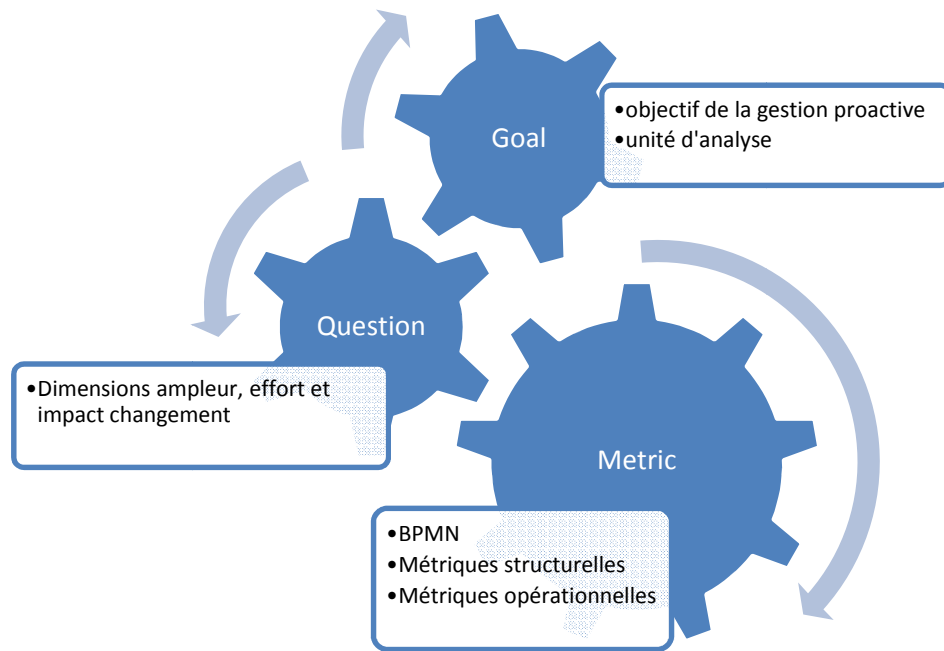


Figure 8 Définition des variables du problème

Chapitre 2.2 Le processus métier dans notre démarche

2.2.1. Introduction

Ce chapitre correspond à la phase de compréhension des données du processus de datamining. Il nous permet de comprendre les sources de données que nous utiliserons pour définir nos variables. Nous allons nous intéresser à l'ingénierie des processus métier. L'objectif est d'identifier les données qui présentent le processus au niveau opérationnel et structurel. Nous présenterons la notation BPMN que nous avons choisie comme langage de modélisation de processus pour définir nos variables structurelles. Le cycle de vie des processus dans notre démarche sera présenté. Parmi les dimensions de l'ampleur, de l'effort de gestion et de l'impact du changement, celles de niveau processus seront présentées. Les relations entre ces dimensions seront aussi déterminées. Nous allons donner une présentation générale de la définition de nos variables en utilisant la méthode GQM. Cette dernière permet de montrer la pertinence des variables par rapport à l'objectif de notre étude.

2.2.2. Ingénierie des processus métier

Le BPM³⁶ est l'ingénierie des processus métier des organisations à l'aide des technologies de l'information (Debauche, et al., 2004). Un processus métier est une suite d'activités qui, à partir d'une ou de plusieurs entrées (inputs), produit un résultat (output) représentant une valeur pour le client (Hammer, et al., 1995). Debauche, et al. (2004 p. 50), distinguent trois phases dans l'ingénierie des processus métier :

- analyse des processus métier,
- informatisation des processus métier,
- supervision des processus métier.

Analyse des processus métier³⁷. Durant cette phase, l'entreprise prend connaissance de ses processus métier et devient capable de les améliorer. Les techniques suivantes sont utilisées dans cette phase :

1. la modélisation des processus métier consiste à les identifier (cartographier) et à les décrire de façon formelle en appliquant une méthode donnée (Debauche, et al., 2004 p. 51) ;
2. la simulation fournit une première analyse de la performance des processus métier, au sens de leur pertinence par rapport aux objectifs de l'organisation et de leur efficacité

³⁶ Business Process Management

³⁷ Business Process Analysis (BPA)

par rapport aux ressources disponibles (Debauche, et al., 2004 p. 51) ;

3. le prototypage permet la validation fonctionnelle de la définition du processus par les utilisateurs opérationnels en charge de l'exécuter (Debauche, et al., 2004 p. 52) ;
4. l'optimisation des processus métier part des résultats de la simulation, du prototypage, ou des mesures effectuées sur les processus opérationnels, et identifie des améliorations et optimisations possibles (Debauche, et al., 2004 p. 52) ;

Informatisation des processus métier³⁸. Lorsqu'une organisation connaît et maîtrise ses processus métier clés, elle peut décider de les informatiser pour accroître sa productivité. Pour ce faire l'entreprise est confrontée aux choix suivants (Debauche, et al., 2004 p. 55):

1. le développement sur mesure,
2. la sélection d'un progiciel applicatif du marché tel que les ERP,
3. la mise en oeuvre d'un système de gestion de processus (Business Process Management System BPMS)

Supervision des processus métier³⁹. La troisième phase consiste, pour une entreprise, à superviser l'exécution des processus déployés dans des applications et à mesurer leurs performances opérationnelles (Debauche, et al., 2004 p. 55). La première activité, nécessaire et préalable à la capacité même d'observer l'exécution des processus, consiste à prélever des mesures et produire des métriques de toutes sortes. Une fois les métriques disponibles, elles sont publiées dans un rapport, qu'il soit en ligne dans une console de supervision fonctionnelle des processus en cours d'exécution ou hors ligne, dans un document fournissant des statistiques d'exécution des processus.

Ces tableaux de bord permettent alors, à ceux qui portent dans l'organisation, la responsabilité de l'efficacité opérationnelle des processus métier, d'analyser les indicateurs ainsi remontés, d'identifier les problèmes, de diagnostiquer les causes, et de proposer des améliorations (Debauche, et al., 2004 p. 56). La dernière activité consiste à agir en conséquence de l'analyse effectuée, c'est à dire réaliser les améliorations décidées, en agissant soit directement en ligne, lorsqu'il faut débloquer les instances de processus en cours d'exécution, soit hors ligne, en modifiant par exemple le modèle du processus métier et en redéployant une nouvelle version optimisant les délais par exemple.

Le projet ERP correspond à un cycle d'ingénierie des processus dans lequel, l'informatisation des processus métier se fait à travers un progiciel. Les phases d'analyse et d'informatisation des processus métier du cycle se déroulent dans l'implémentation du progiciel. La phase de supervision des processus du cycle est réalisée dans la *post* implémentation.

³⁸ Business Process Integration (BPI)

³⁹ Business Activity Monitoring (BAM)

Dans le cas où la réingénierie des processus utilisée lors de l'implémentation du progiciel est « déclenchée par la technologie », la phase d'analyse du cycle dépend du progiciel. Dans le cas où une ingénierie « à partir d'un tableau vierge » est utilisée, la phase d'analyse du cycle est complètement indépendante de l'outil qui n'est pas encore choisi à ce stade. Dans les deux cas, les phases d'informatisation et de supervision des processus métier dépendent de l'ERP lui-même qui exécute les processus.

Par exemple, les étapes d'« analyse opérationnelle » et d'« adéquation et de configuration », de la méthodologie d'implémentation de progiciel proposée par Tomas (2000), correspondent à la phase d'analyse du cycle d'ingénierie de Debauche, et al. (2004). Les étapes de « simulation taille réelle », de « fermeture des trous fonctionnels », de « modification spécifique », de « création des liens avec l'environnement », de « mise en production » et de « déploiement » de Tomas (2000) correspondent à la phase d'informatisation du cycle d'ingénierie (Debauche, et al., 2004).

L'analyse des processus métier est portée par les éditeurs comme Mega, Casewise, ou IDS, et reste l'apanage des fonctionnels (OCTO, 2007). L'informatisation des processus est le domaine des moteurs d'exécution, de l'orchestration des services, de la connexion au SI (OCTO, 2007). C'est le champ du Business Process Management contrôlé par l'informatique. Cela crée un gap technique entre le format de la modélisation initiale du processus métier et le format de langages, tel que le BPEL4WS⁴⁰, qui vont exécuter ces processus (OMG, 2007B). L'émergence de standards partagés par ces deux mondes, notamment le BPEL⁴¹ et le BPMN, pourrait conduire à une relative automatisation du passage de relais (OCTO, 2007). L'organisation BPMI⁴² s'est penchée sur la définition d'une notation standard pouvant être exploitée par les utilisateurs métier et les informaticiens, pour représenter les processus métier et permettre la génération de langages exécutables : le BPMN (White, 2007). Le BPMN fournit une mise en correspondance formelle vers un langage d'exécution d'un système de gestion de processus (BPEL4WS) (OMG, 2007B).

La supervision des processus métier peut être assurée directement par le biais d'un outil de Business Process Management, ou supportée via un produit tiers dit de BAM (Crochet-damais, 2006). Le BAM suppose qu'on intègre des événements générés par le moteur de processus (BPMS) ou des événements directement issus des applications (ERP) (Debauche, et al., 2004). Ces événements sont ensuite traités et interprétés par le système de supervision. Des indicateurs-clés de performance (Key Performance Indicator, KPI) sont construits à partir de ces événements (Debauche, et al., 2004).

⁴⁰ Business Process Execution Language for Web Services

⁴¹ Business Process Execution Language

⁴² Business Process Management Initiative

Dans le cycle d'ingénierie des processus, ce qui nous intéresse le plus, pour cette étude, c'est l'étape de modélisation de la phase d'analyse des processus du cycle. L'autre étape qui nous intéresse est celle de prélèvement des mesures et de production des métriques de la phase de supervision. Ce qui est important pour nous est de mesurer le changement introduit sur la structure du processus, et l'impact de ce dernier sur les indicateurs de la performance opérationnelle du processus. Notre objectif est d'accumuler de la connaissance sur ces deux phases afin de fournir des prédictions sur un processus en cours de restructuration. Nous ne nous intéressons pas à la phase d'exécution que nous considérons comme une boîte noire dont nous utilisons les entrées et les sorties. Nous allons donc présenter la modélisation des processus. En effet, une mesure du changement correcte repose sur l'analyse et la comparaison des versions « en-cours » et « en-devenir » des modèles des processus restructurés.

2.2.2.1. Modélisation des processus

La modélisation des processus est une phase clé du BPM, dont le résultat impactera considérablement les autres phases (Debauche, et al., 2004 p. 77). La modélisation des processus permet de faire apparaître les éléments constructifs d'un processus (Morley, et al., 2007 p. 111). Les modèles de processus reproduisent des tranches de réalité des processus. La modélisation s'intéresse à de nombreux aspects du processus comme, la structure des processus, l'organisation hiérarchique, la qualification des employés, les documents (données), les prestations internes et externes ainsi que les ressources de production et de technique informatique. Les concepts nécessaires à la modélisation d'un processus métier peuvent être représentés par un méta-modèle (Morley, et al., 2007 p. 111). L'élaboration des modèles de processus poursuit les objectifs suivants (IDS, 2002) :

- l'optimisation des modifications de l'organisation dans le cadre de BPR,
- le stockage des connaissances concernant l'organisation, par exemple sous forme de modèle de référence ;
- l'utilisation de la documentation des processus à la certification ISO 9000,
- le calcul des coûts des processus,
- l'utilisation des informations sur les processus pour implémenter et adapter les programmes informatiques standards ou les systèmes de *workflow*.

Ce paragraphe propose de mentionner quelques méthodes de modélisation aujourd'hui disponibles. Pour la modélisation des processus, on trouve des méthodes comme IDEF3, UML, OSSAD, BPMN. La méthode IDEF3 (Process Description Capture Method) par exemple, a été créée spécialement pour représenter des descriptions de séquences d'activités (Mayer, et al., 1995). IDEF3 fait partie de la suite IDEF (Integrated DEFinition) (Hanrahan, 1995). IDEF3 est composé de deux modèles (Hanrahan, 1995):

- le Process Flow Description (PFD) qui décrit comment les choses sont en train de marcher dans l'organisation (le PFD offre une vue centrée processus) ;
- l'Object State Transition Description (OSTD) qui résume les transitions autorisées pour un objet dans un processus particulier : La vue OSTD fournit, entre autre, les critères d'entrées et de sorties.

UML est le langage de notation graphique standardisé bien connu. UML est traditionnellement plus associé au génie logiciel et à la conception de systèmes, qu'à la modélisation des processus (Systems, 2008). Cependant, le standard UML 2.x fournit un ensemble riche de modèles de comportement très adaptés à la modélisation des processus, des activités, des acteurs et de l'information (Systems, 2008).

Le BPMI a développé le standard Business Process Modeling Notation (BPMN) (OMG, 2007B). La publication du BPMN a été un pas majeur vers une notation standardisée et uniforme (itp, 200X). Pour cette raison, nous avons choisi de définir toutes nos métriques à partir de cette notation. Cela nous permet d'appliquer notre démarche sans satisfaire la propriété d'interopérabilité pour nos métriques structurelles. Nous nous attellerons à cette tâche dans les travaux futurs. Nous présenterons cette notation dans la section suivante.





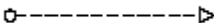

2.2.2.2. Business Process Modeling Notation (BPMN)

L'un des éléments moteurs pour le développement du BPMN est de créer un mécanisme simple de création de modèle de processus, tout en étant capable de gérer la complexité inhérente aux processus métier (OMG, 2007B). L'approche choisie pour gérer ces deux exigences conflictuelles est d'organiser les aspects graphiques de la notation dans des catégories spécifiques. Cela donne un petit nombre de catégories de notation de sorte que le lecteur du diagramme BPMN puisse reconnaître facilement les types de base des éléments (OMG, 2007B p. 17). A l'intérieur de ces catégories basiques d'éléments, des variations ou des informations additionnelles peuvent être ajoutées pour supporter les exigences de complexité sans changer profondément l'allure du diagramme (Tableau 8) (OMG, 2007B p. 24). Le méta modèle du BPMN récapitule ces éléments et leurs relations (Figure 9) (wsper.org, 2007). Les quatre catégories de base sont (OMG, 2007B p. 17) :

- **objets de flux** (événements, activités, porte logiques) : sont les principaux éléments graphiques utilisés pour définir le comportement du processus métier ;
- **objets de connexion** (flux séquence, flux message, associations) : ils correspondent aux manières de connecter les objets de flux entre eux ou avec d'autres informations ;
- **swimlanes** : ils renvoient à deux manières de grouper les objets de modélisation primaires (pools, lanes) ;
- **artefacts** (artefacts courants : objet de donnée, groupe, annotation) : ils sont utilisés pour fournir de l'information additionnelle concernant le processus.

Les types de base des éléments constituent le noyau du BPMN (Tableau 7). L'objectif de l'extension du noyau BPMN est de décliner les types de base des éléments de manière plus détaillée pour tenir compte des différents sous-types qu'ils englobent. Nous ne présenterons dans cette section que les éléments relatifs au parallélisme ou aux points de décisions (Tableau 8).

Tableau 7 Noyau BPMN (OMG, 2007B p. 18)

Elément	Description	Notation
Evénement	Un événement est quelque chose qui se produit durant le cours du processus métier. Ces événements affectent le flux du processus et sont souvent une cause ou un impact.	
Activité	Une activité est un terme générique pour désigner le travail accompli dans l'entreprise. Une activité peut être atomique ou non-atomique. Les types d'activité sont : processus, sous-processus et tâche. Les processus sont contenus dans les Pools.	
Porte logique	Un Porte logique est utilisé pour contrôler la divergence et la convergence d'un flux de séquence. Il détermine ainsi le branchement, le parallélisme, l'unification et la jointure des chemins.	
Flux de séquence	Un flux de séquence est utilisé pour montrer l'ordre dans lequel les activités sont réalisées dans le processus.	
Flux de message	Un flux de message est utilisé pour montrer le flux de messages entre deux participants qui sont préparés à les envoyer ou à les recevoir. Deux Pools distincts dans un diagramme vont représenter deux participants ⁴³ .	
Association	Une association est utilisée pour associer l'information aux flux d'objets. Les objets de type texte et les objets graphiques qui ne sont	

⁴³ Un participant peut être une entité d'affaire spécifique (comme une entreprise) ou un rôle d'affaire plus générale (comme un vendeur ou un fabricant) (OMG, 2007B p. 295).


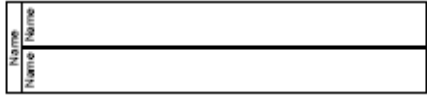


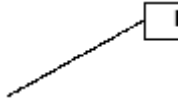
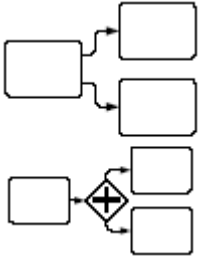
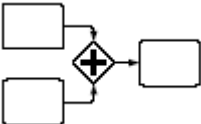
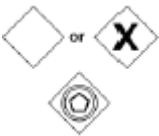
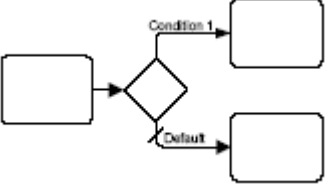
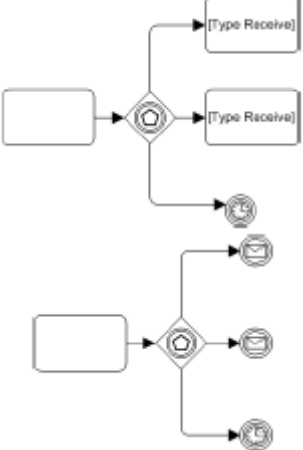
	pas de type flux peuvent être associés aux objets de flux. Une pointe permet d'indiquer la direction du flux, quand cela est approprié.	
Pool	Un Pool représente un participant dans un processus. Il agit aussi, comme un « <i>Swimlane</i> » et un conteneur graphique, pour séparer un ensemble d'activités des autres Pools, souvent dans un contexte B2B ⁴⁴	
Lane	Un <i>Lane</i> est une sous-partition à l'intérieur du Pool. Les <i>Lanes</i> sont utilisés pour organiser et catégoriser les activités. Les <i>lanes</i> sont utilisés pour représenter les rôles internes (ex : manager, associé), les systèmes (ex application de l'entreprise) ou les départements internes (ex : expédition, finance), etc.	
Objet de donnée	Les objets de donnée fournissent l'information sur ce dont les activités ont besoin pour être exécutées ou ce qu'elles produisent.	
Groupe	Un groupe est un regroupement d'activités qui sont dans la même catégorie.	
Annotation textuelle	Les annotations textuelles sont un mécanisme pour fournir de l'information additionnelle au lecteur du diagramme BPMN	

Tableau 8 Extension du noyau BPMN (OMG, 2007B p. 24)

Elément	Description	Notation
---------	-------------	----------

⁴⁴ BusinessToBusiness

Fork	<p>Le terme Fork fait référence à la division d'un chemin en deux chemins parallèles ou plus. Il y a deux options :</p> <ul style="list-style-type: none"> • flux séquence sortants multiples, • une porte logique parallèle peut être utilisée. 	
Join	<p>Le terme Join (AND Join ou synchronisation) fait référence à la combinaison de deux chemins parallèles ou plus. Une porte logique parallèle est utilisé pour représenter la combinaison de plusieurs flux.</p>	
Exclusive	<p>Une porte logique exclusive restreint le flux de telle sorte que seule une alternative est choisie dans l'ensemble. Il y a deux types de portes logiques exclusifs : ceux basés sur les données et ceux basés sur les événements</p>	
Data-based	<p>Cette décision représente un point de branchement dans lequel les alternatives sont basées sur des expressions conditionnelles contenues dans les flux de séquence sortants. Une seule des alternatives sera choisie.</p>	
Event-based	<p>Cette décision représente un point de branchement dans lequel les alternatives sont basées sur un événement qui survient à ce stade du processus. L'événement spécifique, souvent la réception d'un message, détermine lequel des chemins il faut prendre. D'autres types d'événements, tels que les Timer, peuvent être utilisés. Seule une des alternatives est choisie. Il y deux options pour recevoir les messages :</p> <ul style="list-style-type: none"> • des tâches de type « Receive » peuvent être utilisées, • des événements intermédiaires de type Message peuvent être utilisés. 	

<p>Inclusive</p>	<p>Cette décision représente un point de branchement dans lequel les alternatives sont basées sur des expressions conditionnelles contenues dans les flux de séquence sortants. D'une certaine manière, c'est un regroupement de décisions binaires (Oui/Non) indépendantes reliées. Comme chaque chemin est indépendant, toutes les combinaisons de chemins peuvent être prises, de zéro à tous les chemins. Cependant, il faut modéliser la décision inclusive avec une condition par défaut, de sorte qu'au moins un chemin soit pris.</p>	
<p>Merging</p>	<p>Le terme « merge » fait référence à la combinaison exclusive de deux chemins ou plus en un seul chemin. Une porte logique de Merging exclusif est utilisé pour représenter la combinaison de multiples flux.</p>	



BPMN 1.0 Metamodel

Copyright wsoe.org, 2007

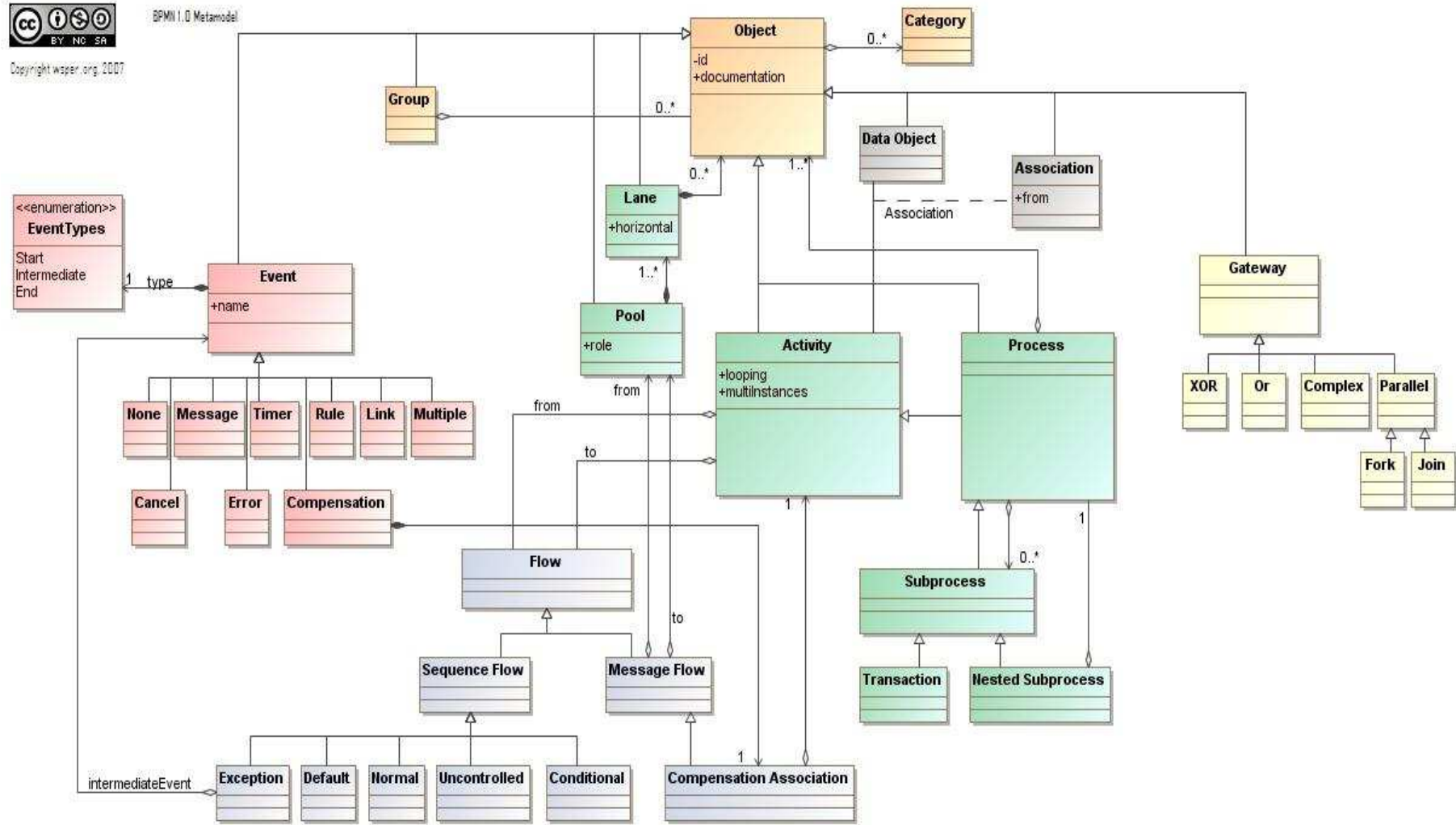


Figure 9 Méta-modèle BPMN

2.2.3. Cycle de vie du processus dans notre démarche

Dans notre étude, les processus auxquels nous faisons référence sont, soit des processus restructurés et implémentés dans des projets ERP antérieurs, soit des processus en restructuration dans des projets ERP en-cours. Ces processus passent par les états correspondant aux domaines de processus de Tomas (2000). Notre apport se trouve dans les nouvelles actions participant à la gestion du changement que nous appliquons à ces processus (Figure 10). Au début d'un projet ERP, les processus existants dans l'entreprise sont les processus « en-cours ». A cet état, toutes les méthodes de l'analyse des processus⁴⁵ (BPA) définies en BPM peuvent leur être appliquées. La supervision aussi leur est appliquée car les processus « en-cours » sont des processus déjà implémentés et exécutés dans l'entreprise. L'activité de réingénierie transforme les processus « en-cours » en processus « en-devenir ». Il est possible d'appliquer aux processus « en-devenir », en plus des méthodes citées pour les processus « en-cours », la méthode d'analyse des écarts (GAP Analysis). Notre démarche propose deux nouvelles actions à réaliser sur les processus « en-devenir » :

- analyse du changement : c'est le calcul des métriques structurelles de changement qui mesurent l'ampleur du changement sur les processus ;
- prédiction de l'impact du changement : cela consiste à prédire les métriques opérationnelles de variation relatives à la performance des processus (Cette activité ne peut être réalisée que si le modèle prédictif est déjà construit) ;
- estimation de l'effort de gestion : l'effort de gestion du changement à déployer sur le processus est estimé à priori par le modèle.

Les processus « implémentés » résultent de l'informatisation des processus « en-devenir ». Ces processus sont supervisés dans la phase « post-implémentation » du projet ERP. L'analyse de la variation opérationnelle des performances se fait par comparaison des données de supervision des processus « en-cours » à celles des processus « implémentés ». Le processus implémenté peut être inséré comme une observation dans l'historique destiné à l'apprentissage du modèle prédictif. Cette observation est décrite par les métriques opérationnelles de variation, les métriques structurelles de changement et de variation, et les métriques d'effort de gestion du changement.

⁴⁵ modélisation, simulation, optimisation et analyse structurelle des processus (ASP)

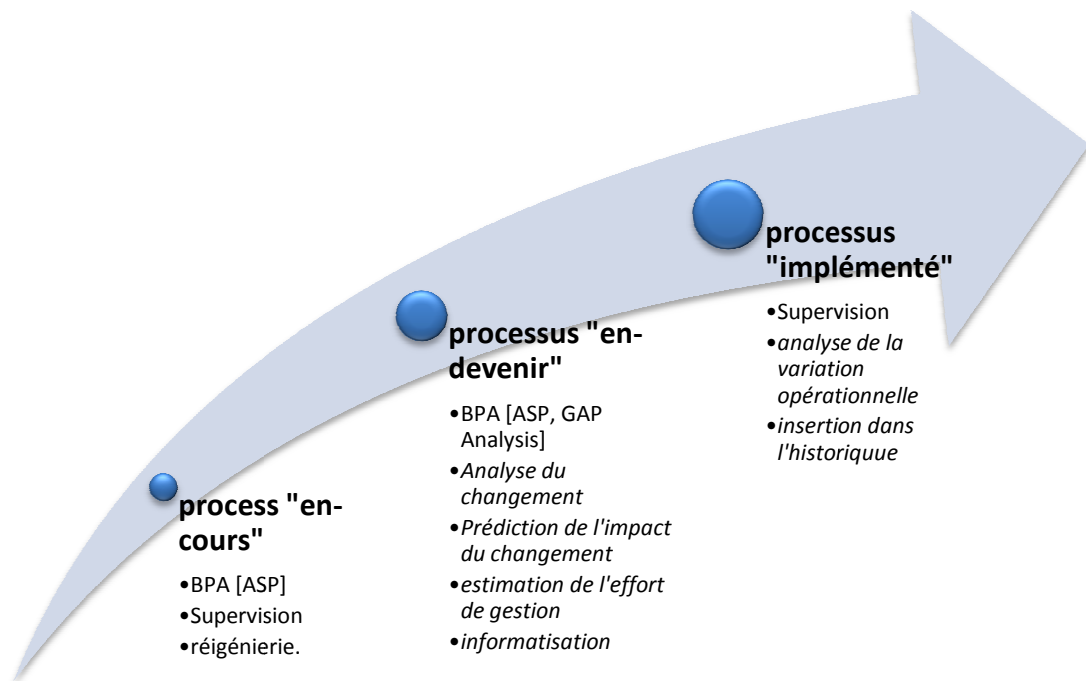


Figure 10 Cycle de vie du processus

Pour illustrer le passage de l'état de processus « en-cours » à celui de processus « en-devenir », nous avons pris en exemple un processus métier de traitement de commandes restructuré que nous avons tiré de Scheer (1998 p. 16). Les versions « en-cours » et « implémenté » de ce processus y sont présentées en « *process chain diagram* ». Nous les avons transcrites en notation BPMN (Figure 11 et Figure 12 respectivement). Ces modèles de processus seront utilisés, par la suite, pour illustrer le calcul des métriques structurelles de changement et de variation. Vu que tout se passe dans une seule entreprise nous n'avons qu'un seul *pool*. C'est pour cette raison que nous n'utilisons pas de flux messages qui sont destinés aux échanges entre entreprise (Pool). Les *lanes* représentent les services de l'entreprise.

Figure 11 Processus traitement de commandes en-cours (Scheer, 1998 p. 16)

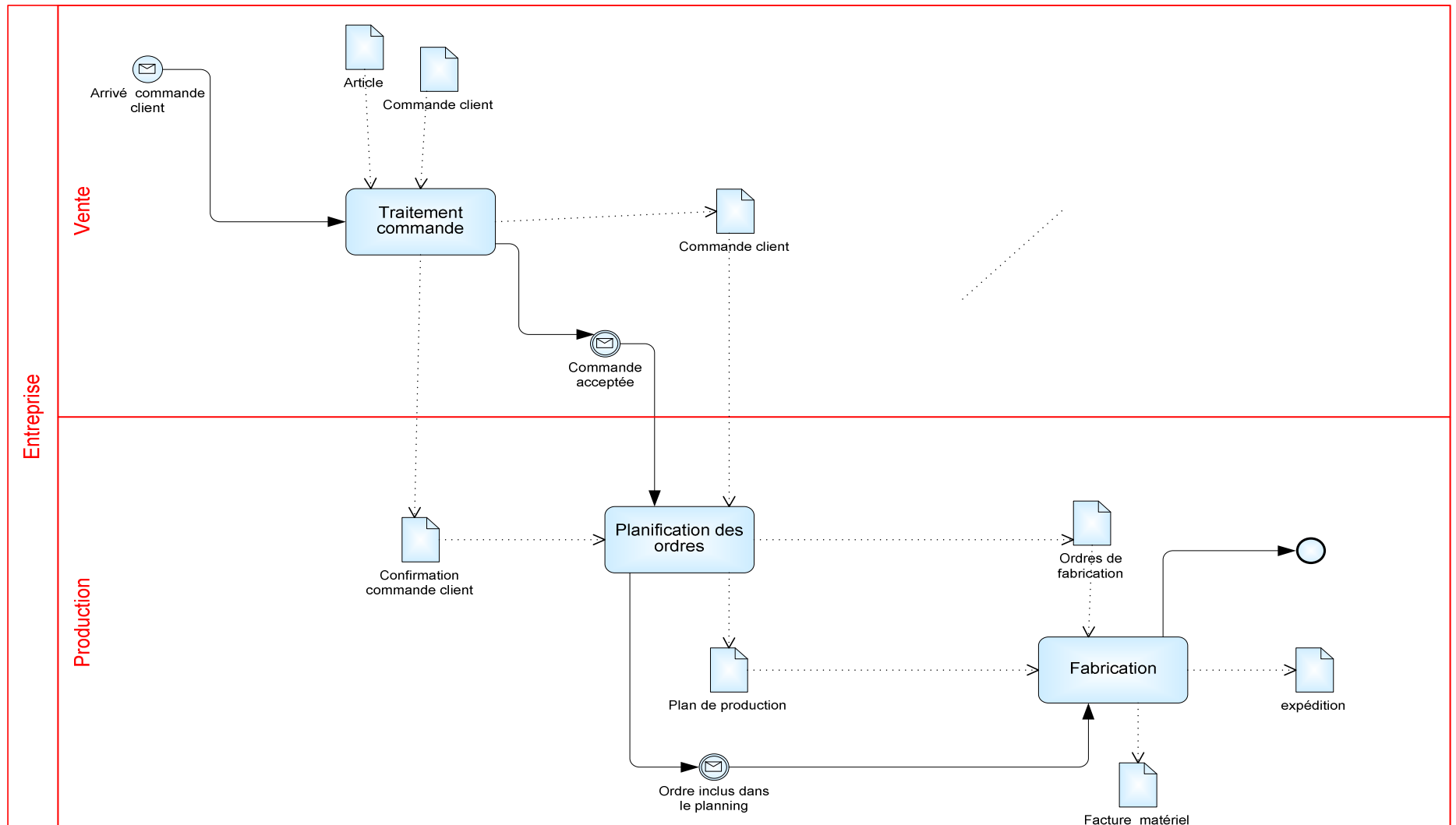
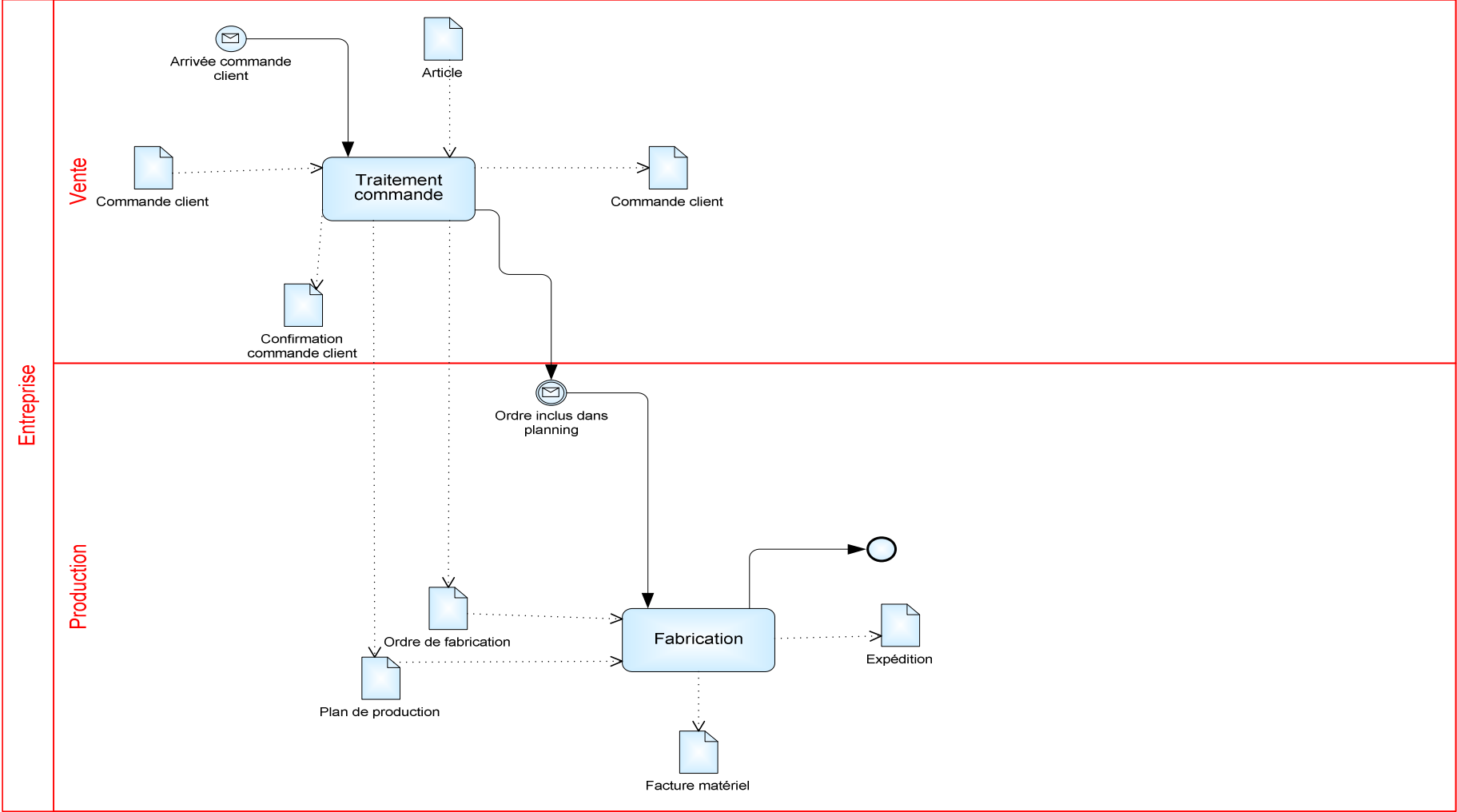


Figure 12 Processus traitement de commandes «implémenté» (Scheer, 1998 p. 17)



2.2.4. Relation d'influence entre les dimensions

Après avoir fixé l'unité d'analyse au niveau processus, nous allons rassembler les dimensions de l'ampleur, de la gestion et de l'impact du changement à ce niveau. Nous allons déterminer ensuite les relations d'influence qui existent entre ces dimensions (Figure 13). Ces relations d'influence sont un des éléments qui guideront la définition des variables de notre modèle. Pour commencer, nous avons cherché, dans la littérature des projets ERP, les hypothèses qui établissent ces relations :

H1 : le Changement Radical (CR⁴⁶) des processus métier permet de réaliser des améliorations spectaculaires de leurs performances (VOP) (Hammer, 1990) ;

H2 : l'Ampleur et la profondeur de ce Changement (AC) peut rendre les Utilisateurs Insatisfait (IU) (Bernard, et al., 2002). L'insatisfaction des utilisateurs (IU) peut se traduire par la diminution de la PROductivité de ces derniers (PROD) (Bernard, et al., 2002) ;

H3 il arrive que cet écart soit faible (AC). Cela réduit le risque, puisque les employés sont plus Facilement à même d'Exécuter le Processus (FEP) (Bernard, et al., 2002) ;

H4 : le degré d'intégration de l'organisation (DIP) évolue en fonction des changements amorcés ou subis par l'organisation (CR), comme notamment dans le cas de la mise en place d'un PGI (Barki, et al., 2003) ;

H5 : le manque de FLeXibilité dans la gestion des erreurs (FLX) (propagation des erreurs potentielles, phénomène accru par l'interdépendance des fonctions et des données) est une conséquence potentielle de l'intégration (DIP) (Perotin, 2004) ;

H6 : l'exigence de simplicité du Reengineering a d'énormes conséquences sur la façon de concevoir les processus : les ROLles évoluent (ROL), le travail est réalisé là où c'est le plus logique (RSP) et les agents assument de nouvelles ReSPonsabilités (Hammer, 1990) ;

H7 : l'Intégration Organisationnelle (IO) est reliée positivement à la performance⁴⁷ (Barki, et al., 2003) ;

H8 : selon Tjaden (2001), les propriétés structurelles des processus influent sur les performances de ce dernier⁴⁸. Dans le cadre de la réingénierie des processus, nous pouvons dériver une nouvelle hypothèse à partir de cette dernière : toute variation dans les propriétés structurelles d'un processus induit une variation dans la performance de celui-ci ;

⁴⁶ Les abréviations sont utilisées pour représenter les différentes dimensions dans le schéma (Figure 13)

⁴⁷ *In general, OI will be positively related to performance) and suggests that the larger the gap between the degree of OI before and after an ERP implementation, the greater will be the benefits obtained*

⁴⁸ *These static, or structural, properties strongly influence the performance of the process, however.*

H9 : les changements dus à la réingénierie avec un ERP (CR) peuvent conduire à des changements dans la manière et la responsabilité d'Entrer les Données (DE)⁴⁹ (O'Leary, 2000 p. 49) ;

H10 : la ForMation (FM) offre une bonne opportunité pour aider les utilisateurs à s'ajuster au changement qui a été introduit par l'ERP (FEP), et aide à construire des attitudes positives à l'encontre du système (IU) (Aladwani, 2001) ;

H11 : le management personnel (ACompagnement AC) est un système de support pour les membres individuels pour s'adapter en douceur (CEP) et participer positivement au projet de changement organisationnel (IU) (Al-Mashari, et al., 2006).

L'objectif de ce travail est de formaliser les relations entre le changement sur les processus, son impact et les moyens de sa gestion. Le choix du processus, comme unité d'analyse, impliquerait principalement de relier les dimensions « quantifiées » de ces trois concepts au niveau processus. Cependant, nous n'avons aucune dimension quantifiée d'un de ces trois concepts niveau processus. Nous allons commencer par représenter les dimensions, qui par le type et le niveau, sont facilement adaptables en dimensions quantifiées de niveau processus. Cela nous a conduit à travailler au niveau processus toujours, avec les dimensions les plus proches des quantifiées : les dimensions concrètes. Nous avons ajouté les dimensions concrètes du changement au niveau employé (rôle, responsabilité, compétences requises) pour faire leur projection en dimensions concrètes au niveau processus.

Parmi les dimensions abstraites, nous n'avons gardé que celles qui nous permettent d'établir une liaison indirecte entre une cause concrète et une conséquence concrète, toutes les deux de niveau processus. La non-satisfaction des utilisateurs (NSU) et la capacité d'exécution du processus (CEP) par exemple, sont des dimensions abstraites. Elles relient des dimensions concrètes comme l'écart entre le processus actuel et le processus cible (EPAC), et la productivité (PROD) (ou la variation opérationnelle de la performance (VOP)). Nous avons regroupé dans l'accompagnement toutes les dimensions de l'effort de gestion du changement différentes de la formation⁵⁰. Par ce regroupement, nous visons à réduire le nombre de variables représentant la gestion du changement.

⁴⁹ As process change, there are concomitant changes in how data is gathered, how often it is gathered, where it is gathered, and who gathers data.

⁵⁰ Communication (Al-Mashari, et al., 2006) (Aladwani, 2001), Management personnel (Al-Mashari, et al., 2006), Mécanismes d'intégration (Barki, et al., 2003), Dispositif favorisant l'intégration (Perotin, 2004)

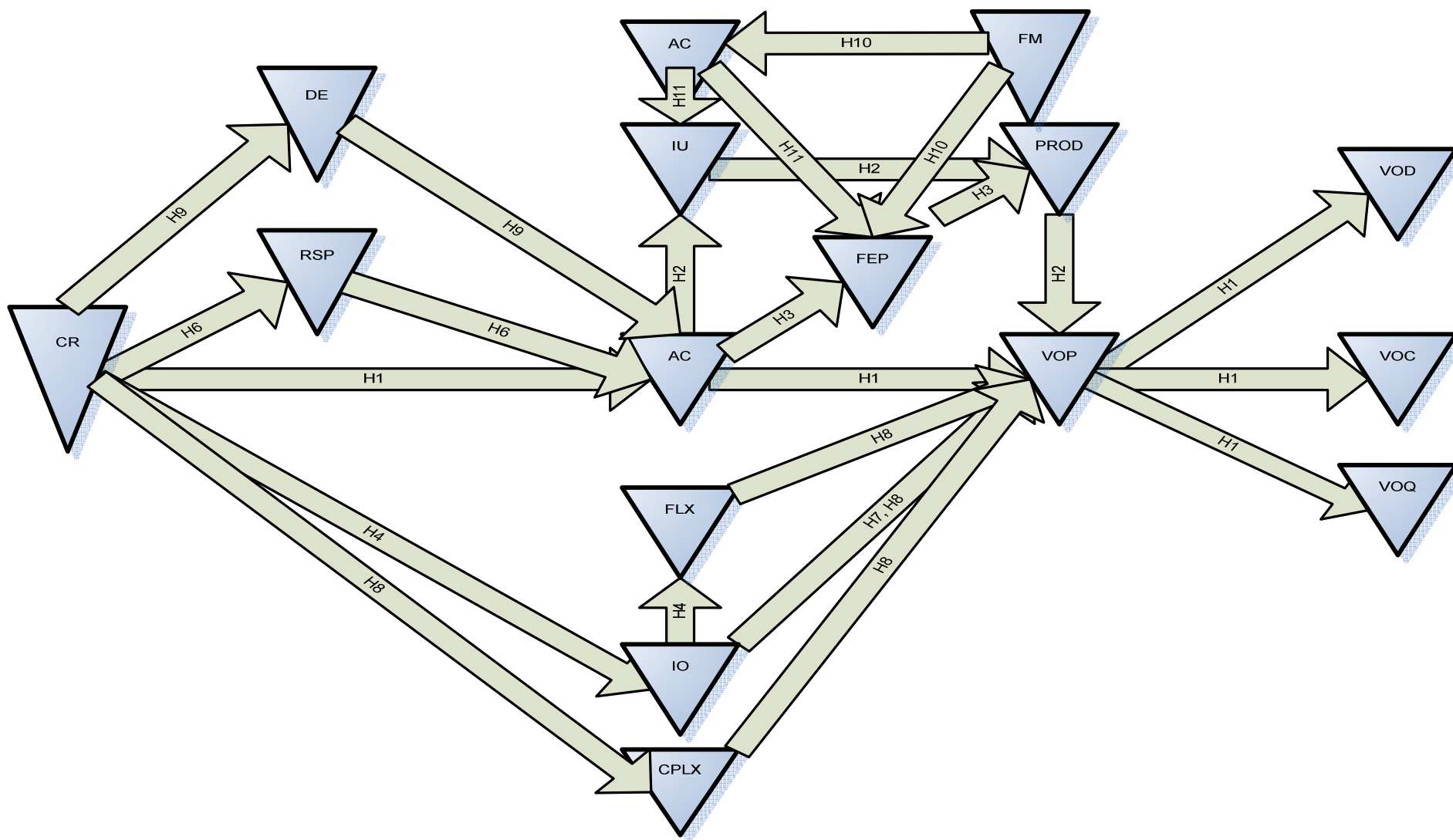


Figure 13 Relation d'influence entre les dimensions

2.2.5. Quelles variables définir sur les processus ?

Une fois que l'unité d'analyse a été fixée, au niveau processus, nous devons maintenant définir les variables du modèle. Les variables à définir doivent répondre à l'objectif de notre étude. En effet, selon Basili, et al. (1994), la quantification doit être définie suivant une approche « *top-down* », c'est-à-dire être centrée sur les objectifs. Ils estiment qu'une approche « *bottom-up* » est inadéquate car elle ne part pas d'un objectif approprié pour définir le contexte (Basili, et al., 1994). Elle ne permet pas de dire quelles métriques utiliser, parmi la multitude de caractéristiques observables⁵¹, et comment les interpréter. C'est pour cela qu'ils ont mis au point le Goal-Question-Metric qui est un mécanisme pour définir des métriques à partir d'objectifs mesurables. Le modèle de quantification du GQM comprend trois niveaux.

Niveau conceptuel (Goal). Un objectif est défini pour un objet, pour une variété de raisons, en relation avec un environnement particulier, à partir de divers points de vues. Les objets de la quantification peuvent être des produits, des processus ou des ressources.

Niveau opérationnel (Question). Un ensemble de questions est utilisé pour caractériser la manière dont un objectif spécifique sera réalisé. Les questions tentent de caractériser l'objet de la quantification.

Niveau quantitatif (Metric). un ensemble de données est associé à chaque question afin d'y répondre de manière quantitative. Les données peuvent être :

- objectives : si elles ne dépendent que de l'objet qui est mesuré et non du point de vue à partir duquel elles sont prises (nombre de versions d'un document, nombre d'heures passées sur une tâche, taille d'un programme) ;
- subjectives : si elles dépendent à la fois de l'objet mesuré et du point de vue à partir duquel elles sont prises (ex : lisibilité d'un texte, niveau de satisfaction utilisateur).

⁵¹ Basili s'intéresse aux caractéristiques des logiciels, mais le constat est aussi valable pour les processus.

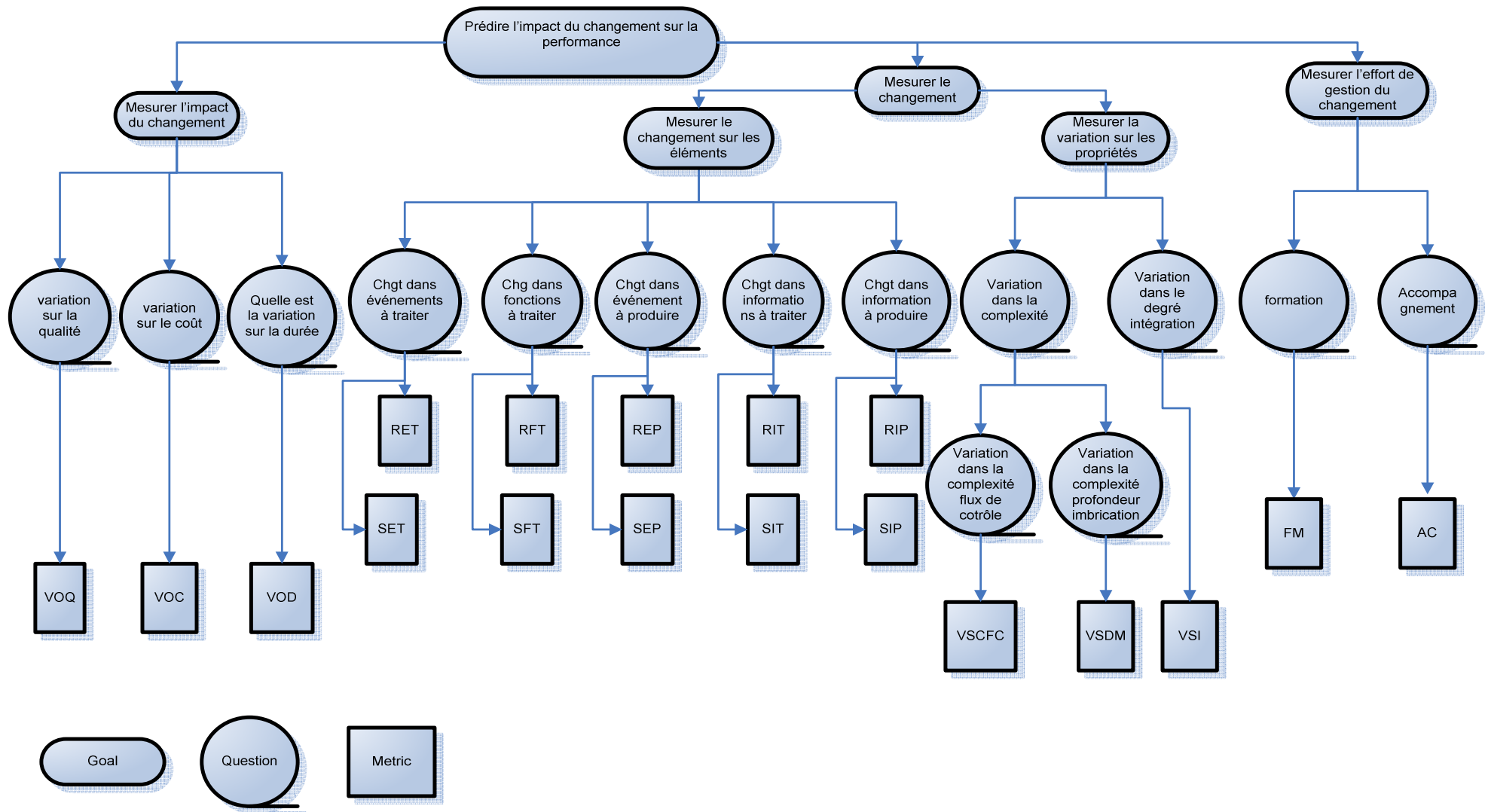


Figure 14 Définition « top-down » des métriques par la GQM

Nous proposons d'appliquer le GQM pour définir les métriques de notre modèle par rapport à l'objectif de l'étude (Figure 14). Nous partons de l'objectif principal de notre étude qui est la gestion proactive du changement. Cet objectif est composé de deux sous-objectifs :

- G prédire l'impact du changement sur la performance et
- G agir sur la base de ces prédictions.

Le sous-objectif de prédiction peut à son tour être décomposé en trois sous objectifs correspondant aux trois concepts principaux de l'étude. Ces trois sous-objectifs sont :

- G mesurer l'impact du changement,
- G mesurer son ampleur et
- G mesurer l'effort de gestion.

Le sous-objectif de « mesure de l'ampleur » du changement est composé de deux autres sous-objectifs :

- G mesurer le changement dans les éléments du processus et
- G mesurer la variation sur les propriétés du processus.

Au niveau opérationnel, les questions sont déterminées par rapports aux sous-objectifs en utilisant les dimensions de l'ampleur, de la gestion et de l'impact du changement définies à la section précédente. Pour chacun des sous-objectifs, des questions seront associées en utilisant les dimensions relatives au concept correspondant. Au niveau quantitatif, les métriques sont définies par rapport aux questions en s'inspirant des métriques processus qui existent dans la littérature. Des métriques opérationnelles de variation sont définies pour les questions associées au sous-objectif de mesure de l'impact opérationnel du changement. Ces métriques sont basées sur celles définies dans les approches de supervision des processus métier :

- Q Quelle est la variation dans le coût du processus ?
 - M métrique opérationnelle de variation dans le coût (MOVC)
- Q Quelle est la variation dans la durée ?
 - M métrique opérationnelle de variation dans la durée (MOVD)
- Q Quelle est la variation dans la qualité ?
 - M métrique opérationnelle de variation dans la non-qualité (MOVNQ)

Des métriques structurelles de changement sont définies pour les questions associées au sous-objectif de mesure du changement dans les éléments du processus. Les métriques structurelles de changement définies ici s'inspirent de « métriques de base » définies en analyse structurelle des processus :

- Q Quel est le changement dans les activités à exécuter ?
 - M métrique structurelle de changement dans les activités à exécuter supprimées (MSCAES)

- M métrique structurelle de changement dans les activités à exécuter rajoutées (MSCAER)
- Q Quel est le changement dans les informations à traiter ?
 - M métrique structurelle de changement dans les informations à traiter supprimées (MSCITS)
 - M métrique structurelle de changement dans les informations à traiter rajoutées (MSCITR)
- Q Quel est le changement dans les informations à produire ?
 - M métrique structurelle de changement dans les informations à produire supprimées (MSCIPS)
 - M métrique structurelle de changement dans les informations à produire rajoutées (MSCIPR)
- Q Quel est le changement dans les événements à surveiller ?
 - M métrique structurelle de changement dans les événements à surveiller supprimés (MSCESS)
 - M métrique structurelle de changement dans les événements à surveiller rajoutés (MSCESR)
- Q Quel est le changement dans les événements à générer ?
 - M métrique structurelle de changement dans les événements à générer supprimés (MSCEGS)
 - M métrique structurelle de changement dans les événements à générer rajoutés (MSCEGR)

Des métriques structurelles de variation sont définies pour les questions associées au sous-objectif de mesure de la variation sur les propriétés du processus. Les métriques structurelles de variation définies ici s'inspirent de « métriques dérivées » définies en analyse structurelle des processus :

- Q Quelle est la variation dans la complexité du flux de contrôle ?
 - M métrique structurelle de variation de la complexité du flux de contrôle (MSVCFC)
- Q Quelle est la variation dans la profondeur d'imbrication maximale ?
 - M métrique structurelle de variation de la profondeur d'imbrication maximum (MSVPIM)
- Q Quelle est la variation dans la simplicité ?
 - M métrique structurelle de variation de la simplicité (MSVS)

Des métriques « effort de gestion » seront définies et reliées aux questions correspondantes au sous-objectif de mesure de l'effort de gestion :

- Q Quel est l'effort de formation ?
 - M métrique effort de Formation (MEF)

Q Quel est l'effort d'accompagnement ?

M métrique effort d'accompagnement (MEA)

Dans la section suivante, chaque métrique est présentée plus en détails et les raisons de son choix exposées.

2.2.6. Conclusion

Ce chapitre correspond à l'étape de compréhension des données dans l'application du processus de datamining à notre problématique. Nous avons d'abord commencé par analyser l'ingénierie des processus en nous concentrant sur la modélisation et la supervision des processus. Nous nous sommes intéressés particulièrement à la notation BPMN que nous avons utilisée par la suite pour la définition de nos métriques explicatives. Nous avons montré ensuite l'application des actions de notre démarche à chaque étape du cycle de vie du processus. Les relations entre les dimensions de l'ampleur, de l'effort de gestion et de l'impact du changement ont été présentées. L'approche de GQM a été employée pour la définition des variables du problème. Elle part de l'objectif de gestion proactive. Des sous-objectifs relatifs aux mesures de l'ampleur, l'effort de gestion et l'impact du changement, sont définis. Des questions associées à ces sous-objectifs sont ensuite formulées en utilisant les dimensions correspondantes. Des métriques sont définies pour répondre à chacune de ces questions.

Chapitre 2.3 Définition des variables à expliquer

2.3.1. Introduction

La définition des variables à expliquer représente le sous-objectif de quantification de l'impact du changement dans l'application de la méthode de GQM (Figure 14) qui structure la définition des variables de notre problème. Le changement est introduit à travers la restructuration des processus et son objectif principal est l'amélioration de la performance en termes de délai, coût et qualité. Ces trois indicateurs sont les dimensions de l'impact du changement à travers lesquelles nous voulons mesurer la variation de la performance. Cette variation fait référence à l'amélioration ou la détérioration de ces critères lorsque l'on passe du processus « en-cours » au processus « implémenté ». Le sous-objectif de quantification de l'impact du changement est relié aux questions suivantes :

- Quelle est la variation dans la qualité ?
- Quelle est la variation dans la durée ?
- Quelle est la variation dans le coût du processus ?

La définition des variables à expliquer revient donc à répondre de manière quantitative à ces questions. Dans le cadre du contrôle opérationnel des processus, les performances actuelles du processus sont enregistrées avec la collection des données opérationnelles. Elles sont ensuite comparées à des valeurs cibles pour identifier les écarts à corriger. Pour traiter le problème de l'impact du changement au niveau opérationnel, nous ne pouvons pas seulement comparer les métriques opérationnelles actuelles aux valeurs cibles. Nous devons nous placer dans l'après restructuration et prendre comme référence les valeurs correspondantes d'avant la restructuration. C'est pour cela que nous parlons de variation, car c'est une différence entre deux valeurs ou positions prises par une même donnée. Le problème qui nous intéresse ici est la baisse de la performance des processus représentée par des métriques opérationnelles de variation. Les métriques opérationnelles de variation des processus concernent donc le délai, le coût et la non-qualité. Ces métriques sont calculées en utilisant la valeur de la métrique opérationnelle du processus « implémenté » à laquelle on soustrait la valeur de la métrique correspondante pour le processus « en-cours » (Camara, et al., 2008). Une baisse de performance correspondrait donc à une métrique de variation avec une valeur positive.

2.3.2. Métriques opérationnelles

De manière générale, nous pensons que la mesure de l'impact du changement doit être présentée dans une donnée unique qui rend compte de la variation pour chaque critère de performance. Dans cette section nous allons nous intéresser à toutes les approches qui participent à la détermination de la performance des processus : supervision des processus, simulation, analyse structurelle des processus. Dans la section 1.4.7, nous avons présenté les

limites des métriques opérationnelles définies en simulation, en supervision et en analyse structurelle par rapport à la mesure de l'impact du changement. L'idée est d'utiliser les indicateurs de performance ou métriques opérationnelles en termes de coût, délai et qualité proposés par ces approches, dans la mesure de l'impact du changement.

Au niveau opérationnel, il est important de prendre en considération la notion d'instance et de chemin de processus. L'instance du processus, encore appelée occurrence, correspond à une exécution du processus. L'instance du processus peut être confondue avec son résultat qui peut être physique ou non. Les chemins du processus sont déterminés par les points de décision dans le modèle du processus. Dans le modèle, chaque point de décision est relié à une variable aléatoire qui détermine les probabilités liées aux arcs de sortie du point. La probabilité d'un chemin est calculée à partir des probabilités de ses arcs dans les nœuds de décision par lesquels il passe. On parle cependant de probabilité, que si on s'intéresse à la simulation. Pour la supervision, ce sont des proportions qui sont observées sur les arcs des points de décision et agrégées au niveau de chaque chemin. Pour la simulation, les probabilités des arcs peuvent être fixées en utilisant des proportions issues de supervisions antérieures ou sur l'avis d'un expert. Les indicateurs de performance issus de la supervision ou de la simulation sur les processus dépendent des chemins et de leurs probabilités. Chaque instance du processus correspond à un chemin, celui à travers lequel le processus s'est exécuté.

Le groupe Gartner définit le terme BAM en se référant à un produit logiciel qui a pour but d'accéder en temps-réel aux indicateurs de performance critiques pour accroître la rapidité et l'efficacité des opérations métiers (Gartner Group, 2002). Les outils de pilotage ou de supervision de la performance opérationnelle (Business Activity Monitoring) connaissent un véritable essor. Une des raisons de cet essor est que ces outils, en contrôlant en permanence la performance, permettent de quantifier les gains de productivité et par conséquent le retour sur investissement. Ils ont une fonction d'alerte car ils évitent les dérives synonymes de pénalités ou préviennent, en temps réel, des dysfonctionnements de toute nature. Le BAM concerne les transactions i.e. exécution actuelle d'une instance du processus métier, et se concentre sur des attributs clés tels que les volumes (nombre de transactions, coût, nombre de ticket clôturés, nombre d'articles consommés, coût, ...) ou la vitesse (temps de cycle processus, temps d'attente entre événements, débit processus ...) (webmethods, 2006).

Les outils du BPM (Business Process Management) permettent aux organisations d'améliorer et d'automatiser leurs processus métiers par une exécution des processus et une gestion de flux d'activités. Selon Gartner, le BAM et le BPM sont très complémentaires et doivent être déployés ensemble pour résoudre les problèmes de niveau business. De plus en plus, les entreprises utilisent le BAM pour analyser leurs indicateurs de performance clés de délai, de coût et de qualité.

En analyse structurelle des processus, on parle aussi de métriques opérationnelles des processus. Les métriques opérationnelles traitent donc des propriétés dynamiques du

processus opérationnel. Les métriques opérationnelles sont celles qui mesurent les performances du processus à travers des occurrences d'exécution de ce dernier (Tjaden, 2001). Selon Tjaden (Tjaden, 2001) aussi bien les métriques opérationnelles que les métriques structurelles sont nécessaires pour un benchmark correct du processus métier. Les métriques opérationnelles, selon lui, permettent de déterminer s'il y a des problèmes de performance sur le processus ; et le cas échéant de vérifier si cela est dû à des déficiences sur la structure du processus (Tjaden, 2001). La définition et le calcul des métriques opérationnelles ne sont pas très formalisés dans les travaux d'analyse structurelle des processus qui s'intéressent plus aux métriques structurelles. Cependant, l'identification de la relation d'influence entre les métriques opérationnelles et les métriques structurelles des processus est un aspect très important. Voici quelques exemples de métriques opérationnelles de processus que nous pouvons trouver dans la littérature de l'analyse structurelle des processus :

- unités produits par jour, pourcentage réalisé correctement (Tjaden, 2001) ;
- temps moyen du processus (Aguilar, et al., 2006A) (Tjaden, 2001) ;
- coût du processus (Aguilar, et al., 2006A) (Scheer, 1998).

Nous avons vu que la simulation, l'analyse structurelle des processus et la supervision des processus définissent toutes des métriques opérationnelles. Dans notre étude, les métriques opérationnelles de variation des processus en cours de restructuration sont obtenues par prédiction. Pour un processus déjà implémenté, les métriques de variation opérationnelles sont calculées par comparaison des métriques opérationnelles obtenues par supervision de ses versions « en-cours » et « implémenté ». Nos métriques opérationnelles de variation seront définies à partir des métriques opérationnelles définies en supervision des processus que nous étudierons d'abord.

2.3.3. Quelle est la variation dans le délai ?

La métrique opérationnelle de variation dans le délai doit mesurer la variation dans la durée d'exécution d'un processus suite à sa restructuration. Cette variation peut être positive ou négative. Nous allons d'abord analyser la métrique opérationnelle de délai définie en supervision des processus (BAM) pour éventuellement l'adapter. En BAM, la durée écoulée d'un processus correspond à la durée d'un chemin de processus lorsqu'il démarre à une date et à une heure spécifique (IBM, 200X) :

- pour un chemin donné, la durée moyenne écoulée est donc : la somme de « heure de fin du processus – heure de début de processus » pour toutes les instances de processus correspondant à ce chemin du processus, divisée par, le nombre d'instances correspondant à ce chemin de processus (IBM, 200X).

Pour un chemin de processus, la distribution est : le nombre d'instances du processus correspondant au chemin divisé par, le nombre total des instances du processus. La durée

écoulée moyenne du processus est basée sur les valeurs calculées pour tous les chemins du processus ; chacun de ces chemins étant pondéré en fonction de sa distribution :

- la durée écoulée moyenne pondérée est : la somme de « durée moyenne du chemin multipliée par distribution du chemin de processus divisé par la distribution totale » pour tous les chemins de processus (IBM, 200X).

La collecte des informations relatives aux durées des processus peut se réaliser en utilisant plusieurs méthodes (Borysowich, 2008) :

- « **dates de début et de fin** » : cette méthode est utilisée quand les dates de début et de fin des processus sont documentées dans un système (un progiciel de BAM par exemple) ;
- « **test sur échantillon** » : cette méthode utilise un échantillon et mesure le temps nécessaire pour traiter, produire cet échantillon. Elle doit être utilisée pour les processus qui ont une fréquence de répétition allant de courte à moyenne (i.e., moins de deux semaines) ;
- « **données historiques** » : cette méthode utilise les documents existants (plan exploitation par exemple) pour déterminer les débuts et les fins des processus. Elle peut être utilisée pour les processus qui ne sont pas souvent exécutés (annuel ou trimestriel) ;
- « **estimation** » : cette méthode estime la durée totale du processus en le décomposant en tâches et en estimant la durée de chaque tâche. Elle doit être appliquée seulement si les trois méthodes précédentes ne sont pas applicables.

La métrique opérationnelle de variation dans le délai représente la variation susceptible de se produire sur la durée moyenne du processus après sa restructuration (Camara, et al., 2008). Nous la définissons comme la différence entre les métriques de durée moyenne pondérée (DMP) pour le processus « en-cours » et celui « en-devenir » (ou « implémenté »). Ainsi, la métrique opérationnelle de variation dans le délai est nommée MOVD et a la formule suivante :

Équation 5 Métrique opérationnelle de variation dans le délai

$$\text{MOVD} = \frac{\text{DMP}_{\text{to-be}} - \text{DMP}_{\text{as-is}}}{\text{DMP}_{\text{as-is}}} \times 100$$

Cette métrique est exprimée sous forme de pourcentage par rapport à la DMP du processus « en-cours » pour assurer sa normalisation (Camara, et al., 2008). En effet, l'utilisation d'une valeur relative⁵² pour la variation opérationnelle dans le délai permet de représenter cette variation de la même manière, quelque soit la valeur du DMP du processus d'origine (en-cours). Supposons par exemple, qu'on ait deux processus P1 et P2 dont les DMP pour les versions d'origine sont 20 mn et 100 mn et qui ont respectivement comme

⁵² Exprimée en % d'une autre valeur

DMP des versions cibles 25 mn et 110 mn. Si la MOVD est calculée en valeur absolue⁵³ nous avons +5 pour P1 et +10 pour P2, ce qui revient à dire que la variation opérationnelle dans le délai est plus importante pour P2 comparée à celle P1. Si cette même métrique de variation est exprimée en valeur relative nous avons le contraire, car P1 avec +25% a une variation supérieure à celle de P2 qui n'est que de 10% et cela exprime mieux la réalité, car l'importance de toute variation est relative à la position de départ.

Pour illustrer l'extraction des métriques opérationnelles de variation nous avons repris les versions « en-cours » et « implémenté » du processus de traitement de commandes pour lesquels nous avons généré automatiquement des données d'exécutions. Ces données portent sur sept instances d'exécution de chacune des deux versions du processus. Avant de calculer la métrique opérationnelle de variation dans le délai sur notre processus, nous allons d'abord extraire les métriques opérationnelles DMP sur ses deux versions. Dans l'extraction des métriques opérationnelles, il faut d'abord considérer les différents chemins du processus. Chaque exécution du processus est une instance et donne lieu à un enregistrement des données d'exécution. Une instance correspond à un et un seul chemin parmi ceux disponibles sur le processus. Les données d'exécution correspondant à un même chemin sont regroupées.

Pour l'extraction de la métrique opérationnelle de variation dans la durée nous utiliserons les données relatives aux temps d'exécution des activités. Nous commençons par regrouper les instances correspondant au chemin 1 de la version « en-cours ». Nous calculons la durée écoulée de chaque instance en faisant la somme des durées d'exécution des activités. La durée moyenne écoulée (DME Tableau 9) pour le chemin 1 est obtenue par la moyenne des durées écoulées pour les 7 instances correspondant à ce chemin. La durée moyenne pondérée (DMP Tableau 9) de la version « en-cours » du processus est égale à la DME du chemin 1, car il ne contient pas d'autres chemins. La DMP est calculée pour la version « en-devenir » en suivant le même principe. La métrique opérationnelle de variation dans le délai du processus (MOVD Tableau 9) est obtenue en utilisant sa formule (Équation 5) avec les DMP de ses deux versions.

Tableau 9 Variation opérationnelle durée processus traitement commandes

⁵³ Exprimée sans %

	chemi 1 "en-cours"					chemin 1 "implemeté"		
	traitement commande	planification des ordres	fabrication	durée écoulée		traitement commande	fabrication	durée écoulée
instance 1	3	14	38	55		7	36	43
instance 2	7	12	87	106		8	81	89
instance 3	5	9	83	97		4	72	76
instance 4	5	12	100	117		3	17	20
instance 5	5	5	18	28		8	103	111
instance 6	3	17	110	130		5	75	80
instance 7	4	5	60	69		2	122	124
DME	86					77,57142857		
	"en-cours"					"implémeté"		
DMP	86					77,57142857		
	processus traitement de commades							
MOVD	-9,800664452							

2.3.4. Quelle est la variation dans le coût ?

Comme pour le délai, nous allons dériver une métrique de variation opérationnelle dans le coût par adaptation de la métrique opérationnelle de coût définie de supervision des processus. Pour chaque chemin, le BAM calcule le coût de traitement moyen, le coût d'inactivité moyen, le coût moyen des ressources allouées et le coût total moyen qui est la somme des trois premiers.

- Le coût de traitement moyen est : la somme de « coût de traitement total » (pour toutes les instances de processus correspondant à ce chemin de processus), divisée par, le nombre d'instances correspondant à ce chemin de processus (IBM, 200X).
- Le coût d'inactivité moyen est la somme de « coût d'inactivité total » pour toutes les instances de processus correspondant à ce chemin de processus, divisée par, le nombre d'instances correspondant à ce chemin de processus (IBM, 200X).
- Le coût moyen des ressources allouées est la somme de « coût total des ressources allouées » pour toutes les instances de processus correspondant à ce chemin de processus, divisée par, le nombre d'instances correspondant à ce chemin de processus (IBM, 200X).
- Le coût total moyen est la somme des trois coûts (le coût de traitement moyen + le coût d'inactivité moyen + le coût moyen des ressources allouées).

Une valeur de coût récapitulative est calculée pour le processus : il tient compte de tous les chemins toujours pondérés en fonction de leurs distributions. Le coût total pondéré est : la somme de « coût total moyen du chemin multiplié par la distribution du chemin de processus, le tout divisé par, la distribution totale » pour tous les chemins du processus (IBM, 200X).

Nous venons de présenter une méthode de calcul des coûts de processus utilisée en BAM. Cela a pour intérêt de démontrer qu'il existe des outils qui permettent de déterminer les

coûts des processus. Cependant, pour avoir plus de détails sur la partie liée au calcul des coûts de traitement du coût des processus défini par le BAM, nous allons nous intéresser au calcul prévisionnel des coûts d'un processus ou *Process Costing*. Le principe essentiel du Process Costing est le fait que les coûts sont générés seulement quand les ressources sont utilisées dans le processus pour produire un bien ou un service (Misof, 2007). Le Process Costing garde en permanence la visibilité sur les charges de travail, en utilisant le temps de travail comme unique moyen d'imputation des coûts (ProcessSoft, 200X). En considérant le travail comme seul facteur de production, le Process Costing est particulièrement adapté aux activités administratives et de gestion, de la banque, de l'assurance mais aussi de l'industrie (ProcessSoft, 200X).

Pour le calcul prévisionnel des coûts d'un processus, la description du processus doit d'abord être élaborée (IDS, 2002). Pour chaque centre de coût, sont définis des processus élémentaires typiques qui font l'objet d'une consolidation transversale aux centres de coûts en type de processus. La notion de « processus élémentaire » (sous-processus) indique qu'il constitue le composant d'un processus métier. Lorsqu'il n'est plus subdivisé, il correspond à la notion de « fonction ». Dans l'exemple illustré à la Figure 15, le processus principal « Achats » se compose de quatre processus élémentaires : commande, mise en demeure, vérification de la facture et génération du paiement. A présent, il est possible de déterminer des taux de coût pour chaque sous-processus, c'est-à-dire les coûts générés par chaque commande typique, chaque mise en demeure typique etc. Pour chaque sous-processus, on détermine aussi les valeurs d'usage qui sont ici le nombre de commandes, le nombre de mise en demeure, le nombre de positions de la facture et le nombre de factures. Les coûts par sous-processus sont déterminés par la multiplication des taux de coûts des sous-processus et leurs valeurs d'usage (IDS, 2002). Leur addition permet d'obtenir les coûts de chaque processus principal, c'est-à-dire dans notre exemple 96 UM⁵⁴ pour un processus d'achat typique (Figure 15).

⁵⁴ Unité Monétaire

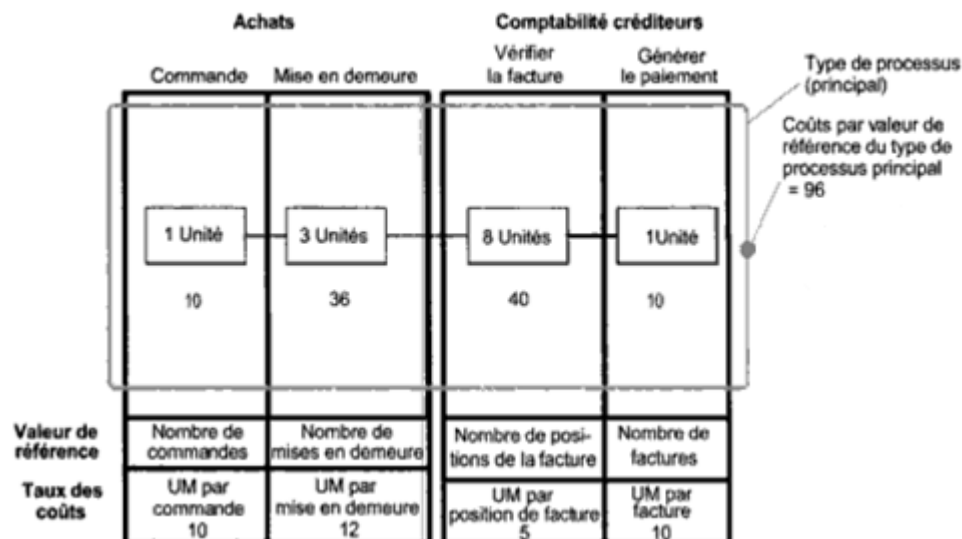


Figure 15 Coût processus

Dans l'exemple précédent, les valeurs de référence des sous-processus sont définies en termes de volume. Les valeurs de référence⁵⁵ les plus fréquentes sont le volume et le temps (Borysowich, 2007). La valeur de référence de type volume est basée sur les unités de travail, comme de nombre d'ordres, ou de commandes (Borysowich, 2007). La valeur de référence de type temps est basée sur la durée nécessaire pour réaliser l'activité. Le coût d'une activité augmente donc par rapport au nombre d'unités traitées ou avec la durée nécessaire pour achever l'activité. Dans cet exemple, le processus étant linéaire, il n'a pas été utile de prendre en compte la manière dont les activités du processus sont séquencées. Un facteur supplémentaire appelé « driver de transaction », directement lié à la structure du processus, a été défini dans Misof (2007). Si, par exemple, une fonction ou un sous-processus est sur un chemin du processus qui a une probabilité totale d'exécution de 15%, par rapport au modèle du processus, une valeur de 0,15 est assignée à la fonction comme « driver de transaction ». Lors du calcul du coût du processus principal, les coûts des sous-processus sont pondérés en utilisant ce « driver de transaction », avant d'être additionnés (Misof, 2007). De même, si un sous-processus a un coût non lié aux ressources, le taux de coût du sous processus est directement multiplié par le « driver de transaction » (Misof, 2007). Dans ce cas, les valeurs de référence de type temps ou volume ne sont pas utilisées.

La métrique de variation opérationnelle dans le coût représente l'augmentation ou la diminution du coût du processus lors de sa restructuration (Camara, et al., 2008). Nous devons d'abord calculer la métrique opérationnelle de coût total pondéré (CTP) pour les processus « en-devenir » et « implémenté », comme présentée ci-dessus. La métrique opérationnelle de variation dans le coût pour un processus métier est exprimée en pourcentage, pour les raisons de normalisation évoquées 2.3.3 :

⁵⁵ Cost drivers

Équation 6 Métrique opérationnelle de variation dans le coût

$$MOVC = \frac{CTP_{to-be} - CTP_{as-is}}{CTP_{as-is}} \times 100$$

Afin d'appliquer le calcul de la métrique opérationnelle de variation dans le coût (MOVC) sur le processus de traitement de commandes, nous avons ajouté aux données sur les durées d'exécution (Tableau 9), des valeurs de référence⁵⁶. Le cout total moyen (CTM Tableau 10) du chemin 1 de la version « en-cours » est obtenu en faisant la somme des coûts des fonctions. Le coût de chaque fonction est calculé en multipliant son temps moyen dans le chemin 1 par sa valeur de référence. Le temps moyen est la moyenne des durées des fonctions sur les 7 instances associées au chemin 1. Comme la version « en-cours » ne contient que le chemin 1, son coût total pondéré (CTP) est égal au CTM du chemin 1. Le CTP est ensuite extrait pour la version « en-devenir » en suivant la même démarche. Pour obtenir la métrique opérationnelle de variation dans le coût, la formule de calcul MOVC (Équation 6) est appliquée, en utilisant comme paramètres les CTP des deux versions.

Tableau 10 Variation opérationnelle coût processus traitement de commandes

	chemin 1 "en-cours"				chemin 1 "implémenté"	
	traitement commande	planification des ordres	fabrication		traitement commande	fabrication
Temps moyen	4,57142857	10,57142	70,8571429		5,2857142	72,2857143
Valeur de référence	8	15	65		8	65
Coût fonction	36,571428	158,5714	4605,71429		42,285714	4698,57143
CTM	4800,857114				4740,857143	
	"en-cours"				"implémenté"	
CTP	4800,857114				4740,857143	
	processus traitement de commande					
VOC	-1,249776229					

2.3.5. Quelle est la variation dans la qualité ?

La métrique opérationnelle de variation dans la non-qualité est principalement basée sur la métrique opérationnelle de succès des processus définie en BAM. En effet, le BAM calcule des métriques pour déterminer le succès ou l'échec général du processus ou de ses activités (IBM, 200X). Le succès est un indicateur signalant, si les instances du processus ont abouti ou échoué. Le BAM définit la métrique opérationnelle pourcentage de succès (PS),

⁵⁶ Attribuées arbitrairement aux fonctions

comme le nombre d'instances réussies multiplié par cent et divisé par le nombre total d'instances. Un échec se produit lorsqu'une ou plusieurs activités du processus échouent pour une des raisons suivantes (IBM, 200X) :

- une ou plusieurs ressources nécessaires ne sont pas disponibles,
- l'activité nécessite une ressource inconnue,
- les ressources disponibles sont insuffisantes pour répondre aux besoins de l'activité,
- la période d'indisponibilité d'une ressource est au-delà du délai d'attente maximum de l'activité.

Les métriques de variation précédemment définies sont basées sur des métriques opérationnelles (délai, coût) qui varient dans le sens de la non-performance. La métrique opérationnelle PS varie dans le sens inverse, celui de la performance. Pour assurer une homogénéité dans la définition de nos variables de réponse, nous devons d'abord créer une métrique opérationnelle de pourcentage de non-succès (PNS) qui varie dans le sens de la non-performance comme les autres :

$$PNS = 100 - PS$$

Nous en déduisons ensuite notre métrique opérationnelle de variation de la non-qualité sans avoir à ré-exprimer cette dernière en pourcentage car elle l'est déjà. En effet, elle est calculée à partir de la métrique opérationnelle PNS qui est elle-même en pourcentage. La métrique opérationnelle de variation dans la non-qualité des processus se calcule comme suit :

Équation 7 Métrique opérationnelle de variation de la non-qualité

$$MOVNQ = PNS_{to-be} - PNS_{as-is}$$

Pour appliquer la métrique opérationnelle de variation dans la non-qualité sur le processus de traitement de commandes, nous avons affecté, arbitrairement, des états finaux d'exécution pour les 7 instances de ses deux versions (Tableau 11). Le pourcentage de non-succès (Tableau 11) de la version « en-cours » est égal à celui du seul chemin qu'elle contient (pourcentage d'échecs pour chemin 1). De la même manière, nous obtenons le PNS pour la version « implémenté » (Tableau 11). La valeur de la métrique opérationnelle de variation dans la non-qualité est obtenue en utilisant la formule correspondante (Équation 7).

Tableau 11 Variation opérationnelle non-qualité processus traitement commandes

	chemin 1 "en-cours"		chemin 1 "en-devenir"
	état final		état final
instance 1	Succes		succes
instance 2	ehec		succes
instance 3	ehec		succes
instance 4	Succes		succes
instance 5	ehec		succes
instance 6	Succes		ehec
instance 7	Succes		ehec
PNS	42,85714286		28,57142857
	"en-cours"		"en-devenir"
PNS	42,85714286		28,57142857
	processus de traitement de commande		
VONQ	-14,28571429		

2.3.6. Conclusion

Ce chapitre fait partie de la phase de préparation des données dans l'application du processus de datamining à notre problématique. L'objectif principal de ce chapitre est la définition des variables à expliquer. Cet objectif est réalisé à travers la quantification de l'impact du changement. Dans le chapitre précédent, nous avons présenté, à travers le GQM, des sous-objectifs, des questions et des métriques relatives à cet impact. Ces différentes métriques ont été présentées, en détail, dans ce chapitre. Ces métriques définies pour répondre aux questions relatives à la mesure de l'impact du changement sont inspirées des métriques opérationnelles, définies en supervision des processus. Les variables de réponse définies sur l'unité d'analyse processus métier sont au nombre de trois (Tableau 12). Ce sont les métriques opérationnelles de variation dans le délai, le coût et la non-qualité. Leurs valeurs ont été extraites pour le processus de traitement de commandes pris en exemple (Tableau 12).

Tableau 12 Synthèse des variables à expliquer

Catégorie	Sous-catégorie	Désignation	Abréviation	VPTC
Variables à expliquer	Impact du changement	Métrique opérationnelle de variation dans le coût	MOVC	-9,8
		Métrique opérationnelle de variation dans le délai	MOVD	-1,24
		Métrique opérationnelle de variation dans la non-qualité	MOVNQ	-14,28

Chapitre 2.4 Définition des variables explicatives

2.4.1. Introduction

Les variables explicatives sont celles relatives à l'ampleur ou à l'effort de gestion du changement. La définition des variables liées à l'ampleur du changement se divise en deux parties : définition des métriques structurelles de changement et celles de variation. Les métriques structurelles de changement répondent aux questions associées au sous-objectif GQM de mesure du changement dans les éléments du processus. Les métriques structurelles de variation répondent aux questions associées au sous-objectif GQM de mesure de la variation dans les propriétés du processus. Nous avons enfin des métriques d'effort de gestion du changement. Chaque variable explicative définie est accompagnée de trois hypothèses. Ces dernières représentent la relation d'influence entre la variable explicative et les variables de réponse. L'objectif de l'apprentissage du modèle prédictif (Chapitre 3.3) est de valider ou de rejeter ces hypothèses.

Nous présenterons d'abord les métriques structurelles définies en Analyse Structurelle des Processus. En effet, chaque métrique structurelle de changement a une métrique de base correspondante en ASP. Chaque métrique structurelle de variation a une métrique dérivée correspondante. Nous présenterons aussi le procédé de calcul des métriques de changement, utilisé en génie logiciel, dont nous nous sommes inspirés dans la définition de nos métriques structurelles de changement. La définition des variables explicatives est appliquée au processus de traitement de commandes pour l'exemple. Les métriques que l'on va définir dans cette section doivent respecter un certain nombre de propriétés identifiées dans la littérature, comme nécessaires pour les processus opérationnels ; telles que :

- **la simplicité** : les métriques doivent être compréhensibles par les utilisateurs finaux (Cardoso, 2005) (Gruhn, et al., 2005),
- **la calculabilité** (Cardoso, 2005),
- **l'automatisation** : il doit être possible d'automatiser la mesure des métriques (Gruhn, et al., 2005) ;
- **la consistance** (Cardoso, 2005) : nécessaire et pertinent
- **l'objectivité** (Cardoso, 2005) ;
- **l'additivité** : si deux structures indépendantes sont mises en séquence, la complexité totale de la structure combinée est au moins égale à la somme des complexités des structures indépendantes (Cardoso, 2005) ;
- **l'interopérabilité** : il existe un grand nombre de langages de spécification de processus provenant, soit de l'université ou de l'industrie. Les métriques doivent être indépendantes du langage de spécification et ainsi avoir la même valeur quel

que soit le langage avec lequel le processus a été modélisé (Cardoso, 2005) (Gruhn, et al., 2005).

Dans cette étude, nous avons fait le choix de ne pas respecter, pour le moment, la propriété d'interopérabilité des métriques. Toute la définition des métriques est basée sur la notation BPMN qui, en plus d'être assez développée, est sensée être une notation standard.

2.4.2. L'analyse structurelle des processus (ASP)

L'analyse structurelle des processus définit deux types de métriques : les métriques opérationnelles et les métriques structurelles (Aguilar, et al., 2006A) (Tjaden, 2001). Les métriques structurelles sont une quantification de la complexité structurelle des processus (Tjaden, 2001). Elles sont donc relatives aux propriétés statiques du processus. Les propriétés statiques ou structurelles influencent fortement les performances du processus. Les métriques structurelles des processus sont, pour la plupart, des adaptations de mesures, définies dans le domaine du génie logiciel (Gruhn, et al., 2005). Les métriques structurelles sont regroupées en deux groupes principaux : les mesures de base et les mesures dérivées (Aguilar, et al., 2006B) (Aguilar, et al., 2006A).

Les mesures de base sont calculées en comptant les différents éléments qui composent le modèle du processus (Aguilar, et al., 2006B). Les mesures dérivées sont des métriques un peu plus complexes, qui résultent de l'agrégation de plusieurs métriques de base (Aguilar, et al., 2006A). Certaines des métriques de base concernent la catégorie objet de flux de BPMN⁵⁷ (Tableau 13) (Aguilar, et al., 2006A) (Cardoso, 2005). La taille d'un modèle de processus peut être mesurée par des mesures, simples similaires au nombre de lignes de code en génie logiciel (Vanderfeesten, et al., 2007). Plusieurs auteurs proposent de compter le nombre d'activités pour établir cette mesure (Vanderfeesten, et al., 2007) (Gruhn, et al., 2005) (Aguilar, et al., 2006A).

Tableau 13 Métriques de base : objets de flux

Activités	TNA	Nombre total d'activités dans le processus
Événements	TNE	Nombre total d'événements dans le processus
Gateway	TNG	Nombre total de Gateway dans le processus

Dans Aguilar, et al. (2006A), d'autres métriques relatives aux objets de connexion (Tableau 14), aux zones de responsabilité (Tableau 15), aux artefacts (Tableau 16), sont également définies.

⁵⁷ activité, événement et porte logique (éléments utilisés pour contrôler les divergences et les convergences dans la séquence du flux)

Tableau 14 Métriques de base : objets de connexion

Flux de message	NSFM	Nombre de flux de message entre les participants du processus
Flux de séquence	NSFA	Nombre de flux de séquence entre activités dans le processus
	NSFE	Nombre de flux de séquence qui venant des événements dans le processus
	NSFG	Nombre de flux de séquence venant des Gateway dans le processus

Tableau 15 Métriques de base : Swimlane

Pool	NP	Nombre de pool dans le processus
lane	NL	Nombre de Lane dans le processus

Tableau 16 Métriques de base : Artéfacts

Object donnée (entrée)	NDOIn	Nombre d'objets de donnée-entrée dans de processus
Object donnée (sortie)	NDOOut	Nombre d'objets de donnée-sortie dans le processus
Object donnée	NDO	Nombre d'objets de donnée dans le processus

Ces métriques n'ont pas été conçues pour mesurer le changement. C'est pour cela que leur définition ne prend en compte qu'un seul modèle de processus. L'évolution du processus n'est pas directement prise en compte dans ces métriques. Nous devons créer de nouvelles métriques sur la structure des processus en tenant compte de l'ampleur du changement. Nous nous intéressons dans la section suivante au procédé de calcul des métriques de changement en génie logiciel.

2.4.3. Procédé de calcul des métriques de changement en GL

Pour Demeyer (2000), les modifications ou changements sont des facteurs critiques pour la compréhension de l'évolution d'un logiciel. Gokhale (1997), considèrent, quant à eux, qu'il y'a une correspondance directe⁵⁸, entre le nombre de changements réalisés sur le code d'un module logiciel et le nombre de défauts que le module contient. Des métriques de changement sont définies dans Gokhale, et al. (1997), pour mesurer ce changement à partir des CRs (Change Reports) où sont notés tous les changements réalisés sur le code des modules. De même, des métriques de changement sont définies dans Demeyer, et al. (2000), sur des logiciels orientés objets. La démarche suivie pour la définition des métriques dans cette étude est le *Goal-Question-Metric Paradigm* (Basili, et al., 1994). L'objectif de cette étude est d'identifier l'évolution dans la conception d'un logiciel orienté objet et d'en déduire les dérives de conception correspondantes (Demeyer, et al., 2000). L'idée est d'expliquer comment et pourquoi l'implémentation s'est éloignée, a dévié, par rapport à sa conception d'origine. La question générée à partir de cet objectif est la suivante : « quand on compare une version d'un logiciel à une autre, quelles caractéristiques peuvent être utilisées comme

⁵⁸ « one-to-one correspondence »

symptômes de *refactorings* ?» (Demeyer, et al., 2000). En développement orienté objet, les changements sont documentés à travers ce qui est appelé *refactorings*. Les opérations de *refactorings* réorganisent une hiérarchie de classes en modifiant les relations entre les classes et en redistribuant les variables d'instance et les méthodes (Demeyer, et al., 2000). Demeyer commence par définir des métriques de complexité logicielle (Tableau 17) qu'il range en trois catégories : celles relatives aux méthodes, celles relatives aux classes et celles relatives à l'héritage (Demeyer, et al., 2000).

Tableau 17 Métriques de complexité logiciel

Abréviation	Référence	Description
Méthode Taille (calculé pour chaque méthode)		
Mthd-MSG	[Lore94a]	Nombre de message envoyés dans le corps de la méthode.
Mthd-NSA	[Lore94a]	Nombre de déclarations dans le corps de la méthode.
Mthd-LOC	[Lore94a]	Les lignes de code dans le corps de la méthode.
La taille des classes (calculé pour chaque catégorie)		
NOM	[Chid94a]	Nombre de méthodes dans la classe.
NIV, PCI	[Lore94a]	Nombre de variables d'instance, le nombre de variables de classe définies par classe.
Héritage		
HNL (DIT)	[Lore94a], [Chid94a]	Niveau de profondeur hiérarchique [Lore94a] ou la profondeur de l'arbre d'héritage [Chid94a]. Nombre de classes dans la chaîne de classes de la superclasse ; en cas de l'héritage multiple, nombre de classes dans la plus longue chaîne.
CNP	[Chid94a]	Nombre d'enfants immédiats de la classe.
NMI	[Lore94a]	Nombre de méthodes héritées.
NMO	[Lore94a]	Nombre de méthodes surchargées.

Des métriques de changement plus globales sont ensuite définies pour le système logiciel en utilisant les métriques ci-dessus (Demeyer, et al., 2000). Ces métriques montrent les changements relatifs aux classes et aux méthodes du logiciel (Tableau 18) (Demeyer, et al., 2000). La première ligne du tableau 18, par exemple, représente les métriques de changement relatives aux classes et aux méthodes pour le logiciel *Visualworks* lors du passage de la version 1.0 à la version 2.0. Notre idée est d'appliquer ce procédé de calcul aux métriques de l'analyse structurelle des processus, pour mesurer le changement.

Tableau 18 Métriques de changement logiciel

(Ini = # initial version; Rmv = # removed; Add = # added; Ret = #retained (same name); Fnl = # final version)

	# classes					# methods				
	Ini	Rmv	Add	Ret	Fnl	Ini	Rmv	Add	Ret	Fnl
<i>Visualworks</i>										
1.0 -> 2.0	593	71	212	522	734	5282	1067	3569	4215	7784
2.0 -> 2.5	734	38	63	696	759	7784	337	858	7447	8305
2.5 -> 3.0	759	8	214	751	965	8305	207	1287	8098	9386
<i>HotDraw</i>										
1.0 -> 2.5	197	2	4	195	199	1425	153	72	1272	1344
2.5 -> 3.0	199	151	67	48	115	1344	1221	707	123	830
<i>Refactoring Browser</i>										
2.0 -> 2.1	277	2	52	275	327	2121	273	786	1848	2634

2.4.4. Définition des métriques structurelles de changement

La définition des métriques structurelles de changement a pour objectif de quantifier le changement dans les éléments du processus lors de la restructuration. Ces métriques correspondent aux questions associées au sous-objectif GQM de mesure du changement dans les éléments du processus. Ce sous-objectif est inclus dans celui de mesure du changement, qui lui-même est inclus dans celui de prédiction de l'impact du changement. Il s'agit d'analyser le changement dans éléments de la notation BPMN appartenant aux catégories objets de flux, objet de connexion, artefacts ou zone de responsabilité (*swimlane*). Les éléments de catégories objets de flux, artefacts ou zone de responsabilité peuvent être considérés comme les nœuds des modèles BPMN. Les objets de connexion (flux message, flux séquence et association) sont les arcs qui relient ces nœuds. Pour un modèle donné, tout changement sur les nœuds se répercute sur les arcs et vice-versa. Nous supposons donc qu'il serait redondant de définir des métriques de changement sur les nœuds et sur les arcs à la fois.

Nous pensons qu'il est plus pertinent de définir les métriques à partir des arcs du modèle. La raison de ce choix est que les arcs contiennent des informations plus complètes sur le changement que les nœuds. Prenons l'exemple du nœud « commande client » de type objet de donnée de la catégorie artefacts du processus de traitement de commandes. Dans le cas où le changement est mesuré sur les nœuds, le nœud « commande client » ne serait comptabilisé, ni dans les objets de données supprimées, ni dans celles rajoutées, donc n'apparaîtrait pas dans le changement. Cela est dû à la présence de ce nœud dans les versions « en-cours » et « implémentée » du processus. Dans le cas où la mesure du changement est réalisée sur les arcs, l'association allant de l'objet de donnée « commande client » vers l'activité « fabrication » serait comptabilisée comme rajoutée. Cette association n'est, en effet, présente que dans la version « implémenté » du processus. Le changement relatif à l'objet de

donnée « commande client » serait ainsi pris en compte dans la mesure du changement sur le processus.

Le mécanisme de définition des métriques structurelles de changement est le même que celui utilisé pour la définition des métriques de changement en génie logiciel. Il s'agit de compter, comme pour les classes et les méthodes en génie logiciel, le nombre de changements sur les arcs BPMN. Comme nous venons de le dire aucune métrique n'est définie à partir des nœuds BPMN pour éviter la redondance. De plus, vu que les processus utilisés sont des processus privés, nous n'utiliserons pas d'arcs de type flux de message. En effet, les messages sont échangés entre pools (différentes entreprises) dans le cadre d'un processus collaboratif (public). Toutes nos métriques seront définies à partir d'arcs de type flux de séquence ou association à l'intérieur d'un processus privé. Les flux de séquence relient les objets de flux (événements, activités et portes logiques) entre eux. Les associations relient les artefacts (objet de données) aux objets de flux et au flux de séquence. Lors du passage d'une version d'un logiciel à une autre, les métriques de changement (en génie logiciel) sont dérivées à partir des métriques de base, en comptant pour chaque type d'éléments (Demeyer, et al., 2000) :

- le nombre d'éléments dans la version initiale,
- le nombre d'éléments dans la version finale,
- le nombre d'éléments supprimés,
- le nombre d'éléments rajoutés, et
- le nombre d'éléments maintenus.

En nous inspirant de cela, nous pouvons déterminer les métriques structurelles de changement à partir des flux de séquence ou des associations dans les modèles de processus. Ces métriques seront définies par comparaison, entre les processus «en-cours » et « en-devenir » (ou « implémentés»). Nous devons aussi chercher à avoir le moins de métriques possibles pour la clarté et la simplicité du modèle. Nous choisissons d'exprimer les rajouts et les suppressions en pourcentage, par rapport au nombre de relations dans le modèle « en-cours ». Cela nous permet, d'une part, d'avoir des métriques normalisées car ayant des valeurs comparables pour tous les processus, comme nous l'avons fait pour les métriques opérationnelles de variation. D'autre part, cela nous permet d'intégrer la taille du processus et éviter ainsi de créer les métriques relatives au nombre d'arcs « retenus », « dans le modèle initial » et « dans le modèle final ».

Pour suivre la logique d'application de l'approche GQM nous ne prendrons pas les métriques de l'ASP pour les adapter afin de répondre aux questions. Nous partirons des questions pour ensuite chercher la métrique ASP la plus proche que nous pourrions adapter, pour mesurer la dimension du changement correspondante et répondre ainsi à la question GQM. Chaque question est reliée à deux métriques structurelles (une pour les suppressions et une pour les rajouts). Les métriques structurelles de changement s'inspirent des métriques de

base de l'ASP, mais ne sont pas calculées à partir de ces dernières. En effet, pour extraire les métriques structurelles de changement, il faut repartir du modèle pour faire un nouveau décompte des éléments qui avaient permis de calculer cette métrique de base en tenant en compte cette fois-ci du changement.

Quel est le changement dans les activités à exécuter ? Nous devons définir une métrique structurelle de changement pour répondre à cette question. Cette question fait référence à la dimension responsabilité (RSP) du changement et est associée au sous-objectif de mesure du changement. Elle renvoie à l'affectation des activités aux différentes zones de responsabilité (*Lane*). En BPMN, cette relation « responsabilité » est représentée implicitement par le fait que l'activité est à l'intérieur de la zone de responsabilité. Il n'existe aucune métrique définie en ASP qui mesure cette affectation. Nous définissons ainsi la métrique structurelle de changement dans les activités à exécuter supprimées (MSCAES).

Équation 8 MSCAES

$$MSCAES = \frac{\text{relation de type responsabilité dans en-cours seulement} \times 100}{\text{relation de type "responsabilité" dans "en - cours"}}$$

Pour appliquer cette métrique au processus de traitement de commandes, nous calculons, d'abord, le nombre de relation du type « responsabilité » dans la version « en-cours » (Figure 16). Nous devons déterminer ensuite le nombre de relations de ce type qui ne figurent que dans la version « en-cours » (Figure 16).

	Relation de type "responsabilité"			
	Dans "en-cours"		Dans "en-cours" uniquement	
	Lane	Activité	Lane	Activité
	Vente	Traitement commande	Production	Planification des ordres
	Production	Planification des ordres		
	Production	Fabrication		
Nombre de raltions	3		1	

Figure 16 Relations "responsabilités" supprimées dans PTC

Nous pouvons déduire la valeur de la métrique pour le processus :

$$MSCAES = \frac{1 \times 100}{3} = 33,33$$

Un triplet d'hypothèses est déduit de la définition de cette métrique MSCAES, dans le contexte de réingénierie d'un processus

- H1c : la suppression d'activités à exécuter va influencer la variation du coût du processus ;
- H1d : la suppression d'activités à exécuter va influencer la variation de la durée du processus ;
- H1nq : la suppression d'activités à exécuter va influencer la variation de la non-qualité du processus.

Nous définissons de la même manière la métrique structurelle de changement dans les activités à exécuter rajoutées (MSCAER).

Équation 9 MSCAER

$$MSCAER = \frac{\text{relation de type responsabilité dans implémenté seulement} \times 100}{\text{relation de type "responsabilité" dans "en - cours"}}$$

Pour le processus de traitement de commandes, il n’y a aucune relation du type « responsabilité » qui ne figure que dans la version « implémentée » (Figure 17).

	Relation de type "responsabilité"			
	Dans "en-cours"		Dans "en-devenir" uniquement	
	Lane	Activité	Lane	Activité
	Vente	Traitement commande		
	Production	Planification des ordres		
	Production	Fabrication		
Nombre de raltions	3		0	

Figure 17 Relations "responsabilités" rajoutées dans PTC

Là aussi la valeur de la métrique est déduite pour le processus de traitement de commandes :

$$MSCAER = \frac{0 \times 100}{3} = 0\%$$

Les trois hypothèses suivantes sont définies pour la métrique MSCAER

- H2c : le rajout d’activités à exécuter va influencer la variation du coût du processus ;
- H2d : le rajout d’activités à exécuter va influencer la variation de la durée du processus ;
- H2nq : le rajout d’activités à exécuter va influencer la variation de la non-qualité du processus.

Quel est le changement dans les informations à traiter ? Pour répondre à cette question, nous devons définir deux métriques structurelles de changement concernant les informations à traiter supprimées ou rajoutées. Cette question est reliée à la dimension du changement « donnée d’entrée » (DE). Nous utiliserons les arcs, de type association, allant des objets de donnée vers les activités. La métrique ASP la plus proche de la mesure à définir est celle d’objet de donnée (entrée) NDOIn. La première métrique est celle de changement structurelle dans les informations à traiter supprimées (MSCITS) :

Équation 10 MSCITS

$$MSCITS = \frac{\text{association objet donnée vers activité dans en-cours seulement} \times 100}{\text{association objet donnée vers activité dans "en - cours"}}$$

La version « en-cours » du processus de traitement de commandes contient cinq associations, allant des objets de donnée vers les activités (Figure 18). Parmi ces associations, une seule ne figure pas également dans la version « implémentée » (Figure 18).

	Association de objet de donnée vers activité			
	Dans "en-cours"		Dans "en-cours" uniquement	
	Objet de donnée	Activité	Objet de donnée	Activité
	Commande client	Traitement commande	Commande client	Planification des ordres
	Article	Traitement commande		
	Commande client	Planification des ordres		
	Ordre de fabrication	Fabrication		
	Plan de production	Fabrication		
Nombre d'associations	5		1	

Figure 18 Associations objets de donnée vers activités supprimées dans PTC

La valeur de la métrique MSCITS peut être calculée pour le processus avec ces données :

$$MSCITS = \frac{1 \times 100}{5} = 20\%$$

Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MSCITS par rapport à notre problématique :

- H3c : la suppression d'informations à traiter va influencer la variation du coût du processus ;
- H3d : la suppression d'informations à traiter va influencer la variation de la durée du processus ;
- H3nq : la suppression d'informations à traiter va influencer la variation de la non-qualité du processus.

La seconde métrique est celle de changement structurelle dans les informations à traiter rajoutées (MSCITR) :

Équation 11 MSCITR

$$MSCITR = \frac{\text{association objet donnée vers activité dans implémenté seulement} \times 100}{\text{association objet donnée vers activité dans "en - cours"}}$$

Une seule association, allant des objets de donnée vers les activités, a été rajoutée dans le processus de traitement de commandes (Figure 19).

	Associatio de objet de donnée vers activité			
	Dans "en-cours"		Dans "implémenté" uniquement	
	Objet de donnée	Activité	Objet de donnée	Activité
	Commande client	Traitement commande	Commande client	Fabrication
	Article	Traitement commande		
	Commande client	Planification des ordres		
	Ordre de fabrication	Fabrication		
	Plan de production	Fabrication		
Nombre d'associations	5		1	

Figure 19 Associations des objets de donnée vers activités rajoutées dans PTC

La métrique MSCITR du processus de traitement de commandes est déduite de la manière suivante :

$$MSCITR = \frac{1 \times 100}{5} = 20\%$$

Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MSCITR par rapport à notre problématique :

- H4c : le rajout d'informations à traiter va influencer la variation du coût du processus ;
- H4d : le rajout d'informations à traiter va influencer la variation de la durée du processus ;
- H4nq : le rajout d'informations à traiter va influencer la variation de la non-qualité du processus.

Quel est le changement dans les informations à produire ? Cette question concerne les informations produites par les activités. Dans BPMN, cette relation est représentée par une association allant d'une activité vers un objet de donnée. Nous définissons la métrique structurelle de changement dans les informations à produire supprimées (MSCIPS)

Équation 12 MSCIPS

$$MSCIPS = \frac{\text{association activité vers objet donnée dans en-cours} \times 100}{\text{association activité vers objet donnée dans "en - cours"}}$$

Les associations allant d'une activité vers un objet de donnée sont au nombre de six, dans la version « en-cours » du processus de traitement de commandes (Figure 20). Deux de ces associations ont été supprimées, lors de la restructuration (Figure 20).

	Association de activité vers objet de donnée			
	Dans "en-cours"		Dans "en-cours" uniquement	
	Activité	Objet donnée	Activité	Objet donnée
	Traitement de commande	Commande client	Planification des ordres	Ordres de fabrication
	Traitement de commande	Confirmation commande client	Planification des ordres	Plan de production
	Planification des ordres	Ordres de fabrication		
	Planification des ordres	Plan de production		
	Fabrication	Facture matériel		
	Fabrication	Expédition		
Nombre d'associations	6		2	

Figure 20 Associations activités vers objets de donnée supprimées dans PTC

La métrique MSCIPS pour le processus de traitement de commandes peut donc être calculée :

$$MSCIPS = \frac{2 \times 100}{6} = 33,33$$

Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MSCIPS par rapport à notre problématique :

- H5c : la suppression d'informations à produire va influencer la variation du coût du processus ;
- H5d : la suppression d'informations à produire va influencer la variation de la durée du processus ;
- H5nq : la suppression d'informations à produire va influencer la variation de la non-qualité du processus.

Nous avons ensuite la métrique structurelle de changement dans les informations à produire rajoutées (MSCIPR) :

Équation 13 MSCIPR

$$MSCIPR = \frac{\text{association activité vers objet donnée dans implémenté seulement} \times 100}{\text{association activité vers objet donnée dans "en - cours"}}$$

Deux associations allant d'une activité vers un objet de donnée ont été rajoutées au processus de traitement de commandes par la restructuration (Figure 21)

	Association de activité vers objet de donnée			
	Dans "en-cours"		Dans "implémenté" uniquement	
	Activité	Object donnée	Activité	Object donnée
	Traitement de commande	Commande client	Traitement de commande	Ordres de fabrication
	Traitement de commande	Confirmation commande client	Traitement de commande	Plan de production
	Planification des ordres	Ordres de fabrication		
	Planification des ordres	Plan de production		
	Fabrication	Facture matériel		
	Fabrication	Expédition		
Nombre d'associations	6		2	

Figure 21 Associations des activités vers objet de donnée rajoutées dans PTC

La métrique MSCIPR est déduite de la manière suivante :

$$MSCIPR = \frac{2 \times 100}{6} = 33,33$$

Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MSCIPR par rapport à notre problématique

- H6c : le rajout d'informations à produire va influencer la variation du coût du processus ;
- H6d : le rajout d'informations à produire va influencer la variation de la durée du processus ;
- H6nq : le rajout d'informations à produire va influencer la variation de la non-qualité du processus.

Quel est le changement dans les événements à surveiller ? Il s'agit pour nous de mesurer le changement dans les événements à surveiller. Ces événements à surveiller sont représentés par les flux séquence allant des événements vers les activités. Ce sont ces

événements qui déclenchent l'exécution des activités. La métrique structurelle de changement dans les événements à surveiller supprimés (MSCESS) est

Équation 14 MSCESS

$$MSCESS = \frac{\text{flux séquence de événement vers activité dans en-coursseulement} \times 100}{\text{flux séquence de événement vers activité dans "en – cours"}}$$

Dans le processus de traitement de commandes, un flux de séquence allant des événements vers les activités a été supprimé sur les trois de départ (Figure 22).

	Flux séquence de événement vers activité			
	Dans "en-cours"		Dans "en-cours" uniquement	
	Événement	Activité	Événement	Activité
	Arrivée commande client	Traitement commande	Commande acceptée	Planification des ordres
	Commande acceptée	Planification des ordres		
	Ordre inclus dans le planing	Fabrication		
Nombre de flux séquence	3		1	

Figure 22 Flux de séquence événements vers activités supprimées dans PTC

La métrique MSCESS prend la valeur suivante pour le processus de traitement de commandes :

$$MSCESS = \frac{1 \times 100}{3} = 33,33$$

Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MSCESS par rapport à notre problématique

- H7c : la suppression d'événements à surveiller va influencer la variation du coût du processus ;
- H7d : la suppression d'événements à surveiller va influencer la variation de la durée du processus ;
- H7nq : la suppression d'événements à surveiller va influencer la variation de la non-qualité du processus.

La métrique structurelle de changement dans les événements à surveiller rajoutés est définie comme suit (MSCESR) :

Équation 15 MSCESR

$$MSCESR = \frac{\text{flux séquence de événement vers activité dans implémenté seulement} \times 100}{\text{flux séquence de événement vers activité dans "en – cours"}}$$

Aucun flux séquence, allant des événements vers les activités, n'a été créé sur le processus de traitement de commandes lors de la restructuration (Figure 23).

	Flux séquence de événement vers activité			
	Dans "en-cours"		Dans "implementé" uniquement	
	Évènement	Activité	Évènement	Activité
	Arrivée commande client	Traitement commande		
	Commande acceptée	Planification des ordres		
	Ordre inclus dans le planing	Fabrication		
Nombre de flux séquence	3		0	

Figure 23 Flux séquence des événements vers les activités rajoutées dans PTC

La métrique MSCESR est extraite pour le processus de traitement de commandes :

$$MSCESR = \frac{0 \times 100}{3} = 0\%$$

Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MSCESR par rapport à notre problématique

- H8c : le rajout d'événements à surveiller va influencer la variation du coût du processus ;
- H8d : le rajout d'événements à surveiller va influencer la variation de la durée du processus ;
- H8nq : le rajout d'événements à surveiller va influencer la variation de la non-qualité du processus.

Quel est le changement dans les événements à générer ? Nous nous intéressons ici aux événements générés par les activités. La notation BPMN représente ces événements par des flux de séquence, allant des activités vers les événements. La métrique structurelle de changement dans les événements à générer supprimés (MSCEGS) est

Équation 16 MSCEGS

$$MSCEGS = \frac{\text{flux séquence de activité vers événement dans en-cours} \times 100}{\text{flux séquence de activité vers événement dans "en - cours"}}$$

La restructuration du processus de traitement de commandes a supprimé les deux flux de séquence allant des activités vers les événements présents dans la version « en-cours » (Figure 24).

	Flux séquence de activité vers événement			
	Dans "en-cours"		Dans "en-cours" uniquement	
	Activité	Évènement	Activité	Évènement
	Traitement des commandes	Commande acceptée	Traitement des commandes	Commande acceptée
	Planification des ordres	Ordres inclus dans planing	planification des ordres	Ordres inclus dans planing
Nombre de flux séquence	2		2	

Figure 24 Flux séquence des activités vers les événements supprimées dans PTC

Pour le processus de traitement de commandes, la métrique MSCEGS est obtenue comme suit :

$$MSCEGS = \frac{2 \times 100}{2} = 100\%$$

Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MSCEGS par rapport à notre problématique

- H9c : la suppression d'événements à générer va influencer la variation du coût du processus ;
- H9d : la suppression d'événements à générer va influencer la variation de la durée du processus ;
- H9nq : la suppression d'événements à générer va influencer la variation de la non-qualité du processus.

La métrique structurelle de changement dans les événements à générer rajoutés (MSCEGR) est définie comme suit :

Équation 17 MSCEGR

$$MSCEGR = \frac{\text{flux séquence de activité vers événement dans implémenté seulement} \times 100}{\text{flux séquence de activité vers événement dans "en – cours"}}$$

Un flux de séquence allant des activités vers les événements a été rajouté par la restructuration du processus de traitement de commandes (Figure 25).

	Flux séquence de activité vers événement			
	Dans "en-cours"		Dans "implémenté" uniquement	
	Activité	Événement	Activité	Événement
	Traitement des commandes	Commande acceptée	Traitement des commandes	Ordres inclus dans planing
	Planification des ordres	Ordres inclus dans planing		
Nombre de flux séquence	2		1	

Figure 25 Flux de séquence des activités vers les événements rajoutées dans PTC

La métrique MSCEGR du processus de traitement de commandes est extraite en utilisant les informations ci-dessus :

$$MSCEGR = \frac{1 \times 100}{2} = 50\%$$

Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MSCEGR par rapport à notre problématique

- H10c : le rajout d'événements à générer va influencer la variation du coût du processus ;
- H10d : le rajout d'événements à générer va influencer la variation de la durée du processus ;
- H10nq : le rajout d'événements à générer va influencer la variation de la non-qualité du processus.

2.4.5. Définition des métriques structurelles de variation

Les métriques structurelles de variation répondent aux questions liées au sous-objectif GQM de mesure de la variation dans les propriétés du processus. Dans cette section, nous n'allons pas nous intéresser au changement dans les éléments à l'intérieur du processus (objet

de la section précédente), mais plutôt sur des propriétés qui caractérisent le processus dans son ensemble. Ces deux points de vue permettent d'avoir une vision complète du changement sur les processus. Ces métriques sont déterminées sur la base des métriques structurelles dérivées de l'analyse structurelle des processus. Elles renvoient à des propriétés plus complexes du processus. Pour ces métriques dérivées, la quantification se fait par calcul, contrairement aux métriques de base obtenues par décompte.

Les métriques structurelles de variation sont obtenues en faisant la différence entre les valeurs des métriques dérivées correspondantes des processus « en-cours » et « en-devenir » (ou « implémenté »). Cette valeur doit être exprimée en pourcentage de la métrique dérivée du processus « en-cours » pour assurer la normalisation des valeurs. Cela équivaut à la division par la métrique opérationnelle ou le nombre de relations du processus de départ utilisé respectivement pour normaliser les métriques opérationnelles de variation et les métriques structurelles de changement. Pour pouvoir déterminer une métrique structurelle de variation, il faut aussi vérifier que la métrique dérivée correspondante est applicable.

2.4.5.1. Variation dans la complexité du flux de contrôle

La complexité du processus est identifiée dans la littérature comme une dimension du changement dont la variation peut conduire à une variation de la performance. C'est pour cela que la question relative à la variation dans la complexité des processus a été reliée au sous-objectif de mesure de la variation sur les propriétés des processus lors de la restructuration. La métrique de Complexité de Flux de Contrôle (CFC) (Cardoso, 2005) est proposée pour mesurer la complexité d'un processus. Il s'agit pour nous de comprendre cette métrique et de l'utiliser dans la mesure de la variation dans la complexité.

La métrique CFC correspond à la définition de la complexité en tant que « difficulté de tester le modèle ». C'est le nombre d'états mentaux à considérer lorsque le designer développe le processus, ou nombre de décisions dans le flux du processus. Chaque point de décision augmente le nombre de décisions possibles comme suit (Cardoso, 2005) :

- **le point de décision AND** : toutes les transitions sortant du nœud AND doivent être exécutées. Le designer n'a besoin de considérer qu'un seul état comme résultat de l'exécution du point AND. Pour cette raison, chaque point de décision AND ajoute 1 à la métrique CFC ;
- **le point de décision XOR (avec n transitions sortantes)** : exactement un des n chemins possibles sera pris. Le designer doit donc prendre en compte n états possibles qui peuvent découler de l'exécution du point de décision XOR. Chaque point de décision XOR (avec n transitions sortantes) ajoute n à la métrique CFC ;
- **le point de décision OR (avec n transitions sortantes)** : il existe $2^n - 1$ possibilités de traiter au moins 1 et au plus n des transitions sortantes du point OR. Chaque point OR avec n transitions sortantes ajoute $2^n - 1$ à la métrique CFC.

Notre première métrique structurelle de variation est la métrique structurelle de variation de Complexité de Flux de Contrôle (MVSCFC). La métrique dérivée CFC n'étant pas exprimée en pourcentage, nous devons donc exprimer notre métrique structurelle de variation du CFC en pourcentage de la CFC du processus d'origine :

Équation 18 Métrique structurelle de variation de la complexité flux de contrôle

$$\text{MSVCFC} = \frac{\text{CFC}_{\text{to-be}} - \text{CFC}_{\text{as-is}}}{\text{CFC}_{\text{as-is}}}$$

Il se trouve, cependant, que cette métrique CFC peut être nulle pour le processus d'origine, si celui-ci ne contient aucun nœud de décision. Cela entraînerait une division par zéro dans le calcul de la métrique structurelle de variation de la CFC. Pour éviter cela, nous appliquerons la règle de non-division par zéro suivante, dans le cas où la CFC du processus « en-cours » est nulle :

- Si la CFC du processus « en-devenir » (ou « implémenté») est nulle, elle aussi: même si nous ne pouvons pas appliquer la formule de calcul, nous savons que la variation est nulle. Nous affecterons donc directement la valeur 0 à la métrique ;
- Sinon : l'extraction de la métrique est impossible et nous avons, dans ce cas, une donnée manquante pour le processus.

Le calcul de la métrique structurelle de variation de la CFC pour le processus de traitement de commandes commence par celui de la métrique dérivée de CFC pour ses versions « en-cours » (Figure 11) et « implémenté» (Figure 12). Ces versions ne comportant aucun point de décision AND, XOR ou OR ; leurs métriques CFC sont égales 0. Comme nous l'avons présenté ci-dessus, dans un tel cas on affecte directement la valeur 0% à la métrique structurelle de variation de la CFC pour le processus. Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MSVCFC par rapport à notre problématique

- H11c : la variation dans la complexité du flux de contrôle va influencer la variation du coût du processus
- H11d : la variation dans la complexité du flux de contrôle va influencer la variation de la durée du processus
- H11nq : la variation dans la complexité du flux de contrôle va influencer la variation de la non-qualité du processus

2.4.5.2. Variation dans la profondeur d'imbrication maximale

Gruhn (2005) propose une métrique dérivée de profondeur de l'imbrication maximale (PIM) en complément de la métrique CFC. Selon lui, deux processus peuvent avoir la même CFC, alors que celui qui est plus linéaire est moins complexe que celui qui comporte plusieurs points de décision XOR imbriqués, par exemple. La profondeur de l'imbrication maximale

étant une mesure de la complexité, sa variation peut avoir un impact sur la performance. C'est pour cette raison que nous avons créé à partir d'elle, une question associée au sous-objectif GQM de mesure de la variation dans les propriétés des processus. La profondeur de l'imbrication maximale pour le processus est la plus grande imbrication parmi les activités du processus. L'imbrication d'une activité est le nombre de décisions dans le flux de contrôle nécessaires pour exécuter cette activité (Gruhn, et al., 2005). La métrique structurelle de variation de la profondeur d'imbrication maximum se calcule de la manière suivante :

Équation 19 Métrique structurelle de variation de la profondeur maximale

$$\text{MSVPIM} = \frac{\text{PIM}_{\text{to-be}} - \text{PIM}_{\text{as-is}}}{\text{PIM}_{\text{as-is}}}$$

Dans le cas où la profondeur d'imbrication maximale du processus « en-cours » est nulle, nous appliquerons la même règle de non-division par zéro définie pour la métrique MSVCFC. La version « en-cours » du processus de traitement de commandes a une PIM nulle car elle est parfaitement linéaire (ne contenant aucun point de décision). Comme sa version « en-devenir » prend elle aussi la valeur nulle, l'application de la règle de non-division par zéro implique l'affectation de la valeur nulle à la métrique structurelle de variation de la profondeur d'imbrication maximum. Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MSVPIM par rapport à notre problématique

- H12c : la variation dans la profondeur d'imbrication maximale va influencer la variation du coût du processus
- H12d : la variation dans la profondeur d'imbrication maximale va influencer la variation de la durée du processus
- H12nq : la variation dans la profondeur d'imbrication maximale va influencer la variation de la non-qualité du processus

2.4.5.3. Variation dans la simplicité

La métrique structurelle dérivée de simplicité a été définie par Tjaden (2001). Elle est associée à la dimension complexité du changement car elle est inversement liée à cette dernière. De plus, elle est calculée avec des informations différentes de celles utilisées pour les deux métriques dérivées de complexité que nous venons de présenter. La mesure de sa variation participe donc à répondre aux questions associés au sous-objectif de mesure de la variation des propriétés du processus. Pour définir cette métrique de simplicité, Tjaden (2001) définit d'abord la complexité basique d'un processus comme la somme du nombre d'activités, du nombre de flux matériels ou informationnels, et du nombre de personnes. La complexité moyenne par activité est alors calculée en divisant la complexité basique du processus par le nombre d'activités.

$$\bar{C}_A = \frac{C_B}{a}$$

Pour un processus, de complexité basique C_B , la valeur maximale pour la complexité moyenne par activité est atteinte lorsque $a=1$:

$$\max_{\bar{C}_A} = C_B$$

Pour ce même processus, de complexité basique C_B , le minimum pour la complexité moyenne par activité est atteint avec le maximum d'activités dans le processus. La formule de ce minimum est détaillée dans Tjaden (2001) et est de la forme suivante:

$$\min_{\bar{C}_A} = \frac{3C_B}{C_B - 1} \approx 3$$

La valeur de \bar{C}_A dans cet intervalle est l'indicateur du degré de simplicité des activités. Ce degré, exprimé en pourcentage, est la métrique de simplicité, définie formellement comme :

$$\text{simplicité} = \frac{\bar{C}_A - \min_{\bar{C}_A}}{\max_{\bar{C}_A} - \min_{\bar{C}_A}} \times 100$$

Équation 20 Simplicité (Tjaden, 2001)

La métrique structurelle de variation de la simplicité est calculée à partir de la métrique de simplicité (S) qui est déjà en pourcentage. Elle s'obtient donc directement par :

Équation 21 Métrique structurelle de variation de la simplicité

$$\text{MSVS} = S_{\text{to-be}} - S_{\text{as-is}}$$

Pour déterminer la métrique structurelle de variation de la simplicité du processus de traitement de commandes, nous avons d'abord estimé les valeurs des métriques structurelles dérivées de simplicité de ses deux versions (Tableau 19). Dans ce calcul, nous avons considéré que les zones de responsabilités (les *Lanes*) de BPMN représentaient les personnes.

Tableau 19 Détails simplicité PTC

	En-cours	En-devenir
Nombre d'activités	3	2
Nombre de personnes	2	2
Flux matériels ou informationnels	8	8
La complexité basique C_B	13	12
La complexité moyenne par activité \bar{C}_A	13/3=4,333333333	12/2= 6

$\max \bar{c}_A$	13	12
$\min \bar{c}_A$	$3 \cdot 13 / 12 = 3,25$	$3 \cdot 12 / 11 =$ $3,272727273$
$\bar{c}_A - \min \bar{c}_A$	$(4,333333333 - 3,25) =$ $1,083333333$	$(6 - 3,272727273) =$ $2,727272727$
$\max \bar{c}_A - \min \bar{c}_A$	$13 - 3,25 = 9,75$	$12 - 3,272727273 =$ $8,727272727$
Simplicité	$1,083333333 \cdot 100 / 9,75 =$ $11,111\%$	$2,727272727 \cdot 100 / 8,727272727 =$ $31,25\%$

Le processus de traitement de commandes a une variation dans la simplicité de :

$$MSVS = S_{\text{to-be}} - S_{\text{as-is}} = 31.25 - 11.111 = 20.13\%$$

Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MSVS par rapport à notre problématique

- H13c : la variation dans la simplicité va influencer la variation du coût du processus
- H13d : la variation dans la simplicité va influencer la variation de la durée du processus
- H13nq : la variation dans la simplicité va influencer la variation de la non-qualité du processus

2.4.5.4. Variation dans l'intégration

L'intégration étant une dimension du changement, nous avons jugé intéressant de définir une question GQM relative à sa variation lors de la restructuration. Cette question est associée au sous-objectif GQM de mesure de la variation sur les propriétés du processus qui participe à mesurer le changement. Une métrique structurelle de variation peut être ensuite définie pour répondre à cette question. Pour calculer cette métrique, il faut d'abord calculer la métrique dérivée d'intégration pour les versions « en-cours » et « implémenté » puis quantifier la variation entre elles. Des définitions pour la métrique dérivée d'intégration ont été proposées par Tjaden (2001) et Aubert, et al. (2003). Tjaden (2001) définit la métrique dérivée d'intégration comme une mesure des relations structurelles qui favorisent la coopération entre processus (ou entre activités au sein d'un processus). Une valeur

d'intégration est assignée à chaque flux matériel ou informationnel selon le type de livraison entre l'activité productrice du flux et l'activité utilisatrice de ce flux (Tjaden, 2001) :

- par événement : 10,
- par demande directe : 7,
- par demande : 4,
- selon un calendrier : 1,
- par commande : 0.

La valeur de l'intégration pour le processus est la somme de valeurs individuelles d'intégration des flux divisée par le nombre de flux matériels ou informationnels du processus (Tjaden, 2001). Cette valeur est exprimée en pourcentage de la valeur individuelle maximale d'intégration des flux (10) et est calculé de la sorte :

Équation 22 Intégration Processus (Tjaden, 2001)

$$Intégration = \frac{\sum_{i=1}^m I_i}{m} \times \frac{100}{10}$$

Aubert, et al. (2003), ont proposé une définition du niveau d'intégration d'un processus métier basée sur quatre propriétés fondamentales (Aubert, et al., 2003) :

- **la transparence** : elle permet l'intégration des données et un partage ouvert. La transparence peut être obtenue par la translation à travers plusieurs langages ou par la standardisation établissant un langage unique ;
- **l'accessibilité** : elle fait référence à la possibilité d'avoir accès aux données à partir de chaque point du processus où c'est nécessaire. Elle permet le partage de l'information et le traitement des ressources comme si elles étaient centralisées, et une capture unique des données ;
- **l'opportunité** : pour permettre la coordination et la synchronisation, l'information doit être à jour. Combinée à l'accessibilité, elle permet à toutes les activités d'être exécutées au moment approprié ;
- **la granularité** : toutes les informations échangées dans le processus doivent être fournies au bon niveau de détail. Une granularité appropriée permet l'élimination des activités sans importance qui serait requises pour décomposer ou résumer l'information.

Ils définissent le niveau d'intégration d'un processus comme la somme des efforts consacrés, par les activités réceptrices, pour atteindre l'accessibilité, la transparence, l'opportunité et la granularité des biens et des données, par rapport à la valeur ajoutée par le processus⁵⁹ (Aubert, et al., 2003):

⁵⁹ Où VA : valeur ajoutée par le processus

$C(a_j)$: Coût pour fournir la propriété x par l'activité j

$$PI = \frac{VA - \sum_{j=1}^n (C(a_j) + C(tr_j) + C(o_j) + C(g_j))}{VA}$$

Un problème de calculabilité se pose, à propos de ces deux définitions de la métrique d'intégration, à cause des informations qu'elles utilisent. En effet, les informations sur le type de livraison dans la définition de Tjaden (2001) ou sur les quatre propriétés d'Aubert, et al. (2003) ne sont pas disponibles dans les modèles de processus BPMN que nous utilisons pour extraire les métriques structurelles. Nous préférons ne pas tenir compte de la mesure de la variation de l'intégration dans notre quantification du changement sur les processus. Nous nous attaquerons à cette question dans les travaux futurs.

2.4.5.5. Variation dans la flexibilité

La question de la variation dans la flexibilité participe à l'objectif de mesure de la variation dans les propriétés du processus. Elle participe donc indirectement à la mesure du changement. La métrique associée à cette question est celle de variation de la flexibilité que nous proposons de définir à partir de la métrique dérivée de flexibilité. Tjaden (2001) définit la flexibilité d'un processus comme la manière dont celui-ci est contrôlé : réactive, préventive, progressive ou dynamique. Il commence par définir des propriétés de flexibilité qui rendent un processus dynamique : les F-propriétés. Chaque flux matériel ou informationnel du processus⁶⁰ est examiné selon ces propriétés ; les valeurs applicables sont additionnées. Chaque flux est lié à une activité utilisatrice, une activité productrice et éventuellement une activité contrôleuse qui assigne les objectifs sur les performances de l'activité productrice du flux. Les propriétés sont définies sous forme de question comme :

- le flux est contrôlé,
- le flux est directement contrôlé,
- l'activité utilisatrice est contrôleuse directe,
- le flux est mesuré,
- l'activité productrice est contrôleuse,
- etc.

a_j Accessibilité pour l'activité j

tr_j Transparence pour l'activité j

o_j Opportunité pour l'activité j

g_j Granularité pour l'activité j

⁶⁰ y compris les entrées venant ou les sorties vers l'environnement

Le total des valeurs de F-propriétés assignées est calculé pour chaque flux matériel ou informationnel. Ces flexibilités individuelles F_{M_i} sont additionnées et divisées par le nombre total de flux matériels ou informationnels m_{op} . Cette quantité est comprise entre 0 et 60 car chacune des flexibilités individuelles est comprise entre 0 et 60 qui est la somme des valeurs maximales des F-propriétés. La métrique de flexibilité du processus est donc exprimée en pourcentage de cette valeur maximale :

Équation 24 Flexibilité (Tjaden, 2001)

$$flexibilité = \frac{\sum_{i=1}^{m_{op}} F_{M_i}}{m_{op}} \times \frac{100}{60}$$

Les informations relatives aux propriétés de flexibilité ne peuvent pas être tirées des modèles de processus tels que ceux de la notation BPMN. Comme pour les mesures relatives à l'intégration, nous rencontrons le problème de calculabilité pour cette métrique de flexibilité. La mise en œuvre de la métrique de variation dans la flexibilité ne peut donc pas être réalisée dans cette étude.

2.4.6. Mesure de l'effort de gestion du changement

La troisième partie de l'application de l'approche de GQM, après la mesure de changement et de son impact, est celle de l'effort de gestion du changement. La mesure de l'effort de gestion du changement est un sous-objectif direct de l'objectif de prédiction de l'impact du changement sur la performance. Comme pour les autres sous-objectifs GQM, nous allons lui associer des questions qui vont nous permettre de définir les métriques. La première question est relative à la formation fournie aux utilisateurs participant à chaque processus. La seconde question est relative à l'accompagnement (toutes les dimensions de l'effort de gestion du changement différentes de la formation). Les variables relatives à l'effort de gestion du changement sont donc la métrique effort de formation (MEF) et celle d'effort d'accompagnement (MEA). Ces variables de mesure de l'effort de gestion du changement sont exprimées en nombre d'heures par acteur du processus. Elles doivent mesurer le temps investi dans les activités de formation ou d'accompagnement vers les acteurs du processus. Ces variables doivent être normalisées afin que leurs valeurs soient comparables entre les processus. En effet, donner 60 heures de formation et 30 heures d'accompagnement ne représente pas la même chose si le processus est exécuté par 5 ou par 10 acteurs. Pour cela, nous choisissons de diviser par le nombre d'acteurs dans le processus « en-devenir ». Nous utiliserons le nombre de zones de responsabilité comme nombre d'acteurs.

Équation 25 Métrique effort de formation

$$MEF = \frac{\text{heures de formations des acteurs du processus}}{\text{nombre d'acteurs du processus}}$$

Équation 26 Métrique effort d'accompagnement

$$MEA = \frac{\text{heures d'accompagnement des acteurs du processus}}{\text{nombre d'acteurs du processus}}$$

Les données relatives aux métriques de l'effort de gestion sont collectées par supervision, en notant, dans les projets antérieurs, le temps alloué à ces deux activités pour les processus. Pour un processus en cours de restructuration, la valeur est fixée au départ et modifiée éventuellement lors des différents passages dans la boucle adéquation-configuration-prototypage de Tomas (2000). A la sortie de cette boucle, le processus « en-devenir » est validé, les prédictions de l'impact du changement sont réalisées. Ces prédictions sont estimées en fonction du changement introduit et de l'effort de gestion à consentir sur le processus. Supposons que 30 heures de formation et 10 heures d'accompagnement aient été données aux acteurs du processus de traitement de commandes. Vu que la version « en-devenir » de ce processus comprend deux zones de responsabilité, nous obtenons les valeurs des métriques d'effort de gestion suivantes :

$$MEF = \frac{30}{2} = 15 \text{ h/acteur}$$
$$MEA = \frac{10}{2} = 5 \text{ h/acteur}$$

Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MEF par rapport à notre problématique

- H14c : l'effort de formation va influencer la variation du coût du processus
- H14d : l'effort de formation va influencer la variation de la durée du processus
- H14nq : l'effort de formation va influencer la variation de la non-qualité du processus

Le triplet d'hypothèses suivant est défini pour tester la pertinence de la métrique MEA par rapport à notre problématique

- H15c : l'effort d'accompagnement va influencer la variation du coût du processus
- H15d : l'effort d'accompagnement va influencer la variation de la durée du processus
- H15nq : l'effort d'accompagnement va influencer la variation de la non-qualité du processus

2.4.1. Conclusion

L'objectif principal de ce chapitre est la définition des variables explicatives de notre problème. Cet objectif est réalisé à travers la quantification de l'ampleur et de l'effort de gestion du changement. Les variables explicatives définies sur l'unité d'analyse sont au

nombre de 15 (Tableau 20). Elles sont composées de 2 métriques d'effort de gestion du changement (formation et accompagnement) et de 13 métriques de mesure de l'ampleur du changement. Trois des mesures de l'ampleur du changement sont des métriques structurelles de variation (dans la complexité du flux de contrôle, dans la profondeur d'imbrication et dans la simplicité). Les dix métriques restantes sont les métriques de changement structurelles mesurées par rapport aux suppressions et rajouts dans les arcs BPMN. Nous avons déterminée la valeur de chacune de ces variables pour le processus de traitement de commandes (Tableau 20)

Tableau 20 Synthèse variables explicatives

Catégorie	Sous-catégorie	Désignation	Abréviation	VPTC
Variables explicatives	Changement	métrique structurelle de changement dans les activités à exécuter supprimées	MSCAES	33
		métrique structurelle de changement dans les activités à exécuter rajoutées	MSCAER	0
		métrique structurelle de changement dans les informations à traiter supprimées	MSCITS	20
		métrique structurelle de changement dans les informations à traiter rajoutées	MSCITR	20
		métrique structurelle de changement dans les informations à produire supprimés	MSCIPS	33
		métrique structurelle de changement dans les informations à produire rajoutées	MSCIPR	33
		métrique structurelle de changement dans les événements à surveiller supprimés	MSCCESS	33
		métrique structurelle de changement dans les événements à surveiller rajoutés	MSCESR	0
		métrique structurelle de changement dans les événements à générer supprimés	MSCCEGS	100
		métrique structurelle de changement dans les événements à générer rajoutés	MSCCEGR	50
		métrique structurelle de variation de complexité de flux de contrôle	MSVCFC	0

		métrique structurelle de variation de la profondeur d'imbrication maximum	MSVPIM	0
		métrique structurelle de variation de la simplicité	MSVS	20
	Efforts de gestion du changement	Métrique effort de Formation	MEF	15
		Métrique effort d'accompagnement	MEA	5

Chapitre 2.5 Conclusion de la deuxième partie

Cette partie « définition des variables » nous a permis de définir les variables explicatives et à expliquer de notre problème. Nous avons, d'abord, analysé le processus opérationnel, tant sur le plan structurel que sur le plan opérationnel. Cela nous a permis de définir des métriques qui quantifient les trois concepts principaux de notre étude. Les variables seront utilisées dans la partie suivante pour l'apprentissage et l'utilisation du modèle. Elles ont toutes été définies au niveau processus. Chaque variable explicative définie donne lieu à la définition d'un triplet d'hypothèses que l'apprentissage se chargera d'accepter ou de rejeter.

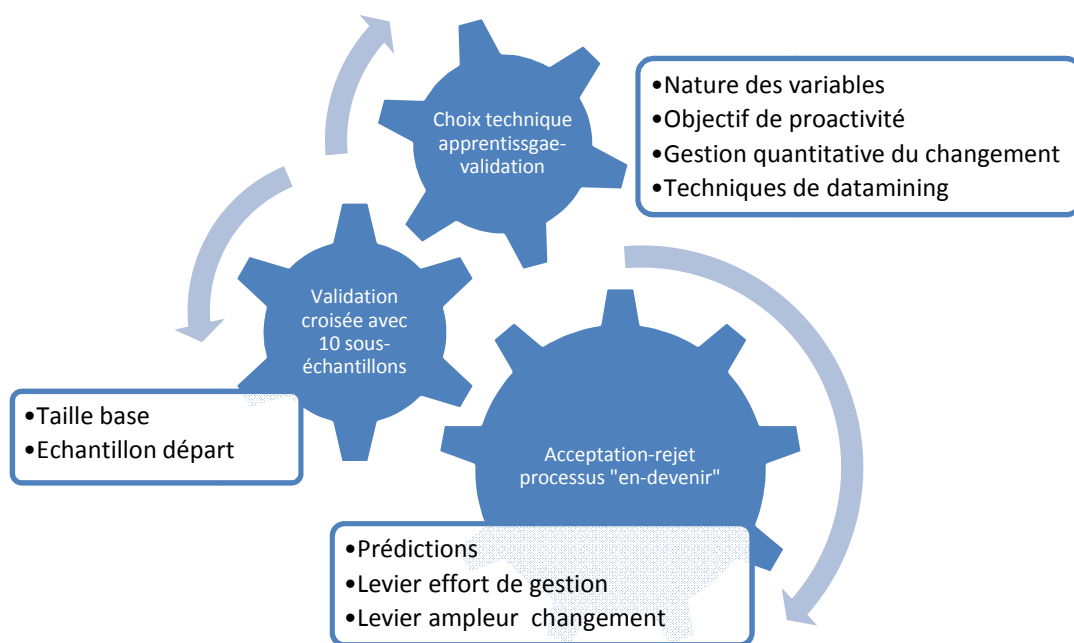
PARTIE 3 APPRENTISAGE ET UTILISATION DU MODELE

Chapitre 3.1 Introduction de la troisième partie

Nous allons présenter, dans cette partie, l'apprentissage et l'utilisation du modèle de notre démarche de gestion proactive du changement dans les projets d'implémentation de progiciel. Dans le Chapitre 3.2, nous allons choisir une technique d'apprentissage et de validation en tenant compte de l'objectif de proactivité, des variables définies dans la partie précédente, et des techniques définies en datamining (Figure 26). Avant de choisir les méthodes à utiliser, nous avons analysé les techniques quantitatives utilisées pour la gestion du changement dans les projets ERP. La technique de modélisation choisie est l'apprentissage d'un réseau bayésien par recherche gloutonne (Camara, et al., 2007a). Pour l'évaluation du modèle, nous avons opté pour la validation croisée avec 10 sous-échantillons. La taille d'échantillon minimum nécessaire est fixée par rapport à ces deux méthodes.

Le Chapitre 3.3 est relatif à l'application des méthodes choisies pour construire et valider le modèle. Dix expériences de la validation croisée y sont exécutées pour obtenir un modèle final de même que l'erreur qui doit être attendue de l'utilisation de ce modèle pour la généralisation (Figure 26). Ce chapitre correspond au regroupement des phases de modélisation et de validation du processus de datamining. Le Chapitre 3.4 porte sur l'utilisation du modèle pour l'acceptation d'une version « en-devenir » d'un processus (Figure 26). Le modèle fournit des prédictions sur les variables de réponse d'un processus en cours de restructuration. Les actions proactives sont décidées sur la base des prédictions obtenues. Ce chapitre représente la phase d'utilisation du modèle dans le processus de datamining.

Figure 26 Apprentissage, validation et utilisation du modèle



Chapitre 3.2 Choix de la méthode de modélisation et de validation

3.2.1. Introduction

Pour choisir une méthode de modélisation et de validation, nous allons d'abord nous intéresser aux méthodes de modélisation utilisées pour la gestion du changement dans les projets ERP ou dans des projets similaires. L'idée est de voir si ces techniques sont adaptées et directement utilisables par rapport à l'objectif et au type des variables de notre étude. Les méthodes de modélisation et de validation retenues sont présentées dans ce chapitre. La taille minimum de la base de cas est fixée en fonction de ces méthodes.

3.2.2. Modélisation dans la gestion du changement dans les projets ERP

Nous avons déjà présenté, dans le chapitre 4 de la première partie, la gestion quantitative en nous intéressant surtout à la définition du niveau d'analyse et des variables. Dans cette section, nous analysons les techniques quantitatives utilisées dans ces travaux pour déterminer leur adéquation à la question du changement dans les projets ERP. Nous étudions aussi la pertinence de ces techniques par rapport à notre objectif de recherche et la nature de nos variables. L'objectif de notre étude est de prédire l'impact du changement sur la performance des processus. Cette prédiction est basée sur la mesure de l'ampleur et de l'effort de gestion du changement.

Siénu, et al. (2006), ont proposé une analyse de la criticité des risques relatifs à des événements qui sont des changements dans l'environnement extérieur ou à l'intérieur de l'entreprise. Ce choix s'explique par la nature des événements traités par les auteurs. Ces événements sont de divers types. Ils se produisent de manière incontrôlée et impactent directement la performance de l'entreprise. C'est pour cela que l'estimation de la criticité utilisée en gestion des risques est appropriée pour l'analyse de ces événements. Le caractère imprévisible des risques est lié, aussi, à la non-limitation de l'étude dans le temps. L'analyse de la criticité définie en analyse du risque et utilisée par les auteurs n'est pas appropriée pour répondre à notre problématique. Elle est nécessaire quand plusieurs types de risques, aux apparitions imprévisibles, sont considérés. Dans notre étude, un seul type de risque est analysé : l'impact du changement sur les processus après leur restructuration. L'étude que nous abordons ici est limitée à la durée du projet et à la période qui suit directement l'implémentation. Dans celle-ci, le début de l'action de restructuration qui va engendrer éventuellement le risque est connu. Tout cela, diminue fortement le caractère imprévisible du risque par rapport au temps.

Raphaëli, et al. (2004) utilisent le raisonnement à base de cas pour définir des actions dédiées aux événements imprévus dans le processus d'implémentation des systèmes d'entreprise. Le raisonnement à base de cas consiste, pour un nouveau cas, à retrouver un autre cas qui lui est similaire, par la description, et lui appliquer la solution de ce dernier. Raphaëli, et al. (2004) définissent la description des cas par les causes de l'événement, les données de contexte et le profil de l'entreprise. Cette méthode établit une relation entre la description du cas et une solution en passant par le cas similaire. Ce sont des connaissances parcellaires, liées à chaque description de cas qui sont extraites, et rassemblées pour couvrir le domaine d'étude. La connaissance n'est pas consolidable en un modèle unique qui montrerait la relation générale entre les descriptions et les solutions. Cela constitue une limite à l'application de cette méthode à notre problématique.

Chapron (2006) applique les méthodes de clusterisation, afin de définir des regroupements entre les processus. La clusterisation⁶¹(ou apprentissage non-supervisé) est une fonction spécifique dans la typologie de datamining (Han, et al., 2001 p. 25) (Rakotomalala, 200B). Elle n'est pas applicable à notre problème car nous n'avons aucun besoin de regroupement.

Parr, et al. (2000) et Barki, et al. (2004) utilisent un tableau équivalent à un arbre de décision. Ce tableau est un modèle de classement dont les variables de réponse sont regroupées en trois classes (durée, effort et budget). Les variables explicatives sont le périmètre BPR, le périmètre physique, le périmètre technique et la stratégie d'implémentation des modules dans Parr, et al. (2000). Dans Barki, et al. (2004), ces variables explicatives sont enrichies puis regroupées en trois variables (étendue, profondeur et magnitude) relatives au projet. Les variables de notre problème étant toutes continues, cette méthode n'est pas adaptée à leur étude.

Anderson (2001) étudie l'impact des changements potentiels dans les variables de structure organisationnelle, de senior leadership et de vision stratégique sur la variable de prise de décision. Il utilise, pour cela, un réseau bayésien discret car toutes ses variables sont discrètes. Là aussi, il y a une incompatibilité par rapport à la nature des variables de notre étude.

Wischnevsky, et al. (2001) ont proposé un modèle de régression pour expliquer la relation entre des variables concernant la performance de l'organisation et des variables explicatives relatives à la transformation organisationnelle et ses conditions d'interaction avec la modification de l'environnement. Le modèle de régression correspond à notre objectif de prédiction et à nos variables continues. La technique utilisée pour construire le modèle de régression dans Wischnevsky, et al. (2001) est la méthode des moindres carrés. Parmi les approches utilisées en gestion quantitative du changement, le modèle le plus proche de notre

⁶¹ Cluster analysis ou clustering

problématique est la régression utilisée dans Wischnevsky, et al. (2001). En effet, la prédiction de variables continues peut être modélisée par les techniques statistiques de régression (Han, et al., 2001 p. 319). Nous allons cependant chercher une technique de détermination du modèle de régression plus optimale que les moindres carrées.

3.2.3. Choix de la technique de sélection des variables

Un réseau bayésien, pour un ensemble de variables aléatoires, est une paire (D, P). Le graphe dirigé sans cycle, ou DAG (Directed Acyclic Graph), D définit la structure du réseau. Il est composé d'un ensemble fini de nœuds et d'un ensemble fini d'arcs dirigés entre les nœuds. L'ensemble des distributions de probabilités locales de tous les nœuds est P. La recherche gloutonne fait partie des méthodes qui vont chercher la structure qui maximise un certain score (Leray, 2006). En effet, le nombre de réseaux possibles augmente plus qu'exceptionnellement avec le nombre de nœuds (Bootcher, et al., 2003A). En principe, il est possible de calculer les scores de tous les réseaux possibles et de les comparer pour trouver la meilleure structure. Cependant, cela n'est pas réalisable informatiquement si le nombre de variables est important (Bootcher, et al., 2003A).

La première raison pour laquelle nous avons choisi la recherche gloutonne des réseaux bayésiens, est la correspondance entre la fonction de prédiction, que ces réseaux permettent de réaliser, et l'objectif de notre recherche. L'objectif de la modélisation, dans notre étude, est la prédiction de l'impact du changement sur les processus. Pour cela, nous devons prédire les valeurs des métriques opérationnelles de variation à partir des métriques structurelles de changement et de variation, et des métriques d'effort de gestion. Nous avons donc besoin de réaliser ce qu'on appelle la fonction d'explication (ou l'apprentissage supervisé) dans la typologie des méthodes de datamining (Han, et al., 2001 p. 24) (Rakotomalala, 200B). La fonction d'explication est composée du classement (classification) et de la prédiction (Prediction). Un réseau bayésien avec des variables continues étant un modèle de prédiction, il relève donc de la fonction d'explication.

L'apprentissage permet de construire un modèle de la réalité à partir de données. Les données sont constituées d'individus (objets, observations, enregistrements) caractérisés par des descripteurs (attributs, variables, champs, caractères). L'apprentissage supervisé consiste à prédire les valeurs d'un attribut (endogène) à partir d'autres attributs (exogènes). L'apprentissage supervisé permet de mettre à jour un modèle M, tel que (Rakotomalala, 1997) : $Y = f(X)$. Le vecteur des attributs prédictifs, ou variables exogènes, est représenté par X. La variable à prédire, ou encore variable endogène, est notée Y.

La seconde motivation de ce choix est l'adéquation entre les types de nos variables et ceux supportés par les réseaux bayésiens. Les variables à prédire dans notre étude sont les métriques opérationnelles de variation qui sont de type continu. Les variables explicatives de notre modèle sont les métriques structurelles de changement et de variation, et les métriques

d'effort de gestion qui sont des variables continues. Pour réaliser un apprentissage sans modifier le type de nos variables, nous pouvons utiliser la régression. L'objectif de l'analyse de régression est de découvrir les relations entre la variable de réponse Y et les variables prédictives x_1, \dots, x_p . Ces relations peuvent être utilisées pour réaliser des prédictions sur Y en se basant sur des observations de x_1, \dots, x_p . Les réseaux bayésiens sont applicables à un ensemble de variables discrètes, continues ou mixtes. Un réseau bayésien, avec des variables continues uniquement, nous permet de réaliser des analyses de régression sur nos trois variables de réponse dans un même modèle.

La troisième raison du choix de l'apprentissage des réseaux bayésiens par recherche gloutonne est la possibilité d'apprendre la structure du réseau à partir des données uniquement (Camara, et al., 2007a). En effet, si une structure renfermant les relations de dépendance entre les variables est disponible, au départ, alors un apprentissage des paramètres est suffisant. Dans le cas contraire, un apprentissage de la structure est nécessaire. Dans notre étude, nous ne disposons à priori que des relations entre les dimensions du changement, de l'impact opérationnel et des moyens de gestion. Ces relations n'étant que des hypothèses issues de la littérature mais non-encore prouvées, nous ne pouvons pas nous en servir pour construire une structure de départ.

La quatrième raison de ce choix réside dans la capacité de cette approche à sélectionner les variables pour avoir une meilleure exactitude prédictive du modèle. Quand le nombre de variables explicatives est grand, il est préférable de ne pas utiliser toutes les variables dans le modèle de régression comme c'est le cas dans la méthode des moindres carrés. Ce problème est plus connu sous le nom de sélection des variables, de sélection de sous-ensemble ou de sélection de modèle. Seber, et al. (2003 p. 392), estiment que la méthode bayésienne de sélection de variables est la meilleure solution à ce problème. Cette décision est basée sur une comparaison avec plusieurs autres approches sur les critères suivants :

Evite de rechercher dans tous les réseaux possibles. L'approche *all possible regressions* est la première solution pour sélectionner un sous-ensemble de variables (Seber, et al., 2003 p. 392). Elle consiste à définir un critère de qualité du modèle, à l'évaluer sur chaque sous-ensemble possible, et à sélectionner celui qui optimise le critère⁶². Evaluer le critère pour tous les sous-ensembles possibles peut demander des calculs intensifs, ou même prohibitifs, si le nombre de variables explicatives est large.

Permet de trouver un modèle optimal selon un critère fixé. La seconde approche de sélection est d'appliquer une séquence de tests d'hypothèses au problème et d'essayer

⁶² Le critère peut être basé sur la mesure standard de qualité d'ajustement, sur l'estimation des erreurs de prédiction, sur l'estimation du nombre de coefficients de régression non-nuls, ou l'estimation d'une certaine mesure de la distance entre le modèle basé sur le sous-ensemble et le modèle réel.

d'identifier les coefficients non-nuls. Ces techniques de tests séquentiels⁶³ sont moins gourmands en calcul mais n'assurent pas une garantie que le modèle trouvé soit optimal par rapport à un critère donné (Seber, et al., 2003 p. 392). Dans les méthodes bayésiennes basées sur le facteur de bayes, il faut juste sélectionner le modèle avec la plus grande probabilité à postériori (Seber, et al., 2003 p. 428).

3.2.1. Apprentissage des réseaux bayésiens par recherche gloutonne

La recherche gloutonne⁶⁴ dans l'espace des réseaux fait partie des algorithmes proposés pour éviter la recherche exhaustive dans cet espace. Dans la recherche gloutonne, les vraisemblances à postériori⁶⁵, $p(D|d)$ et $p(D^*|d)$, sont utilisées pour comparer les scores de deux DAGs, D et D^* (Bootcher, et al., 2003A p. 13) (Bootcher, et al., 2003B). La vraisemblance à postériori ($p(D|d)$ ou $p(D^*|d)$) représente la probabilité qu'une structure donnée soit correcte sachant que les données d ont été observées.

Équation 27 Vraisemblances à postériori et facteur de bayes

$$\frac{p(D|d)}{p(D^*|d)} = \frac{p(D,d)}{p(D^*,d)} = \frac{p(D)}{p(D^*)} \times \frac{p(d|D)}{p(d|D^*)}$$

Les vraisemblances à priori⁶⁶ sont notées $p(D)$ et $p(D^*)$. Lorsque les deux DAG à comparer ont la même vraisemblance à priori, le quotient des vraisemblances à priori est égal à 1. Dans ce cas, le quotient $p(d|D)/p(d|D^*)$, appelé facteur de bayes⁶⁷, peut être utilisé à la place des vraisemblances à postériori pour la comparaison des deux DAGs D et D^* (Bootcher, et al., 2003A p. 13). La probabilité des données d sachant une structure est notée $P(d|D)$ ou $p(d|D^*)$. Sous l'hypothèse d'égalité des vraisemblances à priori, l'algorithme de recherche gloutonne suit les étapes ci-après (Bootcher, et al., 2003A p. 13) :

- 1) sélectionner un DAG initial D_0 à partir duquel la recherche va commencer,
- 2) calculer les facteurs de bayes entre D_0 et de tous les réseaux D possibles, qui diffèrent seulement d'un arc :
 - i. un arc est rajouté à D_0 ,
 - ii. un arc est supprimé dans D_0

⁶³ Parmi ces techniques de tests séquentiels, les plus connues sont le *forward selection*, *backward elimination*, et le *stepwise regression*.

⁶⁴ Greedy Search

⁶⁵ *posterior odds*

⁶⁶ *prior odds*

⁶⁷ *bayes factor*

- iii. un arc dont le sens de la flèche est inversé dans D0
- 3) parmi tous ces réseaux D, sélectionner celui qui augmente le plus le facteur de bayes, c'est celui qui a le meilleur score.
- 4) si le facteur de bayes n'est pas augmenté, ça veut dire que le score non plus n'a pas été amélioré, arrêter la recherche. Sinon le réseau D choisi devient D0 et répéter à partir de 2.

Pour que ces approches de recherche à base de score soient réalisables en pratique, le score doit être décomposable localement, c'est-à-dire s'exprimer comme la somme ou le produit de scores locaux au niveau de chaque nœud (Équation 29)⁶⁸.

Équation 28 Score réseau

$$S(D) = p(D, d) = p(d|D)p(D)$$

Équation 29 Décomposition du score

$$S(D) = \prod_{\delta \in \Delta} S_{\delta}(D) \prod_{\gamma \in \Gamma} S_{\gamma}(D)$$

Le score d'un réseau est théoriquement le produit des scores de ses nœuds (Équation 29). Dans la pratique, le logarithme du score est utilisé pour calculer le score du réseau qui devient la somme des logarithmes des scores de ses nœuds. A Chaque nœud dans D0 et dans tout réseau candidat D (lors de toutes les itérations de l'algorithme de recherche gloutonne), les données suivantes sont affectées :

- une distribution de probabilités à priori,
- des paramètres joints qu'il partage avec les autres nœuds du réseau,
- des paramètres locaux à priori.
- un score local,
- des paramètres locaux à postériori.

Ces données sont présentées dans les cinq phases présentées ci-dessous. Une application de ces différentes phases est présentée à travers un exemple, en annexe 2. Nous n'allons pas nous intéresser aux nœuds discrets car notre modèle n'en contient pas. Nous commencerons par présenter la distribution normale qui est souvent utilisée dans cette section.

3.2.1.1. Distribution normale

La distribution normale est une distribution continue, symétrique et sous forme de cloche (Figure 27). Une distribution normale est complètement définie par deux paramètres : la moyenne μ et la variance σ^2 . L'écart type de la distribution est noté $\sigma = \sqrt{\sigma^2}$. μ caractérise

⁶⁸ Les δ représentent les nœuds discrets et les γ représentent les nœuds continus.

le centre de la distribution alors que σ caractérise l'étalement des valeurs de la caractéristique autour de μ . L'aire qu'on peut obtenir sous la courbe normale permet d'évaluer les chances sur 100, ou encore la probabilité, que la valeur éventuelle de la caractéristique se situe entre deux valeurs données, ou encore qu'elle soit supérieure ou inférieure à une valeur particulière (Baillargeon, 2004). La majorité (environ 68%) des points sont situés entre ± 1 écart-type autour de la moyenne (Figure 27) (Baillargeon, 2004). Un certain pourcentage de points (environ 27%) vont se situer au-delà de 1 écart-type mais n'excèdent pas 2 écarts-types. Quelques points (moins de 5%) vont au-delà de 2 écarts-types mais n'excèdent pas 3 écarts-types.

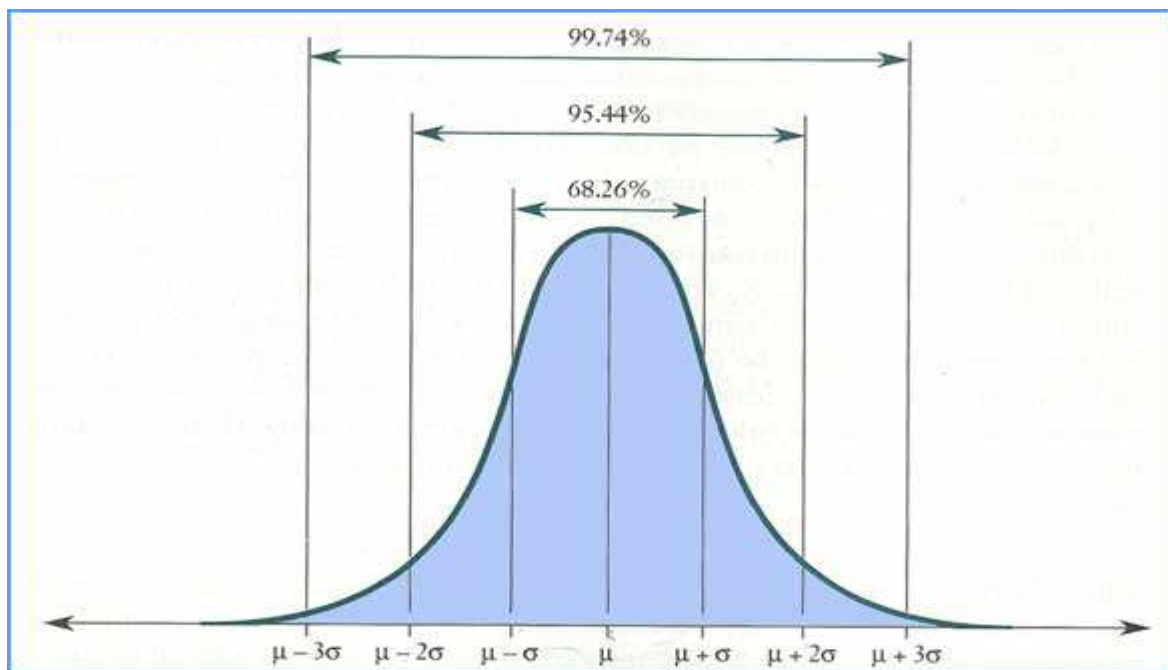


Figure 27 La distribution normale

3.2.1.2. Distributions de probabilités

Pour les nœuds continus (γ), les distributions locales de probabilités sont des régressions linéaires sur les parents continus, avec des paramètres dépendant de la configuration ($i_{pa(\gamma)}$) des parents discrets (Équation 30). Ces paramètres de régression sont obtenus par résolution de l'équation linéaire qui est une représentation mathématique de la relation entre le nœud et ses parents continus. Les $\beta_{\gamma|i_{pa(\gamma)}}$ sont les coefficients de régression, $m_{\gamma|i_{pa(\gamma)}}$ est l'« ordonnée à l'origine », et le $\sigma_{\gamma|i_{pa(\gamma)}}^2$ est la variance conditionnelle (Équation 31). L'« ordonnée à l'origine » est la valeur que prend le « nœud-fils » quand il n'est pas influencé par ses parents. En effet, un nœud parent ne peut influencer son nœud fils que s'il prend lui-même une valeur non nulle.

Équation 30 Paramètres nœud continu

$$\theta_{\gamma|i_{pa}(\gamma)} = (m_{\gamma|i_{pa}(\gamma)}, \beta_{\gamma|i_{pa}(\gamma)}, \sigma_{\gamma|i_{pa}(\gamma)}^2)$$

Équation 31 Calcul de la variance

$$S^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$$

3.2.1.3. Détermination des paramètres joints

Les lois à priori conjuguées sont utilisées dans des environnements où l'information est limitée ; car elles ne nécessitent que la détermination de quelques paramètres (Robert, 2006 p. 124). Le rôle des lois à priori conjuguées est de fournir une première approximation de la distribution à priori adéquate, qui devrait être suivie par une analyse robuste (Robert, 2006 p. 124). Les lois à priori conjuguées sont des paramètres joints partagés entre les nœuds continus d'un réseau dans une configuration donnée des parents discrets. Ils dépendent des relations d'influence entre ces nœuds continus. La moyenne et la variance des lois à priori conjuguées sont supposées suivre une distribution normale (Équation 32) et une distribution inverse-wishart (Équation 33) respectivement (Bootcher, et al., 2003A p. 9) (Colosimo, et al., 2006 p. 27).

Équation 32 Distribution de la moyenne jointe

$$(M_i | \Sigma_i) = \mathcal{N} \left(\mu_i, \frac{1}{v_i} \Sigma_i \right)$$

Équation 33 Distribution de la covariance jointe

$$p(\Sigma_i) = \mathcal{IW}(\rho_i, \Phi_i)$$

Il s'agit de définir, pour chaque configuration i des parents discrets, les paramètres nujoint (v_i), rhojoint (ρ_i), alphajoint (α_i), mujoint (μ_i), sigmajoint (Σ_i) et phijoint (Φ_i) qui permettent de caractériser ces deux distributions. L'alphajoint est le nombre d'observations dans la configuration i des parents discrets. Le rhojoint représente le degré de liberté (Colosimo, et al., 2006 p. 25). Le degré de liberté est égale au nombre d'observations indépendantes qui entrent dans l'estimation d'un paramètre auquel est soustrait le nombre de paramètre estimés⁶⁹ (Legendre, 1998 p. 13). Le nujoint peut être compris comme le nombre d'observations que nous pensons que notre connaissance à priori vaut ; avec $v_i > n$ entraînant une connaissance à priori qui domine les données et vice-versa⁷⁰ (Colosimo, et al., 2006 p. 25). Pour chaque configuration i des nœuds discrets, nujoint, rhojoint et alphajoint sont égales (Équation 34) (Bootcher, et al., 2003A p. 9). La confiance que nous avons dans les relations

⁶⁹ Incluant les paramètres intermédiaires

⁷⁰ n est la taille de la base réelle

exprimées par le réseau à priori, c'est-à-dire le réseau de départ, peut être représentée par la taille N d'une base imaginaire (Bootcher, et al., 2003A p. 8). Nous avons $N = \sum_{i \in I} \alpha_i$ (Bootcher, et al., 2003A p. 8). Le mujoint est le vecteur des moyennes. Il peut être décomposé comme la moyenne du nœud fils et celles des parents continus (Équation 35).

Équation 34 Nujoint, Rhojoint et Alphajoint

$$v_i = \rho_i = \alpha_i$$

Équation 35 Mujoint et sa décomposition

$$\bar{\mu}_i = (\bar{\mu}_\gamma, \bar{\mu}_{pa(\gamma)})$$

La matrice de covariance sigmajoint caractérise les relations d'indépendance entre les variables (Swingler, 1996 p. 28). En générale, les éléments diagonaux d'une matrice de covariance Σ sont les variances $Var(X_k)$ des variables. L'élément situé à l'intersection de la ligne k et de la colonne j est égale à la covariance $Cov(X_k, X_j)$ (Casin, 1999 p. 21). La covariance entre deux variables montre dans quelle mesure l'une des variables varie avec l'autre (Swingler, 1996 p. 28). Deux variables avec une covariance élevée sont dépendantes (Swingler, 1996 p. 28). Bootcher, et al. (2003A), calculent cette matrice d'une manière un peu différente, empruntée de Kenley (1986 p. 108). Cet algorithme repose sur la notion de diagramme d'influence qui est une représentation graphique d'un problème de décision⁷¹ (Kenley, 1989 p. 527). L'avantage de l'utilisation du diagramme d'influence est qu'il y'a moins d'éléments à calculer dans la matrice de covariance (Kenley, 1989 p. 537). Ce procédé de calcul utilise les coefficients de régression et les variances conditionnelles de chaque nœud. Soit K , l'ensemble des nœuds du réseau, chaque nœud j ($j \in K = \{1, \dots, k\}$) possède une moyenne m_j et une variance conditionnelle v_j (Équation 31). Pour chaque couple de nœuds $(i, j) \in K * K$ tel que $i < j$, b_{ij} représente le coefficient de l'arc allant de i vers j . La matrice U_j , des coefficients de propagation des changements dans les moyennes, est calculée pour chaque nœud j (Kenley, 1986 p. 108). Chaque U_{ij} est l'effet sur la moyenne de la variable j par un changement dans la moyenne de la variable i (Kenley, 1989 p. 531).

Équation 36 Matrice des coefficients de propagation des changements

$$U_j = \begin{pmatrix} I_{j-1} & B_{.j} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I_{k-j} \end{pmatrix}$$

Avec

Équation 37 Vecteur des coefficients de régression

$$B_{.j} = (b_{1,j}, \dots, b_{j-1,j})^T$$

⁷¹ Le diagramme d'influence est équivalent au graphe dirigé sans cycle.

Équation 38 Sigmajoint

$$\Sigma = U_k^T \cdots U_1^T \text{diag}(v_1, \dots, v_k) U_1 \cdots U_k$$

La matrice de covariance sigmajoint est déterminée en utilisant les U_j et les variances v_i de tous les nœuds du réseau (Équation 38) (Kenley, 1986 p. 108). Avec cet algorithme, les éléments diagonaux de la matrice s'écrivent de la manière suivante (Kenley, 1989 p. 538) :

Équation 39 Eléments diagonaux de la covariance

$$\Sigma_{ii} = v_i + \sum_{k \in \mathcal{C}(i)} \Sigma_{ik} b_{ki}$$

Le $\mathcal{C}()$ représente l'ensemble des prédécesseurs conditionnels (parents) du nœud en considération. Les éléments non-diagonaux s'obtiennent par (Kenley, 1989 p. 538) :

Équation 40 Eléments non-diagonaux de la covariance

$$\Sigma_{ij} = \Sigma_{ji} = \sum_{k \in \mathcal{C}(j)} \Sigma_{jk} b_{kj} \text{ pour } i = 1, \dots, j-1$$

Un nœud j qui n'a aucun parent, a son $\Sigma_{ij} = \Sigma_{ji} = 0$ car tous ses coefficients de régression b_{kj} qui sont nulles. Cela simplifie la matrice de covariance par rapport à sa représentation standard. Le dernier paramètre issu de la distribution conjuguée est la matrice phijoint (Équation 41). La matrice phijoint permet d'estimer la covariance réelle à travers la distribution inverse-wishart (Bootcher, et al., 2003A p. 9). Elle est déterminée à partir du jointsigama et du nujoint. Elle est décomposable en plusieurs parties selon un nœud γ et ses parents continus $pa(\gamma)$ (Équation 42).

Équation 41 Phijoint

$$\tilde{\phi}_i = (v_i - 1)\Sigma_i$$

Équation 42 Décomposition du Phijoint

$$\tilde{\phi}_i = \begin{pmatrix} \tilde{\Phi}_\gamma & \tilde{\Phi}_{\gamma, pa(\gamma)} \\ \tilde{\Phi}_{pa(\gamma), \gamma} & \tilde{\Phi}_{pa(\gamma)} \end{pmatrix}$$

L'estimation de la variance du nœud continu γ est $\tilde{\Phi}_\gamma$. L'estimation de la variance des parents continus $pa(\gamma)$ du nœud γ est $\tilde{\Phi}_{pa(\gamma)}$. L'estimation de la covariance entre le nœud continu γ et ses parents continus $pa(\gamma)$ est $\tilde{\Phi}_{\gamma, pa(\gamma)}$ ou $\tilde{\Phi}_{pa(\gamma), \gamma}$. Ces parties ainsi extraites du phijoint sont utilisées dans le calcul des paramètres locaux à priori.

3.2.1.4. Calcul des paramètres locaux à priori

Les paramètres locaux à priori ont pour objectif de définir la distribution d'un nœud continu sachant ses parents continus. En effet, le nœud continu suit une distribution normale dont la moyenne conditionnelle et la variance conditionnelle sachant ses parents continus doivent être déterminées (Équation 43) (Bootcher, et al., 2003A p. 10). Les paramètres locaux

à priori sont mu ($\mu_{\gamma|i_{pa(\gamma)}}$) (Équation 45), phi ($\Phi_{\gamma|i_{pa(\gamma)}}$) (Équation 46), rho ($\rho_{\gamma|i_{pa(\gamma)}}$) (Équation 47) et tau ($\tau_{\gamma|i_{pa(\gamma)}}$) (Équation 48) du nœud (Bootcher, et al., 2003A p. 10). Ils sont déterminés à partir des paramètres joints et de leurs décompositions.

Équation 43 Distribution des paramètres de régression

$$\left(m_{\gamma|i_{pa(\gamma)}}, \beta_{\gamma|i_{pa(\gamma)}} \middle| \sigma_{\gamma|i_{pa(\gamma)}}^2\right) \sim \mathcal{N}\left(\mu_{\gamma|i_{pa(\gamma)}}, \sigma_{\gamma|i_{pa(\gamma)}}^2 \tau_{\gamma|i_{pa(\gamma)}}^{-1}\right)$$

$$\sigma_{\gamma|i_{pa(\gamma)}}^2 \sim \mathcal{IG}\left(\frac{\rho_{\gamma|i_{pa(\gamma)}}}{2}, \frac{\Phi_{\gamma|i_{pa(\gamma)}}}{2}\right)$$

La moyenne conditionnelle du nœud sachant ses parents continus est représentée par mu (Équation 44). C'est le vecteur des paramètres de régression sur les parents continus en utilisant la base d'exemples (Bootcher, et al., 2003A p. 10) (Colosimo, et al., 2006 p. 30). Il est de dimension 1+ nombre des parents continus. Il peut être décomposé en utilisant l'ordonnée à l'origine $m_{\gamma|i_{pa(\gamma)}}$ et les coefficients de régression $\beta_{\gamma|i_{pa(\gamma)}}$ (Équation 45).

Équation 44 Mu a priori du nœud

$$\mu_{\gamma|i_{pa(\gamma)}} = \left(m_{\gamma|i_{pa(\gamma)}}, \beta_{\gamma|i_{pa(\gamma)}}\right)$$

Équation 45 Décomposition de mu a priori du nœud

$$\mu_{\gamma|i_{pa(\gamma)}} = \left(\bar{\mu}_{\gamma} - \tilde{\Phi}_{\gamma,pa(\gamma)} \tilde{\Phi}_{pa(\gamma)}^{-1} \bar{\mu}_{pa(\gamma)}, \tilde{\Phi}_{\gamma,pa(\gamma)} \tilde{\Phi}_{pa(\gamma)}^{-1}\right)$$

Une estimation de la variance conditionnelle qui tient compte de la relation d'influence des parents continus est obtenue avec phi (Équation 46) (Bootcher, et al., 2003A p. 10) (Colosimo, et al., 2006 p. 30).

Équation 46 Phi à priori du nœud

$$\Phi_{\gamma|i_{pa(\gamma)}} = \tilde{\Phi}_{\gamma} - \tilde{\Phi}_{\gamma,pa(\gamma)} \tilde{\Phi}_{pa(\gamma)}^{-1} \tilde{\Phi}_{pa(\gamma),\gamma}$$

Les degrés de liberté liés au nœud et à ses parents ($\rho_{\gamma|i_{pa(\gamma)}}$) représentent la quantité d'informations dont on dispose au moment de calculer les paramètres locaux du nœud. Lors du calcul des paramètres joints, nous avons un nombre d'observations ρ_i , correspondant à la taille de la base imaginaire qui représente la connaissance à priori. A cela, nous devons ajouter le nombre de parents continus $|pa(\gamma)|$ car il faut spécifier autant d'indices pour déterminer les paramètres locaux à priori (Équation 47) (Bootcher, et al., 2003A p. 10).

Équation 47 Rho à priori du nœud

$$\rho_{\gamma|i_{pa(\gamma)}} = \rho_i + |pa(\gamma)|$$

La matrice de précision tau (Équation 48) est utilisée dans l'estimation de la dispersion des paramètres de régression (Équation 43). Selon Geiger, et al. (1994 p. 6), il est plus pratique d'utiliser la matrice de précision $\tau_{\gamma|i_{pa(\gamma)}}$ à la place de la matrice de covariance. En

effet, selon ce dernier, spécifier directement la distribution normale d'un réseau bayésien, nécessite de garantir que la matrice de covariance calculée soit définie positive (Geiger, et al., 1994 p. 6). Cette exigence est satisfaite en transformant de manière ad hoc les corrélations fixées par l'utilisateur (Geiger, et al., 1994 p. 6). Cette matrice $\tau_{\gamma|i_{pa(\gamma)}}$ entre dans le calcul du score des nœuds continus ; elle est définie positive (Cain, et al., 1996 p. 166) (Bootcher, et al., 2003A p. 10).

Équation 48 Tau à priori du nœud

$$\tau_{\gamma|i_{pa(\gamma)}} = \begin{pmatrix} \frac{1}{v_i} + \bar{\mu}_{pa(\gamma)}^T \tilde{\Phi}_{pa(\gamma)}^{-1} \bar{\mu}_{pa(\gamma)} & -\bar{\mu}_{pa(\gamma)}^T \tilde{\Phi}_{pa(\gamma)}^{-1} \\ -\tilde{\Phi}_{pa(\gamma)}^{-1} \bar{\mu}_{pa(\gamma)} & \tilde{\Phi}_{pa(\gamma)}^{-1} \end{pmatrix}^{-1}$$

3.2.1.5. Calcul du score local

Le score est le critère de qualité qui permet de comparer deux réseaux bayésiens entre eux. Le score d'un réseau est décomposé en scores locaux calculés pour ses nœuds continus et discrets. Le calcul du score d'un nœud continu (Équation 49) (Bootcher, et al., 2003A p. 12) fait appel à $\mathcal{Y}_{\gamma|i_{pa(\gamma)}}^b$, à $\mathcal{Z}_{pa(\gamma)|i_{pa(\gamma)}}^b$, à n et aux paramètres locaux à priori du nœud présentés à la section ci-dessus. Le vecteur numérique $\mathcal{Y}_{\gamma|i_{pa(\gamma)}}^b$ contient les observations du nœud courant γ sachant la configuration $i_{pa(\gamma)}$ des parents discrets. Le nombre de lignes de la matrice $\mathcal{Z}_{pa(\gamma)|i_{pa(\gamma)}}^b$ est égal au nombre d'observations dans la base d'apprentissage. Le nombre de colonnes de cette matrice est égal à $1+k$, k étant le nombre de parents continus du nœud γ . La première colonne est remplie par des 1. Chacune des k autres colonnes est une observation complète d'un nœud parent. Le n est la taille de la base réelle.

Équation 49 Score nœud continu

$$S_{\gamma}(D) = \prod_{i_{pa(\gamma)} \in \mathcal{I}_{pa(\gamma)}} \frac{\Gamma\left(\frac{\rho_{\gamma|i_{pa(\gamma)}} + n}{2}\right)}{\Gamma\left(\frac{\rho_{\gamma|i_{pa(\gamma)}}}{2}\right) \sqrt{\det(\rho_{\gamma|i_{pa(\gamma)}} S_{\gamma|i_{pa(\gamma)}} \pi)}} \times \left[1 + \frac{1}{\rho_{\gamma|i_{pa(\gamma)}}} \mathbf{a}_{\gamma|i_{pa(\gamma)}} S_{\gamma|i_{pa(\gamma)}}^{-1} \mathbf{a}_{\gamma|i_{pa(\gamma)}}^{\top} \right]^{-\frac{\rho_{\gamma|i_{pa(\gamma)}} + n}{2}}$$

Où

$$S_{\gamma|i_{pa}(\gamma)} = \frac{\phi_{\gamma|i_{pa}(\gamma)}}{\rho_{\gamma|i_{pa}(\gamma)}} \left(I + (Z_{pa(\gamma)}^b) \tau_{\gamma|i_{pa}(\gamma)}^{-1} \tau_{\gamma|i_{pa}(\gamma)} \right)$$

$$a_{\gamma|i_{pa}(\gamma)} = y_{\gamma|i_{pa}(\gamma)}^b - Z_{pa(\gamma)|i_{pa}(\gamma)}^b \mu_{\gamma|i_{pa}(\gamma)}$$

3.2.1.6. Détermination des paramètres locaux à postériori

La détermination des paramètres locaux à postériori correspond à la mise à jour des paramètres à priori en utilisant les données de la base d'apprentissage réelle. La matrice de précision à postériori est $\tau'_{\gamma|i_{pa}(\gamma)}$ (Équation 50). Les paramètres de régression sont mises à jours dans $\mu'_{\gamma|i_{pa}(\gamma)}$ (Équation 51). La taille de la base réelle n est prise en compte dans le calcul des degrés de liberté à postériori $\rho'_{\gamma|i_{pa}(\gamma)}$ (Équation 52). La matrice de covariance est mise à jour dans $\Phi'_{\gamma|i_{pa}(\gamma)}$ (Équation 53). Ces paramètres locaux à postériori sont calculés à partir des mêmes éléments que le score.

Équation 50 Tau a posteriori du nœud

$$\tau'_{\gamma|i_{pa}(\gamma)} = \tau_{\gamma|i_{pa}(\gamma)} + \left(Z_{pa(\gamma)|i_{pa}(\gamma)}^b \right)^T Z_{pa(\gamma)|i_{pa}(\gamma)}^b$$

Équation 51 Mu à posteriori du nœud

$$\mu'_{\gamma|i_{pa}(\gamma)} = \left(\tau'_{\gamma|i_{pa}(\gamma)} \right)^{-1} \times \left(\tau_{\gamma|i_{pa}(\gamma)} \mu_{\gamma|i_{pa}(\gamma)} + \left(Z_{pa(\gamma)|i_{pa}(\gamma)}^b \right)^T y_{\gamma|i_{pa}(\gamma)}^b \right)$$

Équation 52 Rho à posteriori du nœud

$$\rho'_{\gamma|i_{pa}(\gamma)} = \rho_{\gamma|i_{pa}(\gamma)} + n$$

Équation 53 Phi à posteriori du nœud

$$\Phi'_{\gamma|i_{pa}(\gamma)} = \Phi_{\gamma|i_{pa}(\gamma)} + \left(y_{\gamma|i_{pa}(\gamma)}^b - Z_{pa(\gamma)|i_{pa}(\gamma)}^b \mu'_{\gamma|i_{pa}(\gamma)} \right)^T y_{\gamma|i_{pa}(\gamma)}^b + \left(\mu_{\gamma|i_{pa}(\gamma)} - \mu'_{\gamma|i_{pa}(\gamma)} \right)^T \tau_{\gamma|i_{pa}(\gamma)} \mu_{\gamma|i_{pa}(\gamma)}$$

3.2.1.7. Détermination de l'équation de prédiction

Les équations de prédiction des nœuds de réponse sont déterminées sur le réseau obtenu à la fin de l'algorithme de recherche. Le nœud fils ($Y_{\gamma|i_{pa}(\gamma)}$) suit une distribution normale (Bootcher, et al., 2003A) (Équation 54) de moyenne ($m_{\gamma|i_{pa}(\gamma)} + y_{pa(\gamma)} \beta_{\gamma|i_{pa}(\gamma)}$) et de variance ($\sigma_{\gamma|i_{pa}(\gamma)}^2$). Le vecteur ($y_{pa(\gamma)}$) est un composé des valeurs prises par la nouvelle observation sur les nœuds parents continus.

Équation 54 Equation de Prédiction

$$\left(Y_{\gamma|i_{pa}(\gamma)}, y_{pa(\gamma)}, \theta_{i_{pa}(\gamma)} \right) \sim \mathcal{N} \left(m_{\gamma|i_{pa}(\gamma)} + y_{pa(\gamma)} \beta_{\gamma|i_{pa}(\gamma)}, \sigma_{\gamma|i_{pa}(\gamma)}^2 \right)$$

Pour chaque nouvelle observation dont nous possédons les valeurs des variables explicatives, il faut réaliser les actions suivantes sur chaque nœud de réponse :

- calculer la moyenne pour le nœud de réponse qui dépend de l'« ordonnée à l'origine » de ce nœud, des valeurs des variables explicatives de la nouvelle observation et des coefficients de régression sur les parents continus du nœud réponse ;
- la prédiction de la valeur du nœud de réponse consiste tirer aléatoirement une valeur, en suivant une distribution normale avec la moyenne du nœud comme centre et son écart type comme indicateur de dispersion.

3.2.2. Validation croisée avec k sous-échantillons

A l'issue de l'établissement d'un modèle de classement, ou d'un prédicteur, on souhaite en général disposer d'informations relatives à la qualité du modèle qui vient d'être établi. Quand il faut appliquer le modèle sur de nouveaux individus, le biais et la variance sont sources d'erreurs et doivent donc être minimisés (Geurts, 2002). Le biais est l'erreur systématique commise par un modèle, indépendamment de l'échantillon (Geurts, 2004). La variance représente quant à elle, l'erreur due à la variabilité du modèle en fonction de l'échantillon (Geurts, 2004). Comme le biais et la variance réagissent dans des directions opposées à la variation de la flexibilité dans l'espace des hypothèses ou modèles (Geurts, 2002), il se pose un problème plus connu sous le nom de « compromis *Biais/Variance* » (Geman, et al., 1992) :

- si le modèle est très complexe, il mémorise la structure du problème mais aussi la part de bruit dans les données : c'est le sur-apprentissage⁷² (Roweis, 2003).
- si le modèle n'est pas complexe du tout, il n'est pas capable de capter la structure qui est dans les données : c'est le sous-apprentissage⁷³ (Roweis, 2003).

Nous avons choisi, dans cette étude, d'utiliser la validation croisée avec k sous-échantillons à cause des nombreux problèmes qu'elle permet de résoudre. La validation croisée avec k sous-échantillons⁷⁴ est basée sur la propriété selon laquelle, plusieurs estimations sur des portions indépendantes d'un échantillon sont de variance plus faible qu'une estimation sur la totalité. En validation croisée avec k sous-échantillons, l'échantillon D est aléatoirement subdivisé en sous-échantillons mutuellement exclusifs⁷⁵ D_1, D_2, \dots, D_K de tailles approximativement égales (Kohavi, 1995). Le modèle M_t est entraîné et testé k fois. A chaque tour $t \in \{1, 2, \dots, k\}$, il est entraîné sur $D \setminus D_t$ et testé sur D_t . L'estimateur de risque de

⁷² *overfitting* ou *model variance*

⁷³ *underfitting* ou *model bias*

⁷⁴ K-folds cross validation

⁷⁵ Les folds

M_t construit à partir de l'échantillon d'apprentissage $D \setminus D_t$ est déterminé de manière empirique sur le sous-échantillon de validation D_t (Équation 55) (Rakotomalala, 1997) :

Équation 55 Erreur expérience t

$$\varepsilon_{D_t} = \frac{\sum_{\omega \in D_t} q(M_t, \omega)}{\text{card}(D_t)}$$

$q(M, \omega)$ est une fonction indicatrice ou encore fonction de perte telle que⁷⁶ :

Équation 56 Fonction de perte cas qualitatif

$$q(M, \omega) = \begin{cases} 0 & \text{si } M(\omega) = Y(\omega) \\ 1 & \text{si } M(\omega) \neq Y(\omega) \end{cases}$$

dans le cas qualitatif (Rakotomalala, 1997) et :

Équation 57 Fonction de perte cas quantitatif

$$q(M, \omega) = (M(\omega) - Y(\omega))^2$$

dans le cas quantitatif (Chou, 1991). En faisant la moyenne des estimateurs de risque individuels sur les D_t , nous obtenons l'estimateur de risque de la validation croisée (Équation 58) (Rakotomalala, 1997) (Kohavi, 1995):

Équation 58 Erreur globale validation croisée

$$\varepsilon_{cv} = \sum_{D_t} \frac{\text{card}(D_t)}{\text{card}(D)} \varepsilon_{D_t}$$

Après l'évaluation de l'erreur en généralisation, un des modèles des k expériences doit être sélectionné. Le modèle sélectionné au terme de la validation croisée correspond à l'estimateur avec l'erreur minimale⁷⁷ (Dudoit, et al., 2003).

La validation croisée avec k sous-échantillons considère un premier échantillon (Ω^a) qui sert classiquement à construire le modèle et un second, l'échantillon de validation (Ω^v), utilisé pour évaluer les performances du modèle. Cela évite de calculer le taux d'erreurs en resubstitution ; c'est-à-dire directement sur l'échantillon Ω^a [$n = \text{card}(\Omega^a)$] ayant servi à l'apprentissage (Rakotomalala, 1997). En effet, plus le modèle est complexe, plus l'erreur en resubstitution diminue, mais plus l'erreur réelle, sur un nouveau jeu de données, risque d'augmenter pour cause de sur-ajustement

La validation croisée avec k sous-échantillons permet d'éviter aussi des problèmes liés à l'estimation de l'erreur exclusivement sur échantillon de validation. Dans le schéma d'apprentissage-validation, avec le partage courant du fichier en 70%-30% (Breiman, et al., 1984) :

⁷⁶ ω est l'observation, $M(\omega)$ la valeur estimée par le modèle et $Y(\omega)$ est la valeur réelle

⁷⁷ The cross-validated selector corresponds to the estimator with the minimum cross-validated risk.

- Le taux d'erreurs mesuré sur l'échantillon de validation est théoriquement sans biais ; en revanche sa variance est importante (Rakotomalala, 1997). Il faut un grand nombre d'individus que Catlet (1991) quantifie arbitrairement à 1000 pour réduire cette variance à un niveau raisonnable. La validation croisée avec k sous-échantillons, en utilisant toutes les données dans l'apprentissage et dans la validation, propose une solution au problème de la taille trop faible de l'échantillon de départ.
- On constate généralement que le prédicteur est sous-dimensionné à cause de la réduction de l'échantillon d'apprentissage (Rakotomalala, 1997). Le problème se pose surtout quand la taille de l'échantillon de départ est trop faible. Là aussi, l'utilisation de toutes les données dans l'apprentissage permet de dimensionner convenablement le prédicteur.
- Un autre reproche à cette méthode est qu'elle ne prend en compte que la source de variabilité en provenance de l'échantillon de validation en omettant celle issue de l'échantillon d'apprentissage (Rakotomalala, 1997). La validation croisée avec k sous-échantillons, en calculant les erreurs sur l'ensemble des données, permet de tenir compte des deux sources de variabilité.

Enfin, la validation croisée avec k sous-échantillons permet de résoudre le problème du recouvrement des fichiers de validation.

- En effet, dans des techniques comme l'apprentissage-validation répétée⁷⁸ (Rakotomalala, 1997), v modèles sont construits, chacun à partir de $n1$ individus tirés aléatoirement et validés sur les $n-n1$ individus restants. L'estimation de l'erreur est la moyenne des erreurs calculées sur les ensembles de validation (Hissilä, 2006) (Rakotomalala, 1997). A cause des tirages aléatoires, il est possible que les échantillons de validation aient des individus communs (Rakotomalala, 1997) (Hissilä, 2006). L'estimateur est toujours sans biais, mais sa variance est largement sous-évaluée parce qu'elle est calculée sur une série d'échantillons non indépendants (Rakotomalala, 1997). La validation croisée avec k sous-échantillons, par l'exclusivité mutuelle des échantillons de validation, permet de contourner ce problème.

Selon Ian, et al. (2005 p. 151), la validation croisée avec 10 sous-échantillons est la manière classique d'estimer le taux d'erreurs d'une technique d'apprentissage sur une base d'exemples particulière. Des tests étendus sur de nombreuses bases d'exemples, avec des techniques d'apprentissage différents, ont montré que 10 était à peu près le bon nombre de sous-échantillons pour obtenir la meilleure estimation de l'erreur (Ian, et al., 2005 p. 150). Il existe aussi des preuves théoriques qui soutiennent cela (Ian, et al., 2005 p. 150).

⁷⁸ Ou validation croisée de Monte Carlo (Hissilä, 2006)

Les étapes d'apprentissage et de validation du processus de datamining seront réalisées en même temps. Cet apprentissage-validation consiste en une validation croisée avec 10 sous-échantillons sur une recherche gloutonne du réseau bayésien. Nous allons réaliser dix expériences dont chacune est une application de l'algorithme de recherche gloutonne résultant en un modèle et une erreur en généralisation. Nous allons d'abord déterminer le nombre minimum d'observations nécessaires à l'apprentissage-validation avant de générer les données (voir section suivante). L'apprentissage-validation sera réalisé avec Deal, un package dédié à l'apprentissage bayésien utilisable avec R. R est un langage et un environnement logiciel libre pour les calculs et les graphiques statistiques ⁷⁹.

3.2.3. Détermination de la taille de la base

Le processus général de création d'une équation de prédiction implique la collecte de données pertinentes à partir d'un échantillon large et représentatif de la population. Les avis divergent, cependant, sur la taille minimal N requise encore appelé ratio cas-prédicteurs. Selon Tabachnik, et al. (2001 p. 117), tester une corrélation multiple requiert $N \geq 50 + 8m$ (m étant nombre de variables explicatives) et tester les prédicteurs individuellement requiert $N \geq 104 + m$. Les indications sont plus strictes pour la construction d'une équation de prédiction à cause du besoin de généraliser au-delà d'un échantillon donné (Osborne, 2000B) (Tabachnik, et al., 2001). Alors que certains auteurs suggèrent que 15 sujets (cas, observations) par prédicteur (variable explicative) sont suffisants (Park, et al., 1974) (Pedhazur, 1997), d'autres suggèrent un minimum de 40 sujets par prédicteur (Cohen, et al., 2003) (Tabachnik, et al., 2001 p. 117)⁸⁰. D'autres auteurs ont suggéré une taille totale minimum de l'échantillon, 400 par exemple pour Pedhazur (1997 p. 148).

Tabachnik, et al. (2001 p. 117) estiment, sans donner un nombre précis, qu'un échantillon plus large que 40 :1 est nécessaire pour l'analyse de régression si la validation croisée est utilisée pour tester la généralisabilité de la solution. Osborne (2000B) a mené une étude dans laquelle il applique des validations croisées avec des ratios cas-prédicteurs différents (5 :1, 15 :1, 40 :1, 100 :1 et 400 :1). Selon l'auteur (Osborne, 2000B), si l'objectif de la validation croisée est d'avoir une estimation stable et précise de l'équation de régression de la population (ce qui est le cas si l'équation doit être largement utilisée en dehors de l'échantillon original), il apparaît désirable d'avoir au moins 100 observations par

⁷⁹ <http://www.r-project.org/>

⁸⁰ Case-to-IV ratio of 40 to 1 is reasonable because statistical regression can produce solution that does not generalize beyond the sample unless the sample is large.

prédicteur⁸¹. Comme nous utilisons une validation croisée avec comme objectif d'appliquer le modèle sur de nouveaux processus, nous utiliserons ce ratio 100 :1. Notre échantillon de départ devra donc compter 1500 observations (100 pour chacune des 15 variables explicatives). La collecte de ces observations est décrite dans le chapitre suivant.

3.2.4. Conclusion

Dans ce chapitre, nous avons comme objectif de choisir les techniques pour apprendre et valider le modèle de notre étude. Parmi les techniques utilisées en gestion quantitative du changement dans les projets ERP, nous avons identifié la régression linéaire comme étant la plus proche de l'objectif et des variables de notre étude. Nous avons choisi ensuite, sur plusieurs critères, la recherche gloutonne des réseaux bayésiens pour trouver le modèle de régression le plus précis. La validation croisée avec 10 sous-échantillons a été choisie comme technique de validation. Le ratio cas prédicteurs est fixé à 100 :1. A ce stade de l'étude, nous avons toutes les informations nécessaires pour commencer la construction de l'historique et réaliser l'apprentissage et la validation du modèle.

⁸¹ If the goal is to get an accurate, stable estimate of the population regression equation (which it should be if that equation is going to be widely used outside the original sample), it appears desirable to have at least 100 subjects per predictor.

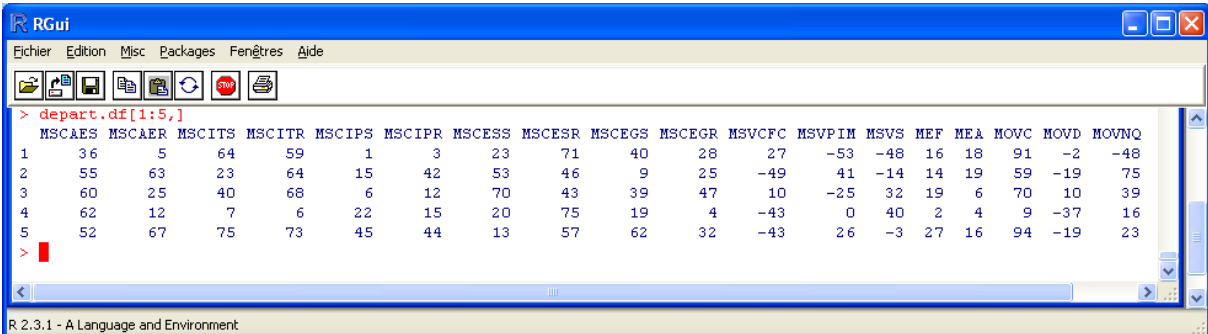
Chapitre 3.3 Collecte des données et apprentissage

3.3.1. Introduction

Ce chapitre correspond aux phases de modélisation et d'évaluation du processus de datamining. Nous présentons d'abord la création de l'échantillon qui sera utilisé pour l'apprentissage et la validation du modèle. Pour contourner les difficultés liées à la collecte des données de notre problème, la création de la base d'exemples repose essentiellement sur la génération automatique des observations. Ce chapitre constitue plus une étude de la faisabilité technique de la démarche. Il conduira à un modèle que nous pourrions utiliser pour montrer les types de décisions qui peuvent être prises. Cependant, le modèle obtenu ne sera pas utilisable dans un projet d'implémentation réel. Une validation croisée avec dix expériences est ensuite réalisée. Chaque expérience est une création d'un réseau bayésien représentant les influences entre les variables de notre étude. Nous calculerons, à la fin, l'erreur globale de la validation croisée et nous choisirons un des dix modèles obtenus dans les expériences.

3.3.2. Création de la base d'apprentissage

La collecte des données consiste à enregistrer les valeurs des métriques structurelles de changement et de variation, des métriques d'effort de gestion et des métriques opérationnelles de variation de plusieurs individus dans un historique. Ces données doivent, en principe, être collectées dans les projets ERP antérieurs. Cependant, les contraintes liées à la supervision des processus et le nombre d'observations requis rendent difficile la collecte de ces données. Pour contourner cette difficulté, nous avons donc eu recours à la création automatique des observations (Camara, et al., 2006). Nous avons donc généré aléatoirement 1499 observations. Nous les avons ensuite complétées avec l'observation correspondant à notre processus de traitement de commandes pour avoir le total de 1500 individus requis (Figure 28 et Figure 29).



The screenshot shows the RGui window with a menu bar (Fichier, Edition, Misc, Packages, Fenêtres, Aide) and a toolbar. The main window displays a data frame named 'depart.df' with 15 columns and 5 rows of data. The columns are labeled MSCAES, MSCAER, MSCITS, MSCITR, MSCIPS, MSCIPR, MSCCESS, MSCESR, MSCCEGS, MSCCEGR, MSVCFC, MSVPIM, MSVS, MEF, MEA, MOVC, MOVD, and MOVNQ. The data is as follows:

	MSCAES	MSCAER	MSCITS	MSCITR	MSCIPS	MSCIPR	MSCCESS	MSCESR	MSCCEGS	MSCCEGR	MSVCFC	MSVPIM	MSVS	MEF	MEA	MOVC	MOVD	MOVNQ
1	36	5	64	59	1	3	23	71	40	28	27	-53	-48	16	18	91	-2	-48
2	55	63	23	64	15	42	53	46	9	25	-49	41	-14	14	19	59	-19	75
3	60	25	40	68	6	12	70	43	39	47	10	-25	32	19	6	70	10	39
4	62	12	7	6	22	15	20	75	19	4	-43	0	40	2	4	9	-37	16
5	52	67	75	73	45	44	13	57	62	32	-43	26	-3	27	16	94	-19	23

Figure 28 Cinq premières observations de l'historique

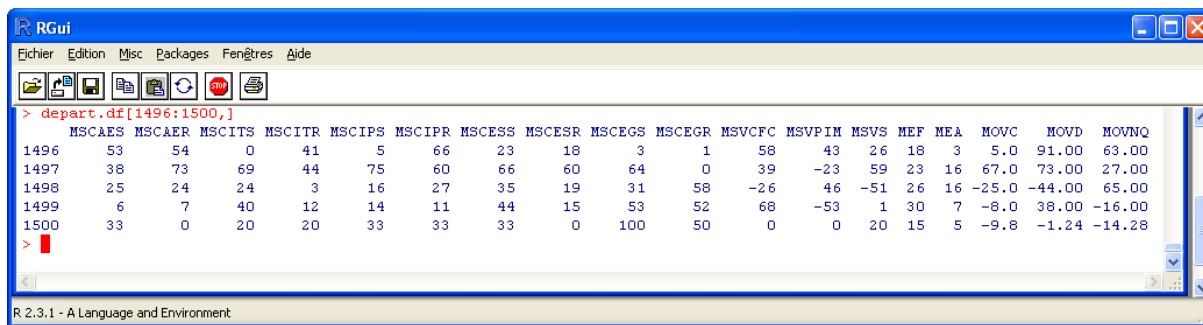


Figure 29 Cinq dernières observations de l'historique

Cet échantillon de départ est divisé en dix sous-échantillons de même taille, numérotés de 1 à 10. A chaque expérience t, le sous-échantillon de numéro t est utilisé pour le test, le modèle étant appris sur les sous-échantillons restants (Figure 30).

	1-150	151-300	301-450	451-600	601-750	751-900	901-1050	1051-1200	1201-1350		test
exp1	10		2	3	4	5	6	7	8	9	1
exp2	1	10		3	4	5	6	7	8	9	2
exp3	1	2	10		4	5	6	7	8	9	3
exp4	1	2	3	10		5	6	7	8	9	4
exp5	1	2	3	4	10		6	7	8	9	5
exp6	1	2	3	4	5	10		7	8	9	6
exp7	1	2	3	4	5	6	10		8	9	7
exp8	1	2	3	4	5	6	7	10		9	8
exp9	1	2	3	4	5	6	7	8	10		9
exp10	1	2	3	4	5	6	7	8	9		10

Figure 30 Création des sous-échantillons à partir de la base de départ

Nous commençons par utiliser la fonction « read.table » de Deal pour créer un « data-frame » et y charger le fichier texte contenant la base de départ (Figure 31). Des commandes R sont exécutées afin de créer, pour chaque expérience, l'échantillon d'apprentissage et celui de test (Figure 31).

```

dispatching - Bloc-notes
Fichier Edition Format Affichage ?
#chargement de la base d'apprentissage de l'expérience exp1
depart.df <- read.table("C:\\AAUBUNTU\\EXCELL FILES\\UBUN\\BASEDEPART.txt", header=TRUE)

#création de la base de test pour l'expérience 1 à partir de la base totale

exp1.df <-depart.df[1:1350,]
exp1.df [1:150,] <-depart.df[1351:1500,]
testexp1<- depart.df[1:150,]

resexp1<-testexp1

#création de la base de test pour l'expérience 2 à partir de la base totale
exp2.df <-depart.df[1:1350,]
exp2.df [151:300,] <-depart.df[1351:1500,]
testexp2<- depart.df[151:300,]
resexp2<-testexp2

#création de la base de test pour l'expérience 1 à partir de la base totale
exp3.df <-depart.df[1:1350,]
exp3.df [301:450,] <-depart.df[1351:1500,]
testexp3<- depart.df[301:450,]
resexp3<-testexp3

```

Figure 31 Création des bases d'apprentissage et de test pour les 10 expériences

La base de test de la première expérience comprend les 150 premiers éléments de la base de départ (Figure 32 et Figure 33).

```

RGui
Fichier Edition Misc Packages Fenêtres Aide
testexp1[1:5,]

```

	MSCAES	MSCAER	MSCITS	MSCITR	MSCIPS	MSCIPR	MCESS	MCESR	MCEGS	MCEGR	MSVFC	MSVPIM	MSVS	MEF	MEA	MOVC	MOVD	MOVNO
1	36	5	64	59	1	3	23	71	40	28	27	-53	-48	16	18	91	-2	-48
2	55	63	23	64	15	42	53	46	9	25	-49	41	-14	14	19	59	-19	75
3	60	25	40	68	6	12	70	43	39	47	10	-25	32	19	6	70	10	39
4	62	12	7	6	22	15	20	75	19	4	-43	0	40	2	4	9	-37	16
5	52	67	75	73	45	44	13	57	62	32	-43	26	-3	27	16	94	-19	23

Figure 32 Cinq premiers éléments base de test expérience 1

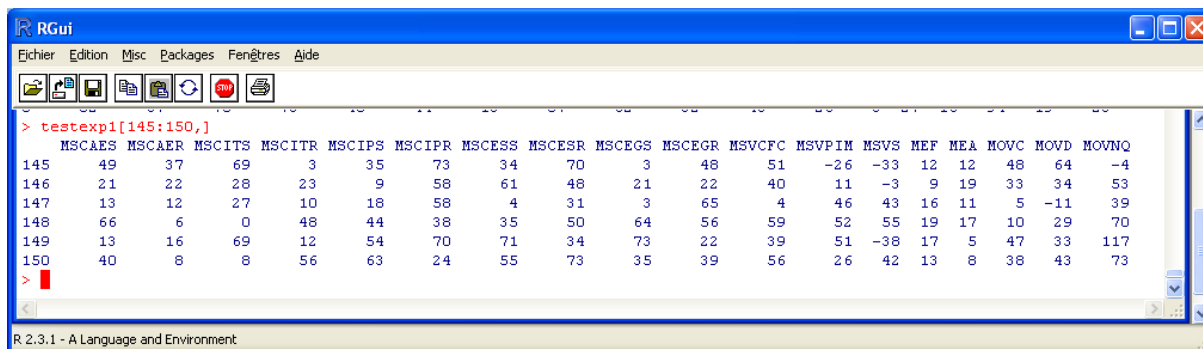


Figure 33 Cinq derniers éléments base de test expérience 1

La base d'apprentissage de l'expérience 1 contient les 1350 observations des sous-échantillons 2 à 10 de la base départ (Figure 34 et Figure 35).

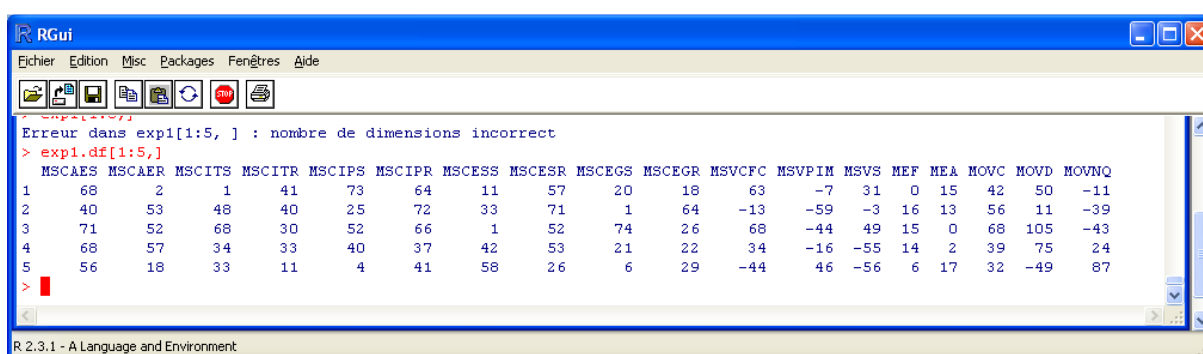


Figure 34 Cinq premiers éléments base d'apprentissage expérience 1

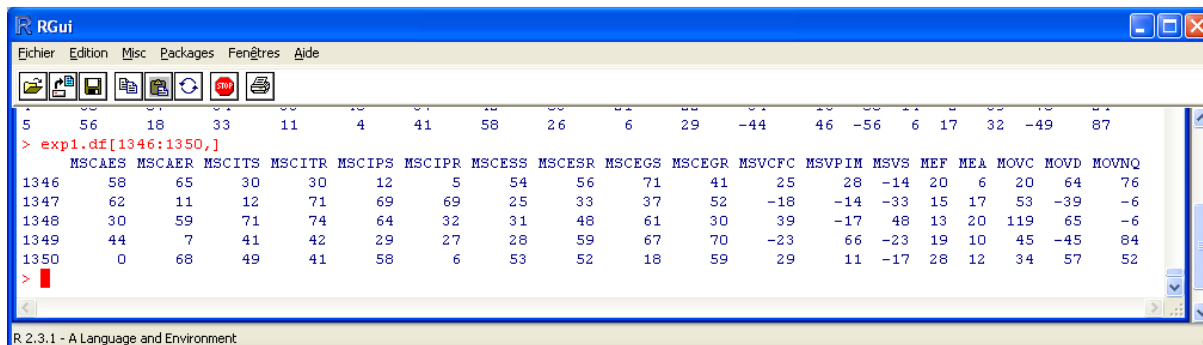


Figure 35 Cinq derniers éléments base d'apprentissage expérience 1

3.3.3. La validation croisée avec dix sous-échantillons

L'objectif de cette section est de fournir un modèle bayésien validé qui pourra être utilisé sur de nouveaux processus. La précision du modèle est assurée par deux niveaux de sélection. Le premier niveau est celui de la recherche gloutonne qui cherche les variables à intégrer au modèle à partir d'une base d'apprentissage donnée. Ce premier niveau représente la modélisation dans le processus de datamining. Le second niveau est celui de la validation croisée qui cherche un modèle parmi plusieurs construits à partir de jeux de données différents. Chacun de ces modèles est construit par une recherche gloutonne. Ce niveau de sélection correspond à la phase de validation dans le processus de datamining. La validation croisée se

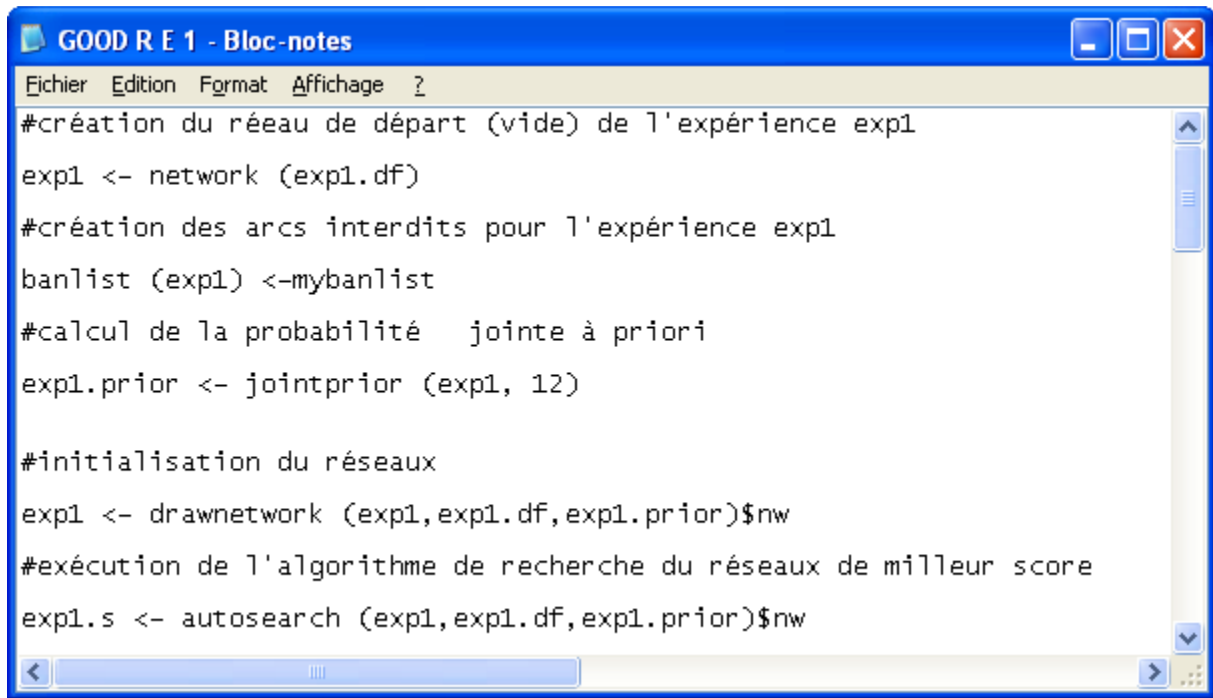
fait en même temps que l'apprentissage qu'elle englobe. Dans cette étude, nous appliquons la validation croisée avec dix sous-échantillons. Nous menons ainsi dix expériences qui donnent lieu, chacune, à la création d'un modèle et au calcul de son erreur. Les bases d'apprentissage et de validation de ces expériences ont été présentées dans la section précédente.

Les métriques opérationnelles de variation dans le coût, le délai et la non-qualité sont les variables de réponse des modèles à apprendre (Camara, et al., 2007a). Les variables structurelles de changement et de variation sont toutes calculables automatiquement à partir des modèles « en-cours » et « en-devenir » (ou « implémentés ») des processus. Les variables correspondant à l'effort de gestion du changement sont fixés par l'utilisateur du modèle. Ces trois catégories de métriques sont les variables explicatives des modèles à construire (Camara, et al., 2007a). Nous n'avons donc pas besoin de chercher l'influence qu'elles subissent, mais plutôt celle qu'elles exercent. Les arcs sont donc dirigés des variables explicatives vers les variables de réponse. Pour spécifier cela, nous avons créé un *banlist* qui est une matrice de deux colonnes spécifiant les arcs interdits dans le graphe. Dans notre *banlist*, les arcs vers les variables structurelles de changement et de variation, et vers les variables d'effort de gestion, sont interdits. Nous l'avons ensuite affecté à notre graphe grâce à la fonction *banlist* (Figure 36). Nous avons écrit une procédure qui permet de créer automatiquement ce *banlist*⁸² (Annexe 3 : Programme *banlist*). Notons que ce *banlist* est valable pour les dix expériences.

3.3.3.1. Déroulement de l'expérience 1

Dans la première expérience, le modèle est appris avec les 9 derniers sous-échantillons (Figure 34 et Figure 35) puis testé sur le premier sous-échantillon (Figure 32 et Figure 33). L'apprentissage de l'expérience 1 commence par la création d'un réseau bayésien vide en utilisant la fonction « *network* » sur la base d'apprentissage (Figure 36). Cela signifie que nous n'avons aucune certitude sur les relations entre les variables au départ et que nous envisageons de tout apprendre à partir des données. Notre réseau de départ contient à présent les nœuds, les arcs interdits et aucun arc normal.

⁸² Le *banlist* lui-même est trop long à afficher



```
Fichier  Edition  Format  Affichage  ?
#création du réseau de départ (vide) de l'expérience exp1
exp1 <- network (exp1.df)
#création des arcs interdits pour l'expérience exp1
banlist (exp1) <-mybanlist
#calcul de la probabilité jointe à priori
exp1.prior <- jointprior (exp1, 12)

#initialisation du réseaux
exp1 <- drawnetwork (exp1,exp1.df,exp1.prior)$nw
#exécution de l'algorithme de recherche du réseaux de meilleur score
exp1.s <- autosearch (exp1,exp1.df,exp1.prior)$nw
```

Figure 36 Recherche gloutonne de l'expérience 1

Nous générons ensuite les paramètres des distributions à priori de chaque nœud avec la fonction *jointprior* (Figure 36). Nous spécifions pour cette fonction la taille de la base de données imaginaire. Nous utilisons ensuite la fonction *drawnetwork* qui permet de spécifier un réseau bayésien à travers une interface graphique. Nous précisons à cette fonction le réseau à éditer, le data-frame utilisé pour l'apprentissage du réseau et les paramètres à priori générés par la fonction *jointprior*. *Drawnetwork* permet d'ajouter des arcs en cliquant successivement sur leurs nœuds source et destination. Nous allons arrêter l'édition sans ajouter un arc car nous n'avons pas de connaissance à priori sur les relations entre les nœuds. Pour terminer, nous allons lancer la méthode *autosearch* qui exécute une recherche gloutonne afin de trouver le réseau le plus probable au vu des données (Figure 36).

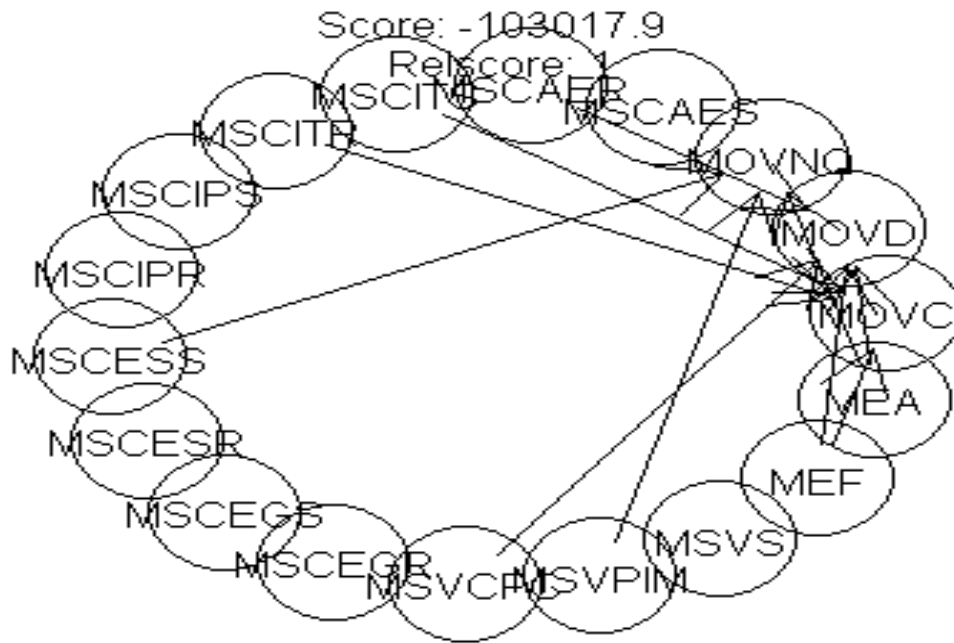


Figure 37 Réseau bayésien expérience 1

Le résultat de la recherche est enregistré dans `expl.s` qui est le modèle de l'expérience 1 (Figure 37) comme spécifié dans la commande. Pour le réafficher nous n'avons plus qu'à utiliser la fonction `plot` sur `expl.s`. « `Relscore = 1` » indique que ce réseau a le meilleur score. La commande composée du nom du réseau uniquement (ici `expl.s`) permet de voir la structure du réseau de manière non graphique, mais très lisible (Figure 38). Pour chaque nœud, nous avons son indice, son nom, son type (continu ou discret) et la liste des indices de ses nœuds parents.

```
> expl.s
## 18 ( 0 discrete+ 18 ) nodes;score= -103017.940127804 ;relscor$
1 MSCAES continuous ()
2 MSCAER continuous ()
3 MSCITS continuous ()
4 MSCITR continuous ()
5 MSCIPS continuous ()
6 MSCIPR continuous ()
7 MSCESS continuous ()
8 MSCESR continuous ()
9 MSCEGS continuous ()
10 MSCEGR continuous ()
11 MSVCFC continuous ()
12 MSVPIM continuous ()
13 MSVS continuous ()
14 MEF continuous ()
15 MEA continuous ()
16 MOVN continuous () 3 4 14
17 MOVD continuous () 2 11 14 15
18 MOVNQ continuous () 7 12 15
>
```

Figure 38 Détails du réseau bayésien expérience 1

La structure du réseau nous renseigne sur les relations de dépendance entre les variables de notre problème. L'interprétation de ce réseau est la suivante :

Pour la variation dans le coût :

- La variable de variation dans le coût a pour parent la variable MSCITS.
 - Cela signifie que l'hypothèse H3c (la suppression d'informations à traiter va influencer la variation du coût du processus) est vraie.
- La variable de variation dans le coût a pour parent la variable MSCITR.
 - Cela signifie que l'hypothèse H4c (le rajout d'informations à traiter va influencer la variation du coût du processus) est vraie.
- La variable de variation dans le coût a pour parent la variable MEF.
 - Cela signifie que l'hypothèse H14c (l'effort de formation va influencer la variation du coût du processus) est vraie.

Pour la variation dans le délai :

- La variable de variation dans le délai a pour parent la variable MSCAER.
 - Cela signifie que l'hypothèse H2d (le rajout d'activités à exécuter va influencer la variation de la durée du processus) est vraie.
- La variable de variation dans le délai a pour parent la variable MSVCFC.
 - Cela signifie que l'hypothèse H11d (la variation dans la complexité du flux de contrôle va influencer la variation de la durée du processus) est vraie.
- La variable de variation dans le délai a pour parent la variable MEF.
 - Cela signifie que l'hypothèse H14d (l'effort de formation va influencer la variation de la durée du processus) est vraie.
- La variable de variation dans le délai a pour parent la variable MEA.
 - Cela signifie que l'hypothèse H15d (l'effort d'accompagnement va influencer la variation de la durée du processus) est vraie.

Pour la variation dans la non-qualité :

- La variable de variation dans la non-qualité a pour parent la variable MSCESS
 - Cela signifie que l'hypothèse H7nq (la suppression d'événements à surveiller va influencer la variation de la non-qualité du processus) est vraie.
- La variable de variation dans la non-qualité a pour parent la variable MSVPIM.
 - Cela signifie que l'hypothèse H12nq (la variation dans la profondeur d'imbrication maximale va influencer la variation de la non-qualité du processus) est vraie.

- La variable de variation dans la non-qualité a pour parent la variable MEA.
 - Cela signifie que l'hypothèse H15nq (l'effort d'accompagnement va influencer la variation de la non-qualité du processus) est vraie.

L'affichage des paramètres d'un nœud donné, nous permet de connaître le poids selon lequel chacun de ses parents l'influence. La commande `expl.s$nodes$MOVC$prob` appliqué aux nœuds réponse nous permet d'accéder à leurs paramètres (Figure 39, Figure 40 et Figure 41). Sur la colonne [1], nous avons la variance conditionnelle du nœud. Sur la colonne [2], nous avons l'« *ordonnée à l'origine* », premier paramètre de la régression linéaire du nœud sur ses parents continus. A partir de la colonne [3], nous avons les coefficients de régression correspondant aux parents continus dans l'ordre présenté dans l'affichage de la structure du réseau (Figure 38).

```

R Console
> expl.s$nodes$MOVC$prob
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 20.42848 0.2977114 0.9919945 0.9924012 -1.982475
  
```

Figure 39 Paramètres nœud MOVC réseau expérience 1

Pour la métrique opérationnelle de variation dans le coût, l'interprétation des paramètres (Figure 39) est la suivante :

- la moyenne attendue pour la variable de variation dans le coût (MOVC) est égale à 0.2977114 avant la prise en compte de l'influence de ses nœuds parents ;
- la moyenne de MOVC évolue de 0.9919945 pour chaque suppression d'information à traiter (MSCITS) ;
- la moyenne de MOVC évolue de 0.9924012 pour chaque rajout d'information à traiter (MSCITR) ;
- la moyenne de MOVC évolue de -1.982475 pour chaque unité d'effort de formation (MEF) ;
- la dispersion des valeurs de MOVC autour de sa moyenne est spécifiée par sa variance 20.42848.

La moyenne de la variable MOVC⁸³ et la variance permettent de déterminer son équation de régression :

$$MOVC \sim \mathcal{N}(0.298 + 0.992 \times MSCITS + 0.992 \times MSCITR - 1.983 \times MEF, 20.428)$$

⁸³ incluant l'influence de ses parents

```

R Console
> exp1.s$nodes$MOVD$prob
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
[1,] 33.42755 0.2217733 0.9907646 0.9918613 -0.9932984 -0.9915775

```

Figure 40 Paramètres nœud MOVD réseau expérience 1

Pour la métrique opérationnelle de variation dans le délai, l'interprétation des paramètres (Figure 40) est la suivante :

- la moyenne attendue pour la variable de variation dans le délai (MOVD) est égale à 0.2217733 avant la prise en compte de l'influence de ses nœuds parents ;
- la moyenne de MOVD évolue de 0.9907646 pour chaque rajout d'activité à exécuter (MSCAER) ;
- la moyenne de MOVD évolue de 0.9918613 pour chaque unité de complexité dans le flux de contrôle (MSVCFC) ;
- la moyenne de MOVD évolue de -0.9932984 pour chaque unité d'effort de formation (MEF) ;
- la moyenne de MOVD évolue de -0.9915775 pour chaque unité d'effort d'accompagnement (MEA) ;
- la dispersion des valeurs de MOVD autour de sa moyenne est spécifiée par sa variance 33.42755.

Les informations précédentes nous permettent de construire l'équation de régression de MOVD :

$$MOVD \sim \mathcal{N}(0.222 + 0.991 \times MSCAER + 0.992 \times MSVCFC - 0.993 \times MEF - 0.992 \times MEA, 33.428)$$

```

R Console
> exp1.s$nodes$MOVNQ$prob
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 31.34093 0.1900695 0.9915345 0.9917212 -0.9868925

```

Figure 41 Paramètres nœud MOVNQ réseau expérience 1

Pour la métrique opérationnelle de variation dans la non-qualité, l'interprétation des paramètres (Figure 41) est la suivante :

- la moyenne attendue pour la variable de variation dans la non-qualité (MOVNQ) est égale à 0.1900695 avant la prise en compte de l'influence de ses nœuds parents ;

- la moyenne de MOVNQ évolue de 0.9915345 pour chaque suppression d'événement à surveiller (MSCESS) ;
- la moyenne de MOVNQ évolue de 0.9924012 pour chaque unité de profondeur d'imbrication maximale (MSVPIM) ;
- la moyenne de MOVNQ évolue de -0.9868925 pour chaque unité d'effort d'accompagnement (MEA) ;
- la dispersion des valeurs de MOVNQ autour de sa moyenne est spécifiée par sa variance 31.34093.

Les informations précédentes nous permettent de construire l'équation de régression de MOVNQ

$$MOVNQ \sim \mathcal{N}(0.190 + 0.992 \times MSCESS + 0.992 \times MSVPIM - 0.987 \times MEA, 31.341)$$

L'apprentissage du modèle pour l'expérience 1 est terminé. Il faut maintenant le tester sur la base de test de l'expérience 1, afin de calculer son erreur. Il s'agit d'abord de réaliser les prédictions sur les observations de cet échantillon pour les trois nœuds de réponse (Figure 43). Pour cela, nous avons écrit une boucle de prédiction en R qui correspond à l'application des équations de régression des nœuds de réponse sur les lignes de la base de test de l'expérience 1 (Figure 42).

```
GOOD R E 1 - Bloc-notes
Eichier Edition Format Affichage ?

#***** prediction des noeuds réponse pour sur la base de test expérience 1*****
for (i in 1:150){
#prediction de MOVc
moyenneMOVc <- exp1.s$nodes$MOVc$prob[2]
if (length(exp1.s$nodes$MOVc$prob)> 2) {
for (x in 3:length(exp1.s$nodes$MOVc$prob)) moyenneMOVc <- moyenneMOVc + exp1.s$nodes$MOVc$prob[x]*resexp1[i, exp1.s$nodes$MOVc$parents[x-2]]}

resexp1[i, 16]<- round(rnorm(1, moyenneMOVc ,sqrt(exp1.s$nodes$MOVc$prob[1])), 1)
#prediction noeud MOVD
moyenneMOVD <- exp1.s$nodes$MOVD$prob[2]
if (length(exp1.s$nodes$MOVD$prob)> 2) {
for (x in 3:length(exp1.s$nodes$MOVD$prob)) moyenneMOVD <- moyenneMOVD + exp1.s$nodes$MOVD$prob[x]*resexp1[i, exp1.s$nodes$MOVD$parents[x-2]]}

resexp1[i, 17]<- round(rnorm(1, moyenneMOVD ,sqrt(exp1.s$nodes$MOVD$prob[1])), 1)
|
#prediction de MOVNQ
moyenneMOVNQ <- exp1.s$nodes$MOVNQ$prob[2]
if (length(exp1.s$nodes$MOVNQ$prob)> 2) {
for (x in 3:length(exp1.s$nodes$MOVNQ$prob)) moyenneMOVNQ <- moyenneMOVNQ + exp1.s$nodes$MOVNQ$prob[x]*resexp1[i, exp1.s$nodes$MOVNQ$parents[x-2]]}

resexp1[i, 18]<- round(rnorm(1, moyenneMOVNQ ,sqrt(exp1.s$nodes$MOVNQ$prob[1])), 1)
}
}
```

Figure 42 Commandes R pour la prédiction sur l'échantillon de test

De nouvelles valeurs de nœuds de réponse sont obtenues après l'exécution de ces lignes de code (Figure 43).

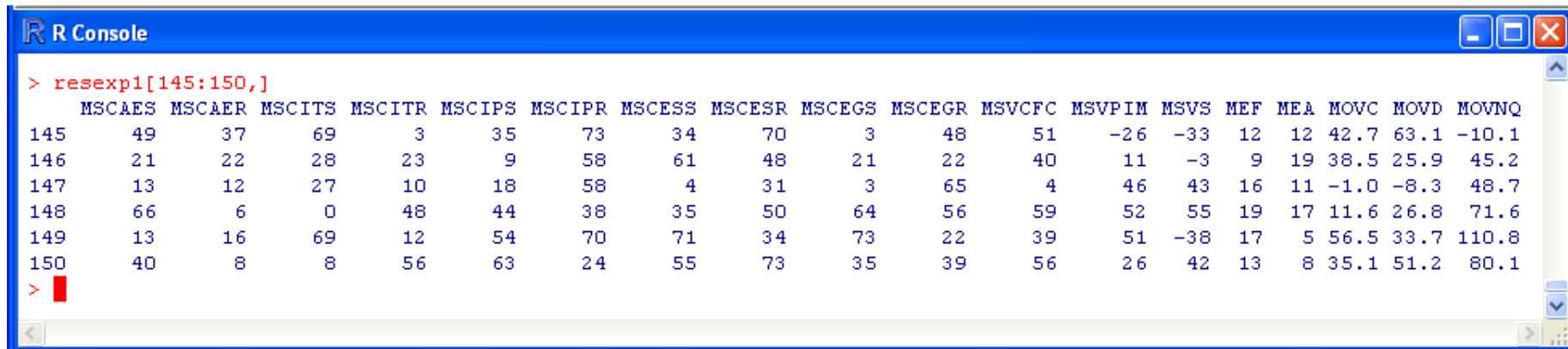


Figure 43 Résultats des prédictions sur l'échantillon de test

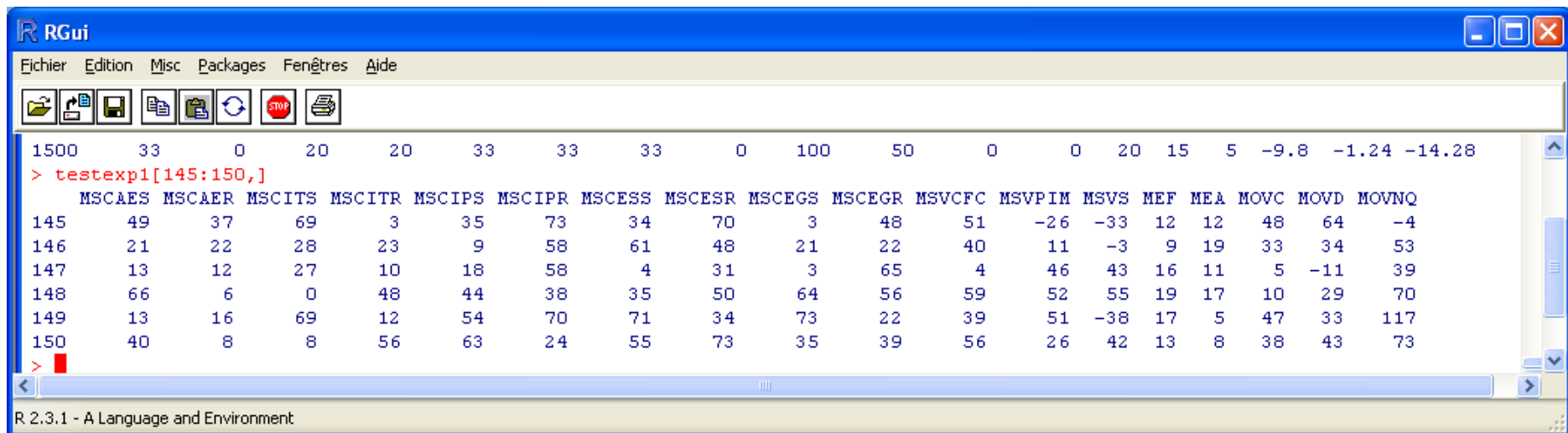


Figure 44 Valeurs réelles dans l'échantillon de test

Dans le cadre de la validation croisée, la finalité de l'évaluation des modèles des différentes expériences est de les comparer pour en choisir un à la fin. Nous avons créé une matrice d'erreurs dans laquelle nous enregistrons, pour chaque expérience, les erreurs des trois nœuds de réponse et l'erreur totale du modèle de l'expérience en question (erreurvc Figure 45). Pour le nœud MOVC par exemple, nous devons considérer d'abord les valeurs d'origine dans l'échantillon de test représentées par la colonne 16 de cet échantillon (testexp1 Figure 44). Nous prenons ensuite la colonne 16 de la matrice de résultat (resexp1 Figure 43) qui représente les valeurs prédites par le modèle.

Le calcul de l'erreur commence par l'application de la fonction de perte qui fait l'estimation de la différence entre la valeur du nœud dans l'échantillon de test et celle prédite par le modèle. Dans la pratique, cela consiste à créer un vecteur par application de l'opérateur de soustraction entre les colonnes 16 de la base de test (testexp1) et celle de la matrice résultat (resexp1). Le produit vectoriel entre ce vecteur différence et sa transposée donne la somme des carrés des erreurs qui divisée par le nombre d'observations donne l'erreur sur le nœud MOVC (Figure 45) (Équation 55 Erreur expérience t). Cette erreur est enregistrée à la colonne erreur_MOVC de la ligne « expérience 1 » de la (Figure 46). Les erreurs des nœuds MOVD et MOVNQ sont déterminées (Figure 45) et enregistrées (Figure 46) de la même manière que pour le nœud MOVC. La moyenne de ces trois erreurs donne l'erreur globale de l'expérience 1 qui elle est enregistrée à la colonne « erreur_expérience » de la ligne « expérience 1 » de la matrice d'erreurs (Figure 45 et Figure 46).

```
GOOD R E 1 - Bloc-notes
Fichier Edition Format Affichage ?
*****
# pour comparer les predictions aux valeurs réelles.

erreurvc["exp1", "erreur_MOVC"] <-t(testexp1[,16]- resexp1[,16])%*(testexp1[,16]- resexp1[,16])/150
erreurvc["exp1", "erreur_MOVD"]<-t(testexp1[,17]- resexp1[,17])%*(testexp1[,17]- resexp1[,17])/150
erreurvc["exp1", "erreur_MOVNQ"]<-t(testexp1[,18]- resexp1[,18])%*(testexp1[,18]- resexp1[,18])/150
erreurvc["exp1", "erreur_expérience"]<- (erreurvc["exp1", "erreur_MOVC"] + erreurvc["exp1", "erreur_MOVD"] + erreurvc["exp1", "erreur_MOVNQ"])/3
erreurvc
```

Figure 45 Enregistrement des erreurs nœuds de réponse et expérience 1

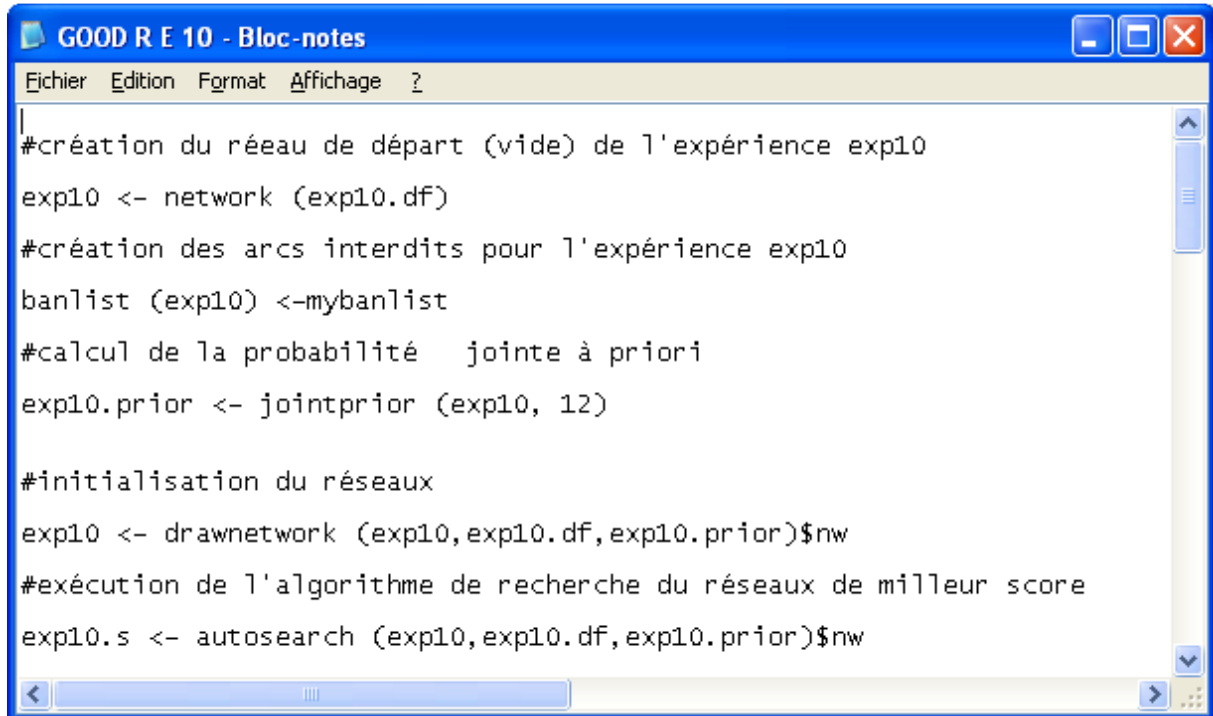
```
RGui
Fichier Edition Misc Packages Fenêtres Aide
[Icons]
> erreurvc["exp1",]
      erreur_MOVC      erreur_MOVD      erreur_MOVNQ erreur_expérience
      19.76300      34.36227      29.33753      27.82093
> █
R 2.3.1 - A Language and Environment
```

Figure 46 Affichage des erreurs nœuds de réponse et expérience 1

3.3.3.2. Déroulement de l'expérience 10

Pour l'apprentissage du réseau bayésien dans l'expérience 10, nous allons appliquer les mêmes commandes R (Figure 47) que pour l'expérience 1. Nous utiliserons le dixième sous-échantillon de l'historique de départ pour tester le modèle. Les données d'apprentissage sont les sous-échantillons restants.

Figure 47 Recherche gloutonne expérience 10



```
GOOD R E 10 - Bloc-notes
Fichier  Edition  Format  Affichage  ?
|
#création du réseau de départ (vide) de l'expérience exp10
exp10 <- network (exp10.df)
#création des arcs interdits pour l'expérience exp10
banlist (exp10) <-mybanlist
#calcul de la probabilité jointe à priori
exp10.prior <- jointprior (exp10, 12)
#initialisation du réseaux
exp10 <- drawnetwork (exp10,exp10.df,exp10.prior)$nw
#exécution de l'algorithme de recherche du réseaux de meilleur score
exp10.s <- autosearch (exp10,exp10.df,exp10.prior)$nw
```

Le réseau obtenu dans l'expérience 10 (Figure 48 et Figure 49) a la même structure que celui de l'expérience 1. Cela signifie que ces deux expériences permettent de valider les mêmes hypothèses sur les variables.

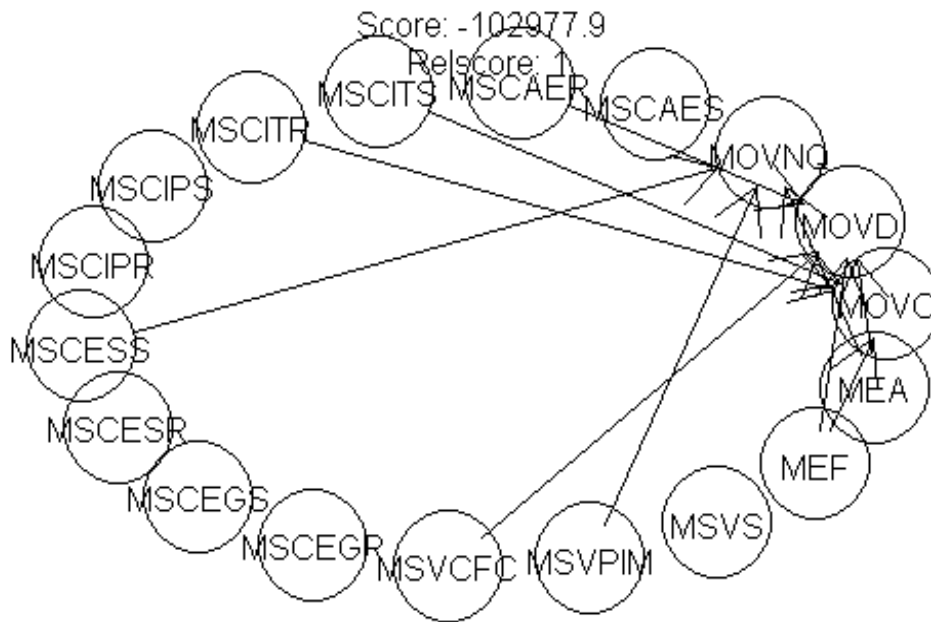


Figure 48 Réseau expérience 10

```

R RGui
Fichier Edition Misc Packages Fenêtres Aide
18 MOVNQ continuous()
> exp10.s
## 18 ( 0 discrete+ 18 ) nodes;score= -102977.928248458 ;relscore= 1
1 MSCAES continuous()
2 MSCAER continuous()
3 MSCITS continuous()
4 MSCITR continuous()
5 MSCIPS continuous()
6 MSCIPR continuous()
7 MSCESS continuous()
8 MSCESR continuous()
9 MSCEGS continuous()
10 MSCEGR continuous()
11 MSVCFC continuous()
12 MSVPIM continuous()
13 MSVS continuous()
14 MEF continuous()
15 MEA continuous()
16 MOVNC continuous() 3 4 14
17 MOVD continuous() 2 11 14 15
18 MOVNQ continuous() 7 12 15
>
R 2.3.1 - A Language and Environment

```

Figure 49 Relations réseau expérience 10

Les paramètres des nœuds de réponse de l'expérience 10 (Figure 50) ont des signes identiques à celles des paramètres de l'expérience 1.

```

RGui - [R Console]
Fichier Edition Misc Packages Fenêtres Aide

> exp10.s$nodes$MOVC$prob
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 20.28931 0.3601684 0.9914481 0.991731 -1.98247
> exp10.s$nodes$MOVD$prob
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
[1,] 32.96347 0.1561402 0.9917087 0.991946 -0.9933697 -0.9901932
> exp10.s$nodes$MOVNQ$prob
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 29.96271 0.2697958 0.9913367 0.9916287 -0.9904522
>
R 2.3.1 - A Language and Environment

```

Figure 50 Paramètres réseau expérience 10

Les équations de prédiction de l'expérience 10 sont déduites des paramètres des nœuds de réponse (Figure 50).

$$MOVC \sim \mathcal{N}(0.360 + 0.991 \times MSCITS + 0.992 \times MSCITR - 1.982 \times MEF, 20.289)$$

$$MOVD \sim \mathcal{N}(0.156 + 0.992 \times MSCAER + 0.992 \times MSVCFC - 0.993 \times MEF - 0.990 \times MEA, 32.963)$$

$$MOVNQ \sim \mathcal{N}(0.2698 + 0.991 \times MSCCESS + 0.992 \times MSVPIM - 0.990 \times MEA, 29.963)$$

Le modèle de l'expérience 10 est évalué sur le dixième sous-échantillon de la base de départ. Les équations de prédiction sont appliquées sur les observations de cet échantillon en utilisant des commandes similaires à celles utilisées pour la première expérience. Pour évaluer le modèle, nous commençons, là aussi, par comparer les valeurs des nœuds de réponse dans la base de test (Figure 51) à celles issues de l'application des équations de prédiction (Figure 52). Les erreurs sur chaque nœud de réponse et sur le modèle dans son ensemble sont calculées suivant le même procédé que pour la première expérience. Ces erreurs sont ensuite enregistrées dans la ligne « expérience_10 » de la matrice d'erreurs (Figure 53).

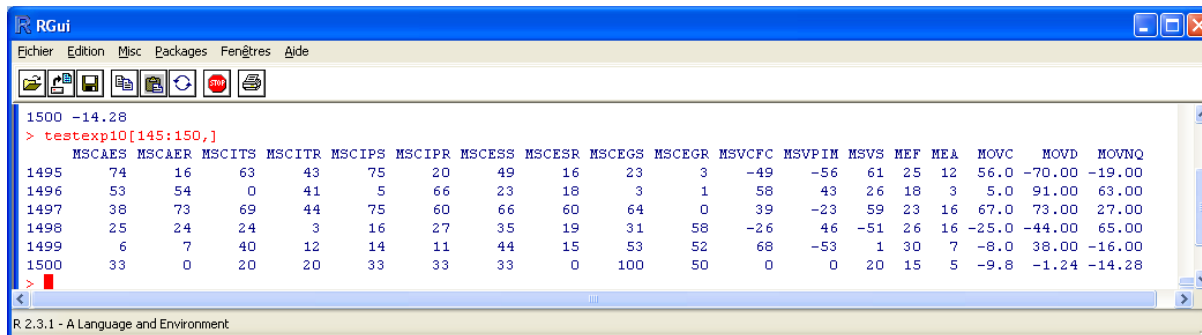


Figure 51 Echantillon de test expérience 10

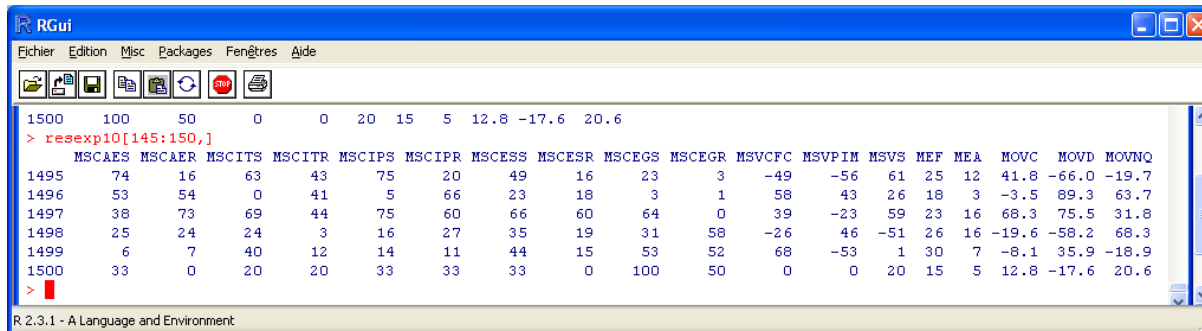


Figure 52 Résultat prédiction expérience 10

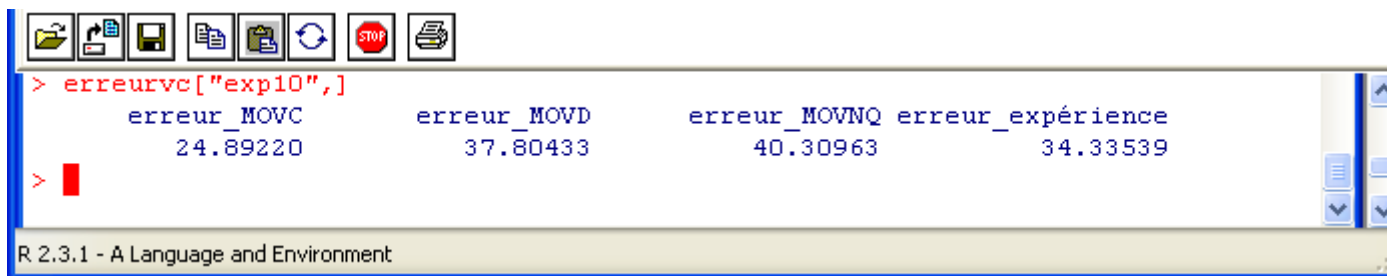


Figure 53 Erreurs expérience 10

3.3.4. Calcul de l'erreur globale et choix du modèle

La dernière étape de la modélisation est le choix du modèle de plus faible erreur parmi ceux générés lors des expériences (Figure 54). Le modèle choisi est le résultat final de l'apprentissage-validation et constitue un outil de prédiction pour les nouveaux individus. Le modèle de plus faible erreur est celui de la première expérience. Il sera donc utilisé dans le chapitre suivant pour montrer l'application du modèle.

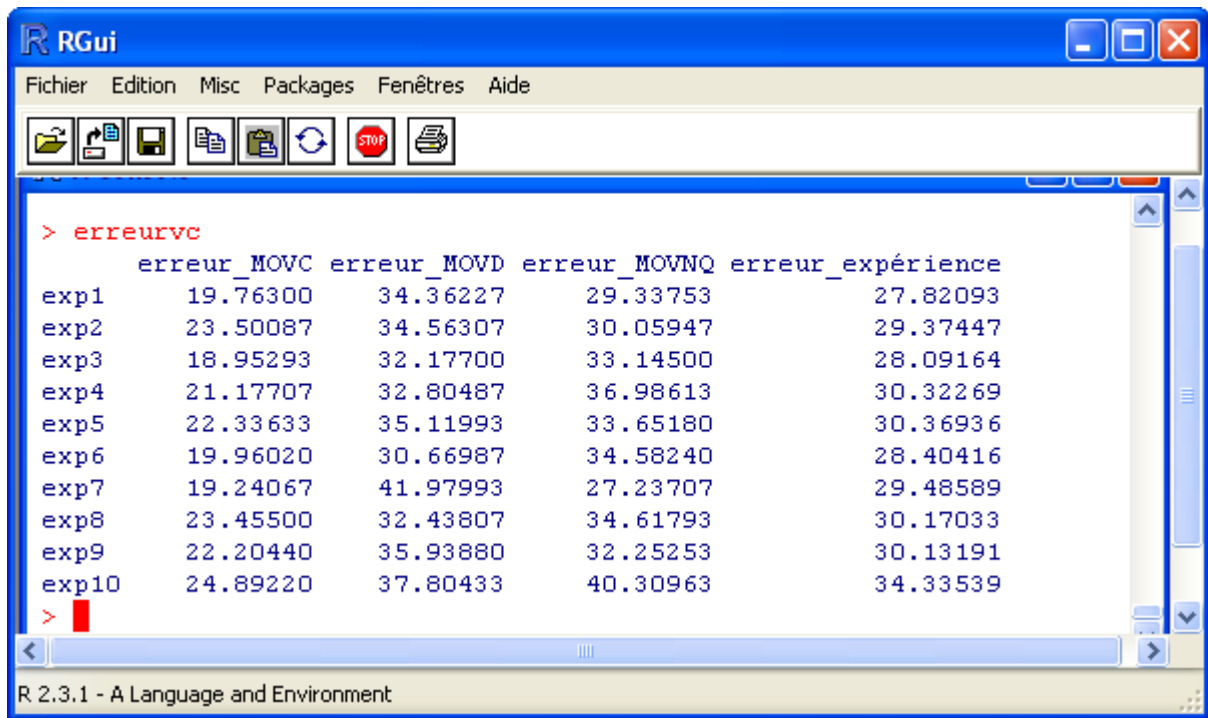
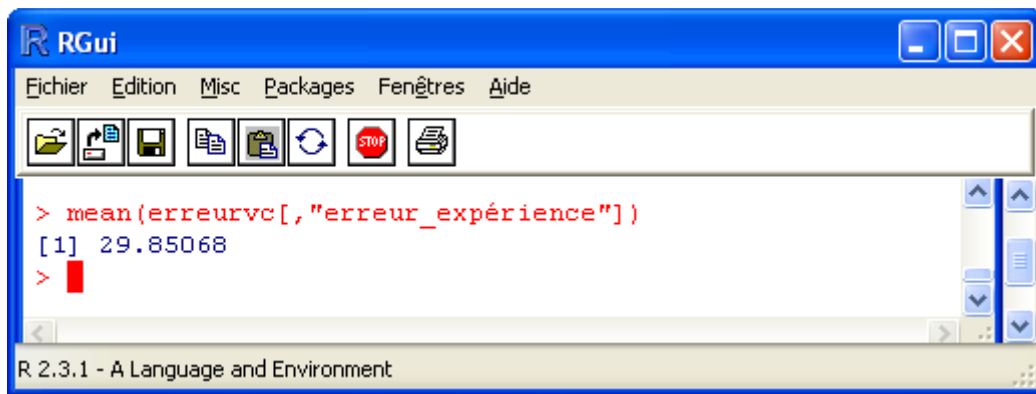


Figure 54 Erreurs des dix expériences

Bien que nous aillions choisi le modèle de la première expérience, l'erreur de ce dernier n'est pas considéré comme erreur globale du modèle. L'erreur globale, dans cette modélisation à partir de l'échantillon de base, est la moyenne des erreurs des modèles des dix expériences (Figure 55). La moyenne sur la colonne « erreur_expérience » de notre matrice d'erreurs nous permet de calculer cette valeur.



```
> mean(erreurvc[, "erreur_expérience"])\n[1] 29.85068\n>
```

Figure 55 Erreur en généralisation validation croisée

3.3.5. Conclusion

La validation croisée avec dix sous-échantillons est utilisée pour la validation du modèle. Nous avons construit un échantillon de départ avec 1500 observations générées aléatoirement. Nous avons travaillé avec le nombre minimum de cas fixé dans la littérature pour l'application de la validation croisée sur une méthode de sélection de variable dans une régression linéaire. Les dix expériences de la validation croisée ont donné chacune un réseau bayésien avec son erreur en généralisation. Le réseau bayésien est une quantification de l'influence de l'ampleur et de l'effort de gestion du changement sur l'impact de celui-ci. Le réseau bayésien de la première expérience présente le taux d'erreurs en généralisation le plus faible. Il a été donc sélectionné comme résultat de l'ensemble de la validation croisée. Il sera utilisé dans la phase d'utilisation du modèle présenté dans le chapitre suivant.

Chapitre 3.4 Utilisation du modèle sur les processus

3.4.1. Introduction

Après l'apprentissage-validation du modèle, nous pouvons l'utiliser pour la validation des processus en cours de restructuration. Un processus en cours de restructuration est un processus dont la restructuration proposée n'est pas encore validée. Cette validation se déroule dans la phase d'adéquation et de configuration, à travers une boucle adéquation-configuration-prototypage (Figure 56). Cette boucle se termine par la décision d'acceptation ou non du processus « en-devenir » proposé. Cette décision fait intervenir des informations issues du prototypage et de la simulation. Le prototypage vérifie que le processus fonctionnera de manière satisfaisante sur l'ERP. Il inclut une identification des trous fonctionnels par l'analyse des écarts. La simulation fournit des prévisions sur les critères de performance du processus en se basant sur la structure de ce dernier et les caractéristiques des ressources.

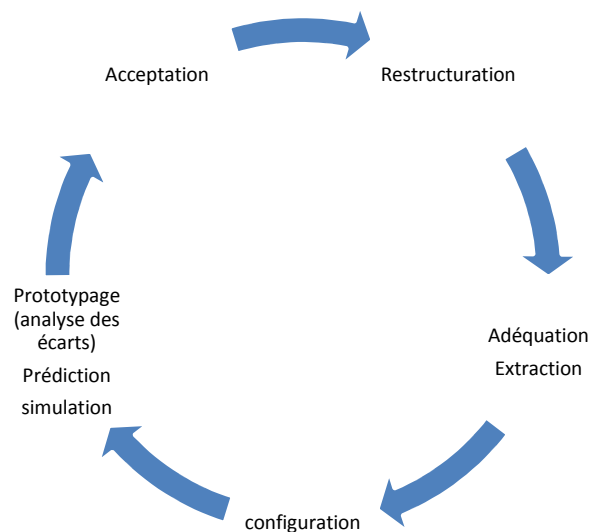


Figure 56 Prise en compte du changement dans l'acceptation du processus

Notre démarche de gestion proactive du changement vise à améliorer la décision d'acceptation du processus « en-devenir ». Elle vient compléter les informations fournies par la simulation et le prototypage en tenant compte de l'impact du changement. Une boucle restructuration-extraction-prédiction-acceptation (Figure 56) est exécutée en parallèle à celle d'adéquation-configuration-prototypage. La restructuration est la proposition d'une version « en-devenir » à partir d'une version « en-cours ». L'extraction est le calcul des variables explicatives par comparaison de ces deux versions du processus. La prédiction est l'application du réseau bayésien issu de la phase d'apprentissage-validation présentée dans le

chapitre précédant. L'acceptation est la validation ou le rejet du processus « en-devenir » en combinant les informations du prototypage, de la simulation et de la gestion proactive. Les différentes phases de cette boucle sont présentées dans les sections ci-dessous en prenant en exemple un processus de gestion de réclamations en cas de décès.

3.4.2. Restructuration

La restructuration consiste à décrire comment le processus est exécuté aujourd'hui dans l'entreprise à travers le processus « en-cours ». Un fonctionnement cible du processus, processus « en-devenir », est déterminé en fonction des améliorations à apporter et des possibilités offertes par l'ERP qui va le supporter (processus « possibles »). Un processus de gestion de réclamations en cas de décès, issu de Küster (2005), est utilisé pour illustrer cette phase de restructuration (Figure 57 et Figure 58). Les versions « en-cours » et « en-devenir » sont décrites avec la notation BPMN. L'étape suivante est l'utilisation de ces deux structures pour extraire les variables explicatives de l'observation correspondant au processus de gestion de réclamations.

Figure 57 Processus d'assurance décès « en-cours » (Küster, et al., 2005)

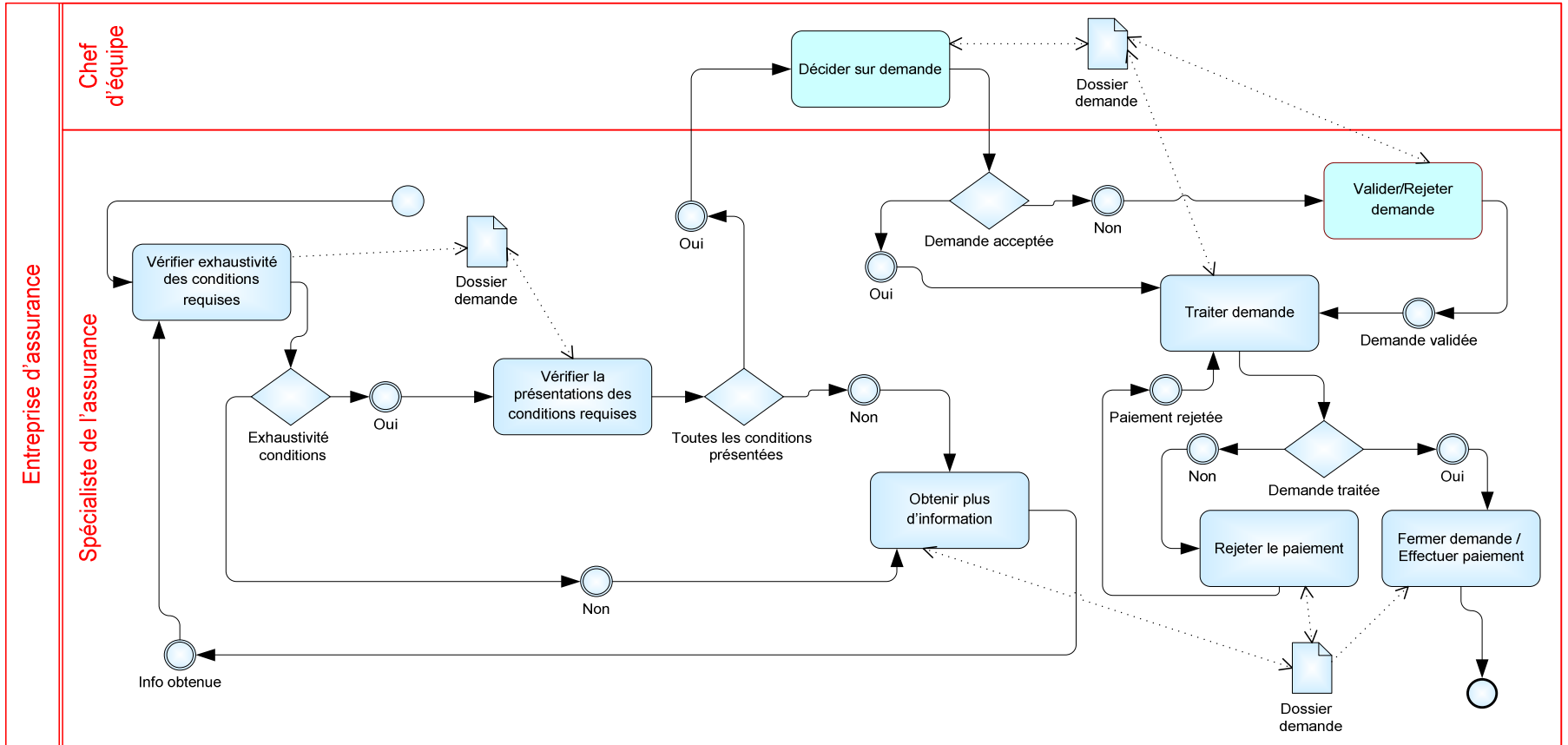
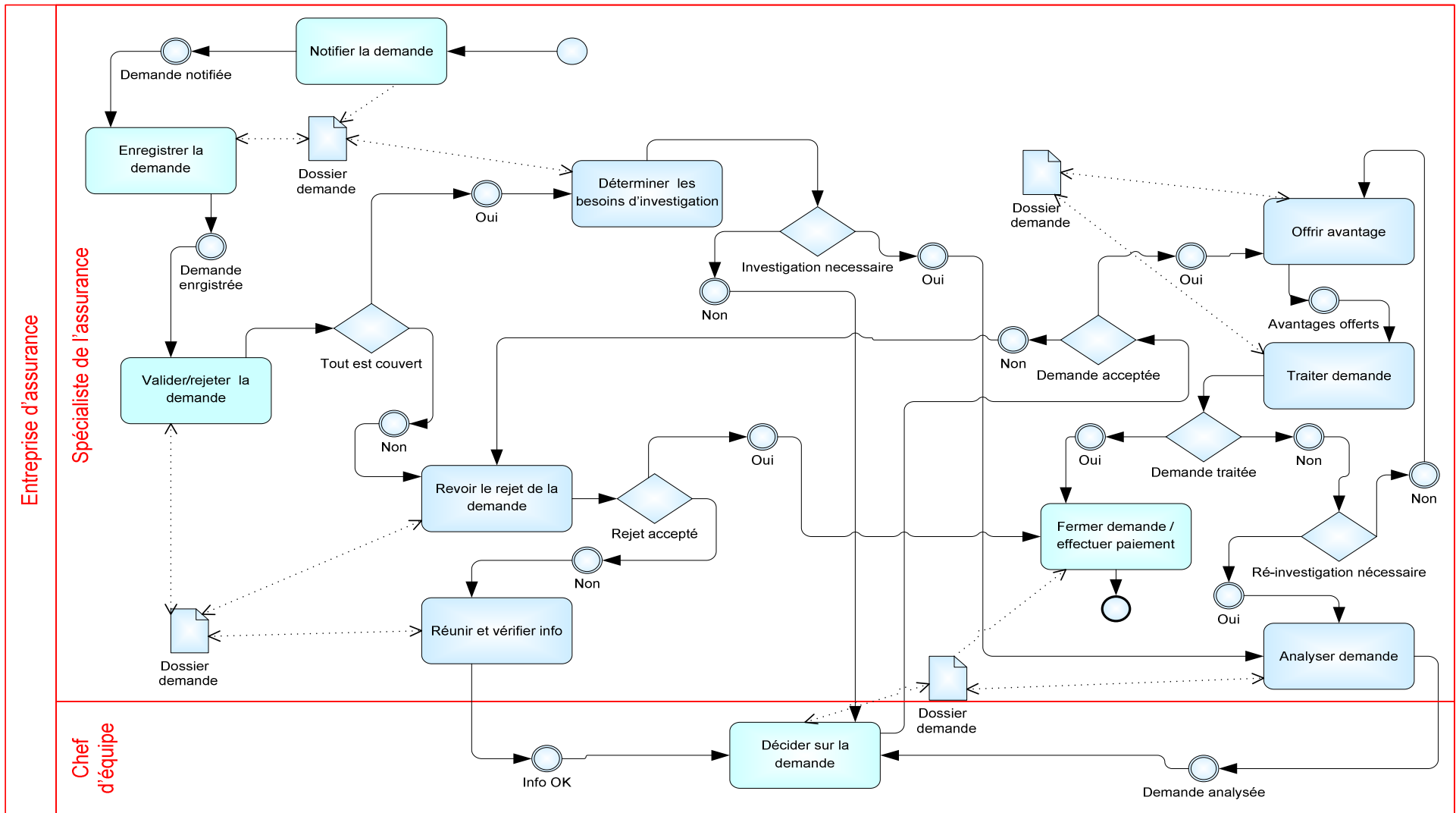


Figure 58 Processus d'assurance décès « en-devenir » (Küster, et al., 2005)



3.4.3. Extraction des métriques explicatives

3.4.3.1. Extraction des métriques structurelles de changement

L'extraction des métriques structurelles de changement est la première partie du calcul des variables explicatives pour le processus de gestion de réclamations en cas de décès. Nous utilisons les équations définies dans la section 2.4.4. Lors de la restructuration de ce processus, les relations de type « responsabilité » supprimées sont au nombre de 4 sur les 8 de départ (Figure 59). La métrique structurelle de changement dans les activités à exécuter supprimées (MSCAES) est calculée comme suit :

$$MSCAES = \frac{\text{relation de type responsabilité dans en-cours seulement} * 100}{\text{relation de type "responsabilité" dans "en - cours"}}$$

$$MSCAES = \frac{4 \times 100}{8} = 50\%$$

		Relation de type "responsabilité"			
		Dans "en-cours"		Dans "en-cours" uniquement	
	Lane	Activité	Lane	Activité	
	Spécialiste de l'assurance	Vérifier exhaustivité des conditions requises	Spécialiste de l'assurance	Vérifier exhaustivité des conditions requises	
	Spécialiste de l'assurance	Vérifier la présentations des conditions requises	Spécialiste de l'assurance	Vérifier la présentations des conditions requises	
	Spécialiste de l'assurance	Obtenir plus d'information	Spécialiste de l'assurance	Obtenir plus d'information	
	Spécialiste de l'assurance	Traiter demande	Spécialiste de l'assurance	Rejeter le paiement	
	Spécialiste de l'assurance	Valider/Rejeter demande			
	Spécialiste de l'assurance	Rejeter le paiement			
	Spécialiste de l'assurance	Fermer demande / Effectuer paiement			
	Chef d'équipe	Décider sur demande			
Nombre de raltions		8		4	

Figure 59 Relations "responsabilités" supprimées dans PGR

Les relations de type « responsabilité » rajoutées sont quant à elles au nombre de 7 (Figure 60). La métrique structurelle de changement dans les activités à exécuter rajoutées (MSCAER) est donc :

$$MSCAER = \frac{\text{relation de type responsabilité dans implémenté seulement} \times 100}{\text{relation de type "responsabilité" dans "en - cours"}}$$

$$MSCAER = \frac{7 \times 100}{8} = 87.5\%$$

		Relation de type "responsabilité"			
		Dans "en-cours"		Dans "en-devenir" uniquement	
	Lane	Activité	Lane	Activité	
	Spécialiste de l'assurance	Vérifier exhaustivité des conditions requises	Spécialiste de l'assurance	Notifier la demande	
	Spécialiste de l'assurance	Vérifier la présentations des conditions requises	Spécialiste de l'assurance	Enregistrer la demande	
	Spécialiste de l'assurance	Obtenir plus d'information	Spécialiste de l'assurance	Déterminer les besoins d'investigation	
	Spécialiste de l'assurance	Traiter demande	Spécialiste de l'assurance	Revoir le rejet de la demande	
	Spécialiste de l'assurance	Valider/Rejeter demande	Spécialiste de l'assurance	Réunir et vérifier info	
	Spécialiste de l'assurance	Rejeter le paiement	Spécialiste de l'assurance	Offrir avantage	
	Spécialiste de l'assurance	Fermer demande / Effectuer paiement	Spécialiste de l'assurance	Analyser demande	
	Chef d'équipe	Décider sur demande			
Nombre de raltions		8		7	

Figure 60 Relations "responsabilités" rajoutées dans PGR

La version « en-cours » du processus de gestion des réclamations contient 7 associations allant des objets de donnée vers les activités. Parmi ces associations, trois ne figurent pas dans la version « en-devenir » (Figure 61). La métrique de changement structurelle dans les informations à traiter supprimées (MSCITS) prend la valeur :

$$MSCITS = \frac{\text{association objet donnée vers activité dans en-cours seulement} \times 100}{\text{association objet donnée vers activité dans "en – cours"}}$$

$$MSCITS = \frac{3 \times 100}{7} = 42.857\%$$

	Association de objet de donnée vers activité			
	Dans "en-cours"		Dans "en-cours" uniquement	
	Objet de donnée	Activité	Objet de donnée	Activité
	Dossier demande	Vérifier la présentations des conditions requises	Dossier demande	Vérifier la présentations des conditions requises
	Dossier demande	Obtenir plus d'information	Dossier demande	Obtenir plus d'information
	Dossier demande	Traiter demande	Dossier demande	Rejeter le paiement
	Dossier demande	Valider/Rejeter demande		
	Dossier demande	Rejeter le paiement		
	Dossier demande	Fermer demande / Effectuer paiement		
	Dossier demande	Décider sur demande		
Nombre d'associations	7		3	

Figure 61 Associations objets de donnée vers activités supprimées dans PGR

Six associations allant des objets de donnée vers les activités ne figurent que dans la version « en-devenir » du processus de gestion de réclamations (Figure 62). La métrique de changement structurelle dans les informations à traiter rajoutées est donc :

$$MSCITR = \frac{\text{association objet donnée vers activité dans implémenté seulement} \times 100}{\text{association objet donnée vers activité dans "en – cours"}}$$

$$MSCITR = \frac{6 \times 100}{7} = 85.714\%$$

	Associatio de objet de donnée vers activité			
	Dans "en-cours"		Dans "implémenté" uniquement	
	Objet de donnée	Activité	Objet de donnée	Activité
	Dossier demande	Vérifier la présentations des conditions requises	Dossier demande	Enregistrer la demande
	Dossier demande	Obtenir plus d'information	Dossier demande	Déterminer les besoins d'investigation
	Dossier demande	Traiter demande	Dossier demande	Revoir le rejet de la demande
	Dossier demande	Valider/Rejeter demande	Dossier demande	Réunir et vérifier info
	Dossier demande	Rejeter le paiement	Dossier demande	Offrir avantage
	Dossier demande	Fermer demande / Effectuer paiement	Dossier demande	Analyser demande
	Dossier demande	Décider sur demande		
Nombre d'associations	7		6	

Figure 62 Association objets de donnée vers activités » rajoutées dans PGR

Le nombre d'associations allant des activités vers les objets de donnée supprimées est de 4 sur un total de 7 au départ (Figure 63). Nous obtenons la métrique structurelle de changement dans les informations à produire supprimées par :

$$MSCIPS = \frac{\text{association activité vers objet donnée dans en-cours seulement} * 100}{\text{association activité vers objet donnée dans "en – cours"}}$$

$$MSCIPS = \frac{4 * 100}{7} = 57.143\%$$

	Association de activité vers objet de donnée			
	Dans "en-cours"		Dans "en-cours" uniquement	
	Activité	Objet donnée	Activité	Objet donnée
	Vérifier exhaustivité des conditions requises	Dossier demande	Vérifier exhaustivité des conditions requises	Dossier demande
	Vérifier la présentations des conditions requises	Dossier demande	Vérifier la présentations des conditions requises	Dossier demande
	Obtenir plus d'information	Dossier demande	Obtenir plus d'information	Dossier demande
	Traiter demande	Dossier demande	Rejeter le paiement	Dossier demande
	Valider/Rejeter demande	Dossier demande		
	Rejeter le paiement	Dossier demande		
	Décider sur demande	Dossier demande		
Nombre d'associations	7		4	

Figure 63 Associations activités vers objets de donnée supprimées dans PGR

La réingénierie du processus de gestion de réclamations a rajouté 7 associations allant des activités vers les objets de donnée (Figure 64). La métrique structurelle de changement dans les informations à produire rajoutées se calcule donc comme suit :

$$MSCIPR = \frac{\text{association activité vers objet donnée dans implémenté seulement} \times 100}{\text{association activité vers objet donnée dans "en - cours"}}$$

$$MSCIPR = \frac{7 \times 100}{7} = 100\%$$

	Association de activité vers objet de donnée			
	Dans "en-cours"		Dans "implémenté" uniquement	
	Activité	Objet donnée	Activité	Objet donnée
	Vérifier exhaustivité des conditions requises	Dossier demande	Notifier la demande	Dossier demande
	Vérifier la présentations des conditions requises	Dossier demande	Enregistrer la demande	Dossier demande
	Obtenir plus d'information	Dossier demande	Déterminer les besoins d'investigation	Dossier demande
	Traiter demande	Dossier demande	Revoir le rejet de la demande	Dossier demande
	Valider/Rejeter demande	Dossier demande	Réunir et vérifier info	Dossier demande
	Rejeter le paiement	Dossier demande	Offrir avantage	Dossier demande
	Décider sur demande	Dossier demande	Analyser demande	Dossier demande
Nombre d'associations	7		7	

Figure 64 Associations activité vers objets de donnée rajoutées dans PGR

Sur les 12 flux de séquence allant des événements vers les activités dans la version « en-cours » du processus de gestion de réclamations, 11 ont été supprimés (Figure 65). La métrique structurelle de changement dans les événements à surveiller supprimés est égale à :

$$MSCCESS = \frac{\text{flux séquence de événement vers activité dans en-cours seulement} \times 100}{\text{flux séquence de événement vers activité dans "en - cours"}}$$

$$MSCCESS = \frac{11 \times 100}{12} = 91.667\%$$

	Flux séquence de événement vers activité			
	Dans "en-cours"		Dans "en-cours" uniquement	
	Événement	Activité	Événement	Activité
Début	Vérifier exhaustivité des conditions requises	Début	Vérifier exhaustivité des conditions requises	
Info obtenue	Vérifier exhaustivité des conditions requises	Info obtenue	Vérifier exhaustivité des conditions requises	
Exhaustivité conditions = oui	Vérifier la présentations des conditions requises	Exhaustivité conditions = oui	Vérifier la présentations des conditions requises	
Exhaustivité conditions = non	Obtenir plus d'information	Exhaustivité conditions = non	Obtenir plus d'information	
Toutes les conditions présentées = non	Obtenir plus d'information	Toutes les conditions présentées = non	Obtenir plus d'information	
Toutes les conditions présentées = oui	Décider sur demande	Toutes les conditions présentées = oui	Décider sur demande	
Demande acceptée = non	Valider/Rejeter demande	Demande acceptée = non	Valider/Rejeter demande	
Demande acceptée = oui	Traiter demande	Demande acceptée = oui	Traiter demande	
Paieement rejeté	Traiter demande	Paieement rejeté	Traiter demande	
Demande validée	Traiter demande	Demande validée	Traiter demande	
Demande traitée = non	Rejeter le paieement	Demande traitée = non	Rejeter le paieement	
Demande traitée = oui	Fermer demande /Effectuer paieement			
Nombre de flux séquence	12		11	

Figure 65 Flux de séquence événements vers activités supprimées dans PGR

Le nombre des flux de séquence allant des événements vers les activités rajoutés par la reconfiguration du processus est de 16 (Figure 66). La métrique structurale de changement dans les événements à surveiller rajoutés est obtenue comme suit :

$$MSCESR = \frac{\text{flux séquence de événement vers activité dans implémenté seulement} \times 100}{\text{flux séquence de événement vers activité dans "en – cours"}}$$

$$MSCESR = \frac{16 \times 100}{12} = 133.334\%$$

	Flux séquence de événement vers activité			
	Dans "en-cours"		Dans "implémenté" uniquement	
	Événement	Activité	Événement	Activité
Début	Vérifier exhaustivité des conditions requises	Début	Notifier la demande	
Info obtenue	Vérifier exhaustivité des conditions requises	Demande notifiée	Enregistrer la demande	
Exhaustivité conditions = oui	Vérifier la présentations des conditions requises	Demande enregistrée	Valider/rejeter la demande	
Exhaustivité conditions = non	Obtenir plus d'information	Tout est couvert = non	Revoir le rejet de la demande	
Toutes les conditions présentées = non	Obtenir plus d'information	Demande acceptée = non	Revoir le rejet de la demande	
Toutes les conditions présentées = oui	Décider sur demande	Rejet accepté = non	Réunir et vérifier info	
Demande acceptée = non	Valider/Rejeter demande	Tout est couvert = oui	Déterminer les besoins d'investigation	
Demande acceptée = oui	Traiter demande	Info OK	Décider sur la demande	
Paieement rejeté	Traiter demande	Demande analysée	Décider sur la demande	
Demande validée	Traiter demande	Investigation nécessaire = non	Décider sur la demande	
Demande traitée = non	Rejeter le paieement	Investigation nécessaire = oui	Analyser demande	
Demande traitée = oui	Fermer demande /Effectuer paieement	Ré-investigation nécessaire = oui	Analyser demande	
		Ré-investigation nécessaire = non	Offrir avantage	
		Demande acceptée = oui	Offrir avantage	
		Avantages offerts	Traiter demande	
		Rejet accepté = oui	Fermer demande /effectuer paieement	
Nombre de flux séquence	12		16	

Figure 66 Flux séquence événements vers activités rajoutées dans PGR

Les flux de séquence allant des activités vers les événements supprimés sont de 7 sur les 12 dans la version « en-cours » du processus de gestion de réclamations (Figure 67). La métrique structurale de changement dans les événements à générer supprimés pour ce processus est de :

$$MSCEGS = \frac{\text{flux séquence de activité vers événement dans en-cours seulement} \times 100}{\text{flux séquence de activité ver événement dans "en – cours"}}$$

$$MSCEGS = \frac{7 \times 100}{12} = 58.334\%$$

	Flux séquence de activité vers événement			
	Dans "en-cours"		Dans "en-cours" uniquement	
	Activité	Évènement	Activité	Évènement
Vérifier exhaustivité des conditions requises	Exhaustivité conditions =oui	Vérifier exhaustivité des conditions requises	Exhaustivité conditions =oui	
Vérifier exhaustivité des conditions requises	Exhaustivité conditions = non	Vérifier exhaustivité des conditions requises	Exhaustivité conditions = non	
Vérifier la présentations des conditions requises	Toutes les conditions présentées = oui	Vérifier la présentations des conditions requises	Toutes les conditions présentées = oui	
Vérifier la présentations des conditions requises	Toutes les conditions présentées = non	Vérifier la présentations des conditions requises	Toutes les conditions présentées = non	
Valider/Rejeter demande	Demande validée	Valider/Rejeter demande	Demande validée	
Traiter demande	Demande traitée = oui	Obtenir plus d'information	Info obtenue	
Traiter demande	Demande traitée = non	Rejeter le paiement	Paiement rejeté	
Décider sur demande	Demande acceptée = oui			
Décider sur demande	Demande acceptée = non			
Fermer demande /Effectuer paiement	fin			
Obtenir plus d'information	Info obtenue			
Rejeter le paiement	Paiement rejeté			
Nombre de flux séquence	12		7	

Figure 67 Flux de séquence activités vers événements supprimées dans PGR

Les flux de séquence allant des activités vers les événements supprimés sont au nombre de 12 (Figure 68). La métrique structurelle de changement dans les événements à générer rajoutés est calculée comme suit :

$$MSCEGR = \frac{\text{flux séquence de activité vers événement dans implémenté seulement} \times 100}{\text{flux séquence de activité ver événement dans "en – cours"}}$$

$$MSCEGR = \frac{12 \times 100}{12} = 100\%$$

	Flux séquence de activité vers événement			
	Dans "en-cours"		Dans "implémenté" uniquement	
	Activité	Évènement	Activité	Évènement
Vérifier exhaustivité des conditions requises	Exhaustivité conditions =oui	Notifier la demande	Demande notifiée	
Vérifier exhaustivité des conditions requises	Exhaustivité conditions = non	Enregistrer la demande	Demande enregistrée	
Vérifier la présentations des conditions requises	Toutes les conditions présentées = oui	Valider/rejeter la demande	Tout est couvert = oui	
Vérifier la présentations des conditions requises	Toutes les conditions présentées = non	Valider/rejeter la demande	Tout est couvert = non	
Valider/Rejeter demande	Demande validée	Déterminer les besoins d'investigation	Investigation nécessaire = oui	
Traiter demande	Demande traitée = oui	Déterminer les besoins d'investigation	Investigation nécessaire = non	
Traiter demande	Demande traitée = non	Revoir le rejet de la demande	Rejet accepté = oui	
Décider sur demande	Demande acceptée = oui	Revoir le rejet de la demande	Rejet accepté = non	
Décider sur demande	Demande acceptée = non	Réunir et vérifier info	Info OK	
Fermer demande /Effectuer paiement	fin	Offrir avantage	Avantages offerts	
Obtenir plus d'information	Info obtenue	Traiter demande	Ré-investigation nécessaire = oui	
Rejeter le paiement	Paiement rejeté	Traiter demande	Ré-investigation nécessaire = non	
Nombre de flux séquence	12		12	

Figure 68 Flux de séquence activités vers événements rajoutées dans PGR

3.4.3.2. Extraction des métriques structurelles de variation

L'extraction des métriques structurelles de variation est la seconde étape dans le calcul des variables explicatives pour le processus de gestion de réclamations. La version « en-cours » de ce processus contient 4 points de décision de type OU exclusif (XOR) avec deux sorties chacun. La complexité du flux de contrôle de cette version est donc de 8. La version « en-devenir » a une complexité du flux de contrôle de 12 car elle contient 6 points de décision XOR à deux sorties. La variation du CFC pour ce processus s'obtient donc par :

$$MSVCFC = \frac{(CFC_{to-be} - CFC_{as-is}) \times 100}{CFC_{as-is}} = \frac{(12 - 8) \times 100}{8} = 50\%$$

Pour la métrique de variation dans la profondeur d'imbrication maximale (MSVPIM), nous devons d'abord calculer la PIM pour chaque version du processus. La PIM du processus « en-cours » est de 4, correspondant aux décisions nécessaires pour exécuter les activités « Rejeter le paiement » et « Fermer demande/Effectuer paiement ». Le processus « en-devenir » a une PIM de 5 qui est le nombre de décisions nécessaires pour atteindre l'activité « Analyser demande ». La métrique structurelle de variation dans la profondeur d'imbrication maximale prend la valeur

$$\text{MSVPIM} = \frac{(\text{PIM}_{\text{to-be}} - \text{PIM}_{\text{as-is}}) \times 100}{\text{PIM}_{\text{as-is}}} = \frac{(5 - 4) \times 100}{4} = 25\%$$

Le calcul de la métrique structurelle de variation de la simplicité dépend des métriques structurelles dérivées de simplicité des processus «en-cours » et « en-devenir » (Tableau 21).

Tableau 21 Détails simplicité PGR

	En-cours	En-devenir
Nombre d'activités	8	11
Nombre de personnes	2	2
Flux matériels ou informationnel	14	20
La complexité basique C_B	24	33
La complexité moyenne par activité \bar{C}_A	24/8=3	33/11= 3
$\max \bar{C}_A$	24	33
$\min \bar{C}_A$	3*24/ (24-1) = 3.130435	3*33/ (33-1) = 3.09375
$\bar{C}_A - \min \bar{C}_A$	(3-3.130435)= -0.130435	(3-3.09375) = -0.09375
$\max \bar{C}_A - \min \bar{C}_A$	24-3.130435 = 20.86957	33-3.09375= 29.90625
Simplicité	-0.130435*100/20.86957 = -0.6250009%	-0.09375*100/29.90625 = - 0.3134796%

La métrique de variation dans la simplicité du processus de gestion de réclamations est obtenue en appliquant la formule correspondante avec les métriques de simplicité des ses deux versions :

$$\text{MSVS} = S_{\text{to-be}} - S_{\text{as-is}} = - 0.3134796 + 0.6250009 = 0.3115213\%$$

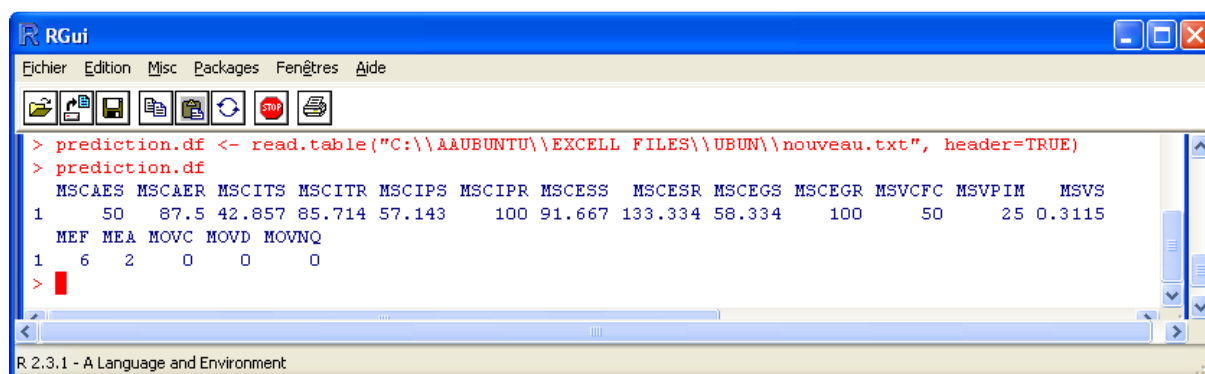
3.4.3.3. Initialisation des métriques d'effort de gestion du changement

La détermination de l'effort de gestion du changement est la dernière partie de l'extraction des variables explicatives. La gestion du changement est composée de l'effort de formation et de celui d'accompagnement. Ce sont des variables dont les valeurs sont fixées au départ, avant la prédiction sur le nouveau processus. Elles peuvent être modifiées par la suite selon l'interprétation des résultats de la prédiction. Les valeurs minimales qui seront affectées à ces deux variables peuvent être utilisées pour les initialiser dans tout nouveau processus. L'objectif est de déterminer si le processus a besoin de plus d'effort de gestion du changement. Cela permet une rationalisation de cet effort au niveau de l'ensemble des processus à restructurer dans le projet en cours. Dans cette étude, nous avons fait la supposition que chaque acteur, dans un processus à restructurer, devrait recevoir au minimum 6 heures de formation et 2 heures d'accompagnement. L'initialisation de la métrique d'effort de formation (MEF) et de la métrique d'effort d'accompagnement (MEA), pour le processus de gestion de réclamations, est faite avec ces valeurs :

$$MEF = 6 \text{ h/acteur}$$

$$MEA = 2 \text{ h/acteur}$$

Les quinze variables explicatives déterminées pour le processus de gestion de réclamations sont chargées dans une base prédiction (Figure 69). Les valeurs des nœuds de réponse sont fixées à zéro, pour spécifier que ce sont des variables continues. Elles seront prédites dans la section suivante.



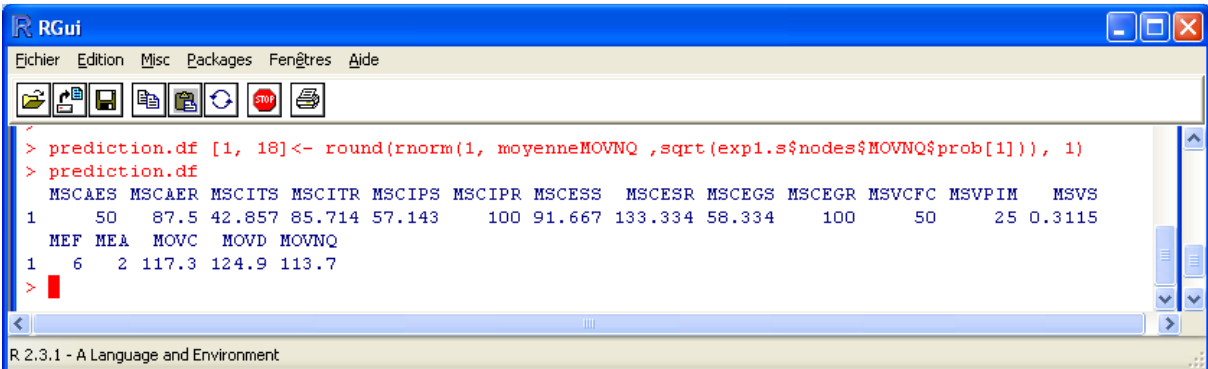
```
> prediction.df <- read.table("C:\\AAUBUNTU\\EXCELL FILES\\UBUN\\nouveau.txt", header=TRUE)
> prediction.df
  MSCAES MSCAER MSCITS MSCITR MSCIPS MSCIPR MSCESS MSCESR MSCEGS MSCEGR MSVCFC MSVPIM MSVS
1      50   87.5 42.857 85.714 57.143    100 91.667 133.334 58.334    100    50    25 0.3115
  MEF  MEA  MOVc  MOVD  MOVNQ
1    6    2    0    0    0
```

Figure 69 Variables explicatives PGR

3.4.4. Prédiction

La première utilisation du modèle obtenu lors de l'apprentissage-validation est la prédiction sur les variations des performances d'un nouveau processus. Le modèle que nous avons choisi à la fin de l'apprentissage-validation est celui de l'expérience 1. Nous appliquons la prédiction sur le processus de gestion de réclamations présenté plus haut. Les commandes similaires à celles utilisées pour les prédictions lors de l'évaluation des modèles des expériences sont appliquées avec la base de prédiction (processus de gestion de réclamations) et les paramètres des nœuds de réponse du modèle 1 (Figure 71). Les prédictions pour les variables de réponse sont relatives à la restructuration proposée et à l'effort de gestion de changement à consentir. Après la restructuration proposée sur le processus de gestion de réclamations, le modèle 1 prédit (Figure 70):

- une augmentation du coût de 117,3 % ;
- une augmentation de la durée de 129,3% ;
- une augmentation de la non-qualité de 113,7 %.



```
> prediction.df [1, 18]<- round(rnorm(1, moyenneMOVNQ ,sqrt(exp1.s$nodes$MOVNQ$prob[1])), 1)
> prediction.df
  MSCAES MSCAER MSCITS MSCITR MSCIPS MSCIPR MSCCESS MSCESR MSCEGS MSCEGR MSVCFC MSVPIM MSVS
1    50    87.5 42.857 85.714 57.143    100 91.667 133.334 58.334    100    50    25 0.3115
  MEF MEA  MOVQ  MOVD MOVNQ
1    6    2 117.3 124.9 113.7
```

Figure 70 Résultats première prédiction PGR

Dans la section suivante nous présentons les types de décision pouvant être prises après l'interprétation des résultats de la prédiction pour un processus particulier.

```
Sans titre - Bloc notes
Fichier Edition Format Affichage ?
#prediction de MOVc
moyenneMOVc <- exp1.s$nodes$MOVc$prob[2]
if (length(exp1.s$nodes$MOVc$prob)> 2) {
for (x in 3:length(exp1.s$nodes$MOVc$prob)) moyenneMOVc <- moyenneMOVc + exp1.s$nodes$MOVc$prob[x]* prediction.df [1,exp1.s$nodes$MOVc$parents[x-2]]}

prediction.df [1, 16]<- round(rnorm(1, moyenneMOVc ,sqrt(exp1.s$nodes$MOVc$prob[1])), 1)

#prediction noeud MOVD

moyenneMOVD <- exp1.s$nodes$MOVD$prob[2]
if (length(exp1.s$nodes$MOVD$prob)> 2) {
for (x in 3:length(exp1.s$nodes$MOVD$prob)) moyenneMOVD <- moyenneMOVD + exp1.s$nodes$MOVD$prob[x]*prediction.df [1,exp1.s$nodes$MOVD$parents[x-2]]}

prediction.df [1, 17]<- round(rnorm(1, moyenneMOVD ,sqrt(exp1.s$nodes$MOVD$prob[1])), 1)

#prediction de MOVNQ

moyenneMOVNQ <- exp1.s$nodes$MOVNQ$prob[2]
if (length(exp1.s$nodes$MOVNQ$prob)> 2) {
for (x in 3:length(exp1.s$nodes$MOVNQ$prob)) moyenneMOVNQ <- moyenneMOVNQ + exp1.s$nodes$MOVNQ$prob[x]*prediction.df [1,exp1.s$nodes$MOVNQ$parents[x-2]]}

prediction.df [1, 18]<- round(rnorm(1, moyenneMOVNQ ,sqrt(exp1.s$nodes$MOVNQ$prob[1])), 1)
```

Figure 71 Commandes de prédiction PGR

3.4.5. Validation du processus « en-devenir »

La prédiction est l'avant dernière étape de la boucle restructuration-extraction-prédiction-validation appliquée aux processus à restructurer. Nous parlons de boucle car ces étapes sont répétées jusqu'à ce que l'étape de validation aboutisse à un succès. Cette étape de validation prend en compte les résultats de la simulation et du prototypage. Le prototypage vérifie que le processus fonctionnera de manière correcte sur l'ERP. Il inclut l'analyse des écarts. La simulation réalise des prédictions sur les performances, mais celles-ci ne prennent pas en compte l'ampleur du changement sur le processus. L'analyse des écarts, quant à elle, met en relief la distance entre les processus « en-devenir » et les « possibles ». Pour l'analyse des écarts, il faut s'assurer que les trous fonctionnels sont justifiés et que les ressources destinées à leur fermeture sont disponibles. Pour la simulation, il faut s'assurer que le processus donnera des performances satisfaisantes une fois implémenté sans tenir compte de l'impact du changement.

La gestion proactive du changement a été proposée pour compléter ces techniques dans cette étape de validation du processus « en-devenir ». La gestion proactive consiste à prendre des décisions relatives à la validation sur la base de prédictions réalisées sur les processus en tenant compte de l'ampleur et de l'effort de gestion du changement. La boucle restructuration-extraction-prédiction-validation est exécutée un nombre fini de fois. Une des deux décisions suivantes peut être prise à chaque tour :

- la décision de ne pas valider la proposition de reconfiguration. Elle intervient lors des passages intermédiaires dans la boucle. Elle est motivée par de mauvaises prédictions dont les causes sont à chercher dans :
 - un effort de gestion du changement insuffisant
 - une ampleur du changement importante
- la décision de valider la proposition qui intervient lors du dernier passage dans la boucle.

Non validation pour effort de gestion du changement insuffisant. Notre modèle contient des variables correspondant aux trois concepts principaux de l'étude : le changement, son impact et sa gestion. Les métriques opérationnelles de variation sont les variables de réponse du modèle prédictif. Elles permettent d'avoir des prévisions sur les conséquences des choix de restructuration et de gestion du changement. Des décisions de validation et de non-validation s'appuient en partie sur ces indications.

Des métriques d'effort de gestion du changement permettent de fixer l'effort à consentir pour la gestion du changement en cas de validation du processus « en-devenir ». L'effort de gestion du changement est une explication et un levier pour la validation de la restructuration du processus. Lorsque par exemple, les résultats de prédiction sur les variables de réponse ne sont pas bons, la cause peut être un effort de gestion du changement faible pour

le processus à reconfigurer. C'est dans ce sens que nous parlons d'explication. Une autre boucle restructuration-extraction-prédiction-validation doit être exécutée après cette décision de non-validation. L'effort de gestion du changement peut alors être utilisé comme levier en augmentant, par exemple, l'accompagnement ou la formation. Ce levier ne doit être activé que dans le cas où l'effort de gestion est considéré comme l'explication de la décision de ne pas valider à la boucle précédente.

Dans la section précédente, nous avons pris en exemple un processus de gestion de réclamations. Ce processus est passé par les trois premières phases de la boucle qui sont la restructuration, l'extraction des variables explicatives et la prédiction des variables de réponse. Il reste à lui appliquer la validation qui est la dernière étape de cette boucle. Les prédictions sont une augmentation de 110% au minimum pour le coût, la durée et la non-qualité⁸⁴. Ces prédictions amènent à ne pas valider le processus « en-devenir » et à s'interroger sur les causes de tels résultats. L'explication réside, sans doute, dans le fait que nous avons fixé le niveau d'effort de gestion du changement au minimum car c'était le premier passage de ce processus dans la boucle. Lors du second passage, nous décidons d'augmenter l'effort de formation et d'accompagnement à engager sur ce processus. Les ressources devant fournir cet effort de gestion n'étant pas infinies, la part accordée à chaque processus doit être limitée à un maximum. L'effort de gestion a été augmenté jusqu'à atteindre ce plafond pour la formation (20 h/acteur) et l'accompagnement (8 h/acteur). Cela n'a pas permis d'arriver à des niveaux acceptables sur les prédictions qui sont inférieures mais proches des +100% (Figure 72). Cela signifie que l'explication des mauvaises prédictions ne se trouve pas dans le manque de gestion du changement.

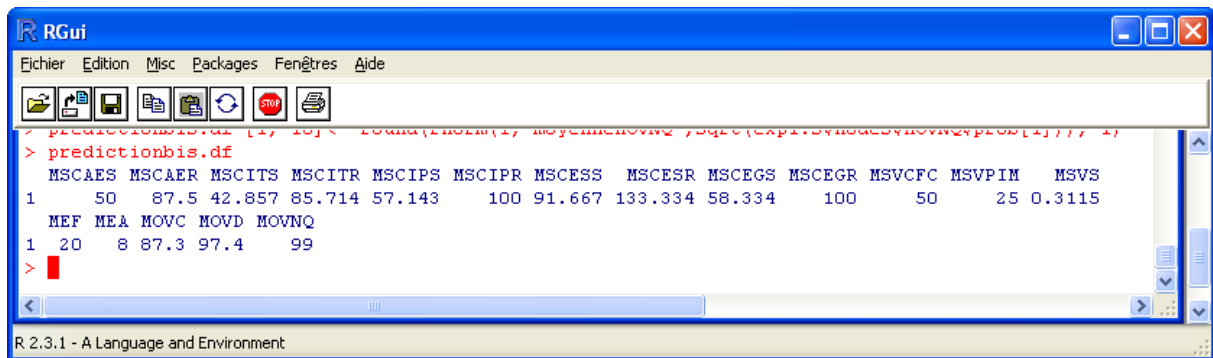


Figure 72 Résultats seconde prédiction PGR

L'activation du levier d'effort de gestion consiste, par rapport à l'impact du changement, à agir sur les conséquences. Cependant, il faut tenir compte des résultats de la simulation et de l'analyse des écarts. L'activation du levier d'effort de gestion du changement n'a aucune incidence sur les résultats venant de l'application de ces deux approches. Supposons que, pour notre processus de gestion de réclamations, nous soyons dans le cas suivant :

⁸⁴ Rappelons que nous cherchons à obtenir des valeurs négatives pour ces variables de réponse

- la simulation du processus « en-devenir » donne des résultats favorables,
- l'analyse des écarts donne des résultats favorables car il n'y a pas de trous fonctionnels à combler sur le processus,
- l'augmentation de l'effort de gestion du changement a permis d'atteindre des prédictions favorables.

L'augmentation de l'effort de gestion étant indépendante de la simulation et de l'analyse des écarts, elle n'oblige pas à reconsidérer leurs résultats qui restent donc favorables. Nous avons donc, sur les trois points de l'analyse, des informations qui permettent de préconiser la validation du processus « en-devenir » et de l'implémenter.

Non-validation pour ampleur du changement important. Des métriques structurelles de changement et de variation mesurent l'ampleur du changement proposé sur le processus. Elles font partie des variables explicatives. L'ampleur du changement est un second champ dans lequel l'explication des mauvaises prédictions peut être trouvée. Dans ce cas, la décision de ne pas valider le processus est motivée par l'inadéquation du changement.

L'ampleur du changement, comme l'effort de gestion, est un levier que l'on peut utiliser pour obtenir la validation d'un processus. L'activation du levier ampleur changement revient à agir sur les causes même du problème de l'impact du changement. Il s'agit de proposer une nouvelle structure pour le processus « en-devenir » lors du passage dans la boucle (restructuration-extraction-prédiction-validation) qui suit la décision de non-validation. Une extraction des variables explicatives est réalisée en tenant compte de cette nouvelle structure. Les valeurs obtenues sont rentrées en paramètres dans le modèle bayésien pour obtenir de nouvelles prédictions sur les variables de réponse.

Là aussi, la décision à prendre dépend des informations obtenues pour la simulation et l'analyse des écarts pour le processus concerné. Contrairement à l'effort de gestion, la restructuration n'est pas indépendante des résultats de ces deux approches. Pour illustrer la relation entre l'ampleur du changement (gestion proactive), les trous fonctionnels (analyse des écarts) et les performances (prototypage et simulation), prenons l'exemple d'un processus « en-devenir » qui, dans un passage dans la boucle, est dans le cas suivant (Figure 73) :

- il est proche des « possibles » (faible distance entre « en-devenir » et « possibles »). Cela réduit les trous fonctionnels et l'analyse des écarts ne signale donc aucun problème ;
- il est optimal du point de vue de la simulation et du prototypage car proche des « possibles » (faible distance entre « en-devenir » et « possibles ») qui représentent les meilleures pratiques ;
- les résultats de la prédiction ne sont pas favorables. L'explication retenue pour cela est l'ampleur du changement (grande distance entre « en-cours » et « en-devenir »).

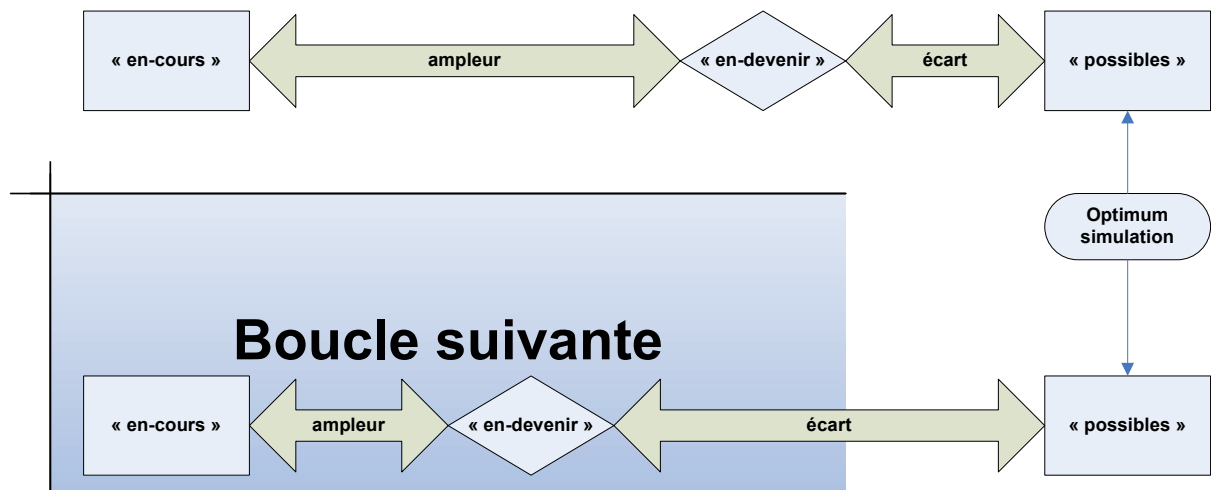


Figure 73 Equilibre entre changement, écart et simulation

Lors du passage suivant dans la boucle, la situation est la suivante (Figure 73) :

- il est décidé d'activer le levier de l'ampleur du changement en proposant une nouvelle version « en-devenir » dans laquelle, le changement est diminué (faible distance entre « en-cours » et « en-devenir »). Cela permettrait de réduire l'ampleur du changement pour obtenir des prédictions favorables ;
- cette décision entraîne un éloignement entre le nouveau processus « en-devenir » et les « possibles ». C'est-à-dire dire que l'on s'éloigne du processus optimal, du point de vue simulation et du prototypage, représenté par les processus « possibles » ;
- cette diminution de l'ampleur du changement augmente aussi les trous fonctionnels que l'analyse des écarts va identifier (éloignement entre processus « en-devenir » et « possibles »).

Nous observons là un équilibre difficile à trouver entre notre démarche proactive et le prototypage (incluant l'analyse des écarts) et la simulation.

Validation. Après un ou plusieurs passages dans la boucle restructuration-extraction-prédiction-validation, nous sortons forcément par une validation effective. La préconisation de sortir de la boucle est donnée sur la base d'un compromis entre les résultats de la gestion proactive, de la simulation et du prototypage. Si l'on s'intéresse particulière à la gestion proactive, la validation doit intervenir quand les prédictions sont favorables. Cela consiste en des valeurs de prédictions négatives pour les trois variables de réponse.

L'acceptation donne la structure du modèle « en-devenir » qui doit être implémentée par la suite. Elle donne aussi les prédictions sur la variation de la performance opérationnelle correspondant au changement apporté. L'effort de gestion à engager envers les acteurs du processus est aussi estimé par le modèle. Les ressources nécessaires à la fermeture des trous fonctionnels doivent être disponibles. Comme pour l'effort de gestion du changement, les

moyens de fermeture des trous fonctionnels sont limités et liés au projet dans sa globalité. Les ressources qui peuvent être ou non engagées sur un processus dépendent de celles déjà engagées ou à engager dans l'ensemble des processus du projet.

3.4.6. Conclusion

Nous avons présenté, dans ce chapitre, l'utilisation du modèle bayésien obtenu pour la validation d'un processus en cours de restructuration. Cette validation se fait dans la phase d'adéquation et de configuration du processus d'implémentation de progiciel. Cette validation est réalisée sur la base d'informations issues du prototypage qui inclut l'analyse des écarts. Elle peut tenir compte aussi de résultats de la simulation sur le processus « en-devenir » proposé. Notre démarche de gestion proactive du changement, par l'utilisation du modèle bayésien, vient compléter ces deux méthodes. L'apport de notre démarche est la prise en compte de l'impact potentiel du changement en fonction de son ampleur et de l'effort de gestion consenti.

Nous avons pris en exemple un processus de gestion de réclamations en cas de décès. Les prédictions sur la variation des critères de performance ont été réalisées sur ce processus. Nous avons présenté les types de décision qui pouvaient être prises sur la base de ces prédictions.

Chapitre 3.5 Conclusion de la troisième partie

Cette partie correspond aux phases de modélisation, de validation et d'utilisation du modèle dans le processus de datamining. La recherche gloutonne des réseaux bayésiens et la validation croisée ont été choisies pour la construction et l'évaluation du modèle respectivement. Leur application a été ensuite présentée sur un échantillon créé automatiquement. Nous avons ensuite montré comment le modèle obtenu pouvait aider dans la prise de décision pour valider un processus en cours de restructuration dans un projet ERP. Au terme de cette partie, nous pouvons dire que la limite principale dans l'application de notre démarche se trouve dans la non-utilisation d'un historique réel. L'utilisation d'un tel historique aurait les avantages suivants :

- le modèle prédictif représentait une réalité avec des relations d'influence entre variables qui ont une signification réelle et visible,
- le retour d'expérience entre phase d'utilisation du modèle et de compréhension des données (Figure 7) pourrait être pris en compte de manière effective :
 - si on observe, par exemple, à partir de données réelles qu'une variable explicative n'est jamais utilisée dans la modélisation, la décision de la supprimer peut être prise ;
 - de même, il se peut que le modèle obtenu ait un taux d'erreurs global trop élevé. Cela signifierait que l'explication des variations de performance ne se trouve pas dans les quinze métriques structurelles de changement, de variation et d'effort de gestion que nous avons définies. Dans ce cas aussi, les phases d'apprentissage et d'utilisation du modèle renverraient un feedback à la phase de compréhension des données en suggérant de chercher d'autres variables explicatives à rajouter.

Ces deux décisions ne peuvent pas être prises si ces mêmes constatations sont faites dans un apprentissage basé sur des données créées automatiquement.

- la validation complète de la démarche proposée dépend de son application sur des données réelles.

CONCLUSION GENERALE

La problématique traitée dans cette étude est le changement dans les projets d'implémentation de progiciel. Il s'agit de proposer une méthode pour gérer le changement de manière proactive. La démarche proposée repose essentiellement sur l'application du datamining. Le datamining nous a permis de faire la combinaison d'éléments issus de la gestion quantitative du changement avec d'autres issus de la gestion non-quantitative. Les autres points centraux de notre démarche sont, l'utilisation du processus comme unité d'analyse et l'évaluation du changement par rapport à l'amélioration de la performance qui est l'objectif principal de la réingénierie dans les projets ERP. Le choix du processus comme unité d'analyse nous a permis d'avoir une cohérence entre la définition des variables, la collecte des données et l'utilisation du modèle obtenu. L'étude est basée sur trois concepts principaux dont nous avons défini les dimensions concrètes au niveau processus. Ces concepts sont l'ampleur, l'effort de gestion et l'impact du changement. Toute notre analyse repose sur l'étude de l'influence subie par le dernier concept de la part des deux premiers.

Nous avons défini des variables à expliquer comme des métriques opérationnelles de variation dans le délai, dans le coût et dans la non-qualité. Les données prédictives sont les métriques structurelles de changement et de variation, et celles d'effort de gestion du changement. La définition des métriques structurelles repose essentiellement sur la modélisation des processus et les métriques définies en analyse structurelle des processus. La définition des métriques opérationnelles de variation est basée sur la supervision des processus et les métriques qui y sont définies. L'apport principal de notre approche se trouve dans la quantification de l'ampleur, de l'effort de gestion et de l'impact du changement dans les variables de notre problème. La méthode Goal Question Metric a été utilisée pour définir ces variables. Une des valeurs ajoutées de ce travail réside dans la transposition d'un procédé d'extraction de métrique utilisé en génie logiciel pour mesurer le changement sur les processus métier.

Pour représenter les relations entre nos variables, nous avons fait le choix d'utiliser un modèle bayésien construit par recherche gloutonne. La validation du modèle est réalisée avec une validation croisée avec dix sous-échantillons. Nous avons défini le nombre de cas nécessaires à l'apprentissage et à la validation en fonction de ces deux méthodes. Un historique a été construit ensuite. Il est composé de processus déjà implémentés pour lesquels les trois variables de réponse et les quinze variables explicatives ont été extraites. Le modèle obtenu est censé être utilisé pour l'acceptation d'une version « en-devenir » proposée pour un processus en cours de restructuration dans un projet ERP. La gestion proactive du changement intervient dans cette acceptation en parallèle avec la simulation et le prototypage. Un consensus ou un équilibre doit être trouvé entre les directions préconisées par ces trois approches. La démarche proposée permet de répondre aux limites constatées dans les

approches quantitatives et non-quantitatives proposées pour la gestion du changement dans les projets ERP. Les limites qui ont été levées par rapport à la gestion quantitative du changement sont :

- notre unité d'analyse étant définie au niveau processus, elle correspond à celle du BPR qui apporte le changement dans les projets ERP ;
- les variables de réponse sont définies par rapport à l'objectif d'amélioration de la performance des processus. En effet, le changement est introduit en projet ERP par la réingénierie des processus dont cette amélioration est l'objectif principal ;
- une définition claire des décisions à prendre à partir des informations obtenues avec l'utilisation du modèle. Notre démarche propose l'activation de deux leviers : celui de l'ampleur du changement et celui d'effort de gestion ;
- la gestion quantitative ne propose que des actions de réduction indirecte de l'impact du changement par les causes. La gestion proactive offre une autre solution par l'activation du levier d'effort de gestion qui est une action de réduction directe de l'impact du changement agissant sur les conséquences.

Les apports de notre démarche par rapport à la gestion non-quantitative du changement sont :

- l'effort de gestion du changement est optimisé car il est quantifié et ciblé sur certains processus ;
- l'exploitation des connaissances formalisées et objectives issues des projets similaires antérieurs ;
- la gestion non-quantitative ne propose que des actions de réduction directes de l'impact du changement qui utilisent l'effort de gestion. La gestion proactive, par le levier d'ampleur du changement, offre une alternative qui consiste à mesurer et à ajuster le changement sur le processus afin de réduire son impact.

L'étude menée a contribué à démontrer la pertinence de la démarche proposée. Le déroulement des étapes clés de la démarche a été présenté à travers des exemples. Cependant, il existe plusieurs contraintes pour passer de la validation théorique à l'application intégrale de la démarche sur des processus réels. Ces contraintes sont liées principalement à la collecte des données. La première contrainte est la nécessité de collecter les données dans des entreprises différentes. En effet, plus les origines des données sont diverses, plus la collecte des observations est aléatoire et plus l'échantillon est représentatif de la population. Ces entreprises doivent toutes avoir implémenté un progiciel en utilisant la stratégie de BPR « déclenchée par la technologie ». La seconde contrainte est liée au respect des périodes fixées pour la supervision des processus et l'extraction des critères de performances. Supposons que nous choisissons par exemple de faire la supervision du processus « en-cours » et « implémenté » le dernier mois avant l'implémentation du progiciel, et le premier

mois après respectivement. La supervision de tout processus à intégrer dans l'historique doit être réalisée pendant ces deux périodes.

Pour contourner cette difficulté de collecter les données nécessaires à l'apprentissage du modèle nous avons généré aléatoirement les observations. Cela constitue l'une des limites principales de l'étude. Une de nos perspectives est de réaliser la mise en œuvre réelle de notre démarche. Par mise en œuvre réelle nous entendons l'apprentissage et l'application du modèle à partir et sur des processus réels. Nous pensons que la démarche peut être mise en œuvre par une entreprise de conseil qui recueillerait les données nécessaires au cours des différents projets ERP auxquels elle participera. Lorsque le modèle sera au point, cette même entreprise pourra l'utiliser dans de nouveaux projets pour valider ses restructurations de processus. Le problème de la confidentialité des données qui empêcherait la collecte des données peut ne pas se poser. En effet, les métriques extraites sur les processus ne sont pas de nature à dévoiler des informations stratégiques sur les processus de l'entreprise.

La seconde limite principale de l'étude est la non-prise en compte de certaines dimensions abstraites du changement et de son impact. En effet, le changement est très lié au facteur humain qui fait intervenir des aspects culturels, sociologiques et psychologiques. Les approches sur-socialisées de la gestion du changement peuvent aider à combler cette limite. Une troisième limite de l'étude est liée à l'absence de formalisation claire de l'évolution des compétences lors de la mise en place du progiciel. En effet, la transformation des compétences est indirectement ou implicitement traitée à travers la mesure de l'ampleur du changement sur les activités à réaliser. Bien qu'il y ait une relation entre ces deux concepts, nous ne pouvons pas dire qu'il y a une correspondance directe entre ceux-ci. Plusieurs activités peuvent, par exemple, être réalisées par un même rôle qui représente la compétence.

Parmi les perspectives à court terme, figure la prise en compte de l'interopérabilité des métriques explicatives. Elle nécessite la définition des variables explicatives indépendantes du langage de modélisation de processus utilisé. La seule notation prise en compte dans notre étude est le BPMN. Une solution peut être trouvée par la mise en correspondance de plusieurs méta-modèles de processus (Darras, 2004) (Roque, 2005) en utilisant les ontologies de processus métier (Camara, et al., 2008). Une autre perspective à court terme est d'inclure l'intégration et la flexibilité dans la mesure de la variation dans les propriétés des processus. Dans la section 2.4.5, nous avons conclu que ces caractéristiques n'étaient pas calculables avec les informations contenues dans le méta modèle BPMN. Cela était dû à l'indisponibilité des informations comme le type de livraison et les propriétés de flexibilité de Tjaden (2001) ou les propriétés d'intégration de Barki, et al. (2003). Trouver une modélisation des processus qui permet de représenter ces informations aiderait à traiter cette perspective.

REFERENCES

Aamodt Agnar and Plaza Enric Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches [Journal] / ed. Press IOS. - [s.l.] : AICom - Artificial Intelligence Communications, 1994. - 1 : Vol. 7. - pp. 39-59.

Adderley Richard, Townsley Michael and Bond John Use of data mining techniques to model crime scene investigator performance [Conference] // the Twenty-Sixth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence. - [s.l.] : Springer, 2006.

Adriaans Pieter and Zantinge Dolf Data mining [Book]. - [s.l.] : Addison-Wesley, 1996.

Aguilar E., R. and Ruiz F., Garcia, F., Piattini, M., Evaluation measures for business process model [Conference] // Symposium on Applied Computing (SAC). - 2006B. - pp. 1567-1568.

Aguilar E.,A. and Ruiz F., Garcia, F., Piattini, M., Applying Software Metrics to evaluate Business Process Models [Journal]. - [s.l.] : CLEI electronic journal, 2006A. - Vol. 9.

Aladwani Adel M. Change management strategies for successful ERP implementation [Journal]. - [s.l.] : Business Process Management Journal, 2001. - 3 : Vol. 7.

Al-Mashari Majed, Zairi Mohamed and Okazawa Koji Enterprise Resource Planning (ERP) implementation: a useful road map [Journal]. - [s.l.] : Int. J. Management and Enterprise Development, 2006. - 1/2 : Vol. 3.

Anderson R., D., Lenz, R., T., Modeling the Impact of Organizational Change: A Bayesian Network Approach [Journal]. - [s.l.] : Organizationnal Research Methods, 2001. - 1 : Vol. 4.

Appleton E. L. How to survive ERP [Journal]. - [s.l.] : Datamation, 1997. - 3 : Vol. 43.

Arrègle Jean-Luc Les modèles linéaires hiérarchiques : principes et illustration [Journal]. - [s.l.] : M@n@gement, 2003. - 1 : Vol. 6.

Aubert A. Benoit, Vandenbosch Betty and Mignerat Muriel Towards tthe Measuement of Process Integration [Report] / Cirano. - [s.l.] : Série Scientifique, 2003.

Baillargeon G. Statistique Appliquée et Outils d'Amélioration de la Qualité [Book]. - [s.l.] : SMG, 2004.

Barki Henri and Pinsonneault Alain The Construct of Organizational Integration: A Research Framework and Its Application to Enterprise Systems Research [Journal]. - [s.l.] : Cahier du GReSI, 2003. - 03.

Barki Henri, Oktamis Sirel and Pinsonneault Alain Dimensions of ERP Implementations and Their Impact on ERP Project Outcomes [Report]. - [s.l.] : Cahier de la Chaire de recherche du Canada en implantation et gestion des technologies de l'information, 2004.

Basili Victor R., Caldiera G. and Rombach H. Dieter The goal question metric paradigm [Journal] / ed. John. - [s.l.] : In Encyclopedia of Software Engineering, 1994. - Vol. 2. - pp. 528–532.

Bea Bea and ProActivity Managing the Business Process Lifecycle. [Report]. - 200X.

Bernard Jean gregoire, Rivard Sjjer and Aubert Bhghdf Évaluation du risques d'implémentation de progiciel [Report] / Hydro-Québec, VRQ (Valorisation Recherche Québec) et Cirano. - 2002.

Bernier Carmen and Roy Vital L'évolution des rôles dans la gestion des projets TI : le cas des ERP [Journal]. - [s.l.] : Revue Internationale de gestion, 2003. - 2 : Vol. 28.

Boar B, H Implementing Client/Server Computing [Journal]. - New York : McGraw-Hill, 1993.

Bootcher Susanne, G and Dethlefsen Claus deal: A Package for Learning Bayesian networks [Journal] // Journal of Statistical Software. - 2003A. - pp. 1-40.

Bootcher Susanne, G and Dethlefsen Claus Learning Bayesian Networks with R [Conference] // The 3rd International Workshop on Distributed Statistical Computing (DSC 2003). - Vienna, Austria : [s.n.], 2003B.

Borysowich Craig Identifying and Measuring Cost drivers [Online] // ITtoolbox. - 08 16, 2007. - 05 21, 2008.

Borysowich Craig Cycle Time Collection Methods [Online] // <http://it.toolbox.com>. - Mars 19, 2008. - Avril 11, 2008. - <http://it.toolbox.com/blogs/enterprise-solutions/cycle-time-collection-methods-23166>.

Boucher Xavier and Crestani Didier Decision Support Systems Issues in Enterprise Change Management [Journal]. - [s.l.] : Journal of Decision Systems (JDS), 2006. - 4 : Vol. 15.

Breiman Leo [et al.] Classification and Regression Trees [Book]. - California : Wadsworth international, 1984.

Bruges Paul ERP Implementation Methodologies [Report] / University of Missouri-St-Louis. - 2002.

Burnes Managing change : A Strategic Approach to Orgaazation dynamics [Book]. - London : pitman, 1996.

Cain Michael and Janssen Christian Bayesian valuation with an elicited nonsymmetric loss function [Book Section] // Bayesian analysis in statistics and econometrics / book auth. Berry Donald A. [et al.]. - [s.l.] : John Wiley & Sons Inc, 1996.

Camara Mamadou, Kermad Lyes and El Mhamedi Abderrahman Analyse des risques dans les projets d'implémentation des ERP : Adaptation de la méthode OSR [Conference] // GI'2005. - Besaçon : [s.n.], 2005.

Camara Mamadou, Kermad Lyes and El Mhamedi Abderrahman Organizational Change Measurement via Change Metrics [Book Section] // Innovations and Advanced Techniques in Computer and Information Sciences and Engineering. - Netherlands : Springer, 2007b.

Camara Mamadou, Kermad Lyes and El Mhamedi Abderrahman Proactive Management of Business Change [Book Section] // Pervasive Collaborative Networks. - Boston : Springer , 2008.

Camara Mamadou, Kermad Lyes and El Mhamedi Abderrahman Proactive Management of organizational change using bayesian network [Conference] // MITIP 2007. - Florence 6-7 September : [s.n.], 2007a.

Camara Mamadou, Kermad Lyes and El Mhamedi Abderrahman Risk Prediction in ERP Projects: Classification of Reengineered Business Processes [Conference] // International Conference on Computational Inteligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06). - Sydney : [s.n.], 2006.

Cardoso J. [et al.] A discourse on complexity of process models. [Conference] // BPM 2006 workshops. - 2006.

Cardoso Jorge Complexity Analusis of BPEL Web Processes [Journal] // Software Process: Improvement and Practice. - 2006.

Cardoso Jorge How to Measure the Control-flow Complexity of Web processes and Workflows [Book Section] // Workflow Handbook / ed. Fischer Layna. - 2005.

Casin Philippe Analyse des données et des panels de données [Book]. - [s.l.] : De Boeck Université, 1999.

Catlet J Megainduction: Machine Learning on Very large Databases [Report] : Phd Thesis / University of Sydney. - Sydned : [s.n.], 1991.

Chapron Julie L'urbanisme organisationnel : méthode et aide à la décision pour piloter l'évolution du système d'information d'entreprise [Report]. - [s.l.] : Ecole nationale supérieure des mines de St Etienne & Université Jean Monnet., 2006.

Chou Philip A Optimal Partitioning for Classification and Regression Trees [Journal] // IeeeTransactions On Pattern Analysis And Machine Intelligence. - 1991. - 4 : Vol. 13.

Chrissis Mary Beth, Konrad Mike and Shrum Sandy CMMI: guidelines for process integration and product improvement [Book]. - [s.l.] : Addison-Wesley, 2003.

Cohen Jacob and Cohen Patricia Applied multiple regression/correlation analysis for the behavioral sciences [Book]. - [s.l.] : Hillsdale, N.J. , L. Erlbaum Associates, 2003.

Collins David Organizational Change : Sociological Perspectives [Book]. - [s.l.] : Routledge , 1998.

Colosimo Bianca M. and Del Castillo Enrique Bayesian Process Monitoring, Control and Optimization [Book]. - [s.l.] : CRC Press, 2006.

Consortium Software Productivity The software measurement guidebook [Book]. - [s.l.] : International Thomson Computer Press, 1995.

Cooke D. and Peterson W. SAP implementation: strategies and results [Report]. - New York : The Conference Board, 1998.

Crochet-damais Antoine BPM pour Business Performance Management [Online] // LeJournalduNet. - Janvier 4, 2006. - Fevrier 28, 2009. - http://www.journaldunet.com/solutions/0601/060104_qr-business-performance-management.shtml.

Darras Franck Proposition d'un cadre de référence pour la conception et l'exploitation d'un progiciel de gestion intégré [Report] / L'Institut National Polytechnique de Toulouse. - 2004.

Debauche Bernard and Patrick Mégar BPM Business Process Management : Pilotage métier de l'entreprise [Book]. - [s.l.] : Hermes science publications, 2004.

Deixonne Jean-Luc Piloter un projet ERP [Book]. - [s.l.] : Dunod, 2001.

Deloitte and Touche survey of 400 U.S and Canadiann CIOs [Journal]. - 1993.

Deloitte ERPs second wave [Report]. - [s.l.] : Deloitte Consulting, 1999.

Deloitte&Touche Danske erfaringer med implementering af integrerede informationssystemer [Report]. - [s.l.] : Deloitte & Touche Consulting Group, 1998.

Demeyer Serge, Ducasse Stéphane and Nierstrasz Oscar Finding Refactorings via Change Metrics [Conference] // the 15 th ACM SIGPLAN conference on Object -Oriented Programming, Systems, Languages, and applications. - 2000.

Draper N. R and Smith H. Applied Regression Analysis [Book]. - [s.l.] : Willey Series in Probability and Statistics, 1998.

Dudoit Sandrine and Laan Mark J. van der Asymptotics of Cross-Validated Risk Estimation in Estimator Selection and Performance Assessment [Report]. - [s.l.] : U.C. Berkeley Division of Biostatistics Working Paper Series, 2003.

El Amrani Redouane and Geffroy Morannat Bénédicte ERP implementation cross-functionality and critical change factors [Book Section] // Enterprise information systems V / book auth. Camp Olivier and Filipe Joaquim B. L.. - [s.l.] : illustrated, 2004.

Eriksson Darek M. Multi-modal investigation of a business process and information system redesign: a postimplementation [Journal]. - [s.l.] : Darek Systems Research and Behavioral Science, 2001. - Vol. 18.

Esteves José and Pastor Joan A A framework to analyse most critical work packages in ERP implementation projects [Conference] // International Conference on Enterprise Information Systems (ICEIS). - 2002A.

Esteves José, Pastor, Joan, Casanovas Josep Monitoring Business Process Redesign in ERP Implementation Projects [Journal]. - 2002B.

Esteves sousa José, Pastor, collado Joan Towards the unification of critical success factors for ERP implementations [Conference] // 10 th Annual Business Information Technology (BIT) 2000 Conference . - 2000.

Forcht Karen A. [et al.] Implementing Enterprise Resource Planning (ERP) for Strategic competitive advantage [Journal]. - [s.l.] : Issues in Information Systems, 2007. - 2 : Vol. 8.

García Díaz and Martha Liliana A methodology to facilitate continuous improvement in the services provided by the facilities department at UPRM [Report]. - [s.l.] : University of Porto Rico, 2004.

Gartner Group Business Activity Monitoring: The Data Perspective. [Report]. - 2002.

Geiger Dan and Heckerman David Learning Gaussian Networks [Report]. - [s.l.] : Microsoft Research Advanced Technology Division , 1994.

Geman Stuart, Bienenstock Elie and Doursart René Neural Networks and Bias/Variance Dilemma [Journal] // Neural Computation. - 1992. - Vol. 4. - pp. 1-58.

Gendron M Learning to live with the electronic embodiment of reengineering [Journal]. - [s.l.] : Harvard Management Update, 1996 .

Geurts Pierre Contribution to Decision Tree Induction: Bias/Variance Tradeoff and Time Series Classification [Report] : Phd Thesis / Université de Liège. - Liège : [s.n.], 2002.

Geurts Pierre Ensembles d'arbres extrêmement aléatoire : Application à la classification d'images [Report] : Séminaire LIP6 / LIP6. - Paris : [s.n.], 2004.

- Giard Vincent** Gestion de projet [Book]. - [s.l.] : Economica, 1991.
- Gokhale S., S. and Lyu M., R.,** Régression Tree Modeling for the Prediction of Software Quality [Conference] // ISSAT'97. - Anaheim, CA : [s.n.], 1997.
- Graham Laurie** How does the Japanese model transfer to the united states? A view from the line [Book Section] // Global Japanization?: the transnational transformation of the labour process / book auth. Elger Tony and Smith Chris. - [s.l.] : Routledge, 1994.
- Greasley Andrew** Effective use of business process simulation [Conference] // Winter Simulation Conference. - 2000.
- Grover Varun and Kettinger William J.** Business process change : reengineering concepts, methods and technologies [Book]. - London : Idea Group Publishing, 1998.
- Gruhn Volker and Laue Ralf** Complexity Metrics for Business Process Models [Conference] // 9th International Conference on Business Information Systems. - Klagenfurt : [s.n.], 2005.
- Hammer Michel and Champy James** Le Reengineering [Book]. - Paris : Dunod, 2003.
- Hammer Michel and Stanton Steven A** The Reengineering revolution [Book]. - [s.l.] : HarperCollins, 1995.
- Hammer Michel** Carnet de route pour le manager [Book]. - Paris : Maxima, 2002.
- Hammer Michel** Re-engineering Work: Don't Automate, Obliterate [Journal]. - [s.l.] : Harvard Business Review, 1990. - Vol. 68.
- Han Jiawei and Kamber Micheline** Data Mining : Concepts and Techniques [Book]. - [s.l.] : Morga Kaufmann Series in Data Management Systems, 2001.
- Hanrahan Robert P** The IDEF Process Modeling Methodology [Report]. - 1995.
- Hastie T., Tibshirani R. and Friedman J.** The Eléments of Statistical Learning [Book]. - New York : Springer, 2001.
- Heili Joseph** Communautés et activités : Deux leviers pour améliorer l'insertion d'un E.R.P. dans une organisation [Report]. - 2004.
- Hissilä Heli** Model validation [Online] // Research and education at CIS. - 10 3, 2006. - Septembre 3, 2007.
- Hitt M.A., Hoskisson E. and Nixon R.D.** A Mid-range Theory of Inter-functional Integration, its Antecedents and Outcomes [Journal]. - [s.l.] : Journal of Engineering and Technology Management, 1993. - 11 : Vol. 10.
- Holland D** It is obvious that ERP (Enterprise Resource Planning) is the big thing in technology enabling solutions. Why are so many companies still reporting marginal results

after implementing such systems? [Online] // The Institute for Change Leadership's, Change Management. - Avril 1998. - Library <http://www.hdinc.com/library.html>.

Hox J. J. Applied Multilevel Analysis [Book]. - Amsterdam : TT Publikaties, 1995.

Humphrey Watts S. Managing the software process [Book]. - [s.l.] : Addison-Wesley, 1989.

Huyet A and Paris J L Optimisation intelligente pour les systèmes de production [Journal] // Journal Européen des Systèmes Automatisés. - 2005.

Ian H. Witten and Eibe Frank Data Mining: Practical Machine Learning Tools and Techniques [Book]. - [s.l.] : Morgan Kaufmann, 2005.

IBM Documetation IBM Websphere Business Modeler Advanced [Online] // publib.boulder.ibm.com. - 200X. - 01 23, 2008.

IDS Scheer ARIS : Des Processus de gestion au Système Intégré d'Applications [Report]. - 2002.

Intermec Aliging Your Data Collection and ERP Implementation Decisions [Report]. - [s.l.] : Intermec Technologies Corporation, 1999.

ISO ISO 9001 VERSION 2000. - 2000.

itp commerce Overview : Process Modeler for Microsoft Visio [Report] / itp commerce ltd. - 200X.

Jenz D. E. BPMO Tutorial : Defining a Private Business Process in a Knowledge base [Report]. - [s.l.] : Jenz & Partner GmbH, 2003A.

Jenz D., E., Business Management Ontology (BMO) [Report]. - 2004.

Jenz D., E., Business Process Ontologies: Frenquently Asqued Questions [Report]. - [s.l.] : Jenz & Partner GmbH, 2003B.

Jin Li-Jie, Casati Fabio and Shan Ming-Chien Business Process Simulation with HP Process Manager [Report]. - [s.l.] : HP Laboratories Palo Alto, 2001.

Kenley C. Robert Gaussian influence diagrams [Journal]. - [s.l.] : Management science, 1989. - 5 : Vol. 35.

Kenley C. Robert Influence diagram models with continous variables [Report] / The departement of engineering-economic systems of stanford university. - 1986.

Kermad L. Roucou P., El mhamedi A., Les conditions de succès d'un projet ERP [Conference] // 5ème Congrès International de Génie Industriel. - Québec : [s.n.], 2003.

Kim Henry M. Representing and Reasoning About Quality Using Enterprise Models [Report]. - [s.l.] : Enterprise Integration Laboratory University of Toronto, 1999.

Kliem Ralph L Risk management for business process reengineering projects [Journal] // Information systems management. - 2000. - 4 : Vol. 17. - pp. 71-73.

Kohavi Ron A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection [Conference] // International Joint Conference on Artificial Intelligence. - 1995.

Küster Jochen [et al.] Methodology and Tooling to combine an existing legacy business process model with best-practice industry reference models for Business Transformation [Report]. - [s.l.] : IBM Zurich Research Laboratory, 2005.

Lawrence P. R. and LORSCH J. W. Differentiation and integration in complex organizations [Journal]. - [s.l.] : Administrative Science Quarterly, 1967. - 1 : Vol. 12.

Least Squares Approximation [Online] // <http://ceee.rice.edu>. - Juin 20, 2007. - <http://ceee.rice.edu/Books/LA/leastsq/index.html>.

Legendre Louis Numerical ecology [Book]. - [s.l.] : Elsevier, 1998.

Legge HRM : rhetorics and realities [Book]. - London : Macmillan, 1995.

Leray P Réseaux bayésiens : apprentissage et modélisation de systèmes complexes [Report]. - 2006.

LERAY P. Réseaux bayésiens : apprentissage et modélisation de systèmes complexes [Report]. - 2006.

Lyonnet Patrick La qualité : outils et méthodes [Book]. - [s.l.] : Tec&Doc Lavoisier, 1997.

Maaloul Imen L'implantation des ERP et l'ingénierie du changement : les déterminants de la satisfaction des utilisateurs d'un ERP dans les entreprises tunisiennes [Conference] // XIIème Conférence de l'Association Internationale de Management Stratégique. - 2003.

Malhotra Yogesh Business Process Redesign: An Overview [Journal] // IEEE Engineering Management Review. - 1998. - 3 : Vol. 26.

Markus M. Lynne and Tanis Cornelis The Enterprise System Experience : From Adoption to Success [Book Section] // Framing the Domains of IT Management : Projecting the Future...Through the Past / book auth. Zmud Robert W.. - [s.l.] : Pinnaflex Education Resource, 2000.

Mayer Richard J. [et al.] IDEF3 process description capture method [Report] / University drive east. - 1995.

Meier Olivier et al. Gestion du changement [Book]. - [s.l.] : Dunod, 2007.

Michalski Ryszard S A Theory and Methodology of Inductive Learning [Journal] // Artificial Intelligence. - 1983. - 2 : Vol. 20. - pp. 111-161.

Millet Pierre-Alain Une étude de l'intégration organisationnelle et informationnelle : Application aux systèmes d'informations de type ERP [Report]. - Lyon : L'institut national des sciences appliquées de Lyon, 2008.

Milutinovic Veljko and Patricell Frédéric E-business and e-challenges [Book]. - [s.l.] : IOS Press, 2002.

Misof Daniel Process Costing with ARIS Business Optimizer [Report]. - [s.l.] : ARIS, 2007.

Morley Chantal [et al.] Processus metier et SI [Book]. - Paris : Dunod, 2007.

Nah Fiona Fui-Hoon and Lee-Shang Lau Janet Critical factors for successful implementation of enterprise systems [Journal]. - [s.l.] : Business Process Management Journal, 2001. - 3 : Vol. 7.

Nah Fiona Fui-Hoon, Zuckweiler Kathryn M. and Lau Janet Lee-Shang ERP Implementation: Chief Information Officers' Perceptions of Critical Success Factors [Journal] // International Journal of Human-Computer Interaction. - 2003. - Vol. 16. - pp. 5-22.

Nakache Jean-Pierre, Canfais, Josiane Approche pragmatique de la classification : Abres hiérarchiques, Partitionnements [Book]. - [s.l.] : TECHNIP, 2005.

O'Leary Daniel E. Enterprise Resource Planning Systems: Systems, Live Cycle, Electronic Commerce, and Risk [Book]. - [s.l.] : Cambridge university press, 2000.

OCTO Le BPM à la recherche du chaînon manquant [Online] // www.octo.com. - OCTO Technology - Communication - Revue de presse, Novembre 29, 2007. - Janvier 15, 2009. - <http://www.octo.com/com/pdf/01info291107.pdf>.

OMG Business Process Maturity Model (BPMM), Beta 1 [Report]. - [s.l.] : OMG, 2007A.

OMG Specification Business Process Modeling Notation Specification [Report]. - 2007B.

Osborne Jason W. Advantages of Hierarchical Linear Modeling [Online] // Practical Assessment, Research & Evaluation. - 2000A. - Janvier 29, 2009. - <http://PAREonline.net/getvn.asp?v=7&n=1>.

Osborne Jason W. Prediction in Multiple Regression [Journal]. - [s.l.] : Practical Assessment, Research & Evaluation, 2000B. - 2 : Vol. 7.

Park Colin N. and Dudycha Arthur L. A cross-validation approach to sample size determination [Journal]. - [s.l.] : Journal of the American Statistical Association, 1974. - 345 : Vol. 69.

Parr A and Shanks G. A Taxonomy of ERP Implementation Approaches [Conference] // 33rd Hawaii International Conference on System Sciences. - [s.l.] : R.H, 2000.

Paulk Mark C. [et al.] The Capability Maturity Model: Guidelines for Improving the Software Process, [Report]. - [s.l.] : Addison-Wesley Publishing Company,, 1995.

Pedhazur Elazar J. Multiple regression in behavioral research [Book]. - [s.l.] : Harcourt Brace, 1997.

Perner Petra and Trautzsch Sascha Multi-Interval Discretization Methods for Decision Tree Learning [Book Section] / book auth. Amin A [et al.]. - [s.l.] : Springer Verlag, 1998.

Perotin Pascal Les Progiciels de Gestion Intégrés, instruments de l'intégration organisationnelle? Etude de cas [Report]. - [s.l.] : Université de Montpellier II, 2004.

ProcessSoft ABC et process costing [Report]. - 200X.

Prosci Prosci's Change Management Maturity Model [Report]. - 2004.

quantlet.org Supervised and Unsupervised Learning [Online] // www.quantlet.org. - 200X. - Octobre 15, 2008.

Quinlan J.R. C4.5: Programs for Machine Learning, . [Book]. - [s.l.] : Morgan Kaufmann, 1993.

Rajagopalan N. and Spreitzer G. Toward a Theory of Strategic Change: A Multi-Lens Perspective and Integrative Framework [Journal]. - [s.l.] : Academy of Management Review, 1997. - 48-79 : Vol. 22.

Rakotomalala Ricco Arbres de classification : Apprentissage non-supervisé ou apprentissage multi-supervisé ? [Report] / Laboratoire ERIC. - 200A.

Rakotomalala Ricco Arbres de décision [Report] / Laboratoire ERIC. - 200D.

Rakotomalala Ricco Arbres de Régression : Prédiction d'une variable continue [Report] / Laboratoire ERIC. - 200C.

Rakotomalala Ricco Graphes d'induction [Report]. - Lyon : [s.n.], 1997.

Rakotomalala Ricco Introduction au Data Mining [Report] / Laboratoire ERIC. - 200B.

Raphaëli Orit, Zahavi Jacob and Kenett Ron Applying Case Based Reasoning Approach in Analyzing Organizational Change Management Data [Conference] // ICDM 2004. - [s.l.] : Springer-Verlag Berlin Heidelberg , 2004.

Ripley B. D., Pattern Recognition and Neural Networks [Article] // Cambridge University Press. - 1996.

Robert Christian P. Le Choix Bayésien: Principes Et Pratique [Book]. - [s.l.] : Springer, 2006.

Romanelli Elaine and Tushman Michael L. Organizational Transformation as Punctuated Equilibrium: An Empirical Test [Journal]. - [s.l.] : Academy of Management Journal, 1994. - 5 : Vol. 37.

Roque Matthieu Contribution à la définition d'un langage générique de modélisation d'entreprise [Report]. - [s.l.] : Université de Bordeaux 1, 2005.

Roweis Sam Overfitting and capacity control [Online]// university of toronto departement of computer science. - Novembre 18, 2003. - Septembre 03, 2007.

Saint-leger G. Neubert G., Pichot L., Projet ERP : Incidence des spécificités des entreprises sur les facteurs clés de succès [Conference] // 7ème colloque de l'AIM 2002. - Hammamet, Tunisie : [s.n.], 2002.

Saint-Leger Guy Quel processus de changement peut permettre une mise en oeuvre et une utilisation efficace et efficiente d'un système d'information de type ERP dans les moyennes structures de production de bien et de service? [Report]. - [s.l.] : Université Jean Moulin Lyon III, 2005.

Scheer A, W Aris - Business Process Modeling - [Book]. - New York : Springer Verlag, 1998. - pp. 16-17.

Scheer August-Wilhelm ARIS Des processus de gestion au système intégré d'applications [Book]. - 2002.

Scott J.E. and Vessey I. "Implementing Enterprise Resource Planning Systems: The Role of Learning From Failure," [Journal]. - [s.l.] : Information Systems Frontiers, 2001.

Seber George A. F. and Lee Alan J. Linear regression analysis [Book]. - [s.l.] : WILEY series in probability as statistics, 2003.

Shields Murrell G. E Business and ERP Rapid Implementation and Project [Book]. - [s.l.] : John Willey & Sons, 2001. - pp. 1-32.

Sienou Amadou, Karduck Achim and Pingaud Hervé Towards a framework for integrating risk and business process management [Conference] // INCOM. - Oxford, UK : Elsevier Science, 2006. - pp. 615-621.

Sienou Amadou, Karduck Achim and Pingaud Hervé TOWARDS A FRAMEWORK FOR INTEGRATING RISK AND BUSINESS PROCESS MANAGEMENT [Conference] // INCOM. - Oxford, UK : Elsevier Science, 2006. - pp. 615-621.

Simon Kai A. Towards a theoretical framework for Business Process Reengineering [Report]. - 1994.

Society The British Computer IT Projects: sink or swim? Computer Bulletin [Online] // www.bcs.org.uk. - 2000. - <http://www.bcs.org.uk/publicat/ebull/jan2000/article1.htm>.

sourceforge Jena A semantic Web Framework for java [Online] // jena.sourceforge.net. - 2008. - <http://jena.sourceforge.net>.

Stanford university protégé [Online] // protege.stanford.edu. - 2008. - <http://protege.stanford.edu>.

Staudenmayer N., Tyre M. and Perlow L. Time to change : Temporal Shifts as Enablers of Orgaizational Change [Journal]. - [s.l.] : Orgaization Science, 2002. - 5 : Vol. 13.

Stefanou CJ A framework for the ex-ante evaluation of ERP software [Journal]. - [s.l.] : European Journal of Information Systems, 2001.

Stefanou CJ Supply chain management (SCM) and organisational key factors for successful implementation of enterprise resource planning (ERP) systems [Conference] // Fifth Americas Conference on Information Systems. - Milwaukee : [s.n.], 1999.

Stoica Mihail, Chawat Nimit and Shin Namchul An Investigation of the Methodologies of Business Process Reengineering [Journal] // Information Systems Education Journal. - 2004. - 11 : Vol. 2.

Systems Sparx Business Process Modelling [Online] // <http://www.sparxsystems.com.au>. - 2008. - Octobre 2008. - http://www.sparxsystems.com.au/business_process_model.html.

Tabachnik Barbara G. and Fidell Linda S. Using multivariate statistics [Book]. - Boston : Allyn and Bacon, 2001.

Tatsiopoulos Ilias P. [et al.] Risk management as a strategic issue for the implementation of ERP systems: a case study from the oil industry [Journal]. - [s.l.] : International Journal of Risk Assessment and Management, 2003. - 1 : Vol. 4.

Tjaden G., S., Business Process Structural Analysis [Report] / Georgia Tech Center for Enterprise Systems. - 2001.

Tomas Jean-Louis ERP et progiciels intégrés : la mutation des systèmes d'information [Book]. - Paris : Dunod, 2000.

Tsai Hui-Liang Information Technology and Business Process reengineering [Book]. - [s.l.] : New perspective and Strategies, 2003.

Tufféry Stéphane Data mining et statistique décisionnelle: L'intelligence des données [Book]. - [s.l.] : Ophrys, 2007.

Turban E, McLean E and Wetherbe J Information Techology for Management [Book]. - [s.l.] : John Wiley&Sons, 2000.

Uschold Mike [et al.] The Enterprise Ontology [Journal] // The Knowledge Engineering Review Vol 13. - 1998.

Vallée F and Ragot A optimisation des tests par la mesure en conception préliminaire [Conference] // SEE Software Testing'97. - 1997.

van der Aalst Wil M.P. The Application of Petri Nets to Workflow Management [Journal] // The Journal of Circuits, Systems and Computers. - 1998.

Vanderfeesten I., Cardoso, J., Mendling, J., Reijers, H., A. and Aalst W.,l van der Quality Metrics for Business Process Models [Book Section] // Workflow handbook. - [s.l.] : Layna Fisher, 2007.

Varlet Drt and Vallée Fdf Evaluation de la fiabilité en phase de conception logicielle [Article] // Genie Logiciel. - 2003.

Vernadat F. Enterprise modeling and integration : Principes and Applications. [Book]. - [s.l.] : Chapman & Hall, 1996.

Villeneuve Arnol Gestion du changement pour les professionnels des technologies de l'information (TI) [Report]. - [s.l.] : Learning Tree International, 2001.

Vose David Quantitative Risk Analysis: A Guide to Monte Carlo Simulation Modelling [Book]. - [s.l.] : John Wiley & Sons, 1996.

webmethods Business Activity Monitoring (BAM) The New Face of BPM [Online]. - June 2006. - http://www1.webmethods.com/PDF/whitepapers/BAM-The_New_Face_of_BPM.pdf.

White Stephen A. Introduction to BPMN Introduction to BPMN [Online] // www.bpmn.org. - Octobre 19, 2007. - Janvier 10, 2009. - <http://www.bpmn.org/Documents/Introduction%20to%20BPMN.pdf>.

Wischnevsky J. Daniel and Damanpour Fariborz Organizational tranformation and firm performance: A conceptual and empirical examination of three perspectives [Report]. - [s.l.] : Proceedings of the Decision Sciences Institute, 2001.

wsper.org BPMN 1.0 [Online] // www.wsper.org. - Octobre 19, 2007. - Octobre 20, 2008. - www.wsper.org/bpm10.html.

yoyodesign owlguide [Online] // [yoyodesign.com](http://www.yoyodesign.com). - 2004. - <http://www.yoyodesign.org/doc/w3c/owl-guide-2040210>.

Zairi M. and Sinclair D. Business process re-engineering and process management: a survey of current practice and future trends in integrated management [Report]. - [s.l.] : Work Study, 1995.

ANNEXES

Annexe 1 : Méthodologie de mise en place d'un ERP

1. la planification (Tomas, 2000 p. 164)
 - a) la constitution des différentes équipes,
 - b) le plan du projet avec les hypothèses,
 - c) les réunions clés de communication
 - d) l'environnement de travail des équipes,
 - e) l'encyclopédie du projet,
 - f) le plan de transition,
 - g) le plan de conversion des données,
 - h) le plan qualité,
 - i) le plan des risques
 - j) le plan sécurité
2. l'analyse opérationnelle (Tomas, 2000 p. 169),
 - a) documentation à haut niveau des processus « en-cours »,
 - b) décomposition des processus « en-devenir » en processus élémentaires,
 - c) documentation des processus « en-devenir » majeurs et élémentaires,
 - d) compréhension des processus « en-devenir » par les consultants,
 - e) acceptation formelle des livrables par le comité de pilotage,
 - f) plan et encyclopédie de projet mis à jour.
3. la formation des équipes projet (Tomas, 2000 p. 175),
 - a) plan de formation,
 - b) environnement de formation,
 - c) feuilles d'évaluation des formations reçues,
 - d) formation des équipes projet—remplaçants inclus,
 - e) plan d'actions d'amélioration des formations,
 - f) acceptation formelle des livrables par le comité de pilotage,
 - g) plan et encyclopédie de projet mis à jour.
4. l'adéquation et la configuration (Tomas, 2000 p. 191),
 - a) connaissance approfondie de l'ERP par les équipes de l'entreprise,
 - b) configuration de l'ERP pour supporter les processus opérationnels,
 - c) liste des trous fonctionnels potentiels (TFP),
 - d) processus modifiés pour une meilleure adéquation à l'ERP,
 - e) scripts ERP supportant les processus opérationnels,
 - f) procédures opérationnelles,
 - g) premier niveau d'adéquation pour les documents et les rapports,
 - h) premier niveau d'adéquation pour les conversions de données et interfaces,
 - i) plan de test et tests unitaires des processus opérationnels,
 - j) machine de production initialisée avec les scripts ERP,
 - k) acceptation formelle des livrables par le comité de pilotage,
 - l) plan et encyclopédie de projet mis à jour.
5. les simulations taille réelle (Tomas, 2000 p. 196),
 - a) confirmation et validation de la phase adéquation-configuration,

- b) scénarios opérationnels correctement implémentés dans l'ERP,
 - c) liste des simulations taille réelle effectuées,
 - d) documentation complète des simulations réalisées,
 - e) plan d'actions pour résoudre les problèmes rencontrés,
 - f) équipes projet confiantes et motivées pour les phases suivantes,
 - g) acceptation formelle des livrables par le comité de pilotage,
 - h) plan et encyclopédie de projet mis à jour.
6. la fermeture des trous fonctionnels (Tomas, 2000 p. 204),
- a) documentation des TFP, les trois solutions alternatives incluses,
 - b) les TFP convertis en trous réels fonctionnels,
 - c) description des solutions systèmes associées à ces derniers,
 - d) les autres TFP et la solution associée : statu quo ou solution opérationnelle,
 - e) acceptation formelle des livrables par le comité de pilotage,
 - f) plan et encyclopédie de projet mis à jour.
7. les modifications spécifiques (Tomas, 2000 p. 211),
- a) méthodologie et standards,
 - b) panorama de l'état d'avancement des modifications spécifiques,
 - c) documents spécifications fonctionnelles,
 - d) documents design système,
 - e) coding et tests des modifications,
 - f) plan et résultats des tests d'acceptation,
 - g) acceptation formelle des modifications,
 - h) documentation utilisateur modifiée,
 - i) modifications spécifiques installées sur la machine de production,
 - j) acceptation formelle des livrables par le comité de pilotage,
 - k) plan et encyclopédie de projet mis à jour.
8. la création des liens avec l'environnement (Tomas, 2000 p. 219),
- a) matrice de correspondance entre les données actuelles et les données futures,
 - b) programmes de conversion testés et prêts à être exécutés,
 - c) plan d'exécution des conversions,
 - d) programmes d'interface testés et prêts à être utilisés,
 - e) définition des niveaux de sécurité par famille d'utilisateurs,
 - f) acceptation formelle des livrables par le comité de pilotage,
 - g) plan et encyclopédie de projet mis à jour.
9. la documentation des utilisateurs (Tomas, 2000 p. 225),
- a) procédures générales de l'entreprise mises à jour,
 - b) documentation de base fournie par l'éditeur,
 - c) procédures opérationnelles normales,
 - d) procédures opérationnelles exceptionnelles,
 - e) description des écrans, rapports, documents, options, paramètres, etc.,
 - f) explication des messages d'erreur et des actions correspondantes,
 - g) matériels de base utilisés par la suite de la formation des utilisateurs,
 - h) documentation technique finalisée (sauvegardes, restaurations, etc.),

- i) acceptation formelle des livrables par le comité de pilotage,
 - j) plan et encyclopédie de projet mis à jour.
10. la formation des utilisateurs (Tomas, 2000 p. 230),
- a) le plan de formation,
 - b) le matériel de formation,
 - c) l'environnement de formation,
 - d) les utilisateurs finaux formés,
 - e) acceptation formelle des livrables par le comité de pilotage,
 - f) plan et encyclopédie de projet mis à jour.
11. la mise en production (Tomas, 2000 p. 248),
- a) plate-forme de production correctement configurée,
 - b) plan des tests d'acceptation,
 - c) procédure de transfert des objets techniques entre environnements,
 - d) plan de restauration en cas d'accident majeur,
 - e) données de base et d'activité converties dans le nouvel environnement,
 - f) ERP opérationnel ainsi que ses systèmes périphériques,
 - g) révision de la stratégie future des solutions systèmes,
 - h) acceptation formelle des livrables par le comité de pilotage,
 - i) plan et encyclopédie de projet mis à jour.
12. le déploiement (Tomas, 2000 p. 254).
- a) stratégie et approche de déploiement,
 - b) ERP préconfiguré quelle que soit l'approche empruntée,
 - c) installation des sites et des secteurs d'activité de la compagnie.

Annexe 2 : Exemple sur les réseaux bayésiens

Dans cette section nous partons d'un exemple contenant les variables X1, X2 et X3 (Figure 74). Nous utilisons une base contenant des observations sur ces variables (trois.df dans Figure 75). Nous lançons l'exécution d'une recherche gloutonne (Figure 75) et nous observons que le meilleur réseau est celui contenant le seul arc X2->X3 (Figure 76). Cela veut dire que l'algorithme n'a effectué qu'une seule itération. Au cours de cette itération, nous passons du réseau sans arc au réseau contenant l'arc X2->X3. Dans ce cas, les paramètres et les contributions au score des nœuds X1 et X2 ne changent pas car leurs parents n'ont pas changé. C'est seulement pour le nœud X3 que les paramètres et le score sont recalculés. Cette contribution au score du nœud X3 est additionnée à celles des deux autres nœuds pour obtenir le score du nouveau réseau. Nous allons montrer donc le calcul des distributions de probabilités, le calcul des paramètres à priori, le calcul du score et la mise à jour des paramètres pour ce nœud X3.

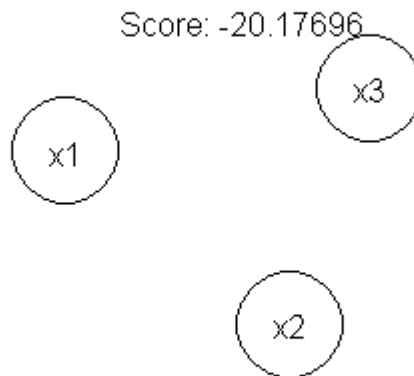


Figure 74 Réseau de départ exemple

```

R RGui
Fichier Edition Misc Packages Fenêtres Aide
[Icons]
[Icons]
> trois.df <- read.table("C:\\EXPERIENCES\\trois.txt", header=TRUE)
> trois <- network (trois.df)
> plot (trois)
>
> trois.prior <- jointprior (trois, 12)
>
> trois <- drawnetwork (trois, trois.df, trois.prior)$nw
>
> trois.s <- autosearch (trois, trois.df, trois.prior)$nw
[Autosearch (1) -19.97769 [x1][x2][x3|x2]
Total 0.08 add 0.05 rem 0.02 turn 0 sort 0 choose 0.01 rest 0 ]
> df
      x1    x2    x3
1 -0.78 -1.55  0.11
2  0.18 -3.04 -2.35
3  1.87  1.04  0.48
4 -0.42  0.27 -0.68
> trois
## 3 ( 0 discrete+ 3 ) nodes;score= -20.1769587516999 ;relscore=
1      x1      continuous()
2      x2      continuous()
3      x3      continuous()
> █
R 2.3.1 - A Language and Environment

```

Figure 75 Greedy search exemple

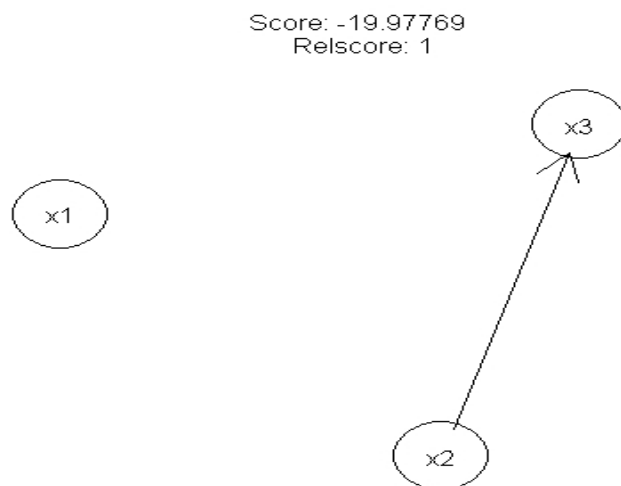


Figure 76 Réseau final exemple

Pour le calcul du score d'un nœud, il faut d'abord définir le nœud en question γ (ici $\gamma = x3$). Ensuite il faut déterminer les différentes configurations des parents discrets du nœud. En effet pour obtenir le score d'un nœud continu (Équation 49), un même terme est calculé pour chaque configuration des parents discrets $i_{pa(\gamma)} \in \mathbf{I}_{pa(\gamma)}$. Vu que notre réseau ne comprend pas de nœud discret donc γ ne peut pas avoir de parents discrets. Par conséquent l'ensemble des configurations est réduite à une seule valeur ($\text{card}(\mathbf{I}_{pa(\gamma)}) = 1$) qui est la première configuration⁸⁵ ($i_{pa(\gamma)} = [1]$)

⁸⁵ Représente l'absence de parents discrets

Paramètres joints

Nous devons commencer par fixer la taille de la base de données imaginaire N et la donner en entrée à la fonction jointprior qui calcule les paramètres joints. En cas d'omission, Deal affecte une valeur raisonnablement faible à N . Dans notre cas nous affectons la valeur 12^{86} . Sachant qu'il n'y a qu'une seule configuration de parents discrets [1] et que la somme des α_i est égale à N , l'alphajoint de cette configuration [1] est égale à N . Cela nous permet de déterminer en même temps nujoint [1] et rhojoint [1]. Le mujoint [1] est obtenue en prenant les moyennes des trois variables.

$$12 = N = \sum_{i \in I} \alpha_i = \alpha_{[1]}$$

$$v_{[1]} = \rho_{[1]} = \alpha_{[1]} = 12$$

$$\bar{\mu}_{[1]} = (\bar{\mu}_\gamma, \bar{\mu}_{pa(\gamma)}) = (0.2125, -0.82, -0.61)$$

Pour calculer le sigmajoint [1], nous devons d'abord calculer les B_j et les U_i pour les trois nœuds. Dans le réseau de départ nous avons $n=3$ et nous n'avons aucun arc donc les b_{ij} sont nuls. Nous obtenons ainsi sigmajoint [1] en utilisant les U_i et les variances v_i (Équation 38). Nous déduisons ensuite phijoint[1] de ce sigmajoint [1] (Équation 41). Nous obtenons ainsi la totalité de nos paramètres joints pour la configuration [1] des parents discrets (Figure 77).

$$v_1 = 1.033369 \quad v_2 = 2.52725 \quad v_3 = 1.18475$$

$$b_{13}=0 \quad b_{23}=0$$

$$B_{.3} = (b_{13}, b_{23})^T = (0, 0)^T \quad B_{.2} = (0, 0)^T \quad B_{.1} = (0, 0)^T$$

Et

$$U_3 = \begin{pmatrix} 1 & 0 & b_{13} \\ 0 & 1 & b_{23} \\ 0 & 0 & 1 \end{pmatrix} \quad U_3 = U_2 = U_1 = I_3$$

⁸⁶ Valeur habituellement donnée dans Bootcher, et al. (2003A) et Bootcher, et al. (2003B)

$$\Sigma_{[1]} = I_3^T I_3^T I_3^T \text{diag}(v_1, v_2, v_3) I_3 I_3 I_3 = \begin{pmatrix} v_1 & 0 & 0 \\ 0 & v_2 & 0 \\ 0 & 0 & v_3 \end{pmatrix}$$

$$\tilde{\phi}_{[1]} = (12 - 1)\Sigma_{[1]} = \begin{pmatrix} 11 \times v_1 & 0 & 0 \\ 0 & 11 \times v_2 & 0 \\ 0 & 0 & 11 \times v_3 \end{pmatrix}$$

```

RGui
Fichier Edition Misc Packages Fenêtres Aide
[Icons]
NULL
> trois.prior$jointnu
[1] 12
> trois.prior$jointmu
      x1    x2    x3
[1,] 0.2125 -0.82 -0.61
> trois.prior$jointrho
[1] 12
> trois.prior$jointsigma
[[1]]
      x1    x2    x3
x1 1.033369 0.00000 0.00000
x2 0.000000 2.52725 0.00000
x3 0.000000 0.00000 1.18475
> trois.prior$jointphi
[[1]]
      x1    x2    x3
x1 11.36706 0.00000 0.00000
x2 0.00000 27.79975 0.00000
x3 0.00000 0.00000 13.03225
R 2.3.1 - A Language and Environment

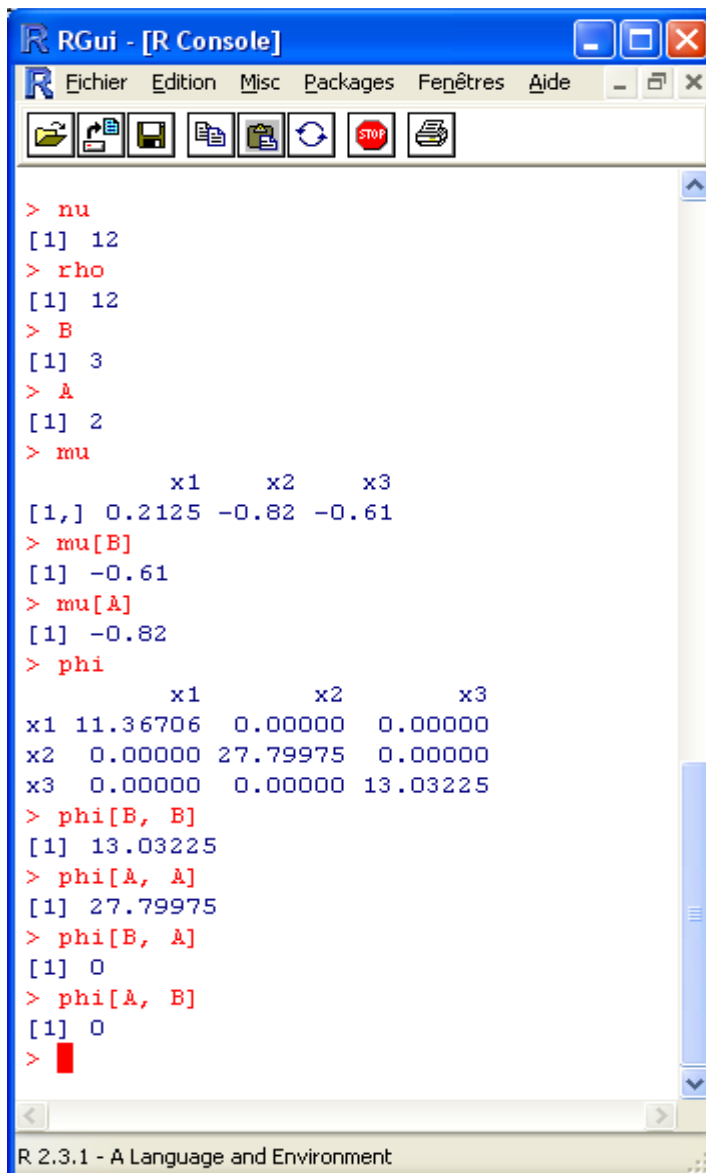
```

Figure 77 paramètres joints pour la configuration [1] des parents discrets

Paramètres locaux à priori

Pour calculer les paramètres à priori, nous devons d'abord déterminer les entrées (Figure 78) de leurs formules. Le nœud X3 a comme parent X2, donc $pa(\gamma) = (X2)$; ce qui donne une taille de ce vecteur $|pa(\gamma)| = 1$. Parmi ces entrées figurent le rhojoint [1] et le nujoint [1] qui ont été calculés dans la section précédente. $\bar{\mu}_\gamma$ et $\bar{\mu}_{pa(\gamma)}$ sont tirés de la décomposition du vecteur mujoint [1] calculé lui aussi. Enfin, de la décomposition de phijoint [1] nous obtenons $\tilde{\Phi}_\gamma$, $\tilde{\Phi}_{pa(\gamma)}$, $\tilde{\Phi}_{\gamma,pa(\gamma)}$ et $\tilde{\Phi}_{pa(\gamma),\gamma}$. Ces décompositions sont réalisées grâce à l'indice B du nœud et à l'indice A de son parent. Les paramètres à priori sont ensuite calculés par la fonction « conditional.cont » avec les lignes de codes (Figure 79) correspondants à leurs formules (Équation 45, Équation 46, Équation 47 et Équation 48)

$$\begin{aligned} \rho_{[1]} &= 12 \text{ et } \nu_{[1]} = 12 \\ \bar{\mu}_\gamma &= \text{mu}[B] \quad \text{et} \quad \bar{\mu}_{pa(\gamma)} = \text{mu}[A] \\ \tilde{\Phi}_\gamma &= \text{phi}[B, B] & \tilde{\Phi}_{pa(\gamma)} &= \text{phi}[A, A] \\ \tilde{\Phi}_{\gamma,pa(\gamma)} &= \text{phi}[B, A] & \tilde{\Phi}_{pa(\gamma),\gamma} &= \text{phi}[A, B] \end{aligned}$$



```
> nu
[1] 12
> rho
[1] 12
> B
[1] 3
> A
[1] 2
> mu
      x1      x2      x3
[1,] 0.2125 -0.82 -0.61
> mu[B]
[1] -0.61
> mu[A]
[1] -0.82
> phi
      x1      x2      x3
x1 11.36706 0.00000 0.00000
x2 0.00000 27.79975 0.00000
x3 0.00000 0.00000 13.03225
> phi[B, B]
[1] 13.03225
> phi[A, A]
[1] 27.79975
> phi[B, A]
[1] 0
> phi[A, B]
[1] 0
> █
```

R 2.3.1 - A Language and Environment

Figure 78 Entrées des paramètres à priori

```

RGui - [R Console]
Fichier Edition Misc Packages Fenêtres Aide

[1] 0
> if (length(A) < 1)
+   A <- TRUE
> rho.B1A <- rho + length(A)
> phi.AA.inv <- solve(phi[A, A])
> phi.tmp <- phi[B, A] %**% phi.AA.inv
> phi.B1A <- phi[B, B] - phi.tmp %**% phi[A, B]
> mu.B1A <- c(mu[B] - phi.tmp %**% mu[A], phi.tmp)
> tau.B1A.inv.11 <- 1/nu + t(mu[A]) %**% phi.AA.inv %**% mu[A]
> tau.B1A.inv.22 <- phi.AA.inv
> tau.B1A.inv.12 <- -t(mu[A] %**% phi.AA.inv)
> tau.inv <- rbind(cbind(tau.B1A.inv.11, t(tau.B1A.inv.12)),
+   cbind(tau.B1A.inv.12, tau.B1A.inv.22))
> tau <- solve(tau.inv)
> tau
      [,1]      [,2]
[1,] 12.00 -9.84000
[2,] -9.84 35.86855
> phi.B1A
      [,1]
[1,] 13.03225
> mu.B1A
[1] -0.61  0.00
> rho.B1A
[1] 13
> █

R 2.3.1 - A Language and Environment

```

Figure 79 paramètres à priori

Calcul du score et des paramètres à postériori

Le calcul du score et des paramètres à postériori (Figure 81) se fait dans la même fonction deal « postc » avec les mêmes arguments (Figure 80) : les paramètres à priori (μ , τ , ρ , ϕ), y et z . Les lignes de codes allant jusqu'à la ligne « loglik<- loglik + mscore » (Figure 81) correspondent à la formule de calcul du score du nœud X3 (Équation 49). Les lignes restantes permettent de réaliser la mise à jour des paramètres en appliquant les formules de calcul des paramètres à postériori (Équation 50, Équation 51, Équation 52 et Équation 53). Nous pouvons observer que la somme des logarithmes des scores des nœuds X1, X2 et X3 donne le logarithme du score du réseau (Figure 82).

```

RGui - [R Console]
Fichier Edition Misc Packages Fenêtres Aide
[Icons]

> node<- nodes(trois.s)$x3
> nw<-trois.s
> df<-trois.df
> mu<-node$condprior[[1]]$mu
> tau<-node$condprior[[1]]$tau
> rho<-node$condprior[[1]]$rho
> phi<-node$condprior[[1]]$phi
> y<-df[, node$idx]
> z<-cbind(1, df[, cparents])
> nw
## 3 ( 0 discrete+ 3 ) nodes;score= -19.977692186553$
1      x1      continuous()
2      x2      continuous()
3      x3      continuous()      2
> df
      x1      x2      x3
1 -0.78 -1.55  0.11
2  0.18 -3.04 -2.35
3  1.87  1.04  0.48
4 -0.42  0.27 -0.68
> mu
[1] -0.61  0.00
> tau
      [,1]      [,2]
[1,] 12.00 -9.84000
[2,] -9.84 35.86855
> rho
[1] 13
> phi
      [,1]
[1,] 13.03225
> y
[1] 0.11 -2.35  0.48 -0.68
> z
      [,1] [,2]
[1,]  1 -1.55
[2,]  1 -3.04
[3,]  1  1.04
[4,]  1  0.27

```

R 2.3.1 - A Language and Environment

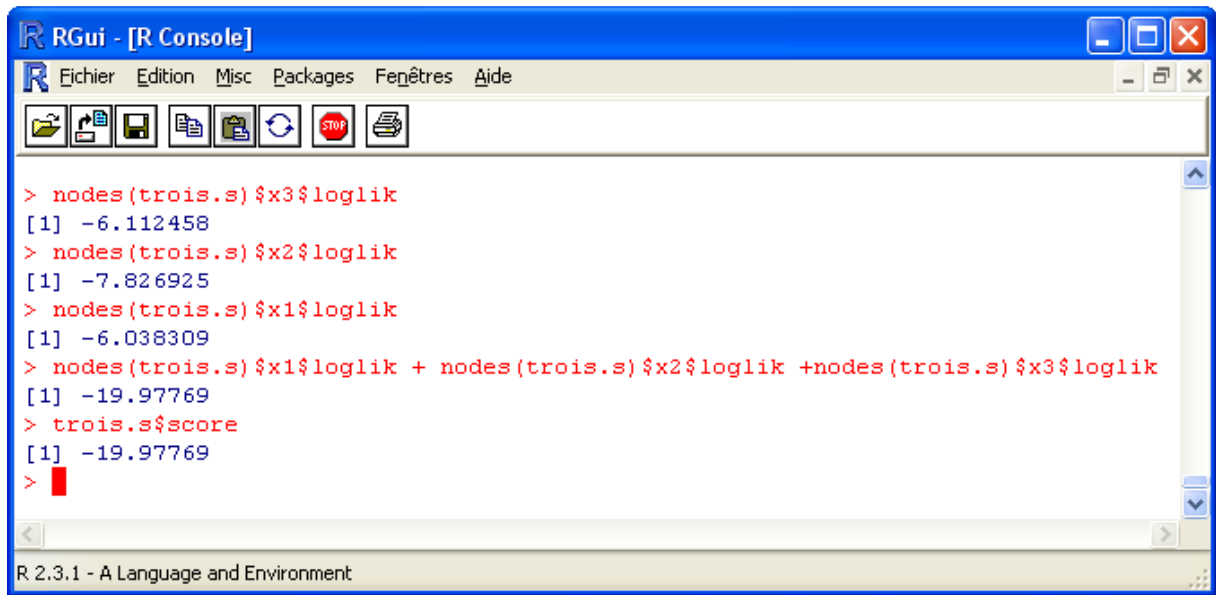
Figure 80 Entrées calcul du score et des paramètres à postériori


```

RGui - [R Console]
Fichier Edition Misc Packages Fenêtres Aide
[2,] 1 -3.04
[3,] 1 1.04
[4,] 1 0.27
> loglik <- 0
> for (i in 1:length(y)) {
+   logscale <- log(phi) + log(1 + t(z[i, ])** solve(tau)**
+     z[i, ])
+   logk <- lgamma((rho + 1)/2) - lgamma(rho/2) - 0.5 * (logscale +
+     log(pi))
+   mscore <- logk - 0.5 * (rho + 1) * log(1 + ((y[i] - z[i,
+     ])** mu)^2)/exp(logscale))
+   loglik <- loglik + mscore
+   oldtau <- tau
+   oldmu <- mu
+   tau <- tau + z[i, ]** t(z[i, ])
+   mu <- solve(tau)** (oldtau** mu + z[i, ] * y[i])
+   rho <- rho + 1
+   phi <- phi + (y[i] - t(z[i, ])** mu) * y[i] + t(oldmu -
+     mu)** oldtau** oldmu
+ }
>
> loglik
      [,1]
[1,] -6.112458
> nodes(trois.s)$x3$loglik
[1] -6.112458
> node$condposterior[[1]]$mu
[1] -0.4956094 0.1395008
> node$condposterior[[1]]$tau
      [,1]      [,2]
[1,] 16.00 -13.12000
[2,] -13.12 48.66715
> node$condposterior[[1]]$rho
[1] 17
> node$condposterior[[1]]$phi
[1] 17.03353
> █
R 2.3.1 - A Language and Environment

```

Figure 81 Calcul du score et des paramètres à postériori



The image shows a screenshot of the R GUI console window. The window title is "RGui - [R Console]". The menu bar includes "Fichier", "Edition", "Misc", "Packages", "Fenêtres", and "Aide". The toolbar contains icons for file operations and execution. The console displays the following R code and its output:

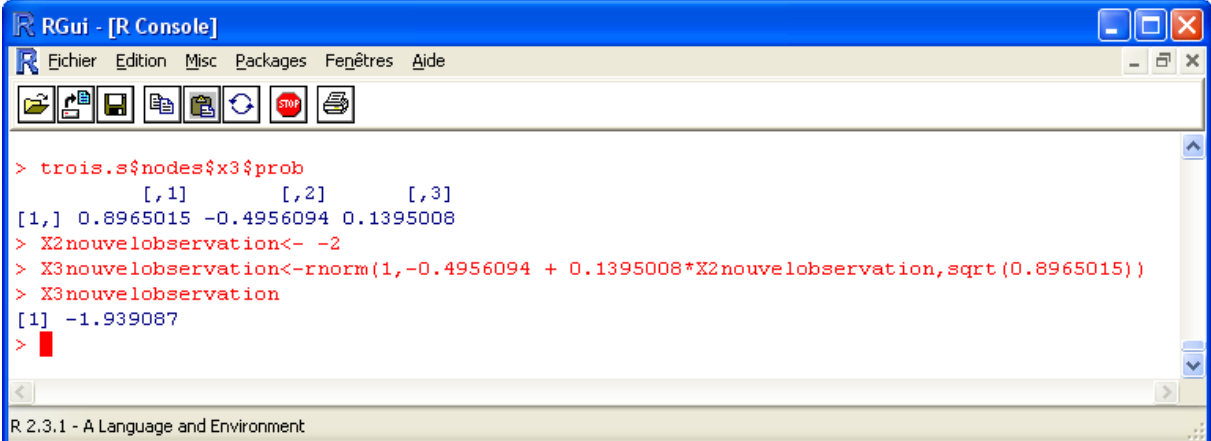
```
> nodes(trois.s)$x3$loglik
[1] -6.112458
> nodes(trois.s)$x2$loglik
[1] -7.826925
> nodes(trois.s)$x1$loglik
[1] -6.038309
> nodes(trois.s)$x1$loglik + nodes(trois.s)$x2$loglik + nodes(trois.s)$x3$loglik
[1] -19.97769
> trois.s$score
[1] -19.97769
> █
```

The status bar at the bottom of the window reads "R 2.3.1 - A Language and Environment".

Figure 82 Somme des log des scores des trois nœuds

Equation de prédiction

La distribution de probabilité est recalculée sur chaque nœud par rapport à la nouvelle configuration de ses parents dans le réseau obtenu à la fin de la recherche gloutonne (Figure 76). La propriété `prob` du nœud à prédire `X3` permet d'afficher ses paramètres de probabilités (Figure 83). Sur la colonne [1], nous avons la variance conditionnelle du nœud `X3`. Sur la colonne [2] nous avons son « *ordonnée à l'origine* ». A la colonne [3], nous avons le coefficient de régression correspondant à son nœud parent continu `X2`. La valeur de la variable `X3` peut être prédite pour une nouvelle observation si nous possédons sa valeur de `X2` (Figure 83). Nous utilisons pour cette prédiction la fonction « `rnorm` » (Figure 83) qui applique la formule de l'équation de prédiction (Équation 54). Cette fonction prend comme argument le nombre de valeurs souhaitées (ici 1), la moyenne conditionnelle et l'écart type qui est la racine carrée (`sqrt`) de la variance du nœud (Figure 83).



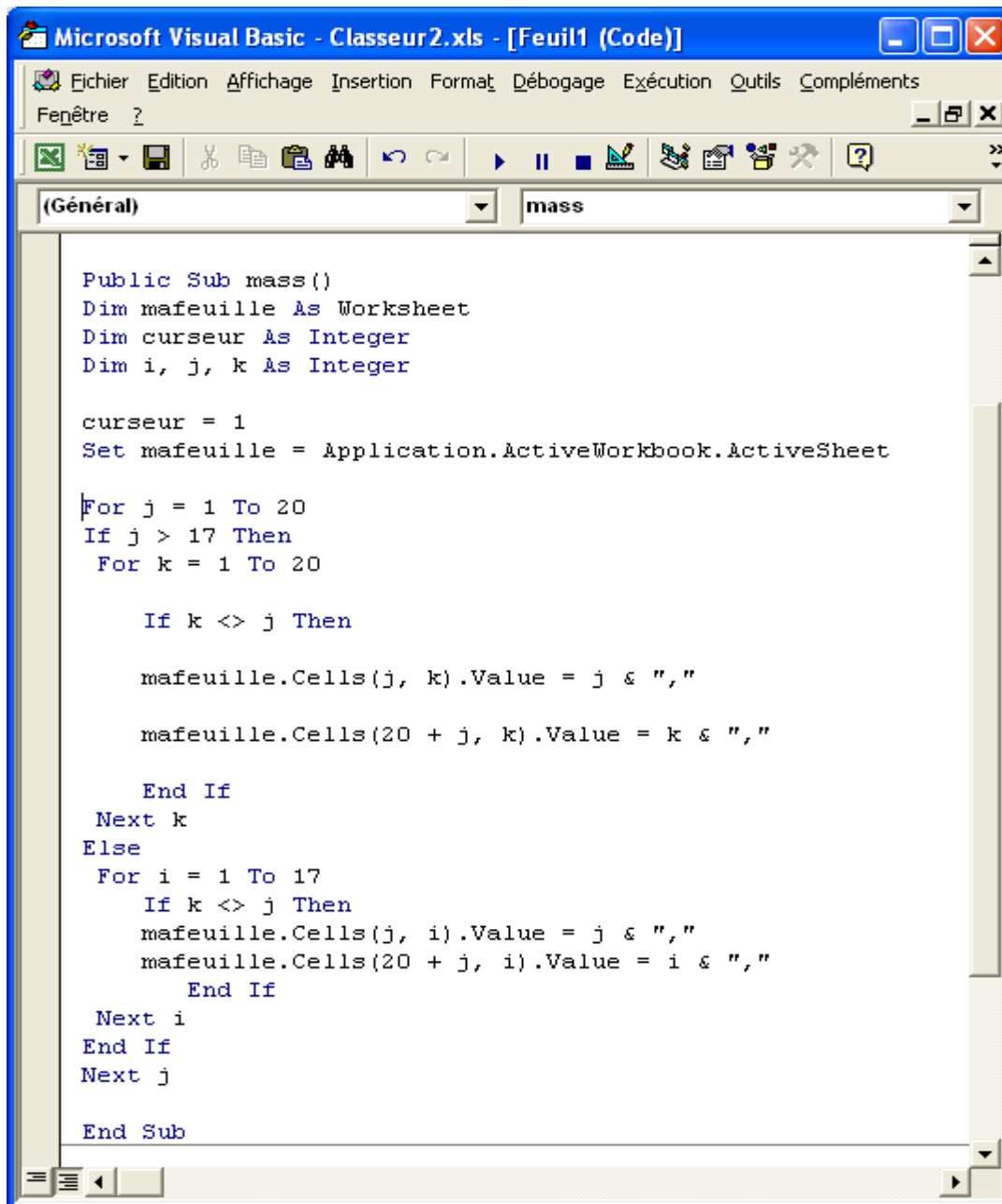
```
RGui - [R Console]
Fichier Edition Misc Packages Fenêtres Aide

> trois.s$nodes$x3$prob
      [,1]      [,2]      [,3]
[1,] 0.8965015 -0.4956094 0.1395008
> X2nouvelobservation<- -2
> X3nouvelobservation<-rnorm(1,-0.4956094 + 0.1395008*X2nouvelobservation,sqrt(0.8965015))
> X3nouvelobservation
[1] -1.939087
>
```

R 2.3.1 - A Language and Environment

Figure 83 Prédiction avec la distribution de probabilités mise à jour

Annexe 3 : Programme banlist



The image shows a screenshot of the Microsoft Visual Basic editor window. The title bar reads "Microsoft Visual Basic - Classeur2.xls - [Feuil1 (Code)]". The menu bar includes "Fichier", "Edition", "Affichage", "Insertion", "Format", "Débogage", "Exécution", "Outils", and "Compléments". The toolbar contains various icons for file operations, editing, and execution. The main area shows a VBA subroutine named "mass" with the following code:

```
Public Sub mass()  
Dim mafeuille As Worksheet  
Dim curseur As Integer  
Dim i, j, k As Integer  
  
curseur = 1  
Set mafeuille = Application.ActiveWorkbook.ActiveSheet  
  
For j = 1 To 20  
If j > 17 Then  
For k = 1 To 20  
  
If k <> j Then  
  
mafeuille.Cells(j, k).Value = j & ","  
  
mafeuille.Cells(20 + j, k).Value = k & ","  
  
End If  
Next k  
Else  
For i = 1 To 17  
If k <> j Then  
mafeuille.Cells(j, i).Value = j & ","  
mafeuille.Cells(20 + j, i).Value = i & ","  
End If  
Next i  
End If  
Next j  
  
End Sub
```

Résumé :

La problématique de cette étude est la gestion du changement dans les projets d'implémentation de progiciel. L'approche proposée est la gestion proactive du changement. Elle est basée sur des actions à priori pour réduire l'impact négatif du changement. Ces actions sont optimisées et ciblées sur certains processus de l'entreprise identifiés par prédictions. Notre objectif est de combiner la gestion quantitative et non-quantitative du changement dans les projets ERP afin de lever leurs limites. En effet, la littérature montre que les méthodes non-quantitatives proposent des actions de gestions du changement, sans pour autant accompagner ces actions d'une optimisation ou d'un ciblage formalisé. Les techniques quantitatives, quant à elles, proposent des modèles mais ne définissent pas des actions claires.

La démarche proposée repose sur l'application du processus de datamining. Nous nous sommes intéressés à trois concepts principaux qui sont l'ampleur, l'effort de gestion et l'impact du changement. L'idée est de tirer parti des informations collectées dans des projets d'implémentation de progiciel antérieurs, pour établir des relations quantitatives et objectives entre ces trois concepts. La première composante de l'application du datamining est l'utilisation d'un modèle pour faire des prédictions sur l'impact potentiel du changement. Cet impact est mesuré de manière prédictive par rapport à la variation des critères de performance du processus après sa restructuration et son implémentation dans le progiciel. L'unité d'analyse dans l'étude est le processus métier. La seconde composante de l'application de processus de datamining est la prise de décision basée sur les prédictions du modèle. L'ampleur et l'effort de gestion du changement sont considérés comme des explications pour l'impact du changement. Ils sont utilisés aussi comme levier dans les actions à définir pour la gestion de cet impact.

Notre approche intervient principalement dans la phase d'adéquation et de configuration dans laquelle les processus sont restructurés. Elle vient compléter la simulation, le prototypage (qui inclut l'analyse des écarts) pour l'acceptation d'un processus « en-devenir ».

Mots clés : Réingénierie Organisationnelle, Changement Organisationnel, Progiciel de Gestion Intégré, Statistique Bayésienne, Gestion des Risques

Abstract:

The issue of this work is change management in enterprise resources planning projects. The proposed approach is called proactive change management. It is based on prior actions to reduce negative impact of change. These actions are optimized and directed on some business processes identified by prediction. Our goal is to combine quantitative and none-quantitative change management in ERP projects in order to overcome their limits. Indeed, the literature shows that none-quantitative methods have proposed change management actions, but do not give means for optimizing and targeting these actions.

The quantitative techniques used in ERP projects have proposed models but do not define clear change management actions. The proposed approach is based on the application of data mining process. We are interested by three main concepts : magnitude, management effort and impact of change. The idea is to capitalize on information from previous ERP project to establish quantitative and objective relations between the three concepts. The first component of the data mining process application is the use of the model for predictions on potential impact of change. The impact was evaluated predictive way against the variation on key performances indicators of business process, after reengineering and implementation of ERP. The unit of analysis of our study is business process. The second component of the data mining process application is decision making based on predictions given by the model. Magnitude and management effort of change are considered as explications of change impact. They are also used as lever in action to define the management of this impact.

Our approach intervenes principally in the business process reengineering step of ERP project where processes are reconfigured. It comes to complete simulation and prototyping (including gap analysis) for the acceptance of a "to-be" process.

Key words: Business Process Reengineering, Organizational Change, Enterprise Resource Planning, Bayesian Statistic, Risk Management